



FORENSIC SPEAKER IDENTIFICATION BY THE LINGUISTIC-ACOUSTIC METHOD IN KEÚ AND IES*

Katarzyna KLUS, Agata TRAWIŃSKA

Institute of Forensic Research, Krakow



Education and Culture DG
Lifelong Learning Programme
Leonardo da Vinci

Abstract

The paper presents selected issues associated with the linguistic-acoustic method employed at the Speech and Audio Laboratory in Bratislava (Kriminalistický a expertízny ústav Policajného zboru, henceforth referred to as KEÚ) and the Section for Speech and Audio Analyses, Institute of Forensic Research in Krakow (henceforth referred to as IES). In the first part of their report, the authors define differences between the Polish and Slovak languages, concentrating in particular on such phenomena that should be regarded as the most useful in the process of speaker identification. Next, the report presents similarities and differences between ways of performing identification studies, as well as describing an attempt at evaluating the effectiveness of the adopted solutions. Becoming acquainted with the methodology of recording examinations employed in the Bratislava KEÚ laboratory was possible thanks to the Leonardo da Vinci exchange program, in which one of the authors of this paper participated.

Key words

Speaker identification; Reference material; Linguistic-acoustic analysis of speech; Polish phonetics; Slovak phonetics.

Received 8 January 2009; accepted 12 February 2009

1. The linguistic part

1.1. The Slovak and Polish vowel system

A very significant element of speaker identification studies is analysis of the vowel system used by the speaker. Vowels are voiced segments with a formant structure. Formants, or acoustic speech parameters, individualise the speaker to a much greater degree than dialect-associated properties, since the values of such parameters result from such factors as the anatomical structure of the speaker and his/her articulative habits. In spite of their marked similarities, the Polish and Slovak vowel systems demonstrate some significant differences.

A phenomenon that is universal for all Slavic languages is a vowel system that is based on the coexistence of two oppositions, i.e. front *versus* back and high *versus* mid *versus* low vowel [1]. The simplest representation of such a model is the vowel system of the Macedonian language (Figure 1). The remaining Slavic languages are characterised by systems that constitute some form of modification of the aforementioned minimal system.

In the Slovak language, the set of vowel phonemes is more than twice as extensive as in the case of the scheme presented in Figure 1; this is a consequence of the quantitative opposition that is noted in this language. Phonemic vowel length or the distinction between long and short vowels is at the same time a major difference between the Slovak and Polish languages. In the Polish language, differences in vowel length disappeared in the 16th century and, therefore,

* Based on research performed at the Speech and Audio Laboratory, Police Institute of Criminalistics and Forensic Expertise (KEÚ) in Bratislava and the Laboratory for Speech and Audio Analysis, Institute of Forensic Research (IES), Krakow.

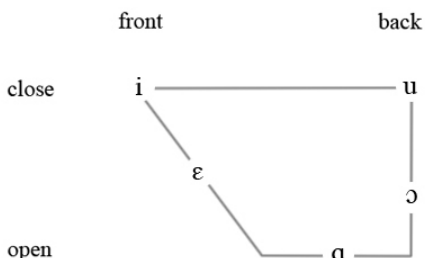


Fig. 1. Minimal vowel system in Slavic languages (the Macedonian language).

in contemporary Polish, the opposition between long and short vowels is not phonological in character. In eastern dialects of Polish, accented vowels are indeed pronounced longer, while non-accented vowels are clearly reduced in duration; nevertheless, describing this phenomenon as phonemic vowel length is erroneous.

Simplified schemes of the vowel system of the Polish and Slovak languages are presented in Figure 2 and 3.

Thus, the Slovak vowel system is considerably richer than the Polish one – in addition to long equivalents of the majority of vowels, the former also has [æ] – a vowel described as front, although more back than [ε], and low, although slightly higher than [a]. The specific character of this vowel consists in, among other things, its restricted distribution – [æ] is a continuant of the former nasal vowel [ɛ̃] and appears only after labial vowels. It should be mentioned here that this phoneme is not known by all Slovak language users – according to the literature on the subject, this allophone is used in everyday speech by approximately 5% of Slovaks, while the rest pronounce it simply as [ε]. Orthoepic standards are not rigorous in this

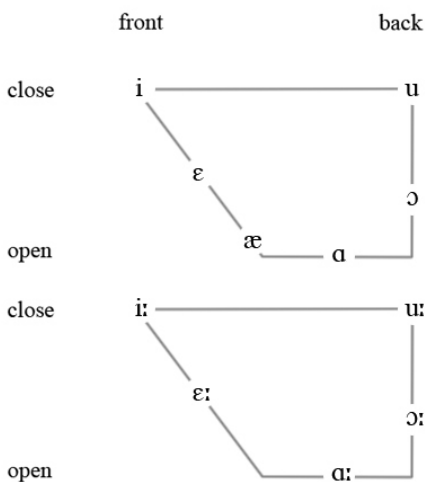


Fig. 2. The vowel system of the Slovak language.

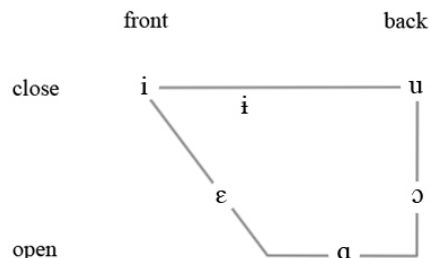


Fig. 3. The vowel system of the Polish language.

respect – in the neutral style, pronouncing [æ] is neither forbidden, nor obligatory. However, the use of this mode of pronunciation is required in the so-called “higher” style, which is employed, for example, by actors when reciting poems or performing parts in classical dramas, where articulation of the [æ] phoneme is a prerequisite for keeping rhythm [8].

A formant analysis of the Slovak vowel system also confirms a very strong tendency towards the [æ] vowel becoming obsolete – the formant values indicate unstable, non-systemic articulation of this vowel [9] (Figure 4).

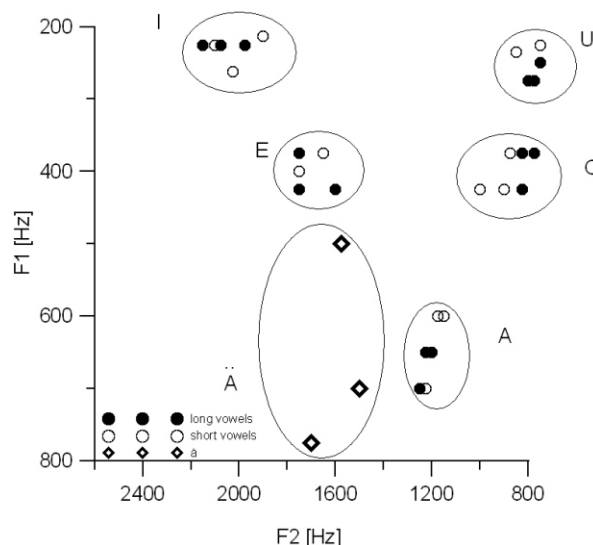


Fig. 4. The mean F1–F2 values for Slovak vowels – 3 speakers [6, 9].

Differences between the Slovak and Polish vowel systems also include the presence of [i] in the Polish language – the vowel is described as front and high, but slightly lower than [i] [3, 15]. The sound, once regarded as a positional variant of [i], is presently much more frequently given the status of a separate phoneme [3]. Although the Slovak language has preserved the grapheme y, and the position at which it occurs is

dictated by the etymological criterion, in pronunciation it is articulated as [i].

Another difference between the systems of the two discussed languages lies in diphthongs, called *dvojhlásky* in the Slovak language, i.e. combinations of [i̯a], [i̯ɛ], [i̯u] or [wo]. They are particularly prone to various types of non-standard articulations, hence their high discriminative value in the process of speaker identification. Diphthongs are articulated with the same duration as long vowels. Polish also has vocoidal groups, or combinations of – for example – the glide [j] and a vowel (e.g. *miasto*, *biel*), but from the phonological perspective, they are not ascribed the status of diphthongs [3].

In turn, the Polish language has so-called nasal vowels, which are unknown in Slovak. They constitute continuants of the proto-Slavic nasal vowels and are articulated in the contemporary general language as combinations of the vowels [ɔ] or [ɛ] and a nasal element, which differs depending on the right-sided phonetic context. To give an example, the nasal vowel transcribed using the grapheme *ą* may be articulated as [ɔ̃m] (e.g. *trąba*), [ɔ̃n] (e.g. *sąd*) or [ɔ̃w] (e.g. *wąs*).

As indicated in studies in the literature on formant frequency for Polish and Slovak vowels [6, 7, 9], the most pronounced differences between the two languages are seen in mid vowels – the Slovak [ɛ] and [ɔ] are more close than in Polish, with the values of the [ɛ] vowel being in some cases closer to the Polish centralised [i] (Figure 5).

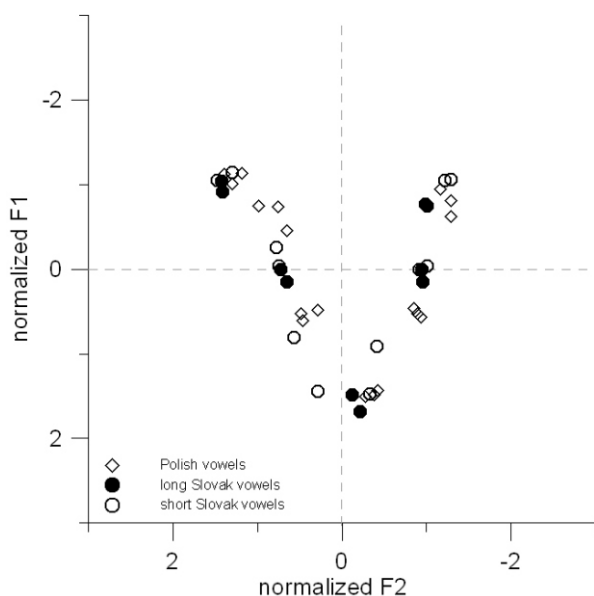


Fig. 5. Comparison of normalised values of F1 and F2 formants – 3 Polish and 2 Slovak speakers [6, 7, 9].

1.2. The Slovak and Polish consonant system

In the Slovak language, there is the [l̥] phoneme, which is absent in the phonological system of the Polish language – in so-called “standard Polish”, the consonant is a positional variant of the [l] phoneme and is characterised by a very limited distribution, occurring solely before the vowel [i] and the sonorant [j]. Another difference is associated with the so-called sonorants – the Slovak language has syllable-forming variants of [r] and [l], which are altogether absent in Polish. In turn, in the Polish language, the sonorant [w] enjoys a phonemic status, while in Slovak [w] is only a positional variant of [v] and is articulated solely when it appears in the medial or final position in the word, before a vowel.

A phoneme which is a source of considerable difficulties experienced by Poles attempting to learn Slovak is the laryngeal [h]. In so-called “standard Polish” there is no discrimination between the pronunciation of *ch* and *h*; the laryngeal variant, which is the execution of the latter grapheme, is present only in some Polish local dialects.

As far as proper consonants are concerned, basic differences are seen in the palatal consonants. In the Slovak consonant system, there are no such consonants as [ç], [ʒ], [tç] and [dʒ]. Diachronic studies of Slavic languages indicate that the Slovak [t,], [d,] are in the majority of cases continuants of the same phoneme combinations, which developed in the Polish language into [tç] and [dʒ].

A significant difference between Slovak and Polish is also noted in the palatalised consonants. The Polish phonetic system distinguishes an opposition of soft versus hard in the vast majority of consonants. Thus, there are palatalised (soft) allophones of particular consonants, articulated when the [i] vowel or the [j] glide occurs in the right-sided phonetic context. No such distinction is observed in Slovak. In the aforementioned positions, consonants are articulated in the same manner with respect to softness versus hardness as when they appear in a neutral context. Here, a question arises as to whether the difference is a result of conventional descriptions in the grammar systems of both languages, or else of genuine discrepancies in the quality of the articulated phoneme combinations. A more comprehensive analysis of the Polish and Slovak phonetic and phonological systems may be found in a paper by Júlia Dudášová-Kriššáková [2].

2. Forensic applications

2.1. Comparative identification studies – theoretical assumptions

The assertion that human speech is a sound representation of a natural language, which – thanks to numerous specific properties – individualises the speaker, is one of the basic premises of criminalistic speaker identification. The speech and audio laboratories worldwide employ various methods, ranging from linguistic, where the analysis focuses solely on linguistic properties, through linguistic-acoustic, which combines elements of linguistic and acoustic analysis, to fully automatic methods, which are usually based on speech signal parameterisation through cepstral coefficients [4, 5, 11, 12]. Identification analyses performed using the linguistic-acoustic method (this method is employed by experts in both the described centres) allow preparation of such a description of the speaker that takes into consideration his/her individual properties resulting from the structure of his/her vocal track and articulators, as well as group properties, for example those common to users of the same dialect.

During speaker recognition examination, the analysis encompasses both individual features, associated with the anatomical structure of the speaker, mainly frequencies of the F1–F4 vowel formants, and group properties acquired in the process of socialisation that takes place in a particular geographic environment through individuals (parents, care-givers), who pass on their own articulatory patterns. The position of dialects is presently much weaker than it was several decades ago. The usage of dialect forms is usually judged as a negative phenomenon, as it is perceived as a limitation of the speaker's linguistic competence; in addition, it is associated with lack of education, sometimes also with a lower social status. This is why speakers strive to eliminate linguistic dialectic variants, especially those that are judged negatively, i.e. those that most obviously differ from the so-called “standard language”. In Polish, such properties include, for example, substitution of dental stops and affricates for alveolar stops and affricates ([t̪s], [z], [s] replacing [tʃ], [ʒ], [ʃ]).

The ability to restrict the use of one of the varieties of language is associated with the speaker's linguistic awareness. The speaker is unable to eliminate properties, of whose existence in his/her phonetic system s/he is completely unaware. It is these types of properties – phonetic properties – that should constitute the basis for a comparative linguistic analysis.

Linguistic dialectic features have a group character, being thus common for language users inhabiting the same geographic area. Theoretically, the discriminative value of a given property is thus greater, the lower the number of users of a given dialect. However, preparing a reliable and complete description of articulative properties of a speaker requires not only demonstration of properties observable in a given material, but also defining of their intensity – or in other words determining whether the tendency to use a particular type of articulation is consistent, strong, moderate or weak.

On the other hand, measurable speech parameters are considered to be properties of a more individual character. Measurements of acoustic speech parameters chiefly focus on the formant structures of vowel segments. To put it simply, formants are frequency bands in which speech signal energy concentrates and which result from acoustic resonance of the sound wave in the supralaryngeal cavities. The values of formants depend on the anatomical structure of the speech organs, but also on the articulatory habits – chiefly on the position of the tongue in the oral cavity. In criminalistic practice, in the majority of cases it is possible to measure the first three, and in some cases four formants.

2.2. Usefulness of evidential material in identification analyses. A preliminary analysis of evidential material

Prior to commencing comparative studies, the expert performs a preliminary assessment of evidential material, which allows determination of whether – from the quantitative and qualitative point of view – utterances included in the evidence provide a basis for performing identification analysis, or in other words, whether the said utterances may be regarded as a representative speech sample of a particular individual. Among other requirements, it is necessary to determine whether the technical parameters of a registered speech fragment allow a reliable extraction of measurable speech parameters. A preliminary analysis is also performed to establish whether distinctive linguistic features can be detected within the utterances, both of a group and individual character. Based on such investigations, the expert also decides on the appropriate formula for the process of collecting reference material in order to get the most adequate sample with respect to the evidence utterances.

At the Speech and Audio Laboratory in Bratislava (Kriminalistický a expertízny ústav Policajného zboru, henceforth referred to as KEÚ) experts have developed

guidelines to be followed when deciding whether given material is suitable for identification studies, and if so – with what degree of probability the final conclusion may be formulated. These guidelines, drawn up on the basis of the long-term experience of the experts, were presented in a paper entitled “Criminalistic phonoscopy” [10]. The authors state that speaker identification with the highest degree of probability is possible when:

- in the case of an analog recording – the analysis is performed using the original tape or a high quality copy;
- the evidence utterances are spontaneous and natural;
- the speaker formulates his/her statements in sentences rather than isolated words;
- the evidence utterances of the speaker, who is to be the subject of an identification analysis, comprise no fewer than 10–15 sentences;
- the utterances are distinctive and understandable;
- the band of speech recording is 80–6000 Hz, while the signal to noise ratio (SNR) is not lower than 45 dB.

If the evidence material does not meet the above specified requirements, then conclusions drawn from identification analyses are formulated with a lower degree of probability, and in some cases, the recordings may provide no grounds at all for commencing any studies.

However, observance of the above-presented criteria does not lead to elimination of a subjective aspect of decisions concerning whether to accept material for investigation or to reject it – for example, determining whether utterances are sufficiently distinctive and understandable continues to be based on a subjective assessment and some decisions are debatable and subject to differences of opinions. It should also be borne in mind that performing identification analyses necessitates collection of appropriate reference material. The quality and amount of this material constitute factors that significantly affect our ability to perform investigations and formulate final conclusions.

2.3. Collection of reference material in the form of speech

The process of collecting reference material should be carried out by the expert who will subsequently perform comparative examination. It is the expert who is responsible for analysing the evidence material with respect to its usefulness in identification analyses. In the case of high quality recordings, which, nevertheless, are limited in terms of quantity, the most signifi-

cant element of material collection consists in the so-called sentence tests, or utterances that are identical in their contents with utterances originating from the evidence material. Thanks to sentence tests, it is possible to compare acoustic parameters of particular vowels and vocoid groups occurring in the same phonetic environment. In turn, if the evidence material has revealed numerous non-standard articulative properties of the speaker, then – when collecting the material – the expert should concentrate on obtaining spontaneous utterances, which allow indication of a higher number of speech habits.

Both at KEÚ and at the Laboratory of Speech and Audio Analysis, Institute of Forensic Research (henceforth referred to as IES), a recording that contains spontaneous utterances of the speaker and so-called sentence tests, which are identical in their contents with sentences selected from the evidence material is considered to be complete reference material. Basic assumptions associated with material collection are identical in the two discussed centres; nevertheless, there are some differences at the level of implementation.

Expert practice indicates that an effective method of reference material collection is having a casual conversation with the investigated subject in the initial phase of the process. Then, the expert attempts to broach a subject that is acceptable to the speaker. His/her concentrating on the content of the utterance allows a casual and natural mode of speaking [13, 14].

Another element of the reference material is the so-called sentence tests. As has previously been mentioned, the aim of sentence tests is to obtain utterances corresponding in their contents to utterances originating from the evidence recording. Such material allows analysis of acoustic parameters of selected vowel and vocoid segments occurring in identical phonetic contexts. Similarly to the case of a spontaneous conversation, it is important to ensure that the utterances are as much as possible in keeping with the natural mode of speaking characteristic for a given person. Nevertheless, in the majority of cases, at this stage of examination, the speaker has a sense of a changed degree of formality in the communication situation – oftentimes, this stage is perceived by the speaker as “proper material collection”, as it may relate to the case in question in a more direct way. In consequence, there is some risk that the speaker will control his/her articulation habits to a greater degree than happens in the course of a casual conversation, or will even strive to articulate hyper-correctly.

Speech and Audio Laboratories employ various techniques of collecting sentence tests material.

A method relying on the speaker reading a provided text which contains sentences selected from the reference material entails a risk of obtaining monotonous and slow utterances, which in consequence are unnatural and thus inadequate from the qualitative point of view. That is why this method of material collection has been rejected by the two described centres. On the other hand, another technique is widely employed in phonoscopic practice; this technique allows monotonous reading of a text to be avoided – the speaker repeats selected utterances originating from the evidence material after the expert. This method of material collection is employed in KEÚ, as well as in some Polish laboratories. In discussing the technique, one should also point out its disadvantages. When repeating the sentences, the speaker follows the intonation structure that is characteristic for the expert's utterances, and in some cases, s/he may unconsciously or consciously adopt other linguistic properties, which in reality are not typical for his/her speech patterns and are a consequence of a passive copying of the perceived model. As a result, side by side with speech habits typical for the speaker, properties appear in the material that are solely an effect of his/her imitating the expert's mode of speaking. In such instances, carrying out a reliable comparative analysis requires a conscious separation of the linguistic properties specific not only for the suspect, but also for the expert.

In striving to eliminate the negative aspects of the above-discussed methods, the IES experts have employed a slightly different method of reference material collection. Selected sentences are presented to the speaker in written form and s/he is given time to quietly familiarise him/herself with the text. Prior to articulating consecutive sentences, the speaker is instructed by the expert as to the expected mode of articulation – e.g. s/he is asked to utter the question, order or other utterance in such a way as if s/he was directly addressing the expert. In this way, the risk of the speaker copying the expert's mode of speaking is avoided, and at the same time, thanks to the speaker having been previously familiarised with the text, the number of hyper-correct utterances that are characteristic for texts read aloud decreases. Thanks to the way in which the communication situation – i.e. the speaker-listener relation – has been arranged, the utterance is more natural in character.

There are also differences in the time spent on material collection – at KEÚ, the total time of collection is usually below twenty minutes, while at IES, the activity takes approximately an hour. A longer period devoted to material collection, although associated with a greater effort on the part of both the expert and the

speaker and capable of evoking impatience in the latter, nevertheless, allows a casual conversation to be conducted, alternating with sentence tests, which usually results in the expert obtaining more natural utterances. In addition, more extensive material allows analysis of a greater number of specific articulative properties of the speaker and assessment of their representativeness, thus making it possible to prepare a more complete description of his/her speech habits.

Both laboratories adhere to the principle that the reference material should be characterised by technical parameters that are as good as possible, and, therefore, the material is collected in a direct manner, i.e. not by telephone. Recording phone conversations leads to distortions, and anyway, a comprehensive reconstruction of the technical conditions of the evidence material recording is not possible. Registering reference phone conversation recordings results in comparing two materials that are distorted in different manners, which hinders interpretation of the results of investigation.

It is worth mentioning formal issues associated with securing reference material. Material collection is an act performed in connection with legal proceedings and for this reason it is necessary to draw up appropriate records. In both institutions, records pertaining to reference material collection include information on individuals executing the collection (one expert conducts the conversation, while another is responsible for operating the equipment), data on the individual from whom the material is collected, and on the document confirming his/her identity, the date of material collection, the file number of the case in association with which the investigation was commissioned, the signature of the individual from whom the material was collected, along with his/her possible comments on the mode of material collection. The individual from whom the material is collected is entitled to hear the entirety or a fragment of the recorded material.

The most desirable situation consists in material collection being performed by the same expert who is to perform identification analysis. However, in practice, instances have been known where an individual from whom material was to be collected, refused to participate in an investigation and therefore, the ordering party submitted material recorded without the speaker's knowledge for analysis (e.g. originating from interview) [13]. It should be emphasised that a necessary prerequisite for such recordings to acquire the status of reference material is unequivocal determination of the speaker's identity. This type of material is usually characterised by a considerably poorer quality of recording than recordings made by experts under

laboratory conditions. In addition, in such cases it is not possible to collect sentence tests and, consequently, to perform analysis of sounds occurring in phonetic contexts that are identical with the evidence contexts, which may in effect lead to a decreased degree of final conclusion probability and prolong the duration of investigations.

Experts from KEÚ also collect reference material outside their laboratory, i.e. chiefly at police stations in other cities and towns. The facilities designated for this purpose have to meet certain criteria and requirements – the room must be carpeted and furnished, the windows have to have curtains and should not look out onto a busy street. Even then, the material is characterised by a somewhat poorer quality as compared to recordings made under laboratory conditions. Reference material in the form of speech collected by IES experts is recorded solely in the facilities of IES.

2.4. Identification analyses. Interlaboratory differences

Both at IES and at KEÚ, determinations of measurable speech parameters are performed using STx software developed by the Austrian Academy of Sciences in Vienna. In the two centres, one of the stages of speaker recognition is a comparison of dynamic formant structure, or momentary values of formants for vowels extracted from evidence material and reference

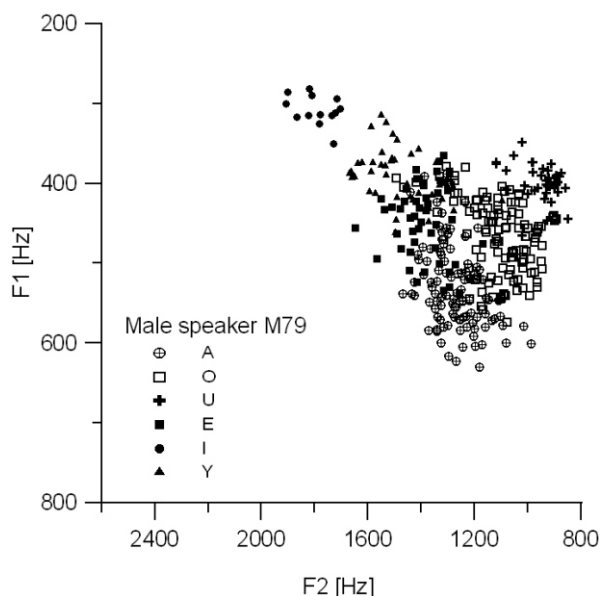


Figure 6. Vowel distribution of a user of the Polish language.

material. At IES, experts additionally plot graphs illustrating the distribution of the mean frequency values

for particular formants, which constitute an acoustic description of the vowels used by a given speaker (Figure 6). A comparison of the values of vowel formants extracted from the evidence material with the frequencies obtained for the reference material allows determination of whether the former lie within the range of the speaker's variability.

As has been mentioned before, experts employed at KEÚ and IES use the acoustic-linguistic method, but some differences are clearly visible at the level of particular stages of their investigations and evaluation of particular properties. In their paper entitled "Criminalistic phonoscopy", the experts from the Slovak laboratory define identification investigations as a combination of linguistic, phonetic and electro-acoustic methods [10]. Thus, the initial part of the investigation – linguistic analysis – should include a description of the lexis, morphology and syntax of the speaker. The authors mention phonetic analysis as the second stage of investigation, which includes segmental (articulation of phonemes and their combinations) and suprasegmental (prosodic properties) phenomena. Phonetic studies are based on an auditive assessment and probe both into phenomena associated with the physiological structure of articulators and properties acquired in the process of socialisation. The third stage is the analysis of measurable speech parameters. The expert extracts speech parameters – mostly formants and the fundamental frequency (F0) – separately for the evidence and reference material, in order to subsequently carry out a comparative analysis. This stage of examination also includes the use of an automatic system of speaker identification, which the experts from the Bratislava centre use to verify findings obtained during previous investigations. The Slovak laboratory follows the rule that the two initial stages of analysis (linguistic and phonetic) are performed by a linguist, while the acoustic analysis – by an engineer-specialist in technical sciences. Such an approach is based on the assumption that an individual carrying out technical investigations has at his/her disposal an extensive body of knowledge on the parameters of a recording, which allows him/her to more comprehensively interpret phenomena observed in the acoustic stratum of the recording – including those associated with measurable speech parameters. The presented procedure also has its drawbacks – separation of phonetic analysis, which is based on a perceptive evaluation, from spectrographic analysis leads to a certain sort of disintegration of acoustic information on the linguistic habits of the speaker. Thus analysis of the amplitude-frequency-temporal course of a speech signal (a spectrogram) and measurements of momentary values of

formants or distribution graphs of mean formant values allow objectivisation of perceptual impressions, as well as relating of particular linguistic implementations to the entire phonetic system of the speaker. That is why at IES, studies of linguistic properties and analysis of measurable speech parameters are carried out by the same individual with a background in linguistic education.

There is also a significant difference between the two laboratories in evaluation of the usefulness of properties originating from various levels of language description in criminalistic speaker identification. One of the assumptions on which the methodology of investigations carried out at IES is based is a contention that phonetics, and thus articulation of particular phonemes or words, differentiates speakers to a higher degree than, for example, lexis or syntax. Selection of a lexical or syntactic item is performed much more consciously by a speaker than use of a given phonetic variant and is, moreover, determined by the communication situation – for example, by its degree of formal character. Distributional instability of lexis and syntax causes their value to be lower in the process of speaker identification. This is why comparative studies performed at IES are mostly based on phonetic analysis, and a description of phenomena from higher language levels is of auxiliary value only.

Identification studies consist in comparing utterances originating from evidence material and reference utterances collected from an individual indicated by the ordering party. At KEÚ, in addition to individual identification analyses, experts also perform group analyses, or so-called profiling. Based on analysis of evidence utterances, the expert defines the geographical area the speaker may come from, his social status, gender and approximate age, and also determines his psychophysical state. The methodological foundations of this type of investigation are debatable. Controversies are mostly sparked off by such issues as opining on the emotional state of the speaker – doubts arise as to the competence of an expert in the field of defining mental states, as well as to his ability to draw conclusions on mental activity based on scant material, which usually does not take into consideration the variability of linguistic behaviour of the speaker in various situations.

3. Summary

At KEÚ and IES, speaker identification investigations are carried out using the linguistic-acoustic method. In both laboratories, conclusions are formulated based

on the results of linguistic analysis and results of acoustic parameter measurements. Differences are seen in the scope of investigations, techniques of sentence test collection and in assumptions concerning the discriminative value of properties originating from particular levels of language description. The relations between the phonetic systems of the Polish and Slovak languages may provide a starting point for further studies on articulative properties of Slovak speakers using the Polish language and *vice versa* – on Poles speaking Slovak, and thus for studies on the effectiveness of the method of identification analysis in relation to non-native speakers of a given language.

Acknowledgements

The author wishes to extend thanks to Miloš Lukovič, M. Eng., and Adriana Farkašová, MA, who presented in detail the methods employed in the Bratislava laboratory – for their involvement in mutual collaboration, valuable comments and constructive discussions.

Katarzyna Klus

The research was financially supported by Leonardo da Vinci programme “Contemporary forensic sciences – chosen issues”.

References

1. Dalewska-Greń H., *Języki słowiańskie*, Wydawnictwo Naukowe PWN, Warszawa 2002.
2. Dudášová-Kriššáková J., Fonologický systém spisovnej slovenčiny a poľštiny z typologického hľadiska, *Slavica Slovaca* 1999, 1, 16–24.
3. Dukiewicz L., Sawicka I., *Fonetyka i fonologia*, Wydawnictwo Instytutu Języka Polskiego PAN, Kraków 1995.
4. Gonzalez-Rodriguez J., Ortega-Garcia J., Sanchez-Boite J. L., Forensic identification reporting using automatic biometric systems [in:] *Biometric solutions for authentication in an E-world*, Zhang D., Sarrukai S. R. [eds.], Kluwer, Dordrecht 2002.
5. Hollien H., *Forensic voice identification*, Academic Press, San Diego 2002.
6. Isačenko A. V., *Spektrografická analýza slovenských hlások*, Vydavateľstvo Slovenskej akadémie vied, Bratislava 1968.
7. Jassem W., *Podstawy fonetyki akustycznej*, PWN, Warszawa 1973.
8. Kráľ A., *Pravidlá slovenskej výslovnosti. Systematika a ortoepický slovník*, Matica Slovenská, Martin 2005.
9. Kráľ A., Saból J., *Fonetyka a fonológia*, Bratislava, Slovenské pedagogické nakladateľstvo 1989.

10. Lukovič M., Snohová D., Trvalec H., *Kriminalistická fonoskopia*, Bratislava 1997.
11. Nolan F., *Speaker recognition and forensic phonetics*, [in:] *A handbook of phonetic science*, Hardcastle W., Laver J. [eds.], Blackwell, Oxford 1997.
12. Rose P., *Forensic speaker identification*, Taylor and Francis, London 2002.
13. Rzeszotarski J., *Kompendium badań fonoskopijnych*, *Prokuratura i Prawo* 2007, 7/8, 181–187.
14. Trawińska A., *Materiały szkoleniowe*, <http://beta.kcskspip.gov.pl/edu/229>.
15. Wierzchowska B., *Fonetyka i fonologia języka polskiego*, Zakład Narodowy im. Ossolińskich, Wydawnictwo PAN, Wrocław 1980.

Corresponding author

Katarzyna Klus
Instytut Ekspertyz Sądowych
ul. Westerplatte 9
PL 31-033 Kraków
e-mail: kklus@ies.krakow.pl

IDENTYFIKACJA MÓWCÓW Z WYKORZYSTANIEM METODY JĘZYKOWO-POMIAROWEJ W KEÚ I IES*

1. Część językoznawcza

1.1. Słowacki i polski system wokaliczny

Bardzo istotnym elementem badań identyfikacyjnych mówcy jest analiza jego systemu samogłoskowego. Samogłoski to segmenty dźwięczne posiadające strukturę formantową. Formanty, czyli parametry akustyczne mowy, indywidualizują mówcę w stopniu większym niż cechy dialektalne, ponieważ ich wartości wynikają m.in. z budowy anatomicznej mówcy i nawyków artykulacyjnych. Między systemami wokalicznymi języków: polskiego i słowackiego istnieją, pomimo znacznych podobieństw, pewne istotne różnice.

Zjawiskiem uniwersalnym dla wszystkich języków słowiańskich jest system wokaliczny opierający się na współwystępowaniu dwóch opozycji, tj. przednia *versus* tylna i wysoka *versus* średnia *versus* niska [1]. Najprostszą reprezentacją takiego modelu stanowi system samogłoskowy języka macedońskiego (rycina 1). Pozostałe języki słowiańskie posiadają systemy będące pewną modyfikacją tego minimalnego systemu.

W języku słowackim inwentarz fonemów samogłoskowych jest ponad dwukrotnie większy niż w przypadku przedstawionego na rycinie 1 schematu, a to za sprawą opozycji ilościowej, która występuje w tym języku. Obecność iloczasu, czyli rozróżnienie między samogłoskami długimi i krótkimi, jest także jedną z podstawowych różnic między językiem słowackim a polskim. W języku polskim iloczas zanikł w XVI wieku, w związku z czym we współczesnej polszczyźnie opozycja między samogłoskami długimi a krótkimi nie jest opozycją fonologiczną. W dialektach wschodnich języka polskiego samogłoski akcentowane wymawiane są co prawda dłużej, zaś nieakcentowane są wyraźnie zredukowane – jednak określanie tego zjawiska mianem iloczasu jest błędne.

Uproszczone schematy systemów samogłoskowych języka polskiego i słowackiego został przedstawiony na rycinie 2.

System wokaliczny języka słowackiego jest więc znacznie bogatszy od polskiego – oprócz długich odpowiedników większości samogłosek występuje w nim jeszcze [ä] – samogłoska opisywana jako przednia, choć bardziej tylna niż [ɛ] i niska, choć nieco wyższa niż [ɑ]. Specyfika tej samogłoski polega m.in. na jej ograniczeniach dystrybucyjnych – [æ] jest kontynuantem dawnej

samogłoski nosowej [ɛ̃] i występuje tylko po samogłoskach wargowych. Trzeba równocześnie zaznaczyć, że głoska ta nie jest znana wszystkim użytkownikom języka słowackiego – według literatury przedmiotu alofonem tym posługuje się na co dzień ok. 5% Słowaków. Normy ortoepiczne nie są w tym względzie rygorystyczne – w stylu neutralnym wymowa [æ] nie jest ani zakazana, ani obligatoryjna. Znajomość tej wymowy jest natomiast wymagana w tzw. stylu wyższym, którym posługują się np. aktorzy podczas recytacji wierszy czy w klasycznych sztukach teatralnych, gdzie artykulacja głoski [æ] jest warunkiem zachowania rymu [8]. Analiza formantowa słowackiego systemu wokalicznego potwierdza również bardzo silną tendencję do wycofywania się samogłoski [æ] – wartości formantów wskazują na niestabilną, niesystemową wymowę tej samogłoski [9] (rycina 4).

Do różnic między słowackim a polskim systemem wokalicznym należy zaliczyć występowanie w języku polskim [i] – samogłoski opisywanej jako przednia i wysoka, lecz nieco niższa niż [i] [3, 15]. Głosce tej, uznawanej niegdyś za wariant kombinatoryczny samogłoski [i], obecnie zdecydowanie częściej przyznaje się status odrębnego fonemu [3]. W słowackim zachował się co prawda grafem y, a o miejscach jego występowania decyduje kryterium etymologiczne, jednak w wymowie jest on realizowany jako [i].

Kolejna różnica między systemami omawianych języków dotyczy połączeń dyftongicznych zwanych w języku słowackim dvojhláskami, czyli połączeń [iɑ], [jɛ], [iu] bądź [uo]. Są one wyjątkowo podatne na różnego rodzaju niestandardowe realizacje, stąd też ich duża wartość dyskryminacyjna w procesie identyfikacji mówców. Połączenia dyftongiczne mają taką wartość iloczasową, jak samogłoski długie. W języku polskim również występują grupy wokoidalne, czyli połączenia np. glajdu [j] i samogłoski (np. miasto, biel), ale w ujęciu fonologicznym nie przypisuje się im statusu dyftongów [3].

Z kolei w języku polskim występują tzw. samogłoski nosowe, których nie zna język słowacki. Są to kontynuanty prasłowiańskich samogłosek nosowych, realizowane we współczesnym języku ogólnym jako połączenie samogłosek [ɔ] i [ɛ] oraz elementu nosowego – różnego w zależności od prawostronnego kontekstu fonetycznego. Przykładowo nosówka zapisywana za pomocą grafemu ą może być wymawiana jako /ɔm/ (np. trąba), /ɔn/ (np. sąd) czy /ɔw/ (np. wąs).

Z przedstawionych w literaturze przedmiotu [6, 7, 9] badań częstotliwości formantowych dla polskich i słowackich samogłosek wynika, że największe różnice między tymi językami dotyczą samogłosek średnich – sło-

* Na podstawie wyników badań przeprowadzonych w Pracowni Fonoskopii Policyjnego Instytutu Kryminalistyki i Ekspertyz (KEÚ) w Bratysławie oraz w Pracowni Analizy Mowy i Nagrań Instytutu Ekspertyz Sądowych (IES) w Krakowie.

wackie [ɛ] i [ɔ] są bardziej zamknięte niż polskie, przy czym wartości samogłoski [ɛ] są w niektórych przypadkach zbliżone do polskiego scentralizowanego [i] (rycina 5).

1.2. Słowacki i polski system konsonantyczny

W języku słowackim występuje fonem [l], którego nie ma w systemie fonologicznym języka polskiego – w tzw. standardowej polszczyźnie głoska ta stanowi wariant kombinatoryczny fonemu [l] o bardzo ograniczonej dystrybucji i występuje jedynie przed samogłoską [i] i sonorantem [j]. Kolejna różnica dotyczy tzw. sonantów – w słowackim istnieją zgłoskotwórcze warianty [r] i [l], które w polskim nie występują w ogóle. Z kolei w języku polskim status fonemu ma sonorne [w], zaś w języku słowackim [w] jest jedynie wariantem kombinatorycznym [v] i realizowane jest tylko w wygłosie lub śródgłosie przed spółgłoską.

Głoską przysparzającą znacznych problemów Polakom uczącym się słowackiego jest krtaniowe [h]. W tzw. standardowej polszczyźnie nie istnieje rozróżnienie między wymową *ch* i *h*; wariant krtaniowy, realizujący ten drugi grafem, obecny jest jedynie w niektórych gwarach polskich.

Jeśli zaś chodzi o spółgłoski właściwe, podstawowe różnice dotyczą spółgłosek palatalnych. W słowackim systemie konsonantycznym nie ma spółgłosek [ç], [z̥], [t̥ç] i [d̥z̥]. Diachroniczne badania języków słowiańskich wykazują, że słowackie [t,], [d,] są w większości przypadków kontynuantami tych samych połączeń głosek, które w polskim rozwinęły się w [t̥ç] i [d̥z̥].

Istotna różnica między językami słowackim a polskim dotyczy także spółgłosek spalatalizowanych. Otóż polski system fonetyczny wyróżnia opozycję pod względem miękkości dla zdecydowanej większości spółgłosek. Oznacza to, że istnieją zmiękczone warianty kombinatoryczne poszczególnych spółgłosek, realizowane wtedy, gdy w prawostronnym otoczeniu znajduje się samogłoska [i] lub glajd [j]. Takiego rozróżnienia nie ma w języku słowackim. W podanych pozycjach samogłoski wymawiane są w ten sam sposób, co w kontekście neutralnym pod względem miękkości. W tym miejscu nasuwa się pytanie, czy różnica ta wynika z konwencjonalnych opisów w gramatykach obu języków, czy też z faktycznych rozbieżności w jakości realizowanych połączeń. Pełniejszą analizę porównawczą polskiego i słowackiego systemu fonetycznego i fonologicznego można znaleźć między innymi w artykule Jülii Dudášovej-Kriššákovéj [2].

2. Część kryminalistyczna

2.1. Porównawcze badania identyfikacyjne – założenia teoretyczne

Twierdzenie, że mowa ludzka jest dźwiękową reprezentacją języka naturalnego, która dzięki szeregowi specyficznych właściwości indywidualizuje osobę mówiącą, to jedno z podstawowych założeń kryminalistycznej identyfikacji mówcy. W laboratoriach fonoskopijnych na całym świecie stosowane są różne metody – od lingwistycznej, gdzie analizowane są wyłącznie cechy językowe, przez językowo-pomiarową, która łączy w sobie elementy analizy językowej i akustycznej po metody w pełni automatyczne, bazujące zazwyczaj na parametryzacji sygnału mowy poprzez współczynniki cepstralne [4, 5, 11, 12]. Badania identyfikacyjne wykonywane metodą językowo-pomiarową (bo właśnie tą metodą posługują się biegli w obu omawianych ośrodkach) pozwalają na sporządzenie takiego opisu mówcy, który uwzględnia cechy osobnicze wynikające z budowy kanału głosowego i artykulatorów oraz grupowe, wspólne np. dla użytkowników tego samego dialektu.

Podczas badań identyfikacyjnych analizowane są zarówno cechy indywidualne, związane z budową anatomiczną mówcy, czyli głównie częstotliwości formantów samogłoskowych F1–F4, jak i cechy grupowe, nabyte w procesie socjalizacji, odbywającym się w konkretnym środowisku geograficznym za pośrednictwem osób (rodziców, opiekunów) przekazujących własne wzorce wymawiania. Współcześnie pozycja dialektu jest znacznie słabsza niż kilkadziesiąt lat temu. Posługiwanie się formami dialektalnymi zazwyczaj wartościowane jest negatywnie, jest bowiem postrzegane jako ograniczenie kompetencji językowej mówcy, ponadto kojarzone jest z brakiem wykształcenia, czasem także niższym statusem społecznym. Dlatego też mówcy dążą do wypierania dialektizmów, zwłaszcza tych ujemnie wartościowanych, a więc w sposób najbardziej wyrazisty różniących się od tzw. standardowego języka. W języku polskim do tego rodzaju cech można zaliczyć np. mazurzenie, czyli wymowę spółgłosek zębowych [t̥s], [z], [s] w miejscu dźwiękowych [t̥], [z], [s].

Zdolność do ograniczenia wymowy dialektalnej związana jest między innymi ze świadomością językową mówcy. Mówca nie jest w stanie wyeliminować tych cech, których obecności we własnym systemie fonetycznym sobie nie uświadamia. Właśnie tego rodzaju cechy – cechy fonetyczne – powinny stanowić podstawę do przeprowadzenia porównawczej analizy językowej.

Cechy dialektalne mają charakter grupowy, a więc są wspólne dla użytkowników języka zamieszkujących ten sam obszar. Teoretycznie wartość dyskryminacyjna danej cechy jest więc tym większa, im mniejsza jest liczba użytkowników określonego dialektu. Sporządzenie wia-

rygodnego i pełnego opisu cech artykulacyjnych mówcy wymaga jednak nie tylko wskazania cech występujących w danym materiale, ale również określenia ich natężenia – czyli sprecyzowania, czy tendencja do konkretnej wymowy jest konsekwentna, silna, umiarkowana czy słaba.

Za cechy bardziej indywidualne uznaje się natomiast mierzalne parametry mowy. Pomiarów parametrów akustycznych mowy dotyczą przede wszystkim struktur formantowych segmentów samogłoskowych. Formanty to – w uproszczeniu – pasma częstotliwości, w których koncentruje się energia sygnału mowy, a które są wynikiem rezonansu akustycznego fali głosowej w jamach nadkrtaniowych. Wartości formantów są zależne od budowy anatomicznej narządów mowy, ale także od nawyków artykulacyjnych – przede wszystkim pozycji języka w jamie ustnej. W praktyce kryminalistycznej najczęściej istnieje możliwość przeprowadzenia pomiarów pierwszych trzech, a w niektórych przypadkach czterech formantów.

2.2. Przydatność materiału dowodowego do badań identyfikacyjnych. Wstępna analiza materiału dowodowego

Przed przystąpieniem do badań porównawczych biegły przeprowadza wstępną ocenę materiału dowodowego, która pozwala stwierdzić, czy wypowiedzi dowodowe pod względem ilościowym i jakościowym dają podstawę do przeprowadzenia badań identyfikacyjnych, czyli czy można uznać je za reprezentatywną próbkę mowy danej osoby. Konieczne jest m.in. stwierdzenie, czy parametry techniczne z jakimi dokonano rejestracji, umożliwiają wiarygodną ekstrakcję mierzalnych parametrów mowy. Wstępna analiza służy również ustaleniu, czy w obrębie wypowiedzi występują dystynktywne cechy językowe – zarówno grupowe, jak i indywidualne. Na podstawie tego rodzaju badań biegły rozstrzyga również, jaką formułę powinna przybrać czynność pobrania materiału porównawczego, aby uzyskane wypowiedzi stanowiły materiał jak najbardziej adekwatny w stosunku do wypowiedzi dowodowych.

W Pracowni Fonoskopii Policijnego Instytutu Kryminalistyki i Ekspertyz w Bratysławie (Kriminalistický a expertízny ústav Policajného zboru, dalej: KEÚ) biegli wypracowali wytyczne, na podstawie których podejmowana jest decyzja o tym, czy materiał pozwala na przeprowadzenie badań identyfikacyjnych, a jeśli tak – to z jakim stopniem prawdopodobieństwa może zostać sformułowany wniosek końcowy. Wytyczne powstałe w oparciu o wieloletnie doświadczenie biegłych przedstawione zostały w pracy pt. „Fonoskopia kryminalistyczna” [10]. Autorzy pracy twierdzą, że identyfikacja mówcy z najwyższym stopniem prawdopodobieństwa jest możliwa wówczas, gdy:

- w przypadku nagrania analogowego do badań przekazany zostanie oryginalny nośnik, względnie kopia dobrej jakości;
- wypowiedzi dowodowe są spontaniczne i naturalne;
- mówca formułuje wypowiedzi za pomocą zdań, a nie pojedynczych słów;
- wypowiedzi dowodowe mówcy, względem którego mają być przeprowadzone badania identyfikacyjne, stanowią nie mniej niż 10–15 zdań,
- wypowiedzi są wyraziste i zrozumiałe;
- pasmo, w jakim zarejestrowana jest mowa, wynosi 80–6000 Hz, zaś stosunek sygnału do szumu (S/N) jest nie mniejszy niż 45 dB.

Jeżeli materiał dowodowy nie spełnia wymienionych warunków, wówczas wnioski z badań identyfikacyjnych są formułowane z niższym stopniem prawdopodobieństwa, a w niektórych przypadkach nagrania mogą w ogóle nie dawać podstaw do podjęcia badań.

Wprowadzenie przedstawionych powyżej kryteriów nie prowadzi jednak do eliminacji subiektywnego aspektu decyzji o przyjęciu materiału do badań bądź jego odrzuceniu – np. stwierdzenie, czy wypowiedzi są wystarczająco wyraziste i zrozumiałe, opiera się nadal na subiektywnej ocenie i niektóre rozstrzygnięcia mogą być dyskusyjne. Nie można także zapominać, że do przeprowadzenia badań identyfikacyjnych konieczne jest zgromadzenie odpowiedniego materiału porównawczego. Jakość i obszerność tego materiału to czynniki mające istotny wpływ na możliwość przeprowadzenia badań i sformułowanie wniosków końcowych.

2.3. Czynność pobrania materiału porównawczego w postaci mowy

Czynność pobrania materiału porównawczego powinna być przeprowadzona przez biegłego, który będzie następnie wykonywał badania porównawcze. To do niego należy również przeanalizowanie materiału dowodowego pod kątem jego przydatności do badań identyfikacyjnych. W przypadku nagrań dobrych jakościowo, jednak szczupłych pod względem ilościowym, najistotniejszym etapem pobrania będą tzw. testy zdaniowe, czyli wypowiedzi tożsame treściowo z wypowiedziami pochodzącymi z materiału dowodowego. Dzięki testom zdaniowym możliwe jest porównanie parametrów akustycznych poszczególnych samogłosek i grup wokoidalnych znajdujących się w takim samym otoczeniu fonetycznym. Z kolei jeśli w materiale dowodowym ujawniono wiele niestandardowych cech artykulacyjnych mówcy, wówczas biegły podczas pobrania materiału powinien kłaść nacisk na uzyskanie wypowiedzi spontanicznych, które pozwalają wskazać większą liczbę nawyków mownych.

Zarówno w KEÚ, jak i w Pracowni Analizy Mowy i Nagrań Instytutu Ekspertyz Sądowych (dalej: IES) za

kompletny materiał porównawczy uznaje się nagranie zawierające wypowiedzi spontaniczne mówcy oraz tzw. testy zdaniowe, tożsame treściowo ze zdaniami wybranymi z materiału dowodowego. Podstawowe założenia dotyczące pobrania materiału porównawczego są wspólne dla obu omawianych ośrodków, jednak istnieją pewne różnice na poziomie realizacji tej czynności.

Praktyka fonoskopijna wykazuje, że efektywnym sposobem pobierania materiału porównawczego jest przeprowadzenie swobodnej rozmowy w początkowej części czynności. Wówczas biegły stara się poruszać tematy, które będą odpowiadać mówcy. Skoncentrowanie się na treści wypowiedzi sprawia, że sposób mówienia jest swobodny i naturalny [13, 14].

Drugim elementem materiału porównawczego są tzw. testy zdaniowe. Jak wspomniano wcześniej, celem pobierania testów zdaniowych jest uzyskanie wypowiedzi odpowiadających treściowo kwestiom pochodzącym z dowodowego nagrania. Taki materiał umożliwia analizę parametrów akustycznych wybranych segmentów samogłoskowych i wokoidalnych występujących w analogicznych kontekstach fonetycznych. Podobnie jak w przypadku rozmowy spontanicznej, istotne jest, aby w miarę możliwości wypowiedzi te cechował naturalny, właściwy danej osobie sposób mówienia. Jednak w większości przypadków na tym etapie badań mówcy towarzyszy poczucie zmiany stopnia oficjalności sytuacji komunikacyjnej – często etap ten jest postrzegany przez mówcę jako „właściwe pobranie materiału”, jako że w sposób bardziej bezpośredni nawiązuje do sprawy. W efekcie istnieje ryzyko, że mówca będzie kontrolował swoje nawyki wymawianiowe w stopniu większym niż w przypadku swobodnej rozmowy, a nawet będzie dążył do wymowy hiperpoprawnej.

W laboratoriach fonoskopijnych praktykowane są różne techniki pobierania testów zdaniowych. Metoda polegająca na odczytywaniu przez mówcę przedłożonego tekstu zawierającego zdania wybrane z materiału porównawczego niesie z sobą ryzyko uzyskania wypowiedzi monotonna, powolnych, w konsekwencji nie naturalnych i tym samym najczęściej nieadekwatnych jakościowo względem wypowiedzi dowodowych. Dlatego ten sposób pobierania materiału został odrzucony przez oba omawiane ośrodki. Szerokie zastosowanie w praktyce fonoskopijnej znajduje natomiast inna technika pozwalająca uniknąć efektu jednostajnego odczytywania tekstu – mówca powtarza za biegłym wybrane kwestie zaczerpnięte z materiału dowodowego. W ten sposób materiał pobierają eksperci z KEÚ, a także z niektórych polskich ośrodków. Omawiając tę metodę, należy wskazać również jej słabe punkty. Otóż mówca, powtarzając zdania, podąża za strukturą intonacyjną charakteryzującą wypowiedzi biegłego, a w niektórych przypadkach może również nieświadomie lub świadomie przejmować inne cechy językowe, które w rzeczywistości

nie są dla niego typowe, a stanowią jedynie wynik biernego powielania percypowanego wzorca. W efekcie obok nawyków właściwych mówcy w materiale mogą pojawić się cechy będące tylko i wyłącznie skutkiem naśladowania sposobu mówienia biegłego. Przeprowadzenie wiarygodnej analizy porównawczej wymaga wówczas świadomego rozdzielania tych dwóch grup cech językowych.

Dążąc do niwelowania negatywnych aspektów opisanych powyżej metod, w IES stosuje się nieco odmienny sposób pobierania materiału porównawczego. Wybrane zdania zostają w formie pisemnej zaprezentowane mówcy, przy czym ma on czas, aby spokojnie zapoznać się z tekstem. Przed wypowiedzianiem kolejnych zdań mówca zostaje poinstruowany przez biegłego odnośnie do oczekiwanego sposobu ich wymówienia – np. zostaje poproszony o zadanie przedmiotowego pytania, polecenia czy innej wypowiedzi w taki sposób, jakby zwracał się do biegłego jako adresata tej wypowiedzi. W ten sposób unika się ryzyka powielania sposobu mówienia biegłego, a równocześnie, dzięki wcześniejszemu zapoznaniu mówcy z tekstem, maleje liczba wypowiedzi hiperpoprawnych charakterystycznych dla wypowiedzi czytanych. Dzięki zaaranżowaniu sytuacji komunikacyjnej, czyli relacji nadawca-odbiorca, charakter wypowiedzi jest bardziej naturalny.

Różny jest także czas poświęcany na wykonanie czynności – w KEÚ łączny czas pobrania wynosi zazwyczaj około kilkunastu minut, podczas gdy w IES czynność ta zajmuje mniej więcej godzinę. Dłuższy czas poświęcany na pobranie materiału, choć łączy się z większym wysiłkiem obu stron i może wywoływać zniecierpliwienie mówcy, umożliwia jednak prowadzenie swobodnej rozmowy naprzemiennie z testami zdaniowymi, co zazwyczaj skutkuje uzyskaniem bardziej naturalnych wypowiedzi. Poza tym obszerniejszy materiał pozwala na analizę większej liczby specyficznych cech artykulatoryjnych mówcy i ocenę ich reprezentatywności, a co za tym idzie – sporządzenie bardziej kompletnego opisu jego nawyków mownych.

Wspólne dla obu laboratoriów jest natomiast założenie, że materiał porównawczy powinny cechować możliwie najlepsze parametry techniczne, w związku z czym jest on pobierany bezpośrednio, czyli bez udziału transmisji telefonicznej. Rejestrowanie rozmów przez telefon prowadzi bowiem do zniekształceń, a i tak kompleksowe odtworzenie warunków technicznych rejestracji materiału dowodowego nie jest możliwe. Rejestrowanie nagrań porównawczych przez telefon skutkuje porównywaniem dwóch materiałów zniekształconych w odmienny sposób, co utrudnia interpretację otrzymanych wyników badań.

Warto również wspomnieć o kwestiach formalnych związanych z zabezpieczaniem materiału porównawczego. Pobranie jest czynnością procesową, dlatego też ko-

nieczne jest sporządzenie odpowiedniego protokołu. W obu omawianych laboratoriach protokoły pobrania materiału porównawczego zawierają informacje o osobach przeprowadzających czynność (jeden z biegłych prowadzi rozmowę, podczas gdy drugi jest odpowiedzialny za obsługę sprzętu), dane osoby, od której pobierany jest materiał oraz dokumentu, na podstawie którego stwierdzono jej tożsamość, datę przeprowadzenia czynności, sygnaturę sprawy, w związku z którą zlecono badania, podpis osoby, od której pobierano materiał oraz jej ewentualne uwagi co do sposobu pobrania materiału. Osoba, od której pobierany jest materiał, ma prawo odsłuchać całość bądź fragment zarejestrowanego materiału.

Sytuacją najbardziej pożądaną jest możliwość przeprowadzenia czynności pobrania materiału przez biegłego, który będzie wykonywał badania identyfikacyjne. Jednak praktyce fonoskopijnej znane są przypadki, że osoba, od której miał zostać pobrany materiał, odmówiła uczestnictwa w czynności, w związku z czym zleceńodawca przekazywał do badań materiał rejestrowany bez wiedzy mówcy (np. pochodzący z przesłuchania) [13]. Należy zaznaczyć, że warunkiem niezbędnym do nadania takim nagraniom statusu materiału porównawczego jest jednoznaczne ustalenie tożsamości mówiącego. Tego rodzaju materiał charakteryzuje zazwyczaj znacznie gorsza jakość niż nagrania sporządzane przez biegłych w warunkach laboratoryjnych. Ponadto nie istnieje wówczas możliwość pobrania testów zdaniowych i, co się z tym łączy, analizy głosek znajdujących się w kontekstach fonetycznych tożsamy z dowodowymi, a to w konsekwencji może doprowadzić do zmniejszenia stopnia prawdopodobieństwa wniosku końcowego i wydłużenia czasu badań.

W KEÚ praktykowane jest również pobieranie materiału porównawczego poza laboratorium, czyli głównie na komisariatach policji w innych miastach. Przeznaczone do tego celu pomieszczenia muszą spełniać określone wymagania – pokój musi być wyposażony w dywan, meble i zasłony, a okna nie powinny wychodzić na ruchliwą ulicę. Nawet wówczas materiał taki cechuje nieco gorsza jakość niż zarejestrowany w warunkach laboratoryjnych. Materiał porównawczy w postaci mowy pobierany przez biegłych IES jest rejestrowany wyłącznie w pomieszczeniach IES.

2.4. Badania identyfikacyjne. Różnice między laboratoriami

Zarówno w IES, jak i KEÚ pomiary mierzalnych parametrów mowy dokonywane są z wykorzystaniem oprogramowania STx opracowanego przez Austriacką Akademię Nauk w Wiedniu. W obu omawianych ośrodkach jednym z etapów badań identyfikacyjnych jest porównywanie dynamicznej struktury formantowej, czyli

wartości chwilowych formantów dla samogłosek wyekstrahowanych z materiału dowodowego i z materiału porównawczego. W IES dodatkowo sporządzane są również rozkłady średnich wartości częstotliwości poszczególnych formantów, które stanowią akustyczny opis samogłosek danego mówcy. Porównanie wartości formantów samogłosek wyekstrahowanych z materiału dowodowego z częstotliwościami uzyskanymi dla materiału porównawczego pozwala na stwierdzenie, czy mieszczą się one w zakresie zmienności mówcy.

Jak wspomniano wcześniej, biegli pracujący w KEÚ oraz IES posługują się metodą językowo-pomiarową, jednak na poziomie poszczególnych etapów tych badań oraz wartościowania analizowanych cech widoczne są pewne różnice. Eksperti słowackiego laboratorium w swojej pracy pt. „Fonoskopia kryminalistyczna” definiują badania identyfikacyjne jako kombinację metod lingwistycznych, fonetycznych i elektroakustycznych [10]. Pierwsza część badań – analiza językowa – ma więc obejmować opis leksyki, morfologii i składni mówiącego. Jako drugi etap badań autorzy skryptu wymieniają analizę fonetyczną, która obejmuje zjawiska segmentalne (wymowa głosek i ich połączeń) i suprasegmentalne (cechy prozodyczne). Badania fonetyki bazują na ocenie audytywnej i dotyczą zarówno zjawisk związanych z budową fizjologiczną artykulatorów, jak i cech nabytych w procesie socjalizacji. Trzeci etap badań to analiza parametrów mierzalnych mowy. Biegły dokonuje ekstrakcji parametrów mowy – przede wszystkim formantów i częstotliwości podstawowej tonu krtańowego – odrębnie dla materiału dowodowego i porównawczego, aby następnie przeprowadzić analizę porównawczą. Do tego etapu badań należy także zaliczyć wykorzystanie automatycznego systemu identyfikacji mówców, za pomocą którego biegli z bratysławskiego ośrodka weryfikują ustalenia poczynione w trakcie wcześniejszych badań. W słowackim laboratorium obowiązuje reguła, zgodnie z którą dwa pierwsze etapy analizy (językowa i fonetyczna) przeprowadzane są przez lingwistę, zaś analiza akustyczna – przez inżyniera nauk technicznych. Takie podejście opiera się na założeniu, że osoba przeprowadzająca badania techniczne dysponuje większą wiedzą na temat parametrów nagrania, co pozwala jej w pełniejszym stopniu interpretować zjawiska obserwowane w akustycznej warstwie nagrania – a więc również dotyczące mierzalnych parametrów mowy. Przedstawiony schemat postępowania ma również słabe strony – oddzielanie analizy fonetycznej opartej na ocenie percepcyjnej od analizy spektrograficznej prowadzi do pewnego rodzaju dezintegracji akustycznej informacji o nawykach językowych mówcy. Otóż analiza przebiegu amplitudowo-częstotliwościowo-czasowego sygnału mowy (spektrogramu), pomiary wartości chwilowych formantów czy rozkłady średnich wartości formantów pozwalają zobiektywizować wrażenia percepcyjne, a także od-

nieść poszczególne realizacje językowe do całości systemu fonetycznego mówcy. Dlatego też w IES badania cech językowych oraz analiza mierzalnych parametrów mowy przeprowadzane są przez tę samą osobę z wykształceniem językoznawczym.

Istotna różnica między omawianymi laboratoriami dotyczy również wartościowania przydatności cech z różnych poziomów opisu języka do celów kryminalistycznej identyfikacji mówcy. Jedno z założeń, na których opiera się metodologia badań w IES, to twierdzenie, że fonetyka, a więc wymowa poszczególnych głosek czy wyrazów, różnicuje mówców w stopniu wyższym niż np. leksyka czy składnia. Dobór środka leksykalnego czy składniowego jest znacznie silniej uświadamiany przez mówców niż posługiwanie się danym wariantem fonetycznym, a ponadto jest zdeterminowany przez sytuację komunikacyjną – np. stopień jej oficjalności. Niestabilność dystrybucyjna leksyki i składni wpływa na ich mniejszą wartość w procesie identyfikacji mówcy. Dlatego badania porównawcze w IES opierają się przede wszystkim na analizie fonetycznej, a opis zjawisk z wyższych poziomów języka ma jedynie wartość pomocniczą.

Badania identyfikacyjne polegają na porównaniu wypowiedzi pochodzących z materiału dowodowego względem wypowiedzi porównawczych pobranych od osoby wskazanej przez zleceniodawcę. W KEÚ, oprócz indywidualnych badań identyfikacyjnych, wykonuje się również badania grupowe, czyli tzw. typowanie. Na podstawie analizy dowodowych wypowiedzi ekspert ustala obszar geograficzny, z jakiego może pochodzić mówca, jego status społeczny, podaje płeć, przybliżony wiek, a także określa stan psychofizyczny. Podstawy metodologiczne tego rodzaju badań są kwestią dyskusyjną. Kontrowersje budzi m.in. kwestia wypowiedziania się na temat stanu emocjonalnego mówcy – powstają tu bowiem wątpliwości dotyczące kompetencji biegłego fonoskopii w zakresie określania stanów psychicznych, a także możliwości wnioskowania na temat aktywności psychicznej w oparciu o szczupły materiał, zazwyczaj nieuwzględniający zmienności zachowań językowych mówcy w różnych sytuacjach.

3. Podsumowanie

Badania identyfikacyjne mówców w KEÚ oraz IES przeprowadzane są z wykorzystaniem metody językowopomiarowej. W obu laboratoriach wnioski formułowane są w oparciu o wyniki analizy językowej i pomiary parametrów akustycznych. Różnice dotyczą zakresu badań, technik pobierania testów zdaniowych oraz odmiennych założeń dotyczących wartości dyskryminacyjnej cech z poszczególnych poziomów opisu języka. Relacje między systemami fonetycznymi języka polskiego i słowackiego mogą stanowić punkt wyjścia do dalszych badań

nad właściwościami artykulacyjnymi mówców słowackich posługujących się językiem polskim i odwrotnie – Polaków mówiących po słowacku, a zatem do badań nad efektywnością metody badań identyfikacyjnych w odniesieniu do mówców niebędących rodzimymi użytkownikami danego języka.

Podziękowania

Pragnę podziękować mgr. inż. Miłošowi Lukovičowi oraz mgr Adrianie Farkašovej, którzy przedstawili mi szczegółowo metody pracy w bratysławskim laboratorium, za ich zaangażowanie we wspólną pracę, za cenne uwagi i konstruktywne dyskusje.

Katarzyna Klus

Niniejsze badania zostały sfinansowane dzięki programowi Leonardo da Vinci „Współczesne nauki sądowe – wybrane zagadnienia”.