SAME TALKER, DIFFERENT LANGUAGE

Verna Stockmal, Danny R. Moates, Z. S. Bond

Ohio University, Athens, Ohio, USA

ABSTRACT

Listeners can easily say whether a language they are hearing is familiar or foreign to them. Infants, young children, and adults are able to make same-language, different-language judgments at better than chance levels. In many of these studies, foreign language samples have been provided by different talkers so that language and talker characteristics have been confounded. We conducted three experiments using the same talker for different pairs of language. Listeners were able to discriminate between two languages they do not know even when spoken by the same talker, suggesting that listeners can distinguish talker characteristics from language characteristics.

1. INTRODUCTION

Experiments examining human performance on language identification tasks, both to provide reference for automatic language identification and to examine language acquisition in multilingual environments, have shown that listeners can discriminate languages they do not know (1, 3, 4, 5, 6, 8, 9, 14, 15). Presumably, listeners rely on the unique acoustic-phonetic properties of each language, its acoustic signature (8), to make discrimination judgments. In some studies (5, 9, 15), listeners reported identifying languages based on voice quality (9) and had difficulty with voice interference in the language decision process (15). Pisoni (12) suggests that talker-specific attributes become an integral part of the perceptual record. Both voice repetition and talker gender have been shown to affect voice recognition in English (11). Arguably voice would be more influential in responding to foreign languages. Foreign language samples provide only phonetic information to listeners. Mullenix (7) suggests that phonetic representations are less detailed than talker voice representations, so investigating foreign language discrimination un-confounded by talker voice quality should address the question: When talker characteristics remain constant, can listeners discriminate between foreign languages?

Three experiments analyzed subject judgments of same-language or different-language pairs, spoken by the same talker. The focus of the first two experiments was on discrimination of language pairs. In Experiment 1, subjects judged language pairs produced by either a male or female talker. In Experiment 2, both genders provided samples to be judged by the listeners. These studies tested listeners' abilities to discriminate between unfamiliar languages unconfounded by talker characteristics. The aim of Experiment 3 was to investigate the salient perceptual categories of judgments about foreign languages.

2. GENERAL METHOD

2.1. Participants

Undergraduate students from linguistics, psychology, and hearing and speech sciences participated in these studies. All were native speakers of American English and had no known deficiency in speech, language, or hearing. No participant had fluency in any of the languages. Sixty-two students participated in Experiment 1, 69 in Experiment 2, and 52 in Experiment 3. No participant served in more than one experiment.

2.2. Materials

Four male and four female bilingual talkers, with native-like proficiency in both languages, recorded passages in each language (see Table 1). The stimulus tape was created by excerpting fluent five-second phrases or sentences from the recorded passages. These excerpts were paired with either different languages recorded in both A,B and B,A orders or same languages. For example, the Japanese-Korean pairs were (a) Japanese-Korean, (b) Korean-Japanese, (c) Japanese-Japanese, and (d) Korean-Korean. Each pair was spoken by the same talker. Phrases were separated by a short beep. Each tape contained a total of 32 pairs; no language samples were duplicates.

Male Talkers	Female Talkers
Arabic-French (Morocco)	Korean-Japanese (Asia)
Hebrew-German (Israel)	Ombawa-French (Cameroon)
Akan-Swahili (Africa)	Latvian-Russian (Europe)
Latvian-Russian (Europe)	Ilocano-Tagalog(Phillip.)

Table 1: Language pairs

For Experiment 1, two stimulus tapes were created containing 32 tokens each, one produced by the male talkers and the other by the female talkers. For Experiments 2 and 3, a test recording was created from the speech samples of the two males and the two females receiving the highest A' scores in Experiment 1. The language pairs for the male talkers were Arabic-French and Hebrew-German and for the female talkers were Korean-Japanese and Ombawa-French. Two different randomizations of this recording were used in Experiment 3.

3. EXPERIMENT 1

Experiment 1 tested listeners' ability to discriminate spoken samples of foreign languages produced by the same talker.

3.1. Procedure

Participants were instructed to listen to the stimulus recording and to judge whether the tokens were in the same language or in different languages. They were not told what the languages were. The 32 experimental pairs were then played on a cassette tape player. Participants checked "same" or "different" on a response sheet for each pair.

3.2. Results

A' was calculated for each listener. A' is a nonparametric measure of sensitivity and is analogous to the d' measure used in signal detection theory. The value of A' ranges from 0 to 1 with 0.5 representing chance performance. Correct "different" judgments were counted as "hits" and same-language pairs judged "different" were counted as "false alarms."

The mean A' value for all male talkers was 0.77, SD=.09 and the mean A' value for all female talkers was 0.71, SD=.11. These values were significantly different from chance for both male talkers, $t(27)=15.88$, $p<.001$ and for female talkers, $t(33)=37.57$, $p<.001$

All different-language pairs for male talkers were discriminable whereas only the Japanese-Korean and French-Ombawa pairs for female talkers were clearly discriminable (see Figs. 1 and 2).

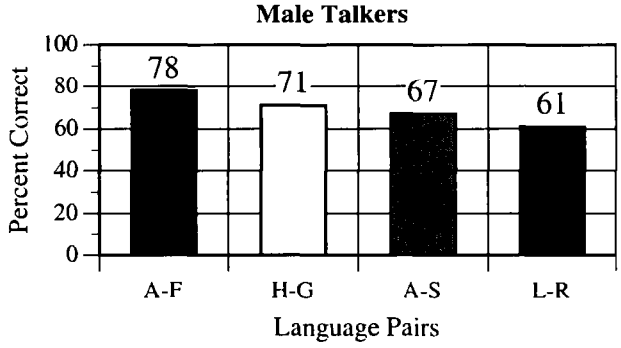

Figure 1: Percent correct judgments for the four language pairs produced by the male talkers.

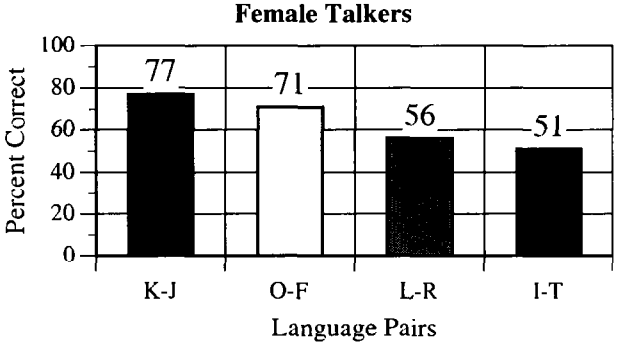

Figure 2: Percent correct judgments for the four language pairs produced by the female talkers.

3.3. Discussion

The linguistic knowledge of these listeners enabled them to discriminate between two languages they did not know even when the languages were spoken by the same talker, though the discriminability varied somewhat among language pairs.

4. EXPERIMENT 2

Experiment 2 was a modified replication of Experiment 1. The experiment assessed listener ability to discriminate between languages produced by both male and female talkers. The procedure was the same as used in Experiment 1.

4.1. Results

As in Experiment 1, A' was calculated for each participant. The mean value of A' for all talkers was 0.71, SD=.11. Thus, listeners were able to discriminate spoken samples of foreign languages even when produced by both male and female talkers, $t(68)=16.15$, $p<.001$.

Percent correct was also calculated for each different-language pair. All pairs were somewhat discriminable, though listener performance was not as accurate as in the first experiment, a difference of approximately 10% (see Fig. 3).

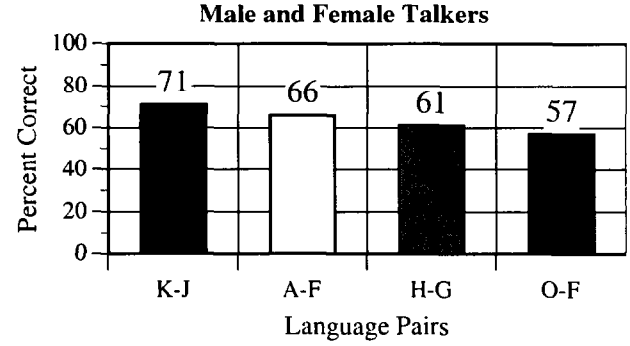

Figure 3: Percent correct judgments for the four language pairs produced by the male and female talkers.

4.2. Discussion

Experiment 2 showed the effect of Experiment 1 to be reliable. Listeners were able to use their knowledge of language to discriminate languages they did not know even when the languages were produced by the same talker. The effect held when individual listeners heard speech samples produced by both male and female talkers. Listeners apparently were able to detect distinctive acoustic patterns for the unknown languages and abstract away from "talker" to the category "language" on the basis of phonetic information present in the speech signal. Overall, listener scores were lower than in Experiment 1, suggesting that listening to male and female talkers was more difficult for the listeners.

5. EXPERIMENT 3

The first two experiments established that listeners can discriminate between foreign languages, relying solely on the properties of the languages. Experiment 3 used a multidimensional scaling task (MDS) of listener ratings to investigate the perceptual dimensions that participants were using to make these judgments.

5.1. Procedure

Participants were instructed to listen to the language pairs and rate the similarity of the languages in each pair using a 7-point scale (13, 16). On this scale a rating of 1 was given if the participant thought the two languages were very dissimilar and a rating of 7 if the languages were very similar.

Several practice trials using Latvian-English pairs were given until participants were comfortable with the procedure. Each participant then rated the 64 experimental language pairs.

5.2. Results

The mean rating for same-language samples was 5.19, SD=.51, significantly different from the different-language value of 3.45, SD=.57 ($t(51)=20.43$, $p<.001$.

An MDS analysis (13) was performed on the similarity ratings for both same- and different-language pairs (see Fig. 4). The best solution for the dissimilarity matrix for the eight languages was computed by ALSCAL as a two-dimensional Euclidian space. The stress value for this solution was 0.29, RSQ = 0.997.

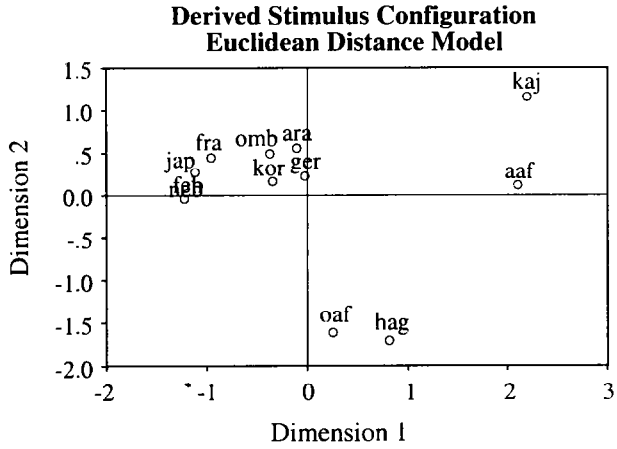

Figure 4: Two –dimensional solution for the multidimensional scaling analysis of judgments on the same and different language pairs.

5.3. Discussion

Neither talker gender nor voice emerged as a dimension in the MDS analysis. Dimension 1 can be interpreted as a same-different factor reflecting participants' perception of language similarity. All language samples clustered on the left are same-language pairs, and all samples on the right are different-language pairs. At the same time the dimension reveals that language familiarity and prosodic pattern influenced perceptual organization. Within the same-language configuration, familiar languages and unfamiliar languages clustered with each other. French, Japanese and Hebrew, somewhat familiar to our listeners, clustered on the left. Arabic and German clustered with the less familiar Korean and Ombawa on the right. Prosodic pattern is seen in the grouping of Arabic with German, as stress-timed, and Korean and Ombawa as syllable-timed. Although Hebrew is considered a stress-timed language and French a syllable-timed language, their proximity may result from final primary stress placement.

Dimension 2 is not as readily interpretable for the entire set of language pairs. For the different-language pairs, those most easily identified as different cluster in the upper half while those less easily identified as different cluster in the lower half of the map. Within this dimension, pairs containing two syllable-timed languages (French-Ombawa) or two stress-timed languages (Hebrew-German) were more difficult to discriminate than those pairs containing stress-timed vs. syllable-timed languages (Arabic-French) or those with highly distinctive pitch excursion (Japanese-Korean).

Participants must have used the same perceptual strategies for rating both types of language pairs. These strategies relate to the prosodic patterns but the salience of specific properties is still elusive. Perception of differences in rhythm may also depend on a language's segmental inventory and phonotactics (2).

6. GENERAL DISCUSSION

Voice quality and talker characteristics have been shown to be salient to listeners and may be encoded with phonetic information. When discriminating unknown foreign languages on the basis of phonetic cues, talker voice characteristics, including affective qualities such as rate, are incorporated into classification strategies (15). Experiments 1 and 2 showed that, even when voice characteristics are controlled, listeners were able to discriminate unknown languages at better than chance rates. Listeners are thus able to abstract away from "talker" to the category "language" on the basis of phonetic information present in the speech signal. The results of the MDS analysis (Experiment 3) suggest that listeners were employing prosodic information within the context of language familiarity.

Although early automatic language identification studies have had only marginal success using prosodic information (8), language acquisition studies indicate that infants rely on prosodic cues to discriminate languages (10). We find that adults also employ prosodic information. Future studies should attempt to fill in the details that listeners use to differentiate languages. The next step in this line of inquiry should explore the nature of the perceptual processes that let listeners distinguish languages they do not know.

7. REFERENCES

1. Bond, Z. and Fokes, J. "Identifying foreign languages". *Proc. XII Cong. Phon. Sciences, 1991, p 198-201.*

2. Dauer, R.M. "Stress-timing and syllable-timing reanalyzed", *J. Phonetics, Vol. 11, 1983, p 51-62.*

3. Fucci, D., McCall, D., Bond, Z, and Stockmal, V. "Perceptual dimensions underlying judgment of languages", *Perceptual and Motor Skills, 1997, p 1468-1470.*

4. Lorch, M, and Meara, P. "How people listen to languages they don't know", *Language Sciences, II(4), 1989, p 343-353.*

5. Lorch, M, and Meara, P. "Can people discriminate languages they don't know? *Language Sciences, XVII(1), 1995, p 65-71.*

6. Mehler, J., Dupoux, E., Nazzi, T., & Dehaene-Lambertz, G. "Coping with linguistic diversity: The infant's viewpoint." In J. L. Mogan and K. Demuth (Eds.), *Signal to Syntax*. L. E. A. Publ.: Mahwah, New Jersey, 1996.

7. Mullenix, J.W., Johnson, K.A., Topcu-Durgun, M., & Farnsworth, L.M. "The perceptual representation of voice gender", *J. Acoust. Soc. Amer., 84, 1995, p 3080-3095.*

8. Muthusamy, Y.K., Bernard, E., and Cole, R. "Automatic language identification: A review tutorial" *IEEE Magazine, 1996, p 1-16.*

9. Muthusamy, Y.K., Jain, N., and Cole, R. "Perceptual benchmarks for automatic language identification", *Proc. IEEE International Cong. on Acoustics, Speech and Signal Processing, 1994, p 1333-1336.*

10. Nazzi, T. "Du rythme dans l'acquisition et le traitement de la parole", Ph.D. dissertation, Ecole des Hautes Studes et Sciences Sociales, Paris, 1997.

11. Palmeri, T.J., Goldinger, S.D., and Pisoni, D.B. "Episodic encoding of voice attributes and recognition memory for spoken words" *J. Experimental Psychology: Learning, Memory, and Cognition, 19, 1993, p 309-328.*

12. Pisoni, D.B. "Effects of talker variability on speech perception: Implications for current research and theory", *Proc. Internat. Cong. on Spoken Language Processing, 1990(2), p 1399-1407*

13. Schiffman, S., Reynolds, M.L., and Young, F., *Introduction to Multidimensional Scaling.* Academic Press, Orlando, Florida, 1981.

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

15. Stockmal, V., Muljani, D., and Bond, Z. "Perceptual features of unknown foreign languages as revealed by multidimensional scaling." *Proc. Int. Cong. on Spoken Language Processing, 1996, p 1748-1751.*

16. Young, F.W., and Harris, D.F. *Multidimensional Scaling in SPSS Professional Statistics 7.5,* SPSS INC., Chicago, 1997.