SPEAKER SEX IDENTIFICATION AND VOWEL IDENTIFICATION IN THE ABSENCE OF GLOTTAL SOURCE CHARACTERISTICS

AN HONORS THESIS (ID 499)

by

ROBIN E. RADEMAKER

THESIS DIRECTOR

Thomas F. Hemeyer

DR. THOMAS HEMEYER

BALL STATE UNIVERSITY

MUNCIE, INDIANA

SEPTEMBER, 1978

TABLE OF CONTENTS

CHAPTER I

STATEMENT OF THE PROBLEM

According to Coleman (1971), "Differences in vocal pitch have been generally accepted as the acoustic cue that distinguishes between male and female speakers." A speaker's fundamental frequency is the primary determinant of the perceived pitch of his or her voice. The average fundamental frequency of the female voice is 260 Hertz and the average fundamental frequency of the male voice is 130 Hertz. "...a listener presumably makes a probability estimate of the speaker's sex based on the pitch of that particular voice as compared to what he already knows about the pitch ranges of men and women." (Coleman, 1971) These presumptions, however, have been disputed by Coleman (1971, 1973), who found that removal of laryngeal cues, i.e., pitch cues, has little effect on the identification of the speaker's sex, especially with regard to the male voice.

Part I of this study will attempt to further investigate Coleman's findings by having listeners decide whether a tape recorded voice (with laryngeal cue removed) is male or female. This will be accomplished by having a group of listeners identify the sex of four speakers from a tape recording of sentences produced by male and female speakers using an artificial larynx.

Studies of Peterson and Barney (1952), Lehiste and Meltzer (1973), and others, have looked at the importance of formants in identifying spoken vowels. Ingemann (1968) and Schwartz and Rine (1968) have studied the ability of listeners to distinguish voiceless fricatives auditorily and have discussed the influence of the size of the vocal tract in the identification of voiceless fricatives.

Part II of this study will attempt to determine if the size of the vocal tract in front of the constriction made in the formation of vowels has an influence on the accuracy of vowel identification. More specifically, an attempt will be made to determine whether or not, with the removal of normal laryngeal cues, back vowels are easier to recognize than front vowels in isolation, and which vowels are of greatest assistance in identification of the speaker's sex. This will be accomplished by having a group of listeners identify the vowel being said and the sex of the speaker from a tape recording of isolated vowels produced by male and female speakers using an artificial larynx.

CHAPTER II

REVIEW OF RELATED LITERATURE

Speaker Sex Identification

Coleman (1971) found, "...it is possible to recognize the sex of a speaker when a single frequency sound source is substituted for the laryngeal tone." In Coleman's study (1973), the subjects performed a discrimination task, deciding whether two utterances were made by the same person. The results indicated that it is possible to achieve 90% correct discrimination in the absence of glottal source cues. Analysis of the errors indicated that most of the errors were in the identification of female voices. Coleman concluded, "...it is easier to avoid errors /in speaker identification/ when the speakers are male." Coleman attributed this difference in identification to the following possibilities: 1) Females were better at disguising their voices than males. 2) Males differed more among themselves than did the females in terms of speech characteristics. 3) The quality of the substituted glottal tone was more like the male voice and added to the distortion of the female voice.

The results of an unpublished master's project by Hoover (1975) indicated that the male voice tends to be more easily identified and that this identification and that of vowels depends on vowel formants. Hoover's results led him

to conclude, "...the lower the fundamental frequencies, the more consistently the sound is labelled as being produced by a male." Hoover's conclusion was contrary to the results of Coleman (1971, 1973) who established the fact that the fundamental frequency of a speaker's voice has little effect on a listener's identification of speaker sex. Coleman's studies (1971, 1973) attempted to establish that listeners do not rely solely, or even heavily, upon the fundamental frequency of a speaker's voice to identify the speaker's sex.

Schwartz (1968) investigated the ability of listeners to identify the sex of speakers from voiceless fricatives. The results of this investigation indicated that listeners most often correctly identified the sex of the speaker from /s/ and /ʃ/ and least often from /f/ and /θ/. The author explained the implications of the results in terms of frequency displacement. From a spectropraphic analysis, /f/ and /θ/ can be seen to have flat broad spectral peaks as opposed to /s/ and /ʃ/ which have steep, more distinct peaks. Because of this, frequency displacement, or the difference between the frequencies of the male and female voice of /f/ and /θ/, is more difficult to detect.

It is also worth noting that /s/ and /ʃ/ are produced further back in the mouth than /f/ and /θ/. As a consequence, in /s/ and /ʃ/ there is a larger portion of the oral cavity through which the sound must travel. As it travels through the oral cavity, the sound will be altered by resonance characteristics of the cavity. Since females have smaller oral

cavaties than males, their resonance effects will be different and provide a possible sex discriminating cue in the sound. In the case of /f/ and /θ/, the noise is made right at the front of the oral cavity and resonance effects (and resonance associated sex cues) are minimal.

Using the vowels /ɑ/ and /i/, Schwartz and Rine (1968) studied the affect of whispered vowels on speaker sex identifications when /ɑ/ and /i/ were whispered. Out of a total of 160 identifications, only four errors were made and no listener made more than one error. The authors concluded, "...the primary acoustic cue that underlies the distinction appears to be the upward frequency displacement of the resonance peaks of female vowels."

As indicated above, Schwartz (1968) explained the identification of voiceless fricatives from a spectrographic analysis. Although the female sound pattern is similar in overall general shape to the male sound pattern, the female spectral peaks are displaced upward in frequency due to females' smaller oral cavities. Schwartz and Rine (1968) hypothesized that the same thing happens with vowels. A whispered vowel, although lacking the conventional spectrum source /glottal tone/ retains "sufficient transfer-function to permit correct sex identification." That is, the female vowel spectral peaks (formants) will be higher because of their smaller vocal tracts.

Ingemann (1968), in a related study, obtained similar results. The author concluded that as the portion of the vocal

tract in front of the constriction made in the formation of the fricative diminishes, the accuracy of speaker sex identification decreases. This is consistent with the results of Schwartz' study (1968) and the explanation given above.

Vowel Identification

Peterson and Barney (1952), in their study of vowels, established the fact that /i/, /æ/ and /u/ are generally quite well understood. The high rate of correct identification of /i/ and /u/ was attributed to their terminal positions in the speaker's mouth. That is, /i/ is the highest front sound in English and /u/ is the highest back sound in English. These terminal, or end positions, are such that they are less likely to be displaced in production. This would mean that these sounds should be more consistenly recognized by listeners.

The authors explained the consistently correct identification of /æ/ from the vowel loop. When listeners incorrectly identified a vowel, the substitution was a vowel adjacent to the correct vowel on the vowel loop. The tendency of listeners is to shift toward a lower speech sound on the vowel loop. In other words, an /ɪ/ would be perceived as an /ɛ/ or an /u/ as a /ʊ/. Since there is a wide gap between /æ/ and /ɑ/, the vowel most directly below /æ/, this might account for the high degree of correct identification. Also, in most vowel diagrams, the /æ/ is the lowest front vowel and would thus be a vowel with a terminal position and the rational for terminal vowels, as indicated above, would be applied to /æ/.

Stevens, et. al. (1968), in a comparison study of spectrographic and auditory identification of speakers, found that it is easier to identify a speaker when he says a word containing a front vowel than when he utters a word containing a back vowel.

Lehiste and Meltzer (1973), in their comparison study of the identification of normal versus synthetic vowels, found that listeners more consistently identified male vowel sounds spoken under normal conditions (as opposed to the synthetic vowel sounds) than female vowel sounds spoken under the same conditions. The results indicated that /ʊ/ was the most difficult vowel to identify under normal speaking conditions and that the high front vowel sounds were most consistently correctly identified under the same conditions. The authors concluded that "formants are more important than fundamental frequencies" in the identification of vowels.

CHAPTER III

METHOD

Speakers and Listeners

The speakers were two male and two female college students. They were between the ages of 19 and 21 years. They spoke with a Standard American Dialect and had no speech impairments. The 20 listeners were college students who were undergraduate majors in Speech Pathology & Audiology. They were unfamiliar with the speakers.

Part I, Procedure

The four speakers were instructed in the use of an Aurex 2 electrolarynx. They practiced the test sentences using the electrolarynx at a fundamental frequency of approximately 100 Hertz. One at a time, the four speakers read five phonetically varied sentences, approximately five seconds in length. All speakers used the same phrasing in each sentence to eliminate individual differences in phrasing. The sentences were tape recorded on a Roberts Tape Recorder, Model X771X. The five sentences of all four speakers were then randomly dubbed across speakers onto a second tape with a five second pause between each segment. This resulted in a listening tape of 20 sentences with a total time of approximately three and a half minutes.

The listeners were given an answer sheet on which they were

instructed to mark the sex of the speaker and rate the confidence of their judgement on a scale of 1-to-10 (1 = totally confident, 10 = no confidence.) All 20 participants listened to the tape recording simultaneously.

Part II, Procedure

The four speakers were instructed in the pronunciation of the vowels /i, ɪ, ɛ, æ, ɑ, o, ʊ, and u/. One at a time, the speakers recorded the vowels with about a two second duration. The vowels of all four speakers were then randomly dubbed, across vowels and across speakers, onto a second tape with a five second pause between each vowel. This resulted in a listening tape of 32 vowels with a total listening time of approximately four minutes.

The listeners did this listening task directly following Part I. They were given an answer sheet on which they were instructed to mark the sex of the speaker, the vowel being said, and to rate the confidence of their choice of vowels on a scale of 1-to-10 (1 = totally confident, 10 = no confidence.) Since the results of the study did not depend on phonetic transcription, the listeners were given the list of vowels to be said.

Part I, Statistical Analysis

A test of significance between an observed and a specified proportion was used to determine whether or not the correct identification of male speakers was significantly better than chance. Since the chance proportion of correct identifi-

cation was 50% (male or female only possible response), the statistical test was done to determine if correct sex identification was significantly better than 50% at the .05 or .01 level of confidence. The same test was run to determine whether or not the correct identification of female speakers was significantly better than chance. Then a test of significance between two specified proportions was used to determine whether or not the correct identification of male speakers was significantly better than the correct identification of female speakers.

Finally, a t-test was used to determine whether or not the listeners' mean judgement of confidence for correct identification of male speakers was significantly greater than the mean judgement of confidence for correct identification of female speakers.

Part II, Statistical Analysis

A test of significance between an observed and a specified proportion was used to determine whether or not the correct identification of males from vowel utterances was significantly better than chance and whether or not the correct identification of females from vowel utterances was significantly better than chance. Since the chance proportion of correct identification was one out of two (either male or female), the statistical test was done to determine whether or not correct identification of sex from vowels was significantly better than 50%. Then a test of significance between

two specified proportions was used to determine whether or not the correct identification of males from vowel utterances was significantly better than the correct identification of females from vowel utterances.

A test of significance between an observed and a specified proportion was used to determine whether or not the correct identification of front vowels was significantly better than chance and whether or not the identification of back vowels was significantly better than chance. Since eight vowels were used in the study, and the listeners had a list of the eight vowels (i.e., closed set response), the number of correct vowel identifications from chance would have been one of eight (or 12.5%). The statistical test was done to determine if the number of correct vowel identifications was significantly better than 12.5%.

When dividing the eight vowels into front and back, Peterson and Barney's vowel loop (1952) was used to determine the dividing point. These two groups were as follows: front vowels (/i, ɪ, ɛ, æ/); back vowels (/ɑ, o, ʊ, u/).

Then a test of significance between two specified proportions was used to determine if there was a significant difference between the following:

1) Correct front vowels and correct back vowels.
2) Correct male front vowels and correct female front vowels.
3) Correct male back vowels and correct female back vowels.
4) Correct male front vowels and correct male back vowels.
5) Correct female front vowels and correct female back vowels.

CHAPTER IV

RESULTS

Part I

The purpose of Part I of this study was to further investigate speaker sex identification from sentence length utterances in the absence of glottal source characteristics, and to discover if there was a significant difference between a listener's ability to identify a male voice versus a female voice with normal laryngeal cues removed.

A one-tailed test of significance between an observed and a specified proportion was used to determine whether or not the correct identification of the male speakers from sentence length utterances was significantly better than chance. Out of 200 tests, 144 were correctly identified as male speakers. The z value of 6.22 was significant at the .01 level. The results of the test of significance for the correct identification of the female speakers also showed significance at the .01 level with a z value of 3.11. Out of 200 trials, 122 were correctly identified as female speakers.

The results indicate that listeners could identify a speaker's sex in the absence of glottal source characteristics, whether the speaker was male or female. Since the proportion of chance was one out of two (male or female), the test results show that the correct identification of the speakers' sex was significantly greater than 50%.

Since both the correct identification of the male speakers and the correct ifentification of the female speakers were significantly better than chance, a test of significance between two specified proportions was used to determine whether or not the correct identification of male speakers was significantly better than the correct identification of female speakers. The resulting z value of 2.33 was significant at the .05 level, but not at the .01 level. This, however, does indicate that listeners could correctly identify male speakers significantly better than female speakers.

The listeners were asked to rate the confidence of their choice of speaker sex on a scale of 1 to 10 (1 = totally confident; 10 = no confidence). The mean confidence rating of listeners for the correct identification of male speakers was 5.27. The mean confidence rating of listeners for the correct identification of female speakers was 5.41.

Since the listeners correctly identified the male speakers significantly better than the female speakers, a t-test was used to determine whether or not the listeners' mean judgement of confidence for male speakers was significantly different than the mean judgement of confidence for female speakers. The t-score of .430 was not significant at the .01 level or the .05 level. This indicated that, although listeners identified male speakers significantly better than female speakers (without glottal source characteristics), they were not significantly more confident of their judgement of male speakers.

Part II

The purpose of Part II of this study was to determine whether or not, with the removal of normal laryngeal cues, sex could be identified from vowel utterances; and to determine if vowels produced without normal laryngeal cues would be intelligible. The intelligibility of the vowels was broken down into front versus back vowels and male versus female vowels.

A one-tailed test of significance was used to determine whether or not the correct identification of vowels spoken by males was significantly greater than chance, and whether or not the correct identification of vowels spoken by females was significantly greater than chance. The z value of 2.01 for the correct identification of vowels spoken by males was significant at the .05 level. Out of 320 trials, 178 were correctly identified as male speakers. Likewise, the z value of 1.67 for the correct identification of female vowels was significant at the .05 level. Out of 320 trials, 145 were correctly identified as female speakers.

This means that listeners could correctly identify the sex of a speaker (without normal laryngeal tones) from a spoken vowel. Since the chance proportion was one out of two (male or female), the results indicate that the listeners correctly identified the speakers' sex significantly greater than 50% of the time.

When comparing the correct identification of males from vowel cues without normal laryngeal cues with the correct identification of females from vowel cues without normal laryngeal

cue, a z value of 2.609 was obtained. This value was significant at the .01 level which indicates that listeners correctly identified the male speakers from vowels significantly better than the female speakers.

A one-tailed test of significance between an observed and a specified proportion was used to determine whether or not the correct identification of front vowels was significantly better than chance. Since there were eight vowels from which to choose, the number of correct identifications from chance was one out of eight. The test was used to determine whether or not the correct identification of front vowels was significantly greater than 12.5%. Out of 320 trials, 163 were correctly identified as front vowels. The z value of 20.79 was significant at the .01 level, indicating that listeners could identify front vowels significantly better than 12.5% of the time. The same test was used to determine whether or not the correct identification of back vowels was significantly better than chance. Out of 320 trials, 160 were correctly identified as back vowels. The z value of 20.28 was significant at the .01 level, indicating that listeners could correctly identify back vowels significantly better than chance, 12.5%.

A two-tailed test of significance between two specified proportions was used to analyze the remainder of the data.

1) A z value of .237 indicates that there was no significant difference between the correct identification of front vowels and the correct identification of back vowels, indepen-

dent of speaker sex identification. This indicated that there was no significant difference in the listeners' ability to identify front vowels versus back vowels in the absence of normal laryngeal cues.

2) There was no significant difference between the correct identification of front vowels spoken by males and the correct identification of front vowels spoken by females. Out of 160 trials each, 85 male front vowels and 78 female front vowels were correctly identified with respect to vowel and sex. The z value of .782 was not significant at the .01 level or thr .05 level.

3) The two-tailed test of significance was used to determine whether or not there was a significant difference between the correct identification of back vowels spoken by males and back vowels spoken by females. Out of 160 trials each, 93 male back vowels and 67 female back vowels were correctly identified with respect to the vowel and the sex. The z value of 2.90 was significant at the .01 level, indicating that the listeners identified the back vowels spoken by the male speakers significantly better than the back vowels spoken by the female speakers.

4) When comparing the correct identification of front vowels spoken by males and back vowels spoken by males, a z value of .9 was not significant at the .01 level or the .05 level. This indicated that there was no significant difference in the listeners' ability to identify front vowels and back vowels when spoken by a male.

5) There was no significant difference between the correct identification of front vowels spoken by females and back vowels spoken by females. The z value of 1.23 was not significant at the .01 level or the .05 level. This indicated that there was no significant difference in the listeners' ability to identify front vowels and back vowels when spoken by a female.

CHAPTER V

DISCUSSION AND SUMMARY

The purpose of this study was twofold. Part I attempted to investigate the affect of a single frequency glottal tone on speaker sex identification from sentence length utterances. The purpose of Part II of this study was to determine whether or not, with the removal of normal laryngeal cues, sex could be identified from vowels to determine if back vowels are easier to identify than front vowels.

Analysis of Part I

The listeners were able to identify both male and female speakers in the absence of glottal source characteristics. The correct identification of the speakers was significantly greater than chance, which means that, when using sentences as the test materials, the listeners did not rely on the pitch or fundamental frequency of the speaker's voice to determine the sex of the speaker. This finding was consistent with Coleman (1971, 1973). On the basis of these results alone, one could conclude that other cues, such as the resonance characteristics of voiceless sounds, are the cues which aid in speaker sex identification. This conclusion was consistent with the findings of Ingemann (1968) and Schwartz and Rine (1968). These studies found that listeners could correctly identify speaker sex from voiceless fricatives. These resonance characteristics were dependent on the size of the speaker's oral

cavity. Since a normal male's oral cavity would be larger than a normal female's oral cavity, one would expect differences in resonance characteristics independent of normal laryngeal tones.

The z value obtained from the comparison of the correct identification of male speakers and the correct identification of female speakers resulted in a significant difference between the two groups of speakers. The listeners correctly identified male speakers using a single glottal tone significantly more often than female speakers under the same conditions. These results were consistent with the findings of Coleman (1971, 1973), Hoover (1975), Schwartz and Rine (1968), and Lehiste and Meltzer (1973).

There are various reasons for this difference in speaker sex identification. The fundamental frequency of the electrolarynx used more closely approximated the fundamental frequency of the male voice than that of the female voice. This explanation was consistent with the results obtained. This was probably the most likely reason why the listeners misidentified a speaker's sex (they more often identified a female voice as a male voice than a male voice as a female voice.) However, there could have been more differences in the speech characteristic between the men which made them easier to recognize than the women, as suggested by Coleman (1973).

Additional research needs to be done using a substitute glottal tone which more closely approximates the fundamental frequency of the female voice. One could hypothesize that

listeners would more often identify a speaker's sex as that of a female.

The t-score obtained from the analysis of the confidence ratings of correct male speaker identification and correct female speaker identification did not result in a significant difference between the mean confidence ratings for each group of speakers. Interestingly, the mean confidence ratings for each group was slightly more than five, which is the midpoint between totally confident and no confidence. There could be several explanations for this occurrence.

First, it was possible that the listeners did not have enough time between speakers to really decide how confident they were. As a result, they were more likely to choose a middle number at random. Secondly, a range of 1 to 10 may have been too broad a spectrum from which to choose. The results may have been more accurate had the listeners only a choice of two confidences, thosebeing "confident" and "not confident."

If, however, these results actually reflect the male and female voices in the absence of normal laryngeal cues, they reveal an interesting fact about the voice identification. Listeners could identify a speaker's voice without glottal source characteristics; therefore, fundamental frequency of a speaker's voice is not the most important factor in a speaker sex identification. However, since the average confidence rating for the correct identification of a speaker's sex was approximately the midpoint between totally confident and no confidence, the distinguishing factors of a male and a female voice were not

readily identifiable by the listeners.

If there were some noticeable differences between the voices (with the removal of differences in fundamental frequencies) one could expect that the listeners would be more confident of their judgement of speaker sex. The hypothesis is, therefore, that with the removal of glottal source characteristics, the distinguishing factors between voices are not identifiable by the listeners, although the listeners are still able to distinguish between a male and a female voice.

Analysis of Part II

The listeners were able to correctly identify male and female voices from vowels in the absence of normal laryngeal cues significantly better than chance. In addition, the listeners were able to correctly identify vowels spoken by males significantly better than vowels spoken by females. These results were consistent with the results from Part I. Listeners could identify speaker sex in the absence of glottal source characteristics and the male voice was correctly identified with significantly less errors than the female voice. These results indicate that very few speech cues are necessary for speaker sex identification. It has been hypothesized by Hoover (1975) that formant frequencies are the most important characteristic in the identification of vowels. Formant frequencies and the difference in spectral peak displacement between the male and female voice combine to aid in the consistently correct identification of speaker sex from vowels.

The fact that sex could be identified from vowels as well as sentences support the theory of characteristics of the formant frequencies of vowels as opposed to phrase structure or other prosodic cues.

Analysis of the z values obtained in the test of significance comparing the correct identification of front vowels and back vowels to chance, resulted in a significant difference between both vowel groups and chance. The z values were extremely significant, indicating that the listeners had no trouble identifying the vowels in the absence of normal laryngeal tones. One factor in the test which added to the listeners ability was the fact that they were given a closed set of eight vowels from which to choose for each vowel that was said. They were not, however, given an indication as to how often each vowel would be said. It would be interesting to see what would happen to the results if the listeners were not given a closed set from which to choose. The results of this test indicated that normal fundamental frequency cues are not needed for correct identification of vowels. The obvious explanation for the correct identification of vowels is the resonance characteristics of the formant frequencies of the vowels. These results were consistent with the findings and conclusion of Hoover (1975).

There was no significant difference between the correct identification of front vowels and the correct identification of back vowels. This was an indication that there was no factor in the differences of vowel production that aid or

hinder a listener's ability to identify the vowel. When analyzing the raw scores of each vowel, the results of this study indicated that /i/, /ɛ/, and /ɑ/ were most often correctly identified while /æ/, /o/, and /u/ were most often mistakenly identified. These results were inconsistent with the results of Peterson and Barney (1952) who found that /i/, /æ/, and /u/ were most often correctly identified and that /ɑ/ and /ʊ/ were among the least correctly identified. The correct identification of /i/ can be explained in terms of its terminal position in production. That is, it is the highest front vowel and, as a result, there is less of a possibility of error in speaker production.

In the study of Peterson and Barney (1952), /u/ was one of the vowels most often correctly identified. The results of this study were the opposite. Because of the choice of vowels in this study (/o/, /ʊ/, /u/), there was a possibility that in the production of the vowels, there was some overlapping which made the three vowels difficult to distinguish. The possibility existed that the sample of listeners was not large enough to obtain a true indication of vowel intelligibility in the absence of glottal source characteristics. Had there been a larger group of listeners, the results may have been more consistent with previous findings.

There were no significant differences in the correct identification of front and back vowels as a function of speaker sex with the exception of back vowels spoken by males and back vowels spoken by females. These results indicated that the

size of the oral cavity has no affect on the intelligibility of vowels. The difference in the correct identification of back vowels spoken by males and back vowels spoken by females was a result of speaker sex identification rather than the influence of vowels. The results indicated that while the removal of glottal source characteristics had an affect on speaker sex identification, there was little difference in the identification of front versus back vowels.

Summary

The purpose of Part I of this study was to investigate speaker sex identification in the absence of normal laryngeal cues. Part II of this study was intended to investigate the affect of vowel intelligibility in the absence of glottal source characteristics, and the affect of these conditions on speaker sex identification. Two male and two female college students were used as the speakers and a group of 20 undergraduate students in Speech Pathology and Audiology were used as the listeners.

Each listener heard a tape of 20 sentences and were required to record the sex of the speaker and the confidence of their choice. Then the listeners heard a recording of 32 vowels and marked the vowel heard and the sex of the speaker.

The results showed that male speakers were more often correctly identified than female speakers when using sentences and vowels as test material. The results also showed that there is no significant difference between the identification

of front vowels and back vowels; but when there is, it is probably a function of speaker sex identification rather than vowel identification. Additional research could be done in the area of vowel intelligibility with the removal of glottal source characteristics using a larger group of listeners and having the substitute glottal source of a fundamental frequency more characteristic of female speech as well as male speech.

BIBLIOGRAPHY

Coleman, R. Male and female quality and its relationship to formant frequencies. Journal of Speech and Hearing Research, 14, 565-577 (1971).

Coleman, R. Speaker identification in the absence of inter-subject differences in glottal source characteristics. Journal of the Acoustical Society of America, 53, 1741-1743 (1973).

Hoover, H. Determination of sex from sublaryngeal cues. An unpublished master's project, Ball State University, SPAA 563 (1975).

Ingemann, F. Identification of the speaker's sex from voiceless fricatives. Journal of the Acoustical Society of America, 44, 1142-1144 (1968).

Lehiste, I. and Meltzer, D. Vowel and speaker identification in natural and synthetic speech Language and Speech, 16, 356-364 (1973).

Peterson, G. E. and Barney, H. L. Methods used in a study of vowels. Journal of the Acoustical Society of America, 24, 175-184 (1952).

Schwartz, M. Identification of speaker sex from isolated voiceless fricatives. Journal of the Acoustical Society of America, 43, 1178-1179 (1968).

Schwartz, M. and Rine, H. Identification of speaker sex from isolated whispered vowels. Journal of the Acoustical Society of America, 44, 1736-1737 (1968).

Stevens, K. N. et. al. Speaker authentication and identification: a comparison of spectrographic and auditory presentation of speech materials. Journal of the Acoustical Society of America, 44, 1596-1607 (1968).

APPENDIX I

PART I

Test Sentences*

1. Once,/a long time ago,/there was a young rat named Arthur/ who could never make up his flighty mind.
2. Whenever his small friends/used to ask him to go out to play with them,/he would only answer airily.
3. He would always shirk/from making a specific choice.
4. "Now look here,"/she stated./ "No one is going to aid or care for you/if you carry on like this."
5. One of them slipped back a broken board/and saw a squashed young rat,/quite dead,/ half out of his hole.

Correct Identifications -- Raw Scores

Male Speakers	Female Speakers
19	10
17	12
15	16
14	16
18	12
12	13
11	13
16	9
9	13
13	9
144	122

*/ indicates phrasing

Confidence Ratings -- Raw Scores

Confidence Ratings	Male Speakers	Female Speakers
1	10	7
2	8	8
3	20	16
4	21	15
5	32	31
6	16	5
7	9	14
8	3	5
9	4	3
10	21	18
Mean	5.27	5.41

Listen to each segment. Record the sex of the speaker in the blank provided -- male (M) or female (F).

Rate the confidence of your choice on a scale of 1 to 10. (1 = totally confident of choice; 10 = no confidence in choice, i.e., a guess.)

	SEX M or F	CONFIDENCE 1 to 10
1.	_____	_____
2.	_____	_____
3.	_____	_____
4.	_____	_____
5.	_____	_____
6.	_____	_____
7.	_____	_____
8.	_____	_____
9.	_____	_____
10.	_____	_____
11.	_____	_____
12.	_____	_____
13.	_____	_____
14.	_____	_____
15.	_____	_____
16.	_____	_____
17.	_____	_____
18.	_____	_____
19.	_____	_____
20.	_____	_____

APPENDIX II

PART II

Correct Identifications -- Raw Scores

Front Vowels

Vowel	Male Speakers	Female Speakers
i	20	23
ɪ	23	18
ɛ	23	21
æ	19	16
	85	78

Back Vowels

Vowel	Male Speakers	Female Speakers
ɑ	26	24
o	26	11
ʊ	22	18
u	19	14
	93	67

Listen to each vowel. Record the sex of the speaker -- male (M) or female (F).

Record the vowel being said -- /i,ɪ,ɛ,æ,ɑ, o, u,ʊ/.

Rate the confidence of your choice of vowels on a scale of 1 to 10. (1 = totally confident of choice; 10 = no confidence in choice, i.e., a guess.)

	Sex	Vowel	Confidence
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			
9.			
10.			
11.			
12.			
13.			
14.			
15.			
16.			
17.			
18.			
19.			
20.			
21.			
22.			
23.			
24.			
25.			
26.			
27.			
28.			
29.			
30.			
31.			
32.			