Research Unit. Dept. of Psychiatry & Neurology, New York University School of Medicine
New York City (U.S.A.)
Dept. of Psychiatry, Yale University School of Medicine, New Haven, Conn. (U.S.A.)

Spectral Characteristics of the Voices of Twins[1]

M. Alpert, R. L. Kurtzberg, M. Pilot, A. J. Friedhoff[2]

In addition to the semantic content of a spoken message, extralinguistic information is transmitted by the voice. Vocal characteristics may distinguish the geographical origin, emotional state, and even anatomical structure of the speaker. Many investigators have attempted to identify the cues by which this information is coded. Evaluations of voice have been conducted primarily through the use of judges instructed to rate personality, emotional, or anatomical characteristics of a speaker solely on the basis of his speech. Although judges were often found to be in agreement with each other, their judgments were not necessarily valid ones. Sanford (13) discusses the ability of judges to agree among themselves, even when their common judgment is wrong. He believes that this results because of prejudices and sterotypes which the judges hold in common. Comprehensive reviews of this literature have been prepared by Sanford (13), by Moses (10), and more recently by Kramer (7).

Another approach to the study of extra-linguistic vocal characteristics has compared voice and anatomical and personality traits in twins. Luchsinger (8) studied fundamental frequency ranges in monozygotic and dizygotic twin pairs. He found monozygotic twin pairs to be more similar to each other in vocal range than were the dizygotic twin pairs. Gedda, Bianchi, and Neroni (5) had monozygotic (MZ) and dizygotic (DZ) twins attempt to distinguish their own voice from the voice of their co-twin. These workers reported that monozygotic twins had great difficulty in carrying out this task, while most dizygotic twins easily performed this discrimination. The findings of Gedda suggest that the voices of monozygotic twins may be more similar than are those of dizygotic twins.

As part of a further study, Gedda et al. (6) compared the vocal spectra of MZ, DZ, and a non-twin control group in a pre-pubescent population. They found that the MZ twin pairs seemed to have very similar spectra while twin pairs in the DZ group were less alike. They attributed the greater similarity in the voice of the MZ twins to structural similarities in the larynx of MZ twins.

[1] This research was supported by U. S. P. H. S. Research Career Development Award MH-K3-14,0,24 to Dr. Friedhoff and U. S. P. H. S. Grant MH-04664.

[2] The authors wish to express their appreciation to L. S. Kogan for his suggestions on the treatment of the data.

Ricevuto il 12. IX. 1963.

In contrast with Gedda, et al. (6), Ostwald, Freedman, and Kurtz (11) were unable to find an association between monozygosity and similarity of vocal spectra in a study of the spontaneous cries of infant twins. The difference in the findings of the two studies may be related to the difference in the age of the subjects. Gedda, et al., studied pre-adolescents and Ostwald, et al., studied infants. A second difference between the studies may also have contributed to the difference between the results. Gedda, et al., averaged relatively long samples of speech. An individual spectrum represented the integration of a large population of sounds. Ostwald, et al., analyzed single cries but could not ensure that his intratwin comparisons were for the same sounds. The spectrum of the voice depends on the location of the articulators as well as on the structure of the vocal tract. It may be that Ostwald would have obtained results similar to those of Gedda, et al., if he had used averaging techniques.

The present study was initiated to evaluate voice characteristics of twins and non-twins. Because of the relation between voice quality and the speech spectrum it was felt that spectral analysis might be a fruitful method for the investigation of similarities and differences in the voices of similar individuals. We hoped, further, to learn which sections of the spectrum are most determined by the anatomy of the vocal apparatus.

In earlier studies in this laboratory (1, 4), it was found that changes in voice energy in the region of the fundamental frequency accompanied changes in the emotional state of the speaker. The higher vocal frequencies, however, remained stable despite transient alterations in the emotional state of the speaker. The relative stability of higher portions of the speech spectrum may reflect the greater dependence of this portion of the spectrum on the anatomy of the larynx or the structure of the resonating cavities. Inasmuch as MZ twins have similar anatomy including the anatomy of the vocal tract, the higher portions of the speech spectrum would also be expected to be similar in this group. The lower vocal frequencies being less dependent on the characteristics of the resonating cavities and more sensitive to transitory stimuli would be less likely to reveal similarities between MZ twins. In the same way, fraternal twins, being less alike anatomically, would be expected to be somewhat less alike vocally as well.

Method

Six MZ and 6 same sex DZ twin pairs were interviewed in a quiet recording studio at the Yale University School of Medicine (9). The ages of the MZ twin pairs ranged from 16 to 50 years of age. The DZ twin pairs ranged from 17 to 33 years of age. There were 3 male and 3 female MZ pairs and 4 male and 2 female DZ twin pairs.

As part of a standard interview each twin was asked to say the 5 (popular) vowel sounds, " a ", " e ", " i ", " o ", and " u ". After a rehearsal the vowel series was recorded. The interviews were recorded through an Altec 21 B microphone onto an

Ampex Consul 300 tape recorder. After an initial adjustment during the rehearsal, the input gain was kept constant.

The signal was drawn from the preamplifier of the playback tape recorder (Tapesonic 70-C modified for full track operation), passed through a rectifier filter circuit (0.05 sec. time constant), and amplified and recorded on a direct-writing polygraph (Offner Type R). In this configuration the deflection of the polygraph pen is proportional to the sound pressure level of the voice of the subject. In order to measure the energy in specific portions of the speech spectrum a band-pass filter (Krohn-Hite 330-M) was inserted between the tape recorder and demodulator.

The analysis of the higher frequencies of the vocal sample of each twin was accomplished by recording the vocal energy at each of 4 spectral points. For each measurement the high pass and low pass sections were set to the same frequency. The four points studied were 670, 1000, 1500, and 2000 cycles per second. Thus, for each spectral point of each vowel, a measure of the energy in the voice was obtained. The obtained readings for each vowel were inserted into the following formula and thus converted into a proportional measure which is independent both of the overall speech levels, and of changes which might be occurring in the lower frequencies.

$$* \; I_{670} = \frac{V_{670}}{V_{670} + V_{1000} + V_{1500} + V_{2000}}$$

* I = Proportional spectral energy at the specific frequency (in this case 670 cps)
V = Voltage. Subscript number refers to frequency setting in cps.

Four filter settings were employed in the analysis of the lower frequencies. To obtain these values the high pass and low pass sections of the filter were set to the same value. Values used were 150, 200, 300 and 400 cps. Index values for each point were obtained as in formula 1 using the 4 lower frequency points.

The 24 individuals composing the 12 twin pairs were combined into 3 groups. One group consisted of the 6 MZ twin pairs. A second group consisted of the 6 DZ pairs. The third group was assembled by drawing 6 non-twin pairs of individuals from either the MZ or the DZ twin pools. For this group a pair consisted of individuals who were of the same sex and not more than one year different in age. In this way a comparison group of 6 non-twin pairs controlled for age and sex was assembled.

The proportional spectral values were computed according to formula (1) and the values obtained by the two members of a pair were inserted into the following formula to obtain a difference score for the pair at that spectral point (2).

$$D_{670} + (I_{670 \text{ for twin}} - I_{670 \text{ for co-twin}})^2$$

Differences for each twin pair were obtained similarly. Difference scores were obtained in this way for each spectral point for all five vowels. The results of the comparisons between the MZ and DZ twin groups and the non-twin comparison group on the higher and lower frequency analyses may be found in Fig. 1. Scores were squared to eliminate negative signs.

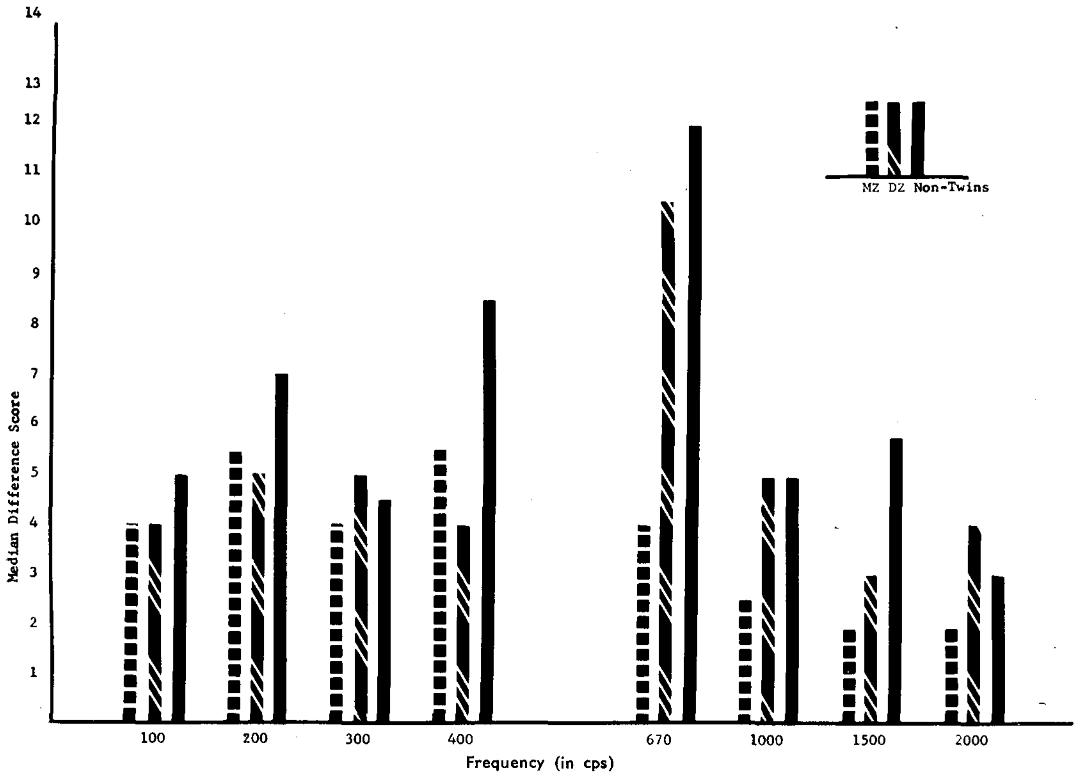

Fig. 1. Median difference scores of 6 monozygotic twin pairs, 6 dizigotic twin pairs, and 6 non-twin pairs saying vowel sounds "a", "e", "i", "o", and "u"

In Fig. 1 the median difference scores for the various spectral points are presented. The higher difference scores are associated with greater differences between the pairs. The non-twin group has the highest range of median difference scores while the MZ group has the lowest range — the highest score obtained by both the non-twin and DZ twin groups is at 670 cps. It is not immediately apparent why this spectral point should yield such relatively high difference scores. The results for the individual vowels show the same trends as the combined data presented in Fig. 1. The MZ twin group does not show increased difference scores at this point.

In Table 1 are presented the results of the statistical comparison of the differences between the groups, the data on which Fig. 1 is based. Non-parametric comparisons (Mann-Whitney) and probability levels were computed following Siegel (10). The probability level reflects the confidence with which we may reject the hypothesis of no difference between the groups, that is no greater similarity in the voices of members of a pair in one group than in another. For the four lower frequency points no differences were found between the groups.

Tab. 1. Statistical comparison of the group data for the monozygotic (MZ), dizygotic (DZ), and non-twin (NT) groups. Mann Whitney tests were used

	Groups compared					
Frequency (cps)	MZ vs DZ		MZ vs NT		DZ vs NT	
	Z score	P<	Z score	P<	Z score	P<
a. Lower —						
150	0.21	.420	1.58	.060	1.23	.110
200	0.66	.260	0.42	.340	1.18	.120
300	0.85	.200	0.54	.300	0.30	.390
400	0.71	.240	0.87	.200	1.20	.120
b. Higher —						
670	3.24	.001	4.04	.001	0.07	.600
1000	2.59	.005	4.05	.001	2.03	.025
1500	1.66	.050	2.62	.005	0.91	.200
2000	1.56	.060	2.65	.005	0.71	.250

Analysis of the higher frequency points, however, revealed greater similarities in the voices of the MZ twins than either the DZ or non-twin pairs. At all four spectral points the MZ twins were found to more closely resemble their co-twins than did pairs in the other groups — only for the 1000 cps band was the voice of the DZ twins found to be more like his co-twin's than the pairs in the non-twin control group. For the remaining three higher frequency points no greater similarity of voice was found within DZ pairs than within non-twin pairs.

The results should be considered in relation to the method of analysis. The spectrum was separated into two segments, the lower frequencies and the higher frequencies, and difference scores were computed separately for each segment. Because the bulk of the energy is in the lower frequency segment it would be inappropriate to compare the magnitude of a higher frequency difference score with one from the lower frequencies. This may, in part, explain the large difference scores obtained by the DZ and non-twin groups at 670 cps. There is more energy at 670 cps than at the other three higher frequency points and differences between pair members would thus tend to be magnified. This characteristic of the data treatment will not affect the statistical comparisons, since the tests were always across groups but within spectral points.

The results support the primary hypothesis that the voices of MZ twins are more similar than DZ twins at higher spectral points but not at lower spectral points. Whether this reflects a greater indebtedness of the vocal spectrum to nature rather than nurture is, perhaps, moot. Gedda, et al. (6) felt that greater similarity in the spectra of MZ twins follows from greater similarity in the anatomy of the larynx. While the structure of the larynx must contribute importantly to the shape of the vocal spectrum, the structure of the resonating cavities is also significant, especially for the shorter wavelengths (the higer frequencies). For the lower frequencies the wavelength is long relative to th edimensions of the cavities so that the output of the vocal tract is relatively independent of the dimensions of the resonating cavities. The cavity characteristics, however, take on increasing importance in determining upper formant resonances (3). Inasmuch as the lower frequency analysis failed to show inter-twin similarities, this range would appear to be subject to less stable factors than anatomy.

In this laboratory we have shown that the emotional state of the speaker is one of the factors associated with energy shifts in the low frequencies (1, 4). Subjects tend to show characteristic differences in the low frequencies of their spectra when responding to emotional stimuli. In our experience, this section of the spectrum is most subject to change from one moment to the next. In line with this, it is of some interest that despite its rather high density of energy, this section of the spectrum is relatively unimportant insofar as the intelligibility of speech is concerned. The moment to moment changes in the lower frequencies may be related to speaker characteristics or may modify, in some way, the words being spoken.

Summary

A study is described in which the spectral similarities in the voices of monozygotic, dizygotic and non-twin controls is compared. The results are based on a statistical comparison of the differences between pairs of individuals at a number of spectral points. The data are interpreted as showing that monozygotic twins are more similar in the higher but not the lower spectral ranges. It is suggested that the lower frequencies are more sensitive to transitory factors while the higher frequencies are more likely to reflect anatomic similarity.

References

1. Alpert M., Kurtzberg R. L. & Friedhoff A. J.: Transient voice changes associated with emotional stimuli. AMA Arch. Gen. Psychiat., 8: 362-365, 1963.
2. von Bracken H.: Psychologische Untersuchen an Zwillingen. Ber. Kongr. dtsch. Ges. Psychol., Leipzig, 13, 117-119, 1934.
3. Fletcher H.: Speech & hearing in communication, Princeton, N. J.: Nostrand, 1953.
4. Friedhoff A. J., Alpert M. & Kurtzberg R. L.: An effect of emotion on voice. Nature, 193, 357-358, 1962.
5. Gedda L., Bianchi A., & Neroni L.: Voce dei gemelli. A. Ge. Me. Ge., 4, 121-130, 1955.
6. Gedda L., Fiori-Ratti L. et Bruno G.: La voix chez les jumeaux monozygotiques. Folia phoniat., 12: 81-94, 1960.

7. Kramer E.: The judgment of personal characteristics and emotions from nonverbal properties of speech. Psychol. Bull., 60, 408-420, 1963.
8. Luchsinger R.: Die Sprache und Stimme von ein- und zweieiigen Zwillingen in Beziehung zur Motorik und zum Erbcharakter. Arch. Klaus-Stift. Verarb-Forsch., 15, 461-527, 1940.
9. Mahl G. F., Dollard J. and Redlich F. C.: Facilities for the sound recording and observation of interviews. Sci., 120: 235-239, 1954.
10. Moses P. J.: The voice of neurosis, New York: Grune & Stratton, 1954.
11. Ostwald P. F., Freedman D. G. and Kurtz J. H.: Vocalization of infant twins. Folia phoniat., 14:37-50, 1962.
12. Pilot M. L., Lewkoski L. D., Spiro H. M. & Shafer R.: Duodenal ulcer in one of identical twins. Psychosomat Med., 19, 221-227, 1957.
13. Sanford F. H.: Speech and personality. Psychol. Bull., 39, 811-845, 1942.
14. Siegel S.: Nonparametric statistics, New York, McGraw Hill, 1956.

RIASSUNTO

Si riferisce su di un'indagine in cui vengono raffrontate le somiglianze spettrali nelle voci di gemelli monozigotici, dizigotici e controlli nongemelli. I risultati sono basati su di un raffronto statistico delle differenze fra coppie di individui a un dato numero di punti spettrali. Secondo gli Autori i dati dimostrano che i gemelli monozigotici sono più somiglianti nei gradi spettrali più alti, ma non nei più bassi. Si ritiene che le frequenze più basse siano più sensibili ai fattori transitori, e che le frequenze più alte riflettano più facilmente le somiglianze anatomiche.

RÉSUMÉ

L'on rapporte sur une investigation conduite sur les ressemblances spectrales dans les voix de jumeaux MZ et DZ et de contrôles non-jumeaux. Les résultats se basent sur une comparaison statistique des différences entre couples d'individus à un certain nombre de points spectraux. D'après les Auteurs les données démontrent que les jumeaux MZ sont plus ressemblants dans les degrés spectraux plus élevés, mais pas dans les plus bas. L'on déduit que les fréquences plus basses sont les plus sensibles aux facteurs transitoires, tandis que les fréquences plus élevées réflètent plus probablement les ressemblances anatomiques.

ZUSAMMENFASSUNG

Es wird über eine Untersuchung berichtet, bei der die Ähnlichkeit in den Stimmspektren von EZ, ZZ und Kontrollgeschwistern verglichen werden. Die Ergebnisse gründen sich auf einen statistischen Vergleich zwischen je zwei Menschen mit einer bestimmten Anzahl von Spektrumspunkten. Den Verf. gemäss beweisen die Ergebnisse, dass die EZ in den höheren, aber nicht in den niedrigeren Spektrumsstufen ähnlicher sind. Man nimmt an, dass die niedrigeren Frequenzen für vorübergehende Faktoren empfindlicher seien, während die höheren Frequenzen eher die anatomischen Ähnlichkeiten wiederspiegeln.