Audio Engineering Society

Convention Paper 7274

Presented at the 123rd Convention
2007 October 5–8 New York, NY, USA

QUANTIFYING THE SPEAKING VOICE: GENERATING A SPEAKER CODE AS A MEANS OF SPEAKER IDENTIFICATION USING A SIMPLE CODE-MATCHING TECHNIQUE

Peter S. Popolo[1,2], Richard W. Sanders, [1,3], and Ingo R. Titze,[2,3]

[1]National Center for Voice and Speech, Denver, CO, 80204, USA
ppopolo@dcpa.org

[2]University of Iowa, Iowa City, 52422, USA
richard.sanders@cudenver.edu AES Member

[3]University of Colorado at Denver and Health Sciences Center, Denver, CO, 80204, USA
ititze@dcpa.org

ABSTRACT

This paper looks at a methodology of quantifying the speaking voice, by which temporal and spectral features of the voice are extracted and processed to create a numeric code that identifies speakers, so those speakers can be searched in a database much like fingerprints. The parameters studied include: (1) average fundamental frequency (F0) of the speech signal over time, (2) standard deviation of the F0, (3) the slope and (4) sign of the FO contour, (5) the average energy, (6) the standard deviation of the energy, (7) the spectral energy contained from 50 Hz to 1,000 Hz, (8) the spectral energy from 1,000 Hz to 5,000 Hz, (9) the Alpha Ratio, (10) the average speaking rate, and (11) the total duration of the spoken sentence.

Introduction

This paper reports on a study to determine if a numeric code generated from a speech sample of an individual can be used for the purpose of speaker identification (SID), defined as an automatic process which finds a match between an unknown speaker and an identity from a database of known speakers [1].

The design of this study was based on the idea of pattern matching that comes from the work of Lawrence Kersta and Oscar Tosi, which forms the basis of the Aural/Spectrographic Method of voice identification [2, 3]. In the spectrographic analysis part of the Method, visual patterns in the voice spectrogram created by vowel formants, formant transitions, consonant features such as fricatives and plosive stops, and other features related to intensity and spectral distribution, are matched to determine if two speech samples were spoken by the same person. This pattern matching is only effective if the patterns being matched are comparable, thus the known and unknown spectrograms must be of the same phonetic utterance [4]. A complication of using this method is that there is a certain amount of intra-speaker variability, which is the variability between different utterances by the same individual, in addition to inter-speaker variability between different individuals. Figure 1 (at the end of this paper) illustrates the differences between the spectrograms obtained from the speech samples of two different speakers saying the same thing. Each speaker repeated the phrase twice. The figure shows that inter-speaker variations are more significant (to the eye) than the intra-speaker variations.

The numeric code chosen for this study was in the form of an 11-element vector, with each element corresponding to the value of an acoustic parameter extracted from the speech sample, and a code-matching program using simple matrix computations was used to perform the SID task. The speech samples from each individual consisted of that person speaking his or her own name, as well as a common false name spoken by each person. The parameters were extracted from multiple repetitions of the same phrase, either true name or false name, and averaged together in an attempt to factor out the intra-speaker variability. Male and female speakers were considered separately, in order to determine if the code-matching technique worked better for one group versus the other. For instance, it is known that female utterances are more difficult to analyze than male utterances when using the method of linear prediction coefficients (LPC) to determine formant frequencies from a prolonged vowel phonation. This is because the fundamental frequencies and first formant frequencies are much closer together in the speech of females than in males [5]. For this study, it was hypothesized that if a set of acoustic parameters were used that were not related to formant frequencies or did not depend heavily on vocal tract characteristics, the technique for performing the SID task would be more robust across males and females.

The reason for having the subjects speak their own name is that such a speech sample already has a pattern to it that can be matched from one instance to the next. It was hypothesized that the manner in which the subject spoke their names, characterized by a set of acoustic cues extracted from the speech samples, in combination with the phonetic content of their spoken name, could also be used to identify the speaker more accurately than if he or she were speaking an arbitrary phrase. To test this hypothesis, each subject was made to speak the same foil name, and the same speaker code and code matching techniques were applied to these samples.

The questions to be answered by this study were: (1) Can a numeric code consisting of acoustic parameters extracted from the speech samples of individuals speaking their names be used for the purpose of speaker identification, with reliable results? (2) Is the identification accuracy different for males versus females? (3) Is the identification accuracy different when all subjects are made to use a false name?

1. METHODOLOGY

1.1 Subjects

Forty-six volunteer subjects were recruited for this study, consisting of 26 males between the ages of 21 and 70 (average age = 31 years) and 20 females between the ages of 21 and 36 (average age = 27 years). All subjects were native American English speakers, with no known speech pathologies.

1.2 Sample Recording

Subjects were seated in an Industrial Acoustics Company sound isolation booth, approximately 2.3 m high x 2.3 m wide x 2.3 m deep, and spoke into a Neumann U-87 microphone with an omni-directional pick-up pattern, at a mouth-to-microphone distance of approximately 2 inches (5 cm). A pop screen was used to prevent clipping due to air turbulence from the lips at this close distance. Subjects' speech samples were recorded using Digidesign Pro Tools (version 5.2) audio capture software on a Macintosh G4 computer and a 24-channel Yamaha O2R Digital Recording Console. Subjects wore AKG K240 Monitor semi-open headphones, which allowed them to receive instructions from investigators outside the booth while still allowing them to hear their own voice, for more natural voice production.

For this study, subjects were asked speak their names as they would normally introduce themselves, using the carrier phrase "My name is ______," in their normal speaking voice, at the pace of their own choosing, and at a comfortable pitch and loudness. The phrase was repeated three times, with sufficient pauses between each repetition so as to prevent them from becoming a single phrase. Subjects were instructed to use the same inflection for each repetition. As a foil, each subject was also asked to speak the phrase "My name is Joseph Stalin," repeating it three times as above. This name was selected for its potential to evoke an emotional response in most subjects that could possibly create greater intra-speaker variability than a person's actual name. This could be useful in future SID investigations.

All the spoken material from each subject was recorded to a *.wav* file at a sampling rate of 48K with a 16-bit quantization level. Recording reference levels were approximately -18dB with respect to full scale. No equalization or dynamic range compression was used. The Pro Tools software was used to create a separate *.wav* file for each repetition, yielding 6 speech samples for each individual.

1.3 Analysis

Each speech sample was analyzed to extract a set of 11 parameters: (1) average fundamental frequency (F0) of the speech signal over time, (2) standard deviation (SD) of the F0 over time, (3) the approximate absolute value of the slope of the F0 contour, (4) the sign (positive or negative) of the slope of the F0 contour, (5) the average energy (related to the sound pressure level (SPL), of the speech signal) over time, (6) the SD of the energy of the speech signal over time, (7) the spectral energy contained in the band from 50 Hz to 1000 Hz, (8) the spectral energy contained in the band from 1000 Hz to 5000 Hz, (9) the dB ratio of the energy in 50-1000 Hz band to the energy in the 1000-5000 Hz band (also known as the alpha ratio), (10) the average speaking rate, calculated as the inverse of the average vowel duration in the spoken sentence (a rough measure of the number of syllables per second), and (11) the total duration of the spoken sentence.

The Pitch Contour, Energy Contour, FFT Power Spectrum analyses were done in MultiSpeech (Kay Elemetrics Corp., Lincoln Park, NJ, Model 3700, version 2.5, copyright 2000-2002), to yield the intermediate data for the first 9 parameters above. Figures 2 through 4 show examples of the Pitch Contour, Energy Contour, and FFT Power Spectrum analyses, respectively. The numerical results of each MultiSpeech analysis were saved into ASCII text files for subsequent processing using Matlab (The MathWorks, Inc., Natick, MA, Version 6.5, Release 13, copyright 1984-2002).

Approximate vowel durations and total duration of the spoken sentences were obtained using a custom program to detect phoneme boundaries and durations from the *.wav* file of each sample and a transcript of the speech. The output of the program was an ASCII text file with the list of phonemes in each speech sample and their durations, which were processed using Matlab to extract the speaking rate and total duration values. Figure 5 shows an example of a speech waveform and its corresponding spectrogram, with the phonetic transcription below the spectrogram and the phoneme boundaries overlaid on the figure.

1.4 Speaker code generation and code-matching program

Once all of the parameters were extracted for each sample, a composite 11-element code vector was created for each individual, consisting of the average of each parameter from the three repetitions of the phrase "My name is _______," when the individual was speaking their true name. This vector is the numerical composite of the person's speech waveform. An example of the code vector is:

V = [113.97, 7.50, 0.06, 1, 59.95, 9.19, 63.22, 43.73, 19.49, 10.23, 1.39] where the elements correspond to the eleven parameters listed in section 1.3.

A Matlab script was written to test whether these composite code vectors could be used as a template for matching an unknown speaker to one of a set of known speakers. The task was to compare the code vector of a single speech sample from an arbitrarily chosen individual to the composite code vectors of all individuals. To accomplish this, an eleven-column matrix was constructed from all the subjects' composite code vectors, with the number of rows corresponding to the number of subjects. Four separate cases were considered: (1) All Males, speaking their true names, (2) All Males, speaking the name "Joseph Stalin," (3) All Females, speaking their true names, and (4) all Females, speaking the name "Joseph Stalin." Male and female subjects were considered separately in order to test whether the matching technique and the parameters chosen would work better for males or females. Cases (1) and (3) represent situations in which the individual is being cooperative; i.e., he or she wants to be correctly identified. Cases (2) and (4) might represent situations the individual is being uncooperative, i.e., using a false name to subvert the identification procedure.

For the first case, the number of male subjects was 26, therefore a 26 x 11 element matrix *M* was constructed having the form

$$M = \begin{bmatrix} S1param1 & S1param2 & S1param3 & \ldots & S1param11 \\ S2param1 & S2param2 & S2param3 & \ldots & S2param11 \\ S3param1 & S3param2 & S3param3 & \ldots & S3param11 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ S26param1 & S26param2 & S26param3 & & S26param11 \end{bmatrix}$$

where *S1, S2, S3, ..., S26* are the subject designations and *param1, param2, param3, ..., param11* are the parameter designations. The matching process was implemented by subtracting the unknown speaker's single-sample code vector ***V*** from each row of the matrix ***M*** to create a difference matrix ***D***, and then searching for all the elements in ***D*** that were less than a predetermined decision criterion *r*. This decision criterion was selected on the basis of intra-speaker variability observed for each parameter, and for this study a value of $\pm$ 5% was chosen. The elements of ***D*** that were less than r were replaced with a "1", representing a parameter value that was matched to within $\pm$ 5%, and all other values were replaced with a "0." Then the elements in each row were summed to indicate the number of matches for each of the 26 subjects, and the unknown speaker was identified as the subject with the most matched parameters. The program also kept track of whether there were any "ties" for first place, i.e., more than one of the known speakers' composite code vectors matched the unknown speaker's single-sample code vector with the same number of parameters. This situation, although not leading to a single correct answer, could nonetheless be considered as a positive outcome, since it successfully narrowed down the field of many speakers to a few (2-5) possibilities. The program also kept track of how many parameters were matched in a correct identification, and kept a tally of how many times each parameter was matched in the set of correct matches counting ties for first place. Finally, for incorrect matches, the program also determined if the correct choice would have been the second highest match, or among the ties for second place, thus giving some indication of whether the program came "close" to the correct answer.

All the individual code vectors for each of the three repetitions of all 26 male subjects speaking their true names were run through the Matlab program in order to come up with a set of statistics to evaluate the performance of the matching program with the set of parameters selected for this study. The procedure was then repeated for the other cases. For case (1), there were 78 individual samples, but in case (2) there were only 77 individual samples because one of repetitions from one of the subjects was not used due to audio clipping. There were 18 females and 54 individual samples in case (3), and 19 females and 57 samples in case (4). In both of these latter two cases the number of female subjects were less than twenty, because one or two of the subjects' sample sets were incomplete.

2. RESULTS AND DISCUSSION

The results of the matching process were automatically tabulated and printed out by the Matlab program. The results (at the end of this paper) for all males, Cases (1) and (2), are shown in Table 1 and for all females, Cases (3) and (4), in Table 2. Table 3 shows the results for Case (5), the case in which all males speaking the name "Joseph Stalin" were compared to the database of all males speaking their true names.

2.1 Male Subjects, Cases (1) and (2)

For Case (1), the percentage of correct identifications was 92%, counting ties for first place. Ties for first place can be considered a favorable outcome, because at least it narrows down a large field of possibilities (or suspects) to just a few individuals, and detailed analysis can be carried out on a smaller number of individuals to determine the correct identification. When ties for first place were not considered, the percentage of correct identifications is 79%.

For Case (2), the percentage of correct identifications were lower than when the subjects were speaking their true names: 84% counting ties for first place, and 73% not counting ties for first place. A possible reason for the decrease in performance from Case (1) to Case (2) is that the first level of pattern matching, that which was dependent on the same name being spoken by the unknown and the known speaker, was removed in Case (2), where all male subjects spoke the phrase "My name is Joseph Stalin." However, the relatively high percentages of correct matches even without the benefit of the subjects speaking their own names seems to indicate that a high degree of identifying information was captured in the acoustic parameters extracted from the speech samples.

The Matlab program kept track of the number of matching parameters in the correct identifications, and also the number of times each parameter was matched. There was only one instance where all 11 parameters were matched to make a correct identification, in Case (1). Ten parameter matches were seen occasionally among the male speakers, 5 times in Case (1) and 4 times in Case (2), but it was much more common for the number of matching parameters to range from 6 to 9. In Case (1), 8 parameters were matched most often (20 times). In Case (2), 7 parameters were matched most often (21 times). In both cases, 5 parameters were matched only occasionally, 5 times in Case (1) and 3 times in Case (2), and there was never an instance in either case when fewer than 5 parameters were matched. Thus at least 81% of the time in Cases (1) and (2) for all males, a correct identification was made on the basis of 6 or more matching parameters out of 11, indicating that the matches were based on strong similarities between the unknown and known speaker samples, and not just on one or two matching parameters.

Also tabulated by the Matlab program was the number of times out of all the correct identifications each parameter was matched. The average F0 parameter matched nearly every time, i.e., 71 out of 72 times in Case (1) and 63 out of 65 times in Case (2). Average energy matched every time. Other high performers were the sign of the slope of the F0 contour (but not the value of the slope) the spectral energy in the 5000 Hz band, the spectral energy in the 1000-5000 Hz band (but not the alpha ratio), and the total duration of the phrase. Energy standard deviation and speaking rate also matched a large number of times. This data provides a good indication of the discrimination power of the parameters used in the pilot study, and will be useful in determining which parameters to use in the future and which will need to be replaced with other parameters.

By tracking the number of matches when the identifications were incorrect, it was found that one of the second highest matches would have been correct 5 out of 6 times (83% of the time) in Case (1), and 6 out of 12 times (50% of the time) in Case (2). This indicates that even when the matching procedure was not able to make a correct identification, it was still "close."

2.2 Female Subjects, Cases (3) and (4)

For the cases involving female speakers, the percentage of correct matches when speaking their true names were 85% counting first place ties and 69% not counting first place ties. When speaking the name "Joseph Stalin," the percentage of correct matches were 84% counting ties and 63% not counting ties. It can be seen from these results that the code-matching technique performed better for male speakers than for females. The significance of this needs to be determined in the future. It will also be important to study how this methodology performs when male and female speakers are lumped together for the speaker identification task.

The tally of the number of matching parameters in the correct identifications show that overall, correct identifications among female subjects were made on the basis of fewer matched parameters than among the males. There were no instances in which all 11 parameters were matched, or even 10 parameters. There were 2 instances where 9 parameters were matched to make a correct identification, in Case (3). For the most part, the number of matching parameters ranged from 5 to 8. In Case (3), 7 parameters were matched most often (13 times), while in Case (4), 6 parameters were matched most often (16 times). In both cases, 4 parameters were matched only occasionally, once in Case (3) and 4 times in Case (4), and there was never an instance in either case when fewer than 4 parameters were matched. Thus for both cases involving female subjects, a correct identification was made on the basis of 6 or more matching parameters out of 11 only about 60% of the time, compared to about 80% of the time for the cases involving male subjects.

As in the cases with the male subjects, the parameters that were matched most often in correct identifications were average F0, average energy, and spectral energy in the 50-1000 Hz band. The average F0 parameter matched 44 out of 46 times in Case (3), but only 38 out of 48 times in Case (4). Average energy matched every time, as did the spectral energy in the 50-1000 Hz band. Spectral energy in the 1000-5000 Hz band, and the total duration of the phrase were also high performers. In contrast to the cases with male subjects, the sign of the slope of the F0 contour was not matched very often.

For incorrect identifications, it was found that one of the second highest matches would have been correct 6 out of 8 times (75% of the time) in Case (3), and 5 out of 9 times (56% of the time) in Case (4).

2.3 Male Subjects, Case (5)

For Case (5), in which all males speaking the name "Joseph Stalin" were compared to the database of all males speaking their true names, the percentage of correct identification counting ties for first place was only 51%, about the same as chance. When ties for first place were not considered, the percent correct was a meager 27%.

This case illustrates the importance of using the same pattern for matching one sample of a person's speech to another; i.e., only recordings of the exact same utterance can be used for the identification process. A negative consequence of this is that a person using a shortened form of his or her name or a nickname would likely not be identified by this process, because he or she would not be recognized as a member of the group of known speakers.

3 CONCLUSION

The initial results of the code-matching program indicate that a numeric code made up of the values of acoustic parameters extracted from the speech samples of individuals speaking their names can indeed be used for the purpose of speaker identification with reasonable reliability. Of the extracted parameters tested, the most discriminating were average fundamental frequency (F0), standard deviation of F0 over time, the average energy of the speech signal (related to SPL) over time, the standard deviation of the energy of the speech signal over time, the alpha ratio, the average speaking rate, and the total duration of the spoken sentence. Spectral energy in the band from 50 Hz to 1000 Hz and the spectral energy in the band from 1000 Hz to 5000 Hz will most likely be subject to amplitude variations and background noise, and a better measure might be the long term average spectra (LTAS) over time. This will need to be tested in future work.

The least discriminating of the extracted parameters tested were F0 slope and the sign of the F0 slope.

The next phase of this project will involve a more complete study of additional parameters, to identify those that are more robust and have intra-speaker variations of less than 5%. It is desirable that all parameters be amplitude-independent and reasonably noise resistant, in order to be more effective in the field, where voice identifications are likely to be made. Such parameters would make this method more applicable to real-life situations involving bandwidth-limited signals (i.e., telephones, cell phones, RF microphones, etc.) with significant background noise.

Further study is needed to determine why a greater percentage of correct identifications were made among male speakers than female speakers. Further research is needed to find parameters that will provide better results for female speakers. Additional testing with a larger number of samples will be done to help determine parameter weightings for more discriminating algorithms.

REFERENCES

[1] Campbell, J.P., Jr. Speaker recognition: A tutorial. Proc. IEEE, Vol. 85, No. 9, Sept. 1997, pp.

[2] Kersta, L.G. Voiceprint identification. Nature 196 (1962).

[3] Tosi, O., Oyer, H., Lashbrook, W., Pedrey, C., Nichol, J. & Nash, E. Experiment on voice identification, J. Acoust. Soc. Am., Vol. 51, No. 6 (Part 2), June 1972, pp. 2039-43.

[4] Posa, F., and Begault, D. Voice identification and elimination using aural-spectrograhic protocols. Proceedings of the AES 26th International Conference, Denver, CO, July 2005, pp. 21-28.

[5] Markel, J.D., Gray, A.H. Linear Prediction of Speech, Communication and Cybernetics Series, 12, Springer-Verlag, Berlin - New York: 1976; pp.

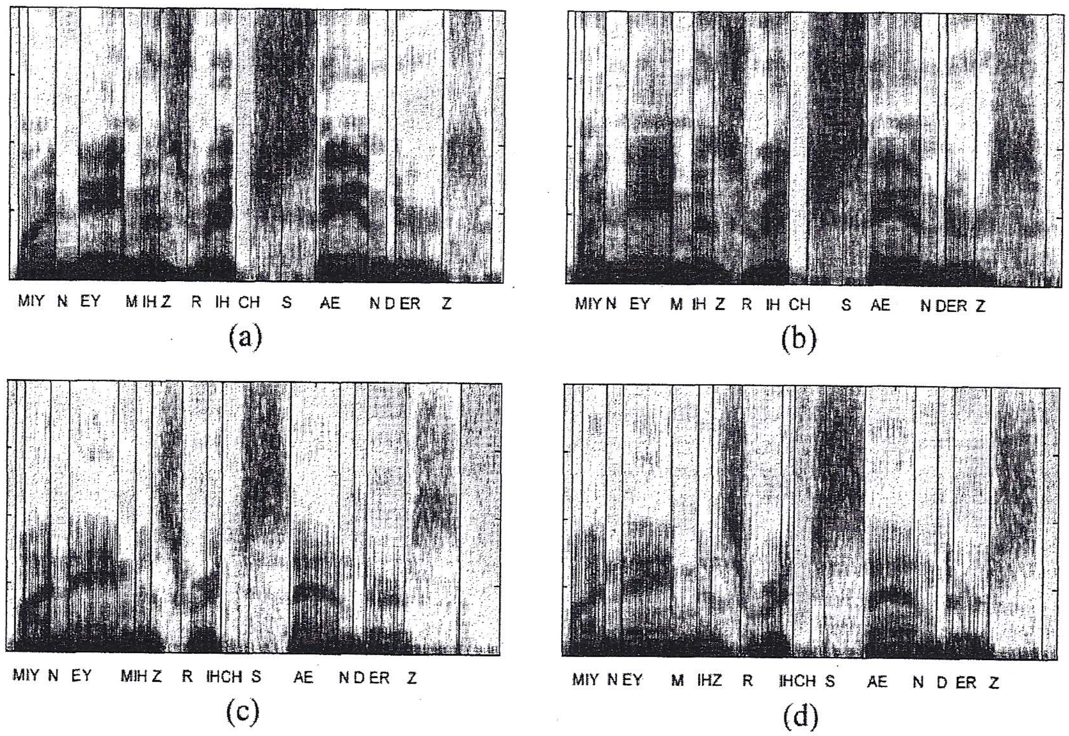

Figure 1. Spectrograms of two different individuals speaking the same phrase, "My name is Rich Sanders." The ones in (a) and (b) are spoken by the true Rich Sanders, and the ones in (c) and (d) are spoken by an imposter.

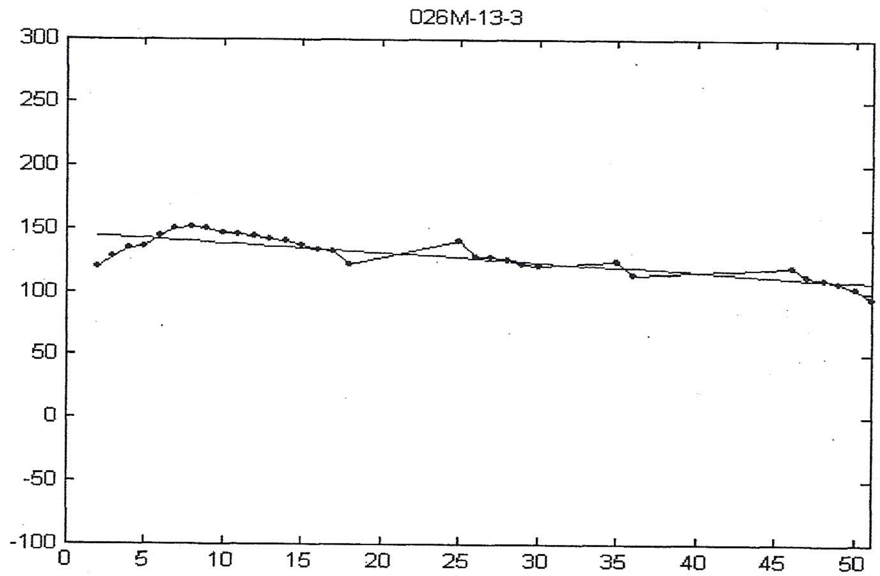

Figure 2. Example of a fundamental frequency (F0) contour from a male subject.

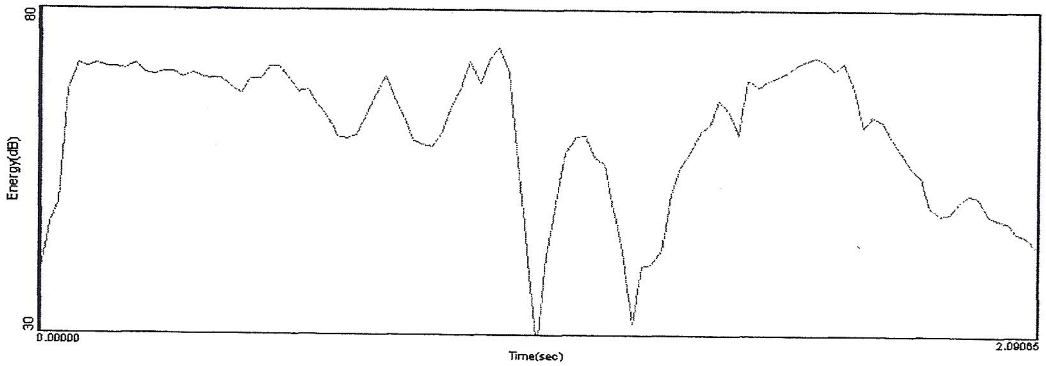

Figure 3. Example of an energy contour from a male subject.

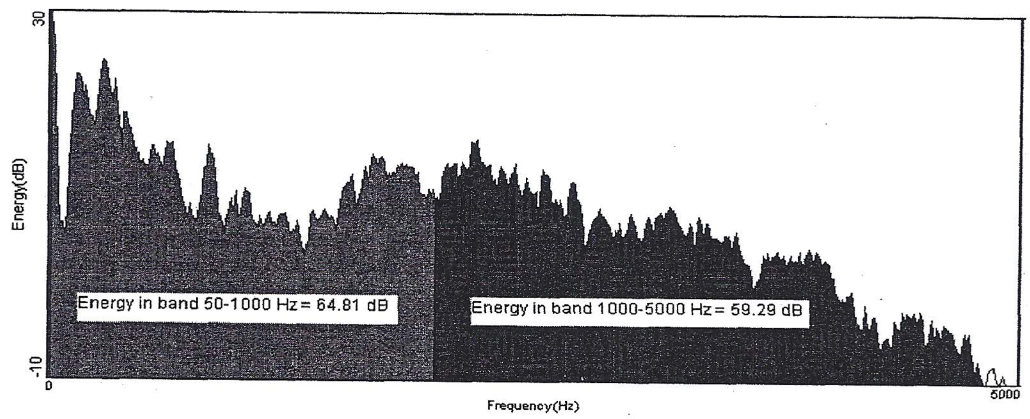

Figure 4. Example of a long-term average spectrum showing how alpha ratio is calculated: alpha ratio = [Energy in band 50-1000 Hz] dB – {Energy in band 1000-5000 Hz] dB

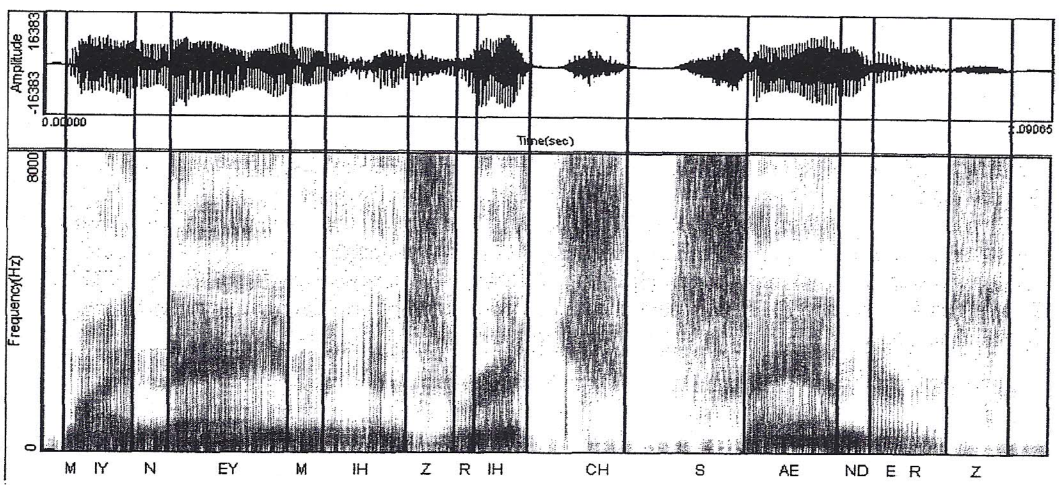

Figure 5. Spectrogram with phoneme boundary identification.

Table 1. Results for Cases (1) and (2) all male speakers

Test condition: All Males, speaking their true names	Test condition: All Males, speaking the name "Joseph Stalin"
Decision criteria = 5% GLOBAL	Decision criteria = 5% GLOBAL
Number incorrect = 6 out of 78	Number incorrect = 12 out of 77
Number correct = 72 out of 78 (counting ties for 1st place) Percent correct = 92%	Number correct = 65 out of 77 (counting ties for 1st place) Percent correct = 84%
Number correct = 62 out of 78 (not counting ties for 1st place) Modified Percent correct = 79%	Number correct = 56 out of 77 (not counting ties for 1st place) Modified Percent correct = 73%
For correct matches (including ties for 1st place): 11 parameters were matched 1 times 10 parameters were matched 5 times 9 parameters were matched 9 times 8 parameters were matched 20 times 7 parameters were matched 17 times 6 parameters were matched 15 times 5 parameters were matched 5 times 4 parameters were matched 0 times 3 parameters were matched 0 times 2 parameters were matched 0 times 1 parameters were matched 0 times	For correct matches (including ties for 1st place): 11 parameters were matched 0 times 10 parameters were matched 4 times 9 parameters were matched 11 times 8 parameters were matched 18 times 7 parameters were matched 21 times 6 parameters were matched 8 times 5 parameters were matched 3 times 4 parameters were matched 0 times 3 parameters were matched 0 times 2 parameters were matched 0 times 1 parameters were matched 0 times
F0av matched 71 times F0sd matched 29 times F0sl matched 11 times F0sn matched 58 times ENav matched 72 times ENsd matched 46 times PSD1 matched 68 times PSD5 matched 55 times ALPH matched 21 times RATE matched 47 times DUR matched 58 times	F0av matched 63 times F0sd matched 25 times F0sl matched 14 times F0sn matched 62 times ENav matched 65 times ENsd matched 40 times PSD1 matched 65 times PSD5 matched 56 times ALPH matched 15 times RATE matched 34 times DUR matched 54 times
For incorrect matches: Correct match was either the speaker with the 2nd highest # of matches, or among the ties for second place, 5 out of 6 times	For incorrect matches: Correct match was either the speaker with the 2nd highest # of matches, or among the ties for second place, 6 out of 12 times

Table 2. Results for Cases (3) and (4), all female speakers

Test condition: All Females, speaking their true names Decision criteria = 5% GLOBAL	Test condition: All Females, speaking the name "Joseph Stalin" Decision criteria = 5% GLOBAL
Number incorrect = 8 out of 54	Number incorrect = 9 out of 57
Number correct = 46 out of 54 (counting ties for 1st place) Percent correct = 85%	Number correct = 48 out of 57 (counting ties for 1st place) Percent correct = 84%
Number correct = 37 out of 54 (not counting ties for 1st place) Modified Percent correct = 69%	Number correct = 36 out of 57 (not counting ties for 1st place) Modified Percent correct = 63%
For correct matches (including ties for 1st place): 11 parameters were matched 0 times 10 parameters were matched 0 times 9 parameters were matched 0 times 8 parameters were matched 12 times 7 parameters were matched 13 times 6 parameters were matched 9 times 5 parameters were matched 11 times 4 parameters were matched 1 times 3 parameters were matched 0 times 2 parameters were matched 0 times 1 parameters were matched 0 times	For correct matches (including ties for 1st place): 11 parameters were matched 0 times 10 parameters were matched 0 times 9 parameters were matched 2 times 8 parameters were matched 8 times 7 parameters were matched 8 times 6 parameters were matched 16 times 5 parameters were matched 10 times 4 parameters were matched 4 times 3 parameters were matched 0 times 2 parameters were matched 0 times 1 parameters were matched 0 times
F0av matched 44 times F0sd matched 20 times F0sl matched 2 times F0sn matched 3 times ENav matched 46 times ENsd matched 21 times PSD1 matched 46 times PSD5 matched 33 times ALPH matched 11 times RATE matched 35 times DUR matched 39 times	F0av matched 38 times F0sd matched 8 times F0sl matched 3 times F0sn matched 4 times ENav matched 48 times ENsd matched 31 times PSD1 matched 48 times PSD5 matched 41 times ALPH matched 14 times RATE matched 28 times DUR matched 37 times
For incorrect matches: Correct match was either the speaker with the 2nd highest # of matches, or among the ties for second place, 6 out of 8 times	For incorrect matches: Correct match was either the speaker with the 2nd highest # of matches, or among the ties for second place, 5 out of 9 times

Table 3. Results for Case (5), all male speakers

Test condition: All Males, speaking the name "Joseph Stalin", compared to the database of their true names
Decision criteria = 5% GLOBAL

Number incorrect = 38 out of 77

Number correct = 39 out of 77 (counting ties for 1st place)
Percent correct = 51%

Number correct = 21 out of 77 (not counting ties for 1st place)
Modified Percent correct = 27%

For correct matches (including ties for 1st place):
11 parameters were matched 0 times
10 parameters were matched 0 times
9 parameters were matched 0 times
8 parameters were matched 7 times
7 parameters were matched 5 times
6 parameters were matched 17 times
5 parameters were matched 7 times
4 parameters were matched 3 times
3 parameters were matched 0 times
2 parameters were matched 0 times
1 parameters were matched 0 times

F0av matched 33 times
F0sd matched 13 times
F0sl matched 4 times
F0sn matched 32 times
ENav matched 36 times
ENsd matched 11 times
PSD1 matched 37 times
PSD5 matched 29 times
ALPH matched 12 times
RATE matched 16 times
DUR matched 17 times

For incorrect matches:
Correct match was either the speaker with the 2nd highest # of matches,
or among the ties for second place, 22 out of 38 times