
Formation of the Final Outcome in Complex Research

Latif P. Aliyev¹, Mammad R. Rajabov², Bakhtiyar G. Pashayev^{*1}

¹Department of Optics and Molecular Physics, Baku State University, Baku, Azerbaijan

²Department of Theoretical Physics, Baku State University, Baku, Azerbaijan

Received 18-Sep-2025; Accepted 07-Nov-2025

DOI: <https://doi.org/10.30546/209501.101.2025.2.04.0117>

Abstract

The study analyzes the separate evaluation of the results obtained from auditory, linguistic, and instrumental analyses applied in forensic phonoscopic examination and identifies the limitations of the existing “Y” probability scale. In order to generalize the results of complex research within a unified mathematical model, a new probability scale termed “X” is proposed, which is constructed based on the theory of mathematical probability. The numerical mean of the results obtained through different analytical methods is calculated. The conducted analysis demonstrates that the proposed probability scale enables an objective and measurable assessment of the results of complex studies. The new approach reduces subjectivity in the final expert conclusion, enhances the level of scientific justification of the findings, and provides a more reliable methodological basis for the practical application of forensic phonoscopic examination.

Keywords: forensic phonoscopic examination, speaker and speech identification, probability scale, complex research

PACS Numbers: 43.75.Wx; 43.60.Tj; 43.60.Hj

1. Introduction

Sound is a mechanical wave that propagates through an elastic medium and is formed as a result of aeroacoustic vibrations generated in the vocal apparatus during speech production. The vibration mode of the speaker’s vocal folds, the resonant characteristics of the vocal tract, and the position of the articulatory organs are among the primary physical factors determining the acoustic structure of the

*Corresponding author – Tel.: (+994) 50 394 78 00

e-mail: bakhtiyarpashaev@bsu.edu.az ; ORCID ID: 0009-0004-1051-8943.

speech signal. Each of these processes produces measurable changes in the frequency spectrum, amplitude distribution, and temporal parameters of the sound wave. Consequently, the speech signal functions as an individual physical-acoustic "trace," becoming a carrier of information for personal identification. Acoustic analysis of speech is performed through instrumental measurement of the waveform, fundamental frequency (ν_0), formant frequencies ($\nu_1 \div \nu_4$) - where ν_1 is mainly related to the degree of oral cavity opening, ν_2 to the front-back position of the tongue and the effective length of the vocal tract, ν_3 to the individual geometric structure of the vocal tract, lateral tongue movement, and dental-labial configuration, and ν_4 primarily to the overall dimensions of the vocal tract and the high-frequency characteristics of the resonator - as well as spectral energy distribution and intensity. These parameters are closely associated with the dimensions and elastic properties of the vocal tract and the dynamics of airflow and are physically deterministic in nature. However, in real forensic phonoscopic materials, recorded signals are often accompanied by noise, reverberation, channel distortions, and non-linear effects, which lead to variations in measured acoustic parameters and increase the uncertainty of the results [1-3].

For this reason, instrumental-acoustic measurements alone are insufficient for speaker and speech identification, and they are applied in combination with auditory and linguistic analyses in a comprehensive manner. In modern forensic expert practice, the application of integrated approaches is of particular importance when examining complex and multicomponent objects [4-10]. Nevertheless, in existing methodologies, the results of these diverse physical and psychoacoustic measurements are often evaluated separately, which introduces subjective elements into the formation of the final conclusion. This represents one of the factors that reduce the objectivity and reliability of expert opinions. In this context, the synthesis of results obtained through various acoustic and non-acoustic analytical methods within a unified mathematical probability model is of significant scientific importance [5-7]. The present study proposes a new probability scale for the objective evaluation of the results of complex speech and voice examinations. The proposed approach is based on the theory of mathematical probability and aims to address the existing methodological gap. Its application enables a more substantiated and transparent formulation of the final expert conclusion.

As noted above, in forensic research, the formation of a final opinion through the synthesis of results obtained using multiple research methods (complex examinations) to address the same issue is one of the key tasks. A complex examination involves a multidimensional study of an object using methods from various scientific disciplines in order to obtain a complete and objective result and to assess the situation as a whole. It represents a unique type of investigation conducted through the simultaneous application of knowledge acquired in science and technology to

establish the facts of a case. Its primary purpose is to resolve complex, multifaceted issues that extend beyond the competence of a single expert or a specific field of study. This approach enables comprehensive problem analysis, identification of cause-and-effect relationships, and the derivation of objective and scientifically substantiated conclusions that are essential in judicial proceedings and pretrial dispute resolution [4, 8]. Unlike commission-based examinations, in which several specialists from the same field jointly produce a common opinion, complex examinations involve experts from different disciplines. For example, in speaker and speech identification, three types of analytical methods are employed: auditory, linguistic, and instrumental analyses [4, 9].

This synthesis of knowledge allows for a comprehensive analysis of all aspects of an event, identification of latent factors, and the provision of objective answers to questions posed by the court or investigator. Complex examinations play a crucial role in establishing the truth, particularly in cases where simple logic or one-sided analysis is insufficient for rendering a fair decision [4, 8, 9, 10]. By combining various methods and approaches, complex examinations contribute to revealing the complete picture of an event, clarifying the most challenging situations, and supporting accurate legal decision-making. They are indispensable in resolving disputed issues that require deep and multifaceted analysis and enable specialists from different fields to unite their efforts in a comprehensive investigation.

The aim of this study is to reduce subjectivity arising in the process of synthesizing results obtained through auditory, linguistic, and instrumental analytical methods in speaker and speech identification and to enhance the objectivity of the final expert conclusion.

2. Experimental Section

By their nature, the objects and materials required for complex examinations are highly diverse. Each component of a complex examination requires the provision of appropriate materials in accordance with legal requirements and the methodologies applicable to the respective type of analysis.

The research objects may include the following:

Documents: electronic documents, images, audio and video recordings, and correspondence (both paper-based and electronic).

Digital storage media: computers, laptops, mobile phones, tablets, servers, flash drives, hard drives, and related devices.

In the auditory analysis process, the expert linguist examines groups of intrinsic factors that influence variability in voice and speech. These factors are associated with the individual anatomical characteristics of the vocal tract organs involved in voice and speech production, as well as the speaker's gender and age. During audi-

tory analysis, the expert identifies speaker-specific features characterizing articulation, voice timbre, melodic patterns, rhythm, speech rate, dialectal characteristics, and individual speech habits, including emotional coloring, fluency, and dynamism. These identifying features are compared with the corresponding parameters in the experimental phonogram, and the results of the auditory examination are formulated. In other words, auditory examination involves comparing corresponding features in the disputed phonogram and the comparative phonogram presented as an experimental voice and speech sample, and determining the proportion of similar and dissimilar features.

Following the same approach, the expert linguist identifies identifying features at the level of linguistic units in the speech material and conducts comparative analysis. Speech in each phonogram submitted for examination is analyzed separately, linguistic features are examined, and the probability that the individuals recorded in the disputed and comparative phonograms are the same person is assessed.

Instrumental analysis of the disputed and comparative phonograms submitted for examination may be carried out using various methods depending on the quality of the examined material. The expert acoustician performs instrumental analysis by measuring the prosodic and spectral characteristics of the speech signals recorded in the disputed and comparative phonograms and determines the degree of similarity or difference between the compared speakers. At the final stage of the examination, the results obtained through auditory, linguistic, and instrumental methods are synthesized and generalized, and the final expert conclusion is formulated.

3. Discussion of Results

In forensic phonoscopic examination, the expert opinion in speaker and speech identification is formed by synthesizing the results of auditory, linguistic, and instrumental analyses and is expressed as follows:

a) the voice and speech of the speaker designated as K1 in the transcript of the disputed phonogram belong to Salamov Salam Salam oglu, whose experimental voice and speech sample was provided;

b) the voice and speech of the speaker designated as K1 in the transcript of the disputed phonogram probably belong to Salamov Salam Salam oglu, whose experimental voice and speech sample was provided;

c) it was not possible to determine whether the voice and speech of the speaker designated as K1 in the transcript of the disputed phonogram belong to Salamov Salam Salam oglu, whose experimental voice and speech sample was provided, or to another person;

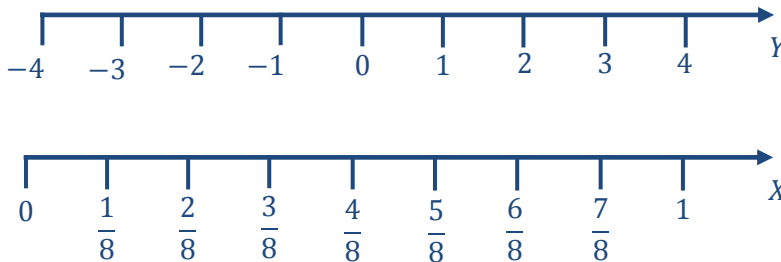
d) the voice and speech of the speaker designated as K1 in the transcript of the disputed phonogram probably do not belong to Salamov Salam Salam oglu, whose

experimental voice and speech sample was provided, but to another person;

e) the voice and speech of the speaker designated as K1 in the transcript of the disputed phonogram do not belong to Salamov Salam Salam oglu, whose experimental voice and speech sample was provided, but to another person.

In speaker and speech identification, the assessment for each analytical method is conducted using the “Y” scale. For example, if an instrumental analysis shows more than 75% similarity between the examined features of the disputed and comparative phonograms (corresponding to values between 3 and 4 on the scale), the conclusion specified in item (a) is drawn for that analytical method. If the similarity ranges from 25% to 75% (values between 1 and 3 on the scale), the conclusion specified in item (b) is formed. If similarity is less than 25% or dissimilarity is less than 25% (values between -1 and 1 on the scale), the conclusion specified in item (c) is obtained, and so forth. However, since existing methodologies do not provide a specific computational procedure for forming a final conclusion through the synthesis of results from all three analytical methods, a degree of subjectivity is inevitably introduced. In complex examinations, replacing the “Y” scale with the newly proposed “X” scale allows, in our view, for an objective assessment in calculating the final conclusion. It should be noted that the relationship between the selected Y and X scales is defined as:

$$X = \frac{Y + 4}{8} \quad (1)$$



Here, the above scale represents a probability scale corresponding to the evaluation of an individual research method considered separately, whereas the scale below represents a new probability scale intended for the formation of the final conclusion through the synthesis of the results of all three analytical methods. As is known, mathematical probability is a quantity measured as the ratio of the number of favorable outcomes for a given event to the total number of equally likely outcomes:

$$P(A) = \frac{m}{n} \quad (2)$$

Here, n – denotes the number of equally likely outcomes, while m – denotes the number of favorable outcomes. Since the number of favorable outcomes for any event cannot exceed the total number of possible outcomes, it follows from the classical definition of probability that mathematical probability varies between zero and one ($0 \leq P(A) \leq 1$). This requirement is satisfied by the probability scale proposed in this study.

Assume that, as a result of instrumental analysis, the identifying features in the disputed and comparative phonograms demonstrate 75% similarity ($Y = 3$), linguistic analysis indicates 25% similarity ($Y = 1$), and auditory analysis indicates 25% dissimilarity ($Y = -1$). According to the Y scale, the probability that the compared speakers are the same person equals 75%, which does not adequately reflect the actual situation. According to the proposed probability scale, the following result is obtained:

Considering that when $Y = 3$, $X_1 = 7/8$; when $Y = 1$, $X_2 = 5/8$, and when $Y = -1$, $X_3 = 3/8$, the final probability obtained using three different analysis methods is determined as the arithmetic mean of the probabilities:

$$\bar{X} = \frac{X_1 + X_2 + X_3}{3} = \frac{7/8 + 5/8 + 3/8}{3} = \frac{15}{24} \quad (2)$$

Thus, the probability that the compared individuals are the same person is $\bar{X} = (15/24) \cdot 100\% = 62.5\%$, which reflects the objective result.

4. Conclusion

The conducted research indicates that, in the identification of an individual based on voice and speech, merely providing a qualitative summary of auditory, linguistic, and instrumental analyses is insufficient for forming an objective final opinion. The newly proposed “X” probability scale (MLB scale) presented in this paper enables the synthesis of results from complex studies based on mathematical foundations. This approach reduces the risk of subjective evaluation and enhances the scientific reliability of the expert opinion. The application of the proposed model in forensic phonoscopic examinations can be considered both practically feasible and promising.

References

- [1] Fant G. Acoustic Theory of Speech Production. Mouton, The Hague, 1960, 328 p.
- [2] Stevens K.N. Acoustic Phonetics. MIT Press, Cambridge, 1998, 604 p.
- [3] Morse P.M., Ingard K.U. Theoretical Acoustics. Princeton University Press, 1986, 927p.
- [4] Aliyev L.P. Fundamentals of Forensic Examination of Audio and Video Recordings (Scientific and Practical Guide). Baku-2025, 171 p.

- [5] Rose P. Forensic Speaker Identification. Taylor & Francis, London, 2002, 360 p.
- [6] Singh S. Forensic and Automatic Speaker Recognition System. International Journal of Electrical and Computer Engineering. 2018, vol. 8, № 5, p. 2804-2811.
<http://doi.org/10.11591/ijece.v8i5.pp2804-2811>
- [7] Rose P. Likelihood ratio-based forensic voice comparison with higher level features: research and reality. Computer Speech & Language. 2017, vol. 45, p. 475-502.
<https://doi.org/10.1016/j.csl.2017.03.003>
- [8] Bali A., Basu N., Weber P., Rosas-Aguilar C., Edmond G., Martire K., Morrison G.S. Speaker identification in courtroom contexts - Part III: Groups of collaborating listeners compared to forensic voice comparison based on automatic-speaker-recognition technology. Forensic Science International. 2024, vol. 360, p. 112048.
<https://doi.org/10.1016/j.forsciint.2024.112048>
- [9] Sergidou E.K., Ypma R., Rohdin J., Worring M., Geradts Z., Bosma W. Fusing linguistic and acoustic information for automated forensic speaker comparison. Science & Justice. 2024, vol. 64, №5, p. 485-497. <https://doi.org/10.1016/j.scijus.2024.07.001>
- [10] David van der Vloed. Data strategies in forensic automatic speaker comparison. Forensic Science International. 2023, vol. 350, p.111790.
<https://doi.org/10.1016/j.forsciint.2023.111790>