THE PERCEPTION OF GENDER DIFFERENCES IN THE SPEECH OF 4½ - 5½ YEAR OLD CHILDREN

Moray J. Nairn
Queen Margaret College, Edinburgh, Scotland

ABSTRACT

Listeners judged the gender of different samples of the speech of 89 4½ - 5½ year old children. Sentences yielded the highest mean rate of identification (76.23%) and isolated vowels yielded the lowest (65.91%). Gender differences in listener response were also reported. The results reflect previously reported rates of accuracy from other countries and are interpreted in a theoretical framework which assumes that gender identification is a universal perceptual ability.

INTRODUCTION

The perceptual distinction between the voices of the adult male and female is in little doubt. When judges have been asked to classify subjects into male or female groups on the basis of voice alone, studies have reported success rates which approach 100% [1]. What is perhaps more surprising is the ability of judges to correctly classify prepubertal children at above chance levels of accuracy in the same way.

The issue of listeners' ability to identify the gender of prepubertal children from their speech has been addressed by several empirical studies [2,3,4,5,6,7]. A number of sample types have been employed in previous research, ranging from extracts of spontaneous speech (with selection criteria designed to minimise potential cues from utterance content) to isolated vowels spoken in a whisper. Among the experiments which used spontaneous speech samples, Weinberg and Bennett [2] demonstrated that judges were able to correctly identify the gender of 5 and 6 year old American children with a success rate of 74%. Meditch [5] achieved a similar level of accuracy using the speech of a small group of children as young as 3 years. Sentences have also been widely used as a sample type. Sachs et al [3] found a success rate of 81% with recorded sentences from a group of children with a rather large age range (4 to 14 years); the success rate dropped to around 66% when the samples consisted of isolated vowels only [4]. Günzburger et al's [7] results followed a similar pattern with 7 and 8 year olds, they found that the average recognition rate of 74% for sentences dropped to 55% when listening instead to isolated vowels. Bennett and Weinberg [6] had listeners judge the gender of 73 six and seven year old children using various types of speech sample. Overall, the identification rate was around 68%, with the success rates for the isolated vowels slightly lower than for the sentences. The identification rates of a quarter of the subjects were above 97% for the isolated vowels; however a significant number of the children were not judged consistently as either male or female. This emphasises the need to consider the distribution of not only mean results but also the identification scores of individual subjects. While the gender of many subjects may be very consistently identified, others may be ambiguous and some may even be consistently identified as having the opposite gender. With regard to possible listener bias in favour of one or the other gender, Meditch [5] found that listeners made significantly more 'male' guesses (i.e. gender judgements in favour of males). However Bennett and Weinberg [6] found no overall bias towards either gender.

The general conclusion that may be drawn from these studies is that perceptible gender differences do exist in the speech of many prepubertal children. Adult listeners are able to correctly identify gender with a success rate of somewhere around 70% from samples of normal speech. Even with samples of isolated vowels the success rate (at around 66%) is not much reduced. However, the identification rate is by no means as high as it is with adult subjects. This experiment is part of a larger study which attempts to address some of the issues surrounding gender identification using, for the first time, Scottish children who are, in addition, younger than many of the subjects used in previous studies.

METHOD

Eighty nine child subjects (46 boys and 43 girls), screened to exclude speech and hearing impairments, were recruited from two Edinburgh primary schools. English was the only language spoken in their homes and the accent used was the local Scottish accent of English. The methodological techniques and broad purpose of the study were explained to parents and children, but they were not informed of the study's relation to gender as it was felt that this might insert bias into the children's responses.

Three sample types were selected:-

Isolated vowels: The children sustained the vowels /i/, /a/ and /o/ at a comfortable pitch and level. These vowels were selected because it was felt that they represented enough of a contrast between high, low, front and back articulations to highlight any spectral gender difference which may exist.

Isolated sentences : The two sentences chosen were : *Rover is a big, brown dog* and *Rover has got a bone in his mouth*. These sentences were elicited by associating each with a picture and then showing the picture to the child who would then speak the sentence. This method avoided verbal or intonational cueing which often accompanies mirrored stimuli.

Spontaneous speech : A period of quasi-spontaneous speech was elicited from each child by means of the 'Bus Story' [8]. This procedure involves the researcher narrating a story which is illustrated by colour pictures viewed by both child and researcher. The child is then asked to re-tell the story from memory using the pictures. The resulting speech sample has the advantage of being of a very similar topic across all children but leaves the final choice of wording open to each individual child.

Recording equipment consisted of a Realistic® PZM Microphone and an AIWA HD-S1 Digital Audio Tape recorder fitted with a HDA-1 A/D converter. The data was recorded directly onto Maxell DM90 DAT tapes at a sampling rate of 48kHz with 16-bit Digital to Analogue (quantization) rate.

The master tapes were edited down to give three experimental tapes, one sample type per tape. On the vowel and sentence tapes, each child's vowels or sentences were kept together as a unit although the order of presentation of children on the tape was randomised.

The panel of listeners consisted of 8 male and 8 female subjects with an age range between 18 and 33 years. All were members of the academic staff and students of Queen Margaret College, Edinburgh and reported themselves to be native speakers of English and free of hearing disorders.

All of the adult subjects judged all of the children's speech samples. The adult judges recorded their responses on an answer sheet by circling 'M' or 'F' (for boy and girl respectively) in the appropriate box. Each judge received one answer sheet which was divided into three sections (one section for each type of speech sample - vowels, sentences, passage).

The judges made one response to each child's set of vowels, one response to each child's set of sentences and one response to each child's spontaneous speech. Therefore there were 3 gender responses to each child. As there were 89 children, this realised a total of 267 gender judgements per judge and a grand total of 4272 gender judgements in the whole experiment.

The purpose of the experiment was explained to the judges before beginning the task and they were each given written and oral instructions on how to record their responses. The adult subjects were seated in a quiet room and listened to each tape through headphones which connected directly to a Sony DTC-690 digital audio cassette player. Judges marked their estimate of the gender of each child on the response sheet. A short pause was included between tapes to allow the judges to rest. The gender of each judge was recorded along with their responses to permit analysis of the effect of the sex of the listener on gender identification.

RESULTS

Bias statistics revealed that one female judge was heavily biased to respond 'male' and so was removed from the analysis. Tables 1-3 show the proportion of correct gender responses made by remaining judges. The differences between these correct gender judgement rates varied between the three speech sample types. Paired two-samples t-tests were

carried out between the correct judgement rates of each speech sample type. Both the sentence and passage samples yielded significantly higher identification rates than the vowel sample ($p<0.05$). There was no significant difference between the identification rates of the sentence and passage samples. All identification rates are significantly above chance ($p<0.05$). In addition, the female judges were better at correctly identifying gender than the male judges ($p<0.001$) and girls were better identified than boys (girls - 76.37%; boys - 67.79%; $p<0.01$).

A total of 1859 (46.42%) 'male' responses and 2146 (53.58%) 'female' responses were made. A greater proportion of female to male responses was made in every condition (judges' sex × children's sex × sample type).

Table 1. Correct gender identification rates from VOWEL *samples*

	Male Judges	Female Judges	Average
Boys	59%	65%	61.96%
Girls	68%	72%	70.22%
Average	63.36%	68.82%	**66.08%**

Table 2. Correct gender identification rates from SENTENCE *samples*

	Male Judges	Female Judges	Average
Boys	71%	72%	71.49%
Girls	80%	82%	81.17%
Average	75.43%	77.22%	**76.33%**

Table 3. Correct gender identification rates from PASSAGE *samples*

	Male Judges	Female Judges	Average
Boys	67%	73%	69.93%
Girls	77%	78%	77.72 %
Average	71.81%	75.85%	**73.83%**

DISCUSSION

The results of this experiment go some way towards answering certain relevant questions and appear to raise some of their own. It is clear, for example, that the gender recognition rates found in this study conform closely to the levels of accuracy reported in previous research from around the world (U.S.A. - [2]; Sweden - [9]; etc.). This finding is interesting in that it suggests a universal ability on the part of the human listener to perceive acoustic gender differences. Existing experimental evidence seems to indicate that the differences in larynx size and vocal tract length which largely account for the adult vocal sex-difference are absent (or at least less influential) in pre-pubertal children. It has been suggested that processes of socialisation, which are responsible for emphasising existing physiological differences, and even separate gender dialects are in use. If this latter proposition, that gender is signalled vocally by learned speech characteristics, is indeed the case some obvious questions present themselves: are there universal aspects of the gender-marking and recognition process; at what age do these dialects emerge and what exactly is the method of gender-marking used by males and females?

Karlsson and Rothenberg [9] found that recognition rates of American and Chinese listeners were highly correlated with the rates displayed by Swedish and Finnish listeners and suggested that "at least some aspects of the gender differentiation in Swedish are not language-specific." [p14]. If the ability of listeners to determine gender in pre-pubertal children is a feature which is present universally, then reported differences in gender recognition rates may be accountable by differences in methods used by children from different cultures to signal gender. It seems unlikely that there are large-scale typological differences in the linguistic features employed across different languages to mark speaker gender, however one possibility is that there is a pool of features (phonetic, phonological, morphological etc.) from which languages are supplied - the precise features present in any given language depending on the configuration of the grammar (i.e. the parameters of Universal Grammar set during language acquisition). In Hebrew, for example, the correct conjugation of verbs depends on the sex of the speaker whilst in Japanese, boys and girls are encouraged to use different forms for the first-person pronoun although the spoken language is syntactically unmarked for gender. Also, listeners find it harder to identify Finnish boys from girls than Swedish boys from girls. The suggestion is that as there is no grammatical gender in Finnish it may be that young children acquiring the language "are less aware of sex-differences...than are Swedish children and therefore less motivated to develop speech patterns [based on] speaker sex." [9]

The age at which these gender dialects start to emerge cannot be answered by the present study, however in the light of these results using 4½-5½ year old children it is clear that there is some information relating to the child's gender present in the voice by the fifth year of life. To ascertain what form this gender information might take in any given language would require an in depth acoustic study of the voice. Such an investigation is currently being carried out by the author using the same children from this study. Previous research has tended to be inconclusive, however the weight of opinion seems to be that fundamental frequency, unlike the situation in adults, is not a major factor in gender marking in English.

Turning now to the differences in listener accuracy between the three different speech sample types, the fact that the passage and sentence conditions yielded significantly higher rates than the vowel condition may be attributed either to the fact that there was more phonetic information present in the former or to the fact that there was a qualitatively different type of information available (intonation, formant transitions, etc.). What is more surprising is that the sentence sample gave a higher correct recognition rate than the spontaneous speech sample (although not significantly so). One possible interpretation is that the sentence sample involved the children speaking exactly the same words every time whereas the passage sample only controlled for the topic of conversation but allowed the children the freedom of choice of words. Therefore, strictly speaking, listeners were not comparing like with like in the passage condition, they were forced to extrapolate each child's speech patterns to a certain extent in order to make mental comparisons. This extra step in the recognition process might adversely influence levels of accuracy relative to the sentence sample.

In summary, listeners were highly successful in determining the gender of 4½-5½ year old Scottish children. The identification rates achieved were very similar to previous comparable studies. The result that girls were better perceived than boys, and female judges were better then male judges appears to contradict some of the previous findings. Further work is needed to supply the missing answers.

REFERENCES

[1] Coleman R.O. (1971) The perception of maleness and femaleness in the voice and its relationship to vowel formant frequencies, *Proceedings of the 7th Int. Congress of Phonetic Sciences.*

[2] Weinberg B and Bennett S.(1971) Speaker sex recognition of 5 and 6 year old children's speech, *Journal of the Acoustical Society of America* **50**:1210-3.

[3] Sachs J, Lieberman P, Erickson D. (1973) Anatomical and cultural determinants of male and female speech. in *Language attitudes: current trends and prospects.* ; Shuy R, Fasold R, editors. Washington D.C. Georgetown,

[4] Sachs J. (1975) Cues to the identification of sex in children's speech, in *Language and sex: difference and dominance.* ; Thorne B, Henley N, editors, Rowley, Mass.

[5] Meditch A.(1975)The development of sex-specific speech patterns in young children, *Anthropological Linguistics* **17**:421-33.

[6] Bennett S and Weinberg B.(1979) Sexual characteristics of preadolescent children's voices, *Journal of the Acoustical Society of America* **65**:179-89.

[7] Günzburger D, Bresser A, and Ter Keurs, (1987) Voice identification of prepubertal boys and girls by normally sighted and visually handicapped subjects, *Language and Speech* **30**(1):47-58.

[8] Renfrew, C. (1969) *The Bus Story - A Test of Continuous Speech,* C.E. Renfrew, Old Headington, Oxford.

[9] Karlsson, I. and Rothenberg, M. (1992) Inter-cultural variations in gender-based language differences in young children, *Quart. Prog. and Status Report 1/1992. Speech Transmission Lab. Institute of Technology, Stockholm.* pp.1-17.