

AGE AND SEX RECOGNITION OF SPEAKERS

NATARAJA, N. P. AND KUSHAL RAJ, P.*

Studies have been conducted to identify speaker's age, Ptacek and Sander, (1966); Shipp and Hollien, (1969); Ryan and Burk, (1972):Burk, Hoyer, Fey and Charlip, (1975); Hartman and Danhauer, (1969); Sex(Schwartz, 1968); Schwartz and Rine (1968); Ingemann, (1968); Coleman, (1971: 1973a; 1976); Lass, Hughes, Bowyer, Waters and Bourne, (1976); Lass, Tecca, Mancuso and Black, (1979c); Lass, Mertz and Kimmel, (1978); Lass, Almerino Jordon and Walsh, (1980d) race (Stroud, 1956; Hibler, 1960; Dickens and Sewyer, 1962; Larson and Larson, 1966; Bryden, 1968; Alverenga, 1970; Abram, 1975; Lass *et al.*, 1979c; Lass, Mertz and Kimmel, 1978; Lass *et al.*, 1980d), Socio-economic status (Harms, 1961; 1963), personality (Stagner, 1936; Eisenbeig and Zalowitz, 1938; Markel, Eisler and Reese, 1967), specific identity (McGehee, 1937; Pollack, Pickett and Sumbly, 1954; Compton, 1963; Voiers, 1964; Clarke Becker and Nixon, 1966; Bricker and Pruzansky, 1966; Holmgeen, 1967; Stevens, Williams, Carbonell and Woods, 1968; Clarke and Becker, 1969; Coleman, 1973b) and facial features (Lass and Harvey, 1976). These studies have been considered to 'provide very useful information in a variety of future theoretical and applied areas of investigation' (Lass *et al.* 1980). For example, the information from such studies will be useful in training listeners in recognizing various characteristics of speakers.

As stated above studies have shown that it is possible to identify speaker's age and sex on the basis of perceptual cues obtained from recorded speech samples. Lass and his associates have conducted a series of experiments in this regard.

Allport (1963), summarizes his research with Cantril (1936), which corroborated studies by Pear (1931) and Herzog (1933), by stating that 'when subjects read aloud the same written passage, age can usually be told within ten years.

Schwartz (1968) and Ingemann (1968), employed isolated voiceless fricatives as auditory stimuli and found that listeners could accurately identify speaker's sex from these stimuli, especially from /h/, /s/ and /f/, since the laryngeal fundamental (fo) was not available to the listeners because of the voiceless condition of the consonants. These findings indicate that accurate sex identification is possible from vocal tract resonance information alone and therefore, that formants are important cues for speaker sex identification.

•Mr Kushal Raj, P. Student, AIISH, Mysore-6

Further, support for this conclusion has come from studies by Schwartz and Rine (1968) and Coleman (1971). In the Schwartz and Rine (1968) study, the ability of listeners to identify speaker sex from two whispered vowels (/i/ and /a/) was investigated. They found 100 per cent correct identification for /a/ and 95 per cent correct identification for /i/ despite the absence of the laryngeal fundamental. In Coleman's study (1971) on male and female voice quality and its relationship to vowel formant frequencies /i/, /u/, and /a/ prose passage was employed to explore listeners ability to identify the sex of the speaker. All stimuli were produced at the same vocal fundamental frequency (85 Hz) by means of an electrolarynx. Coleman (1971), discovered that listeners were capable of accurately recognizing the sex of the speaker, even when the fundamental frequency remained constant for all speakers.

In a later experiment, Coleman (1973), attempted to reduce the influence of possible differences in rate, juncture and inflection between male and female speakers by presenting their voiced productions of a prose backward to listeners. By means of rank order correlation coefficients computed between measures of vocal tract resonance, laryngeal fundamental, and judgement of degree of maleness and femaleness in the voices of the speakers (with degree of correlation indicative of the contribution of each of the vocal characteristics to listener judgements) he found that 'listeners were basing their judgements of the degree of maleness or femaleness in the voice on the frequency of the laryngeal fundamental'.

Coleman (1976), in another experiment has found that 'in natural speech the degree of male and female quality in the voice is a function of the frequency of the laryngeal fundamental and individual vocal tract resonance characteristics, whether male or female, contribute little or nothing to the perception of this particular vocal quality. When a laryngeal vibrator producing a sex specific F_0 is substituted for the normal glottal tone, however, the female F_0 is perceptually weaker than the male vocal tract resonance characteristics, while the male F_0 retains its perceptual prominence'.

Lass *et al.* (1976), compared accuracy of ability to identify sex of the speakers from voiced, whispered and 225 Hz low pass filtered isolated vowels. They found that listener's accuracy was greatest for the voiced stimuli and followed by the filtered stimuli and least accurate for the voiceless vowels. Since the low pass filtered vowels apparently had no formant information, they concluded that the laryngeal fundamental was a more important acoustic cue for speaker sex identification than the speaker's vocal tract resonance characteristics.

On the basis of a fundamental frequency analysis of reading performance, McGlone and Hollien (1963), concluded that there was no evidence of 'senile

voice' among the older women (65-94 years ages) whom they studied. Ptacek and Sander (1966) studied the age recognition from voice, they tried to find the ability F₀ 10 listeners to differentiate the voices of younger adults (under age 35) from older (over age 65) subjects on the basis of a prolonged vowel, a reading sample played backward and a reading sample played forward was tested. They concluded that the listeners were able to differentiate the voices of younger adults from aged speakers with impressive accuracy under each of the three successive listening conditions of decreasing difficulty.

However, no such reports are available regarding Indian population or regarding the ability of listeners (Indian) to identify the age and sex of the speakers. Hence it was considered that it would be interesting to investigate the listener ability to identify the age and sex of young children.

The purpose of the present study is to determine the relative importance of portions of the broad band frequency speech spectrum in speaker age and sex identification in an attempt to provide additional evidence on the specific acoustic cues responsible for identification tasks.

METHODOLOGY

Ten male and ten female children age ranging from 3 to 5 years studying in nursery school were considered as speakers for the present study. These children were normal with respect to speech and hearing. They were asked to say vowels /a/, /i/, /u/, /e/ and also to count numbers one to ten. These were recorded using a National Panasonic Tape Recorder with in-built microphone kept at a distance of 1 \ ft. approximately in a less noisy area of the school. The recordings were made with sex randomly distributed.

A cosmic stereo deck, an amplifier and two speakers which were kept at an angle of 120° and 10 ft. apart from each other were used to play the recordings.

Twenty normals in terms of Speech and Hearing were selected as listeners. The listeners, ten male and ten female in an age range of 18 to 24 years were seated comfortably at a distance of 10 ft. approximately from the speakers. They were instructed to listen to the recordings and to identify and note down the age and sex of the speaker.

Discussion and Results

The judgements given by the above listeners were analyzed. The following table shows the percentage of correct identification of sex and age. Further, it also indicates identification of age within ± 1 year.

Subject No.	Percentage of correct identification of sex	Percentage of correct identification of age	Percentage of correct identification of age within ± 1 year
1	90	25	75
2	85	25	70
3	65	35	80
4	60	10	30
5	45	55	90
6	40	30	65
7	50	25	75
8	50	35	90
9	95	25	65
10	90	30	35
11	100	30	65
12	100	45	75
13	55	25	80
14	70	25	80
15	45	25	70
16	45	25	70
17	70	40	75
18	75	30	75
19	100	30	65
20	50	45	85

The inspection of the table reveals that more than 50 per cent of the judges have identified the sex of thirteen (65 per cent) speakers correctly.

This indicates that it is possible to identify the sex even though not much difference is found in the fundamental frequency of voice and in vocal tract resonance characteristics (Samuel, 1973; Usha, 1979; Gopal, 1980) in the early age group as included in this study. The present investigation has shown results which corroborates with earlier studies.

Further, inspection of the table indicates that none of the speaker's age could be identified by the judges beyond chance level. However, when the data was analyzed to note the identification of the age of the speakers within ± 1 year it was found that age of 18 speakers (90 per cent of the speakers) was identified by 13 judges or more i.e., beyond chance level therefore it can be concluded that, it would be possible to identify the age nearly 90 per cent of the cases within ± 1 year from actual age.

In spite of lack of much difference in terms of fundamental frequency and vocal tract resonance as stated earlier, it is still possible to identify the age. No studies regarding the identification of age in children is available to the present investigators.

However, Ptacek and Sander (1966), have shown that, their listeners were able to identify the ages of young adults and older subjects accurately based on reading samples of these two groups. The present investigation has shown that, it is not possible to identify the age accurately in the case of children (3 to 5 years).

Thus it may be concluded that as found by previous investigators (Schwartz, 1968; Schwartz of Rine, 1968; Ingemann, 1968, Coleman, 1971, 1973 a, 1976; Lass *et al.*, 1976, 1978, 1979, Ptacek of Sander, 1966, Shipp and Hollien, 1969; Rayan and Burk, 1972; Burk *et al.*, 1975; Hartman and Danhauer, 1976) in the present study also the age and sex of the speakers have been identified by listeners, beyond chance level, in spite of over lapse in the fundamental frequency and Vocal tract resonance in the age groups used. These results have increased the need for further studies to find out the acoustic cues required for identification of age and sex.

The information from studies of this kind would be useful ' in a variety of future theoretical and applied areas of investigation ' (Lass *et al.*, 1980).

REFERENCES

- Abramsj A.S., (1975): Auditory Cues and racial identification. Paper presented at the Annual Conference of the American Speech and Hearing Association, 21-4 Nov. Washington, DC.
- Airport, G.W., (1963): Pattern and Growth in Personality. New York: Holt, Rinehart and Winston.
- Alverenga, J.A., (1971): An investigation of the ability of listeners to differentiate race on the basis of tape recorded evidence. Master's thesis (Herbert H., Lehman College) unpublished.
- Bricker, P.D. and Pruzansky, S., (1966): Effects of stimulus content and duration on talker identification. *Journal of the Acoustical Society of America*, 40, 1441-9.
- Bryden, J.D., (1968): An acoustic and social dialect analysis of perceptual variables in listener identification and rating of Negro speakers. Doctoral dissertation (University of Virginia), unpublished.
- Burk, K.W., Hoyer, E.A., Fey, M., and Charlip, W.S., (1975): Perceptual and acoustic correlates of ageing in the female voice. Paper presented at the Annual Convention of the American Speech and Hearing Association, 21-4 Nov., Washington, DC, unpublished.
- Cautril, H. and Allport, G.W., (1935): The Psychology of Radio. New York: Harper.
- Clarke, F.R. and Becker, R.W., (1969): Comparison of techniques for discriminating among talkers. *Journal of Speech and Hearing Research*, 12, 747-61.
- Clarke, F.R., Becker, R.W. and Nixon, J.C., (1966): Characteristics that determine speaker recognition. Report ESD-Tr-66-636, Electronics Systems Division, Air Force System Command, Hanscom field—unpublished.
- Coleman, R.O., (1971): Male and female voice quality and its relationship to vowel formant frequencies. *Journal of Speech and Hearing Research*, 14, 565-77.
- Coleman, R.O., (1973a): A comparison of the contributions of two vocal characteristics to the perception of maleness and femaleness in the voice. Paper presented at the Annual Convention of the American Speech and Hearing Association, Detroit, MI—unpublished.
- Coleman, R.O., (1973b): Speaker identification in the absence of inter-subject differences in ' glottal source characteristics—*Journal of the Acoustical Society of America*, 53, 1741-3.

- Coieman, R.O., (1976): A comparison of the contributions of two voice quality characteristics to the perception of maleness and femaleness in the voice. *Journal of Speech and Hearing Research*, 19, 168-80.
- Compton, A.J.j (1963): Effects of filtering and vocal duration upon the identification of speakers, aurally. *Journal of the Acoustical Society of America*, 35, 1748-52.
- Dickens, M. and Sawyer, G.M., (1962): An experimental comparison of vocal quality among mixed groups of Whites and Negroes. *Southern Speech Journal*, 18, 178-85.
- Eisenberg, P. and Zalowitz, F., (1938): Judging expressive movements III. Judgements of dominance-feeling from phonograph records of voice. *Journal of Applied Psychology*, 22, 620-31.
- Fairbanks, G., (1960): Voice and Articulation Drill Book. New York: Harper and Row.
- Gopal, H.S., (1980): Relationship for locating optimum frequency in the age range of 7 to 25 years. Mysore University—M.Sc, Dissertation.
- George Samuel, (1973): Study of the fundamental frequency of Voice and Natural Frequency of Vocal Tracts on an Indian population of different age ranges—Mysore University M.Sc, Dissertation.
- Harms, L.S., (1961): Listener judgements of status cues in speech. *Quarterly Journal of Speech*, 47, 164-8.
- Harms, L.S., (1963): Listener comprehension of speakers of three status groups. *Language and Speech* 4, 109-12.
- Hartman, D.E. and Danhauer, J.L., (1976): Perceptual features of speech for males in four perceived age decades—*Journal of the Acoustical Society of America*, 59, 713-5.
- Herzog, H., (1933): Stimme and Persönlichkeit—*Z. Psychol*, 130, 300-379.
- Hibler, M.B., (1960): A comparative study of speech patterns of selected Negro and White kindergarten children. Doctoral dissertation (University of Southern California), unpublished.
- Holmgren, G.L., (1967): Physical and psychological correlates of speaker recognition. *Journal of Speech and Hearing Research*, 10, 57-66.
- Ingemann, I., (1968): Identification of the speaker's sex from voiceless fricatives. *Journal of the Acoustical Society of America*, 44, 1142-4.
- Larson, V.S. and Larson, G.H., (1966): Reactions to pronunciations. In McDavid, R.I., and Austin, W.M., Communication Barriers to the culturally deprived U.S. Office of Education Co-op. Res. Proj. No. 2107, Washington, D.C.—unpublished.
- Lass, N.J., Almerino, C.A., Jordan, L.F. and Walsh, J.M., (1980b): The Effect of filtered speech on speaker race and sex identification—*Journal of Phonetics*,
- Lass, N.J.j Beverly, A.S., Nicosia, D.K., and Simpson, L.A., (1978): An investigation of speaker height and weight identification by means of direct estimations. *Journal of Phonetics*, 6, 69-76.
- Lass, N.J., Brong, G.W., Ciccolella, S.A., Walters, S.C., and Maxwell, F.I., (1980a): An investigation of speaker height and weight discriminations by means of paired, comparison judgements. *Journal of Phonetics*, 8.
- Lass, N.J., and Brown, W.S., (1978): Correlational study of speaker's heights, weights, body surface areas, and speaking fundamental frequencies. *Journal of the Acoustical Society of America* (in press).
- Lass, N.J. and Colt, E.G., (1980): A comparative study of the effect of visual and auditory cues on speaker height and weight identification. *Journal of Phonetics* (in press).
- Lass, N.J. and Davis, M., (1976): An investigation of speaker's height and weight identification. *Journal of the Acoustical Society of America* 60, 700-3.
- Lass, N.J., Dicola, G.A., Beverley, A.S., Barbera, C, Henry, K.G., and Badali, M.K., (1979): The effect of phonetics complexity of speaker's height and weight identification. *Language and Speech* 22, 297-309.

- Lass, N.J. and Harvey, L.A., (1976): An investigation of speaker photograph identification *Journal of the Acoustical Society of America* 59, 1232-6.
- Lass, N.J., Hughes, K.R., Bowyer, N.D., Waters, L.T. and Bourne, V.T., (1976): Speaker sex identification from voiced, whispered, and filtered isolated vowels. *Journal of the Acoustical Society of America*, 59, 675-8.
- Lass, N.J., Mertss, P.J. and Kimmel, K.L., (1978): The effect of temporal speech on speaker race and sex identification. *Language and Speech*, 21, 279-200.
- Lass, N.J., Tecca, J.E., Mancuso, R.A. and Black, W.I., (1979c): The effect of phonetic complexity on speaker race and sex identification. *Journal of Phonetics*, 7, 105-118.
- Markel, N., Eisler, R.M. and Reese, H.W., (1967): Judging personality from dialect. *Journal of Verbal learning and Verbal behaviour*, 6, 33-5.
- McGehee, P., (1937): The reliability of the identification of the human voice. *Journal of General Psychology*, 17, 249-71.
- McGloue, R.E., and Hollien, H., (1967): Vocal pitch characteristics of aged women. *J.S.H.R.*, 6, 164-170.
- Pear, T.H., (1931): Voice and Personality. London, Chapman and Hall.
- Pollack, I., Pickett, J.M. and Sumbly, W.H., (1954): On the identification of speakers by voice. *Journal of the Acoustical Society of America*, 26, 403-6.
- Ptacek, P.H., and Sander, E.K., (1966): Age recognition from voice. *Journal of Speech and Hearing Research*, 9, 273-7.
- Ryan, W.J., and Burk, K.W., (1972): Predictors of age in the male voice. *Journal of the Acoustical Society of America*, 53, 345(A).
- Schwartz, M.F., (1968): Identification of speaker sex from isolated, voiceless fricatives. *Journal of the Acoustical Society of America*, 43, 1178-9.
- Schwartz, M.F., and Rine, H.E., (1968): Identification of speaker sex from isolated, whispered vowels. *Journal of the Acoustical Society of America*, 44, 1736-7.
- Shipp, F.T., and Hollien, H., (1969): Perception of the ageing male voice. *Journal of Speech and Hearing Research*, 12, 703-10.
- Stagner, R., (1936): Judgements of voice and personality. *Journal of Educational Psychology*, 27, 272-7.
- Stevens, K.N., Williams, C.E., Carbonell, J.R., and Woods, B., (1968): Speaker authentication and identification: A comparison of spectrographic and auditory presentations of speech material. *Journal of the Acoustical Society of America*, 44, 1596-607.
- Stroud, R.V., (1956): A study of the relation between social distance and speech differences of White and Negro high school students of Dayton, Ohio. Master's Thesis (Bowling Green State University) unpublished.
- Usha Anna Abraham, (1978): Study of fundamental frequency of voice in an Indian population. Mysore University—Dissertation.
- Voiers, W.D., (1964): Perceptual bases of speaker identity. *Journal of the Acoustical Society of America*, 36, 1065-73.