FROM *VISIBLE SPEECH* TO VOICEPRINTS – BLESSING AND CURSE OF SOUND SPECTROGRAPHY

Angelika Braun

University of Trier, Germany
brauna@uni-trier.de

Abstract: Sound spectrography is probably the single most widely used tool in acoustic phonetics today. It was first called *Visible Speech*, thus linking it to Alexander Melville Bell's 1867 publication. At that time, of course, the term only referred to auditory analysis and the iconic documentation thereof. During World War II, sound spectrography as we know it today was developed in the US as part of a war project and was therefore not published until well after the war. A similar device was built in the Soviet Union in the post-war period. Although claims were made on both sides that the device was intended for the benefit of the congenitally deaf, it is quite clear that its primary purpose was in speaker identification. This contribution focuses on the development of the sound spectrograph and outlines its role in (forensic) speaker identification.

1 Introduction – Visualizing speech

1.1 The early days

One of the fundamental problems with human speech and its analysis is that it is volatile. Until the late 19th century, there were no technical means of conserving spoken language. The Edison phonograph dates back to 1877, and the gramophone was not developed until ten years later by Emil Berliner. The next step in the conservation of spoken utterances was once again undertaken by Berliner and consisted in the development of shellac disks in 1896. It was not until the 1930s that magnetic tape was used to record longer stretches of sound than would fit on a long play disk. Reel-to-reel tape recorders were largely replaced by tape cassettes in the 1960s, and the era of digital recordings did not start until the 1980s (cf. [33] for an excellent summary).

Attempts at visualizing speech in terms of acoustic dimensions started even later than the development of devices to record it. In the early days of experimental phonetics, various aspects of speech were made visible by way of kymograms. The kymograph could be used to record just about any kind of event over time, be it respiratory movements or volumes, phonation, or certain aspects of articulation as, e.g., nasality. The device consisted of a rotating drum which was covered by paper that had been smoked with petroleum fumes, and a stylus to record various kinds of signals. After the recording process, which could extend to a few seconds at best (depending on the diameter of the drum and the speed at which it rotated), the trace left on the blackened paper had to be fixated by a substance chemically similar to hair spray. The next step in visualization was the cathode ray oscillograph which could be used to depict the time signal in almost real-time.

10.21437/HSCR.2017-8

1.2 The Sound Spectrograph

The single most important device for the visualization of speech and its quantitative analysis, however, was the sound spectrograph (cf., e.g., [23]). Baken and Daniloff [5] call it "a genuine landmark in the phonetic sciences, a birth announcement of a new era" (p. 1). The device consists of a rotating drum which is covered by electrosensitive paper (reminiscent of kymography). In its analog form it uses electronic filtering of the speech signal in order to display the energy which is present in various frequency components over time by means of a moving filter and a stylus etching a pattern into the paper according to frequency and intensity. (For a detailed description cf. [4], pp. 329-345.) The spectrogram will thus depict time (abscissa), frequency (ordinate), and intensity (blackening or – today – color) in a pseudo-3D plot.

The digital age has added a lot of flexibility to spectrographic analysis – the sampling rate, quantization, windowing and the color scheme, to name only a few key parameters, can be adjusted by mouse click. Above all, the speed of the analysis has tremendously increased from taking 2½ minutes for a 2.4 second recording ([19], p.75) to a fraction of a second for long stretches of recorded speech. However, the three dimensions depicted in a spectrogram (time, frequency, and intensity) are still and will remain indispensable for acoustic phonetic analysis. Baken and Daniloff's ([5], p. x) assessment from the early days of the digital age still holds today: "During recent years, speech sound spectrography has become easier, cheaper, and potentially more useful than ever."

2 *Visible speech* and the teaching of the deaf

Being able to record and – in particular – visualize speech was of prime importance for teaching congenitally deaf individuals to speak before the age of cochlear implants. There is a fairly long tradition of efforts in this area in various countries, e.g., Spain (Pedro Ponce de Leon 1516-1584), The Netherlands [2], and England [12; 41]. Somewhat later, Alexander Melville Bell [6] devised a "phonetic" alphabet which was also known as the physiological alphabet. He called it *Visible Speech*. It symbolizes somewhat iconically the positions of the vocal organs during the production of the respective sounds [1]. It was meant to serve as a tool which would allow deaf patients to learn to articulate properly.

The following illustration is a projection of Bell's symbols onto the current IPA consonant chart:

251658240

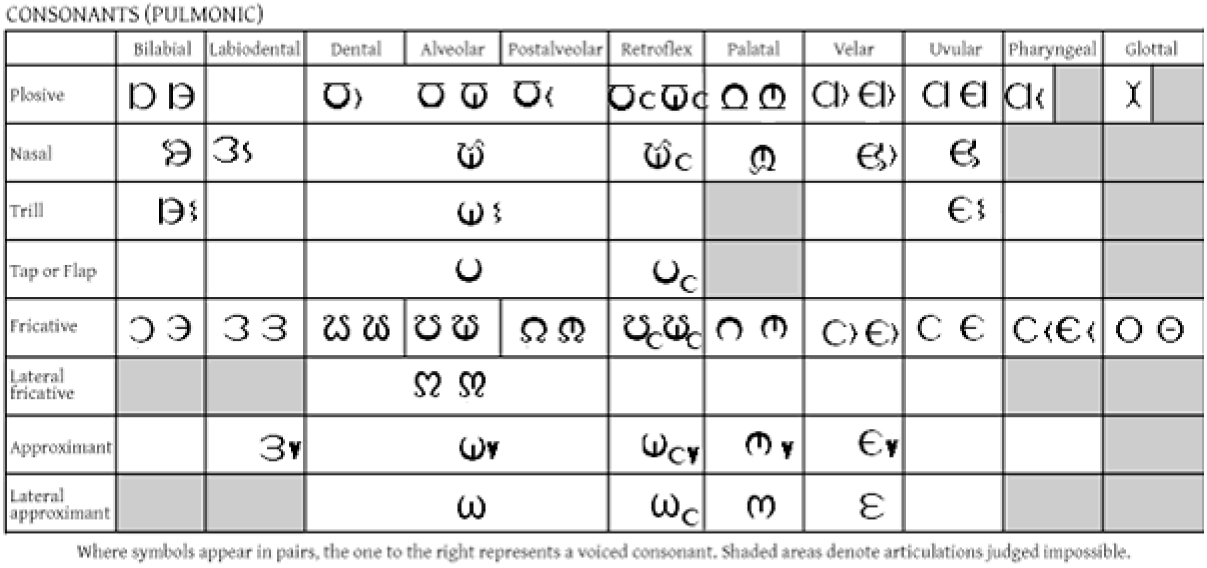

Table 1: IPA consonant chart with *Visible Speech* symbols (www.omniglot.com/writing/visiblespeech.htm)

3 *Visible Speech* revived

3.1 Official reports and publications

The publicly accessible sources on the development of the sound spectrograph are all post-World War II. Those sources [28, 29, 30] demonstrate that even before World War II researchers at Bell Telephone Laboratories in New York, N.Y. (henceforth: Bell Labs) had begun "to replace [...] an extensive array of fixed filters with one adjustably tuned unit" for the purpose of the acoustic analysis of speech ([28], p.1). According to Potter ([29], p.1), "A great deal of patient development work was necessary before the pictures were good enough to show the detail they are able to show today." To be precise, two devices were developed: A "cathode-ray translator [producing] instantaneous but transitory patterns on a phosphorescent screen in a rotating cathode-ray tube" ([19], p. 75; [8], [31]), and the spectrograph producing the kind of pseudo-3D images which are in principle still in use today. The former was not available until "late in 1944" ([30], p.5), whereas "a spectrograph development was started early in 1941" (ibid, p.4). This can be taken as an indication of the relative importance which was attached to those two devices, the spectrograph clearly taking priority.

The latter view is confirmed by Potter et al. ([30], pp.4-5), according to whom the purpose of creating those devices was twofold. "American participation in the war emphasized the military applications of, and needs for, a visual translation of sound, and during the war period major interest was centered upon these military requirements. It was not until military needs had been met that opportunity was offered for a resumption of the development of visible speech toward the goal of visual hearing" ([30], p.5).

Still, the monograph [30] is for the most part devoted to spectrogram reading and thus to the guise of clinical use. All other post-war publications likewise focus on clinical aspects, i.e., they emphasize that the development of the translator in particular occurred with deaf patients in mind. Ostensibly, a prime goal of their war time research was to enable congenitally deaf individuals to use the telephone. In fact, they describe at length how a group of those patients learned to "read" spectrograms ([30], chapters 4, 13). This, however, does not in the least explain the development of the spectrograph. The latter was said to be motivated by "other interests in sound portrayal that are important to those working with sound in all its variations" ([29], p.2). Nowhere in these articles, let alone in the Potter et al. [30] monograph is there any clear indication of a specific use for speaker identification purposes. Instead, there is ample reference to the fact that spectrograms produced by different speakers look alike (e.g., [28], p.465). However, he goes on to remark that "[t]he pattern shapes [...] vary with individual characteristics in much the same way that handwriting varies among individuals" (ibid.). This vague analogy to handwriting, which is much more adequate than that to fingerprints, remains the closest hint at voice comparison efforts in the post-war publications. Kopp and Green ([19] p.87) write: “It is axiomatic that no two individuals have voices that are exactly alike in pitch or vocal quality, but visible speech, in the form considered in this paper, does not emphasize these variables.” Thus, while individuality of speech patterns is assumed, the speaker specificity of sound spectrograms is expressly played down.

3.2 The mysterious report

Contrary to the impression created by, e.g., Kopp and Green [19], it is quite clear that the sound spectrograph was not primarily developed for research purposes. Instead, as early as 1941, a project on speaker identification was initiated by the American Military Forces at that same laboratory. C.H.G. Gray supervised the project until the summer of 1945. According to Potter et al. [30] he was succeeded by J.C. Steinberg, whose work [37] had served as a basis for the development of the spectrograph ([25], p.85).

Potter ([28], p.463) points to the relevance of the project for the war, but does not become specific: "The work here described was begun before the war. Because of related war interests it was given official rating as a war project, [...]". Thus, the project was classified during WW II for obvious reasons, assuming that the rendition given by Tosi ([40], p.67) is correct. At any rate, the early 1940s can be identified as the time when the research unit at Bell Labs working under C.H. Gray succeeded in developing the sound spectrograph.

There are no publications whatsoever currently available from those early days, but reference is made by Tosi [39] and [40] to an internal Bell Labs report from 1944 which apparently summarizes the war-time usage of the spectrograph [10]. Tosi must have had this report at his disposal since he quotes from it extensively ([39], p.415; [40], pp.67f.). Other researchers also mention this source, but it is quite clear that they did not have a chance to look at the original document since the name of the first author is repeatedly and systematically misspelt as Gr**e**y as opposed to Gr**a**y [9, 24]. Hollien [13], p.352) lists a "report *to* Bell Labs" [emphasis mine, AB] by Gray and Kopp [10] among his references, but does not refer to its content in the respective chapter of his book.

In view of the crucial information which, according to Tosi [39; 40], is contained in the alleged report by Gray and Kopp from 1944 [10], the present author has made extensive efforts to acquire a copy of it. Nokia Bell Labs felt unable to either supply a copy or even confirm its existence[1]. Neither the Library of Congress Catalog nor the National Archives and Records Administration (NARA) nor the Naval Research Lab (NRL) nor the NSA Archives Office

[1] Rebecca S. Nadolny of Nokia Bell Labs by email on 01-11-2017.

yielded any result. Thus, the present contribution has to rely on Tosi [39; 40] as a secondary source.

Tosi quickly dismisses the guise of the spectrograph as a device for the congenitally deaf and leaves no doubt as to its military use: "The machine was developed with the primary purpose of helping the military during the war" ([40], p.64). He continues to report that "[h]owever, in 1944, Gray and Kopp found that spectrograms could be applied to the identification of speakers [...]" (ibid, p.67). He attributes the coinage of the term "voiceprinting" to those authors (ibid, p.67) and quotes from their report as follows:

"Much more simple and quantitative than the aural scheme is a technique based upon the Sound Spectrograph, an instrument which makes permanent patterns of speech that can be examined visually in the way fingerprints or handwriting can be examined and can be filed as reference samples." ([10], p.1; after [40], p.67)

According to Tosi ([39], p.415; [40], pp.67f.), the report was concluded as follows:

"Voice print identification seems to offer the[2] possibility of a useful radio intercept feature and one that could be put into use without extensive training or elaborate equipment. As a study of combat conditions in the European and Pacific theaters indicates a need for such identification, it is suggested that a trial voice identification group be established to carry on the work under actual or simulated field conditions to test the conclusions of this laboratory work."

Tosi ([39], p.415) adds that "[t]o the best knowledge of this writer, such a test was not completed because of the end of the war and the problems of recordings."

From these quotations it emerges that the spectrograph was developed primarily with speaker identification in mind. It would seem that the intention was to use it to identify the voices of German radio operators on various vessels in order to be able to track the movements of the enemy's fleet. Based on goniometry and the matching of radio operators with a specific vessel, the movements of the respective vessels could be determined (cf. [40], p.67).

Indirect proof of this use of spectrograms is contained in an internal report [3] which was provided to the present author by Rebecca S. Nadolny, archivist at Bell Labs. This report summarizes the activities related to sound spectrography in the 1940s as follows: "During World War II, acoustics scientists C. H. G. Gray and G. A. Kopp suggested that enemy radio voices might be identified by spectrogram in order to detect the movement of units from place to place" ([3], 214). This entirely confirms the information contained in Tosi [39; 40].

This must have given the American military forces a significant strategic advantage over the enemy and probably formed an incentive for post-war usage as a tool in forensic speaker identification (see below), even though the matter was not publicly followed up in the early post-war years.

Even though this military focus on the development of a device which could display speech in a quasi-three-dimensional way was clearly prevalent during WW II, it was only mentioned in passing in one the – numerous – post war publications on sound spectrography

[2] The definite article is missing in Tosi [39]. Other than that, the wording is identical.

by various researchers from Bell Labs ([28], p.465). An impressive number of publications by different authors ensued in 1946, basically occupying one whole issue of The Journal of the Acoustical Society of America with a preface written by Ralph Potter [29]. They cover various aspects of spectrographic analysis, including technical detail [21; 31; 8] and – once again – the issue of spectrogram reading [38; 19] but there is no mention of using this technique for speaker identification purposes.

Lonnie Smrkovski, one of the prime advocates of voiceprint identification, hints at the voiceprint application in the 1940s, but does not become specific: "World War II provided a catalyst for research in speaker recognition, and Bell Laboratories conducted experiments in an attempt to develop voiceprints as a tool which could be used in the war effort. The war ended before such a method could be developed, and with its termination there did not appear any need for a means of identifying speakers' voices" ([34] pp.5-6).

Meuwly ([25], p.86) reports that precisely this application of sound spectrography was "classified and financed by the army" in the post-war period. Further evidence for the fact that the "voiceprint" technique was indeed commonplace within US Federal agencies shortly after WW II comes from the FBI itself: Koenig ([20], p.2089) remarks that "FBI examiners have used the spectrographic technique since the 1950s [...]." This, of course, leads to the conclusion that the interest in using sound spectrograms as "voiceprints" never ceased to exist, but was kept secret.

4 *Visible Speech* behind the Iron Curtain

It is interesting to note that similar efforts at constructing a device which would allow for comparing visual renditions of verbal utterances were made in Russia at about the same time as in the US. However, the evidence of the construction and use of a sound spectrograph in Russia is of a completely different nature. In his novel *First Circle,* Aleksander Solzhenitsyn describes an episode dating back to December of 1949. The novel is set in a prison camp for experts in various fields, whose assistance was considered useful by the Stalin regime. Therefore, the scientists were held captive but did not face an immediate threat of death.

One of the tasks given to the scientists in 1949 was to develop a device for graphically depicting speech utterances, which could, in turn, be decoded by trained experts or used as an aid for congenitally deaf patients. The description of what they came up with under the guidance of a character called "Rubin" reads like a blueprint of the Bell Labs spectrograph: This starts with its name: *Widimaja retschj* 'Visible Speech'. Even the guise about it being an aid for the deaf is copied. In Chapter 31 of the German edition, the sample utterance which is to be "read" from a spectrogram by the prisoner called Rubin is "*Swukowidy rasreschajut gluchim goworitj po telefonu*" 'Sound diagrams allow deaf persons to speak over the telephone'. This is reminiscent of the description of that same device allowing a congenitally deaf engineer who was trained to read spectrograms in the US to reconstruct the wording: "The translator has [...] provided a means for this congenitally deaf engineer to make and receive telephone calls" ([30], p.25).

It also emerges from Solzhenitsyn's rendition that spectrographic analysis in that particular case was to be used for forensic phonetic purposes. It was called – voiceprint! In chapter 33 Solzhenitsyn writes: "This was really something quite new – catching a criminal using his voiceprint. So far there had only been fingerprints, a science which had taken centuries to develop. Now there was to be a new technique of 'voiceprinting', and it had to be developed in a matter of days." ([36], p.255; [35], p.233; my translation into English, A.B.) Depending on which edition one chooses to look at[3], an anonymous phone call is described to have been made either to the US embassy or a medical doctor, in each case leaking confidential information. The spectrographic analysis served to narrow down the number of suspects from N=5 to N=2 in a closed-set situation, i.e. only a limited number of speakers could have possibly made the phone call in question.

Since this application was hinted at in an unclassified publication by Potter ([28], p.465), it can be speculated how the Russians may have gained access to the technical information which led to the development of the device. At any rate, it does not seem conceivable that Lev Kopelev, who is the scientist represented by the character "Rubin" in the novel, could have built a spectrograph from scratch, let alone within the time span in question. The time at which the novel is set suggests that the Russians were trailing the American scientists in developing a sound spectrograph by about four years if one considers the publicly accessible papers starting with Potter [28] and by eight years if the completion of the device for military purposes is taken into account. This is what gave rise to the above-mentioned speculations on the source of Russian knowledge about the device. It would seem, however, that none of those speculating have ever taken a close look at Solzhenitsyn's original text. He literally says: "Rubin himself had already forgotten that this device had been bootlegged from an American journal." ([36], p.244)[4]. This detail is lacking in the German edition of 1968. It may not be mentioned in the English translation either. A reason for this may be that most translations (the German one for certain) are based on the 1968 edition of the novel, whereas the remark in question is taken from the 1978 edition, which is known to be more comprehensive than the earlier ones. This constitutes indirect proof for the fact that from the start the main purpose of the sound spectrograph was to compare speakers' voices. It finally leaves no doubt as to who first developed the device – these were quite clearly the scientists at Bell Labs, their Russian colleagues lagging behind by close to a decade. Eriksson (n. dat., p. 4) correctly points out that the term "phonoscopy" is still in use in the Eastern European Countries. This is confirmed for Russia by Feruza Bayramova, a Russian speaker identification expert (p.c. April 2017).

5 From *Visible Speech* to "voiceprints"

Up to now, the term "voiceprint" has been closely associated with the name of Lawrence Kersta ([32], p.107; [22], p.58; [13], p.212). And, yet, it may be much older than that, dating back to Gray and Kopp [10], if one chooses to believe Tosi ([39], p.415). Plausibility is added to his version by the fact that it was adopted by Solzhenitsyn before Kersta ever published it. – Still, Lawrence Kersta played a key role in establishing "voiceprints" as courtroom evidence, which merits looking at in more detail.

5.1 The role of Lawrence Kersta

In their preface to the first edition of *Visible Speech*, Potter, Kopp, and Green [30] pay tribute to all members of the group of scientists who took part in the development of the sound spectrograph (p. viii). One name which is notably absent is that of Lawrence Kersta (ibid.,

[3] The original version contained 96 chapters, but Solzhenitsyn did not envisage ever seeing it published and therefore created a second version which consisted of only 87 chapters. The German translation is based on this version. The full version was published in Moscow in 1991.

[4] This can be taken to refer to Koenig et al. [21], which contains a technical description of the Sound Spectrograph.

p.viii). While we do know from Kersta [15] that he indeed worked for Bell Laboratories at the time and was involved in extending the spectrograph, he is neither listed by Potter et al. ([30], p. viii) as being among those who developed the sound spectrograph nor as a member of the translator development team, let alone as a supervisor. His personal merit in the context of sound spectrography obviously consisted of expanding the methodology to display amplitudes [15] including the development a new kind of visualization, i.e., the "contour spectrogram (voiceprint)", in which amplitude "is shown by seven quantized or contour steps" ([16], p.1254). Its use in voice comparison is mentioned in the famous article in *Nature*, but the results for "contour type voiceprints" were in fact slightly worse than those of "bar type voiceprints" (ibid., p.1256).

According to Tosi ([39], p.415), Lawrence Kersta took up the idea of using sound spectrography for speaker identification purposes "at the request of a law enforcement group." In 1962, he published an article and two abstracts in highly reputed Journals, which set off an unfortunate development in forensic speaker identification [16; 17; 18]. These articles established the infamous analogy between voiceprints and fingerprints. Interestingly, his article published in *Nature* [16] does not contain any references to back up his ideas, but relies on personal convictions instead: "My claim to voice pattern uniqueness then rests on the improbability that two speakers would have vocal cavity dimensions and articulator use-patterns identical enough to confound voiceprint identification methods."(p.1255). He goes on to make a strong, yet unsubstantiated claim: "It is my opinion, however, that identifiable uniqueness does exist in each voice, and that masking, disguising, or distorting the voice will not defeat identification, if the speech is intelligible" ([16], p.1257). Kersta even went so far as to use ten words in voice comparison in order to emphasize the fingerprint analogy (ibid., p.1255). The closed-set experiments which he published seemed to support his claim of "voiceprint infallibility". In 1966, Kersta left Bell Labs and founded his own company called "Voiceprint Laboratories Corporation" in Somerville, N.J. ([25], p.87).

It is interesting to note that at that time the support by Bell Labs for the notion of voiceprints seems to have had receded to some extent. An internal report of June 1962 [3] is phrased much more carefully than Kersta's own publications: "He [i.e. Mr. Kersta, AB] believes that voice identification will eventually achieve a degree of accuracy somewhere between the present accuracy of identification by fingerprints and identification by handwriting" (p.215). The allusion to handwriting is most notable at this point in time – it reflects the observation made by Potter 17 years earlier (cf. 3.1 above). One might speculate that by that time the fingerprint analogy had been given up by those who had originally created it but was upheld by Kersta.

5.2 How voiceprints survived

In 1966, the first criminal case was tried in an American court in which voiceprint evidence was introduced (*People vs. Straehle* – a perjury case) with Lawrence Kersta as an expert. This trial resulted in a hung jury. According to Tosi ([40], pp.137f.), Kersta's company *Voiceprint Laboratories* worked on about 200 samples in total, but Kersta only ever testified eight times. Of those, voiceprints were admitted into evidence only once. The company went bankrupt in 1973. Lawrence Kersta passed away in 1994.

In 1970, the Michigan State Police established a *voice identification unit.* This unit worked on about 4,000 samples throughout the 1970s and testified in court about 65 times ([40], p.140). Of all samples, 3,400 (85%) were either eliminated from the process or resulted in a *non liquet*. Koenig ([20], p.2089) reports that "[t]he FBI conducts forensic voice identification examinations using the spectrographic or voiceprint technique for the FBI, other Federal agencies, state and local law enforcement authorities, and many foreign governments." He goes on to remark that those reports are "[...] for investigative support, but [the agents] have not provided expert court testimony on comparison results" ([20], p.2089). In other words, this method is deemed suitable for covert investigations by the FBI and for implicating, if not indicting, people, but the agents will not stand up to its results in court. Of the 2,000 comparisons reported by Koenig, about two thirds (65.2%) yielded a no decision or a low confidence decision. There were only 15.9% "identifications" ([20], p.2090). These numbers do not exactly speak for the practical usage of this method.

In the 1970s, ample evidence was published demonstrating the fallibility of voiceprint identification ([7; 11; 14], to name only a few). Hazen [11] found error rates of up to 52%, which are in sharp contrast to the 1% reported by Kersta ([16], p.1256). On a different strand, Tosi tried to render this procedure respectable by refuting the voiceprint analogy, adding auditory analysis and calling the method "aural and spectrographic examination" ([40], p. ix). In the late 1970s, voiceprint evidence was admissible in 23 states (ibid., p.143). The exact number of states whose judiciary will admit them today is not known to the present author, but voiceprint evidence continues to be used by the FBI (still without testifying to it, though; cf. [20]), and by some private companies, e.g., Owen Forensic Services LLC, run by Thomas Owen. What is worse, “[v]oice fingerprinting” is still considered to be the state-of-the-art technology in voice comparison in certain countries [27].

6 Conclusions

The perspective on *Visible Speech* has shifted from creating a tool for the training of the deaf towards a device depicting the major dimensions of speech: time, frequency, and intensity. While sound spectrograms have become an indispensable research tool which is used world-wide, they are totally unsuitable for the purposes which allegedly and actually led to their development in the first place in the US as well as the USSR, i.e. enabling the congenitally deaf to use the telephone as well as speaker identification. Sonagrams have indeed been used as models for training deaf persons to articulate more clearly, but in the past two decades, cochlear implants have become the method of choice. Whenever CIs did not turn out to be an option, spectrograms have been replaced by, e.g., ultrasound, because it reflects the actual tongue position more directly than the spectrogram.

As far as the forensic perspective is concerned, no properly qualified forensic phonetician would today subscribe to the analogy between spectrograms and fingerprints. The main reason for the latter is what Francis Nolan has termed the "plasticity of the vocal tract" ([26], p. 748], i.e. the inherent intraspeaker variability owing to age, health status, speaking context etc. which may well exceed interspeaker differences as are displayed in a spectrogram. And yet, the concept of voiceprints still seems so attractive to interested parties that it has not to date been eradicated.

Acknowledgements

I would like to thank Natalia Filatkina of the University of Trier for her kind assistance with the translation of the Russian original. I am indebted to Rebecca Nadolny of Nokia Bell Labs

for her patient search for the Gray and Kopp Report and to Harry L. Harting for going through numerous microfilms at NARA trying to find that same document.

References:

[1] Abercrombie, David: Elements of General Phonetics. Edinburgh: Edinburgh University Press 1967.

[2] Amman, John Conrade: The Talking Deaf Man: A Method Proposed, Whereby He Who is Born Deaf, May Learn to Speak. Amsterdam: Westein 1692.

[3] Anon.: Voice Spectrograms Are Unique Personal Identification. Bell Laboratories Record, June 1962, 214-215.

[4] Baken, R.J.: Clinical Measurement of Speech and Voice. London: Taylor & Francis 1987.

[5] Baken, R.J. / Daniloff, R.G. (eds.): Readings in Clinical Spectrography of Speech. San Diego/Pine Brook, N.J.: Singular and Kay Elemetrics 1991.

[6] Bell, Alexander M.: Visible Speech: The Science of Universal Alphabetics; or self-interpreting Physiological Letters for the Writing of all Languages in one Alphabet. London /New York: N. Trübner & Co. 1867.

[7] Bolt, R. H. / Cooper, F. S. / Green, D. M. / Hamlet, S. L. / Hogan, D. L. / McKnight, J. G. / Pickett, J. M. / Tosi, O. / Underwood, B. D.: On the Theory and Practice of Voice Identification. Washington, D.C.: National Academy of Sciences 1979.

[8] Dudley, Homer / Gruenz, Otto O, Jr. (1946): Visible Speech Translators with External Phosphors. The Journal of the Acoustical Society of America 18 (1946), 62-73.

[9] Eriksson, A. (n.dat): Tutorial on *Forensic Speech Science.* 14p.

[10] Gray, C.H. / Kopp, G.A.: Voiceprint Identification. Bell Telephone Laboratories Report 1944, 1-14.

[11] Hazen, B.: Effects of differing phonetic contexts on spectrographic speaker identification. The Journal of the Acoustical Society of America 54 (1973), 650-660.

[12] Holder, William: Elements of Speech: An Essay of Inquiry into the Natural Production of Letters’ with an Appendix concerning persons Deaf & Dumb. London 1669.

[13] Hollien, H.: The Acoustics of Crime. The new Science of Forensic Phonetics. New York etc. Plenum Press 1990.

[14] Hollien, H. / McGlone, R. E.: An evaluation of the "Voiceprint" technique of speaker identification. Proceedings of the Carnahan Conference on Crime Countermeasures. University of Kentucky: University of Kentucky Press, Lexington (1976), 39-45.

[15] Kersta, L.G.: Amplitude Cross-Section Representation with the Sound Spectrograph. The Journal of the Acoustical Society of America 20 (1948), 796-801

[16] Kersta, L. G.: Voiceprint Identification. Nature 196 (1962), 1253-1257.

[17] Kersta, L. G.: Voiceprint Identification. The Journal of the Acoustical Society of America 34 (1962), 725.

[18] Kersta, L. G.: Voiceprint Identification Infallibility. The Journal of the Acoustical Society of America 34 (1962), 1978.

[19] Kopp, G.A. / Green, H.C.: Basic Phonetic Principles of Visible Speech. The Journal of the Acoustical Society of America 18 (1946), 74-89.

[20] Koenig, B. E.: Spectrographic voice identification: A forensic survey. The Journal of the Acoustical Society of America 79 (1986), 2088-2090.

[21] Koenig, W. / Dunn, H.K. / Lacy, L.Y.: The Sound Spectrograph. The Journal of the Acoustical Society of America 18 (1946), 19-49.

[22] Künzel, H. J.: Sprechererkennung. Grundzüge forensischer Sprachverarbeitung. Heidelberg: Kriminalistik Verlag 1987.

[23] Lieberman, Philip / Blumstein, Sheila E.: Speech physiology, speech perception, and acoustic phonetics. Cambridge etc.: Cambridge University Press 1988.
[24] Lindh, J.: Handling the "Voiceprint" Issue. Proceedings FONETIK, Dept. of Linguistics, Stockholm University 2004, 4p.
[25] Meuwly, Didier: Reconnaissance de locuteurs en sciences forensiques: l'apport d'une approche automatique. Thèse de doctorat, Université de Lausanne 2001.
[26] Nolan, F.: Speaker Recognition and Forensic Phonetics. In: Hardcastle, J. and Laver, J. (eds.): The Handbook of Phonetic Sciences. Oxford etc.: Blackwell 1997, 744-767.
[27] Pragnesh, Parmar / Udhayabanu, R. : Voice Fingerprinting: A Very Important Tool against Crime. Journal of Indian Academy of Forensic Medicine 34 (2012), 70-73.
[28] Potter, R.K.: Visible Patterns of Sound. Science 102 (1945), 463-470.
[29] Potter, R.K.: Introduction to Technical Discussions of Sound Portrayal. The Journal of the Acoustical Society of America 18 (1946), 1-3.
[30] Potter, Ralph K. / Kopp, George A. / Green Kopp, Harriet: Visible Speech. New – York: Dover Publications 1947.
[31] Riesz, R.R. / Schott, L.: The Visible Speech Cathode-Ray Translator. The Journal of the Acoustical Society of America 18 (1946), 50-61.
[32] Rose, Philip: Forensic Speaker Identification. London/New York: Taylor & Francis 2002.
[33] Schubert, Hans: Historie der Schallaufzeichnung. Frankfurt am Main: Deutsches Rundfunkarchiv 2002.
[34] Smrkovski, Lonnie L.: Voice Identification. Michigan Department of State Police 1977.
[35] Solschenitsyn, Alexander: Der erste Kreis der Hölle. Frankfurt: S. Fischer 1968.
[36] Солженицын, А.И.: В круге первом. Москва: Панорама 1991.
[37] Steinberg, John C.: Application of Sound Measuring Instruments to the Study of Phonetic Problems. The Journal of the Acoustical Society of America VI (1934), 16-24.
[38] Steinberg, J.C. / French, N.R.: The Portrayal of Visible Speech. The Journal of the Acoustical Society of America 18 (1946), 4-18.
[39] Tosi, Oscar: The Problem of Speaker Identification and Elimination. In: Singh, Sadanand (ed.): Measurement Procedures in Speech, Hearing, and Language. Baltimore etc: University Park Press (1975), 399-431.
[40] Tosi, Oscar: Voice Identification. Theory and Legal Applications. Baltimore etc: University Park Press 1979.
[41] Wallis, John: Grammatica Linguae Anglicanae [with an appendix: De loquela, sive sonorum formatione, tractus grammatico-physicus.] London: Leon Lichfield 1653.