

Polish pronunciation animations developed on the basis of electromagnetic articulography

Anita Lorenc¹

Department of Logopedics and Applied Linguistics, UMCS in Lublin,
trochymiuk@gmail.com

Abstract: Anita Lorenc. *Polish pronunciation animations developed on the basis of electromagnetic articulography*. The Poznań Society for the Advancement of the Arts and Sciences. PL ISSN 0079-4740, ISBN 978-83-7654-384-0, pp. 107–123

This paper discusses the subject of pronunciation visualization, based on methodologies in experimental phonetics. It presents a brief survey of major Polish instrumental research into articulation, focusing primarily on contemporary dynamic visualizations using electromagnetic articulography. The author's own investigations were conducted using an AG500 articulograph, a device which records and visualizes the working and movement of the articulatory organs. Two speakers were recorded: one with standard pronunciation and the other with articulation defects. A multi-specialist team prepared vocal tract models, taking into account the speaker's anatomical conditions as recorded with the articulograph, video recordings, and photographs. Articulographic data enabled the preparation of pronunciation animations of 45 words, which show the standard realization of all vowels and consonants of Polish, and eight animations of non-standard articulation.

Keywords: speech visualization, electromagnetic articulography, standard and non-standard pronunciation animations

1. Introduction

Contemporary studies in articulatory phonetics conducted using advanced instruments permit the verification, updating and structuring of existing knowledge about articulation. The traditional, often subjective techniques of investigation and visualization of articulation are now being replaced by the objective and accurate methods of assessment developed in instrumental phonetics (Code & Ball 1984).

For example, in respect of measuring the movements and positions of articulators, palatography has been replaced by electropalatography (EPG), where the contact of the tongue with the palate is recorded dynamically by means of a sensor-filled artificial palate. X-ray methods have been replaced by magnetic resonance imaging (MRI), while cineradiography has given way to dynamic visualizations using an electromagnetic field to investigate the position of articulators.

¹ Formerly Trochymiuk.

2. Background

The methods of pronunciation visualization used in Polish instrumental phonetic research, which developed throughout the last century, aimed first of all to present and document articulatory movements that were inaccessible to direct observation. Some places of articulation (bilabial, labiodental, interdental) can be easily assessed by sight, and as such they can be photographed or video-recorded. Thus, traditional labiograms (cf. Dłuska 1950; Wierzchowska 1967) have gradually been replaced with video-recordings which permit the researcher to capture the working dynamics of articulatory organs such as the lips, mandible or tongue (in the case of interdental or apico-dental articulations), and freely to stop, repeat, or slow down the pace of selected articulation stages (cf. Figure 1).

a)



b)



Figure 1: Video-recording sequences corresponding to selected stages of articulation of the post-alveolar fricative consonant [ʂ] in Polish: a) front view, b) lateral view. Source: own recordings

One of the earliest phonetic methods of imaging articulations inaccessible to visual assessment is unquestionably palatography, which presents the contact of the tongue with the palate in a horizontal section. Pioneering studies for Polish in this field, which have still not been updated, were conducted in 1931 by the eminent phonetician Tytus Benni (see Figure 2).

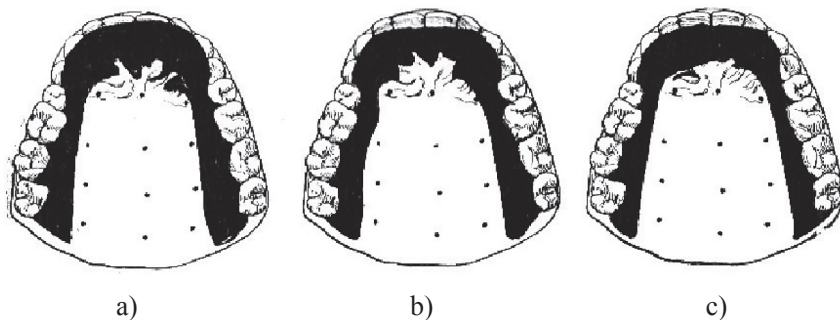


Figure 2: Examples of palatograms of apico-dental consonants: a) [t] in the word *tata*, b) [d] in the word *dama*, c) [n] in the words *nam* and *pan*. Source: Benni 1931: 11

The figures were obtained through copying the marks left by the tongue on an artificial aluminum plate covered with a layer of talcum or chalk and placed in the mouth. In the diagrams T. Benni marked the teeth and three series of reference points, which make it easier to interpret and compare the contact of the tongue with the palate for pronunciations of particular speech sounds.

In the 1970s traditional palatography began to be replaced by electropalatography – EPG (Hardcastle 1972). This procedure uses a thin, light artificial palate filled with electrodes, which record contact with the tongue during speaking at a frequency of from 100 to 200 times per second.² This technique has predominately been used in experimental phonetic research and in the diagnosis and treatment of articulation disorders in children and adults (eg. Hardcastle et al. 1991). Investigations using electropalatography have not yet been conducted in Poland.

Another group of tests used in articulation assessment involves imaging of the vocal tract. The oldest technique utilized X-rays to illustrate the position of the speech organs while pronouncing individual sounds.

Pioneering and invaluable phonetic studies for Polish in this field were carried out by Koneczna & Zawadowski (1951). The investigation consisted in taking X-ray photographs of Polish sounds being pronounced by four speakers. The soft tissues of the tongue, the lips and the palate were covered with a thin layer of radiation-absorbing substance (barium in the form of the insoluble sulfate) in order to obtain a better picture of the positioning of the articulators. Details of interest were copied from the X-ray pictures on to transparent paper. The outline of the tongue in contour projections most often consists of two or three lines which correspond to its central part and margins (cf. Figure 3). Therefore, the bottom line of the tongue in the lateral section is the medial part of its depression, a duct formed during the articulation of most sounds, while one or two upper lines are the higher-situated tongue edges, which produce a joint contour if they are equally elevated; sometimes, however, the left and right sides do not work symmetrically (which is also seen in T. Benni's palatograms).

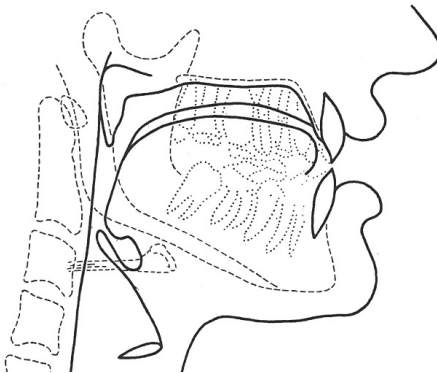


Figure 3: X-ray section of the consonant [s]. Source: Koneczna & Zawadowski 1951: Fig. 117 – speaker III

² The number of electrodes in the artificial palate ranges from 62 to 96; this number, like the sampling frequency, depends on the manufacturer.

At present, a picture of a section of an investigated articulation can be obtained through ultrasonography, which is a non-invasive method (cf. Gick 2002). The method utilizes the phenomenon of propagation, dispersion and reflection of the ultrasonic wave on the border of media. Medical ultrasonography uses ultrasounds at a frequency exceeding 20 kHz directed at the studied anatomical structures, their position being recorded 30 times or more per second. Another non-invasive method currently used in phonetic tests is magnetic resonance imaging (MRI), based on the phenomenon of nuclear magnetic resonance. It is widely used in medicine (in examining practically the whole body), in which it is one of the basic techniques of imaging diagnosis (tomography). MRI is a method of obtaining static images corresponding to a section through a particular structure of the human body (e.g. the vocal tract), at a frequency of more than four times per second, depending on the equipment type. Despite huge technological advances enabling three-dimensional computer modeling of the structures examined, the MRI method still yields unsatisfactory results with regard to image accuracy and speed of sampling (speakers have to prolong a particular articulation, often for many seconds); it is also negatively influenced by the effect of gravity on articulation, because speakers have to lie down during testing (cf. Stone et al. 2007). Neither of the abovementioned methods – ultrasonography or magnetic resonance imaging – has been used in Poland for articulation investigations.

The first attempts to represent continuous speech in phonetic tests involved the use of the cineradiographic method. X-ray techniques made it possible to obtain recordings of speech organs during the pronunciation of single speech sounds, usually prolonged in isolation or in appropriate contexts (cf. Wierzchowska 1980). In contrast, cineradiography, a film method, enabled the recording of articulatory sequences and their changes in continuous speech. Regrettably, the pictures obtained in this way were not as clear as traditional X-ray images, although they are indisputably more natural.

The ambition of phoneticians all over the world was to obtain an objective and at the same time dynamic image of the position of speech sounds during articulation.³ The turning point was the development of a device which used the X-ray microbeam system and systems determining the position of sensors in an electromagnetic field (EMA).

The X-ray microbeam was first used for studies on articulation by Kiritani (1986). This method makes it possible to locate and monitor the position of gold balls placed on the speech organs in two-dimensional space. This type of test has been conducted only in several centers worldwide, mainly because of the high costs. The technique has not yet been successfully used to produce a 3D picture; moreover, despite the reduction in the X-ray exposure time, this method cannot yet be regarded as safe for speakers' health.

Far more widely used are EMA systems, which offer a number of advantages such as high sampling frequency, the possibility of three-dimensional measurement (even 5D-measurement if the position angles of sensors are taken into consideration), safety, high data quality, and unlimited mobility of the speaker's head during investigations. Electromagnetic articulography is used to describe the movements and positions of speech organs by leading phonetic centers all over the world (USA, Canada, Australia, Japan, Russia, China,

³ According to Wierzchowska (1980), as early as 1887 the pioneer of X-ray techniques in phonetics, M. Scheier, emphasized the need to investigate articulation sequences rather than static articulatory systems corresponding to pronunciations of isolated sounds.

Germany, and the Netherlands), e.g. Massachusetts Institute of Technology (MIT), Haskins Laboratories, Queen Margaret University College (Edinburgh), Zentrum für Allgemeine Sprachwissenschaft (Berlin), University of California Phonetics Lab, Speech Dynamics Lab in the Beckman Institute, Max Planck Institutes, and many others. Thanks to articulographic studies, the articulation characteristics of many languages have been described, including English, German and French, as well as Hindi, Arabic and Hungarian. These investigations focus primarily on describing the dynamic characteristics of the movements of speech organs, their coordination and positions in particular stages of the articulations studied. Some research is devoted to the evaluation of relationships between the acoustic signal and the position of speech organs recorded by the articulograph, as well as to the application of the obtained data in speech synthesis and phonological theory. The state of research on the articulography of Polish is decidedly not too impressive. Pioneering articulography studies of Polish were carried out in the Berlin Zentrum für Allgemeine Sprachwissenschaft and focused on a small section of Polish phonetics, i.e. on the articulation of postlexically palatalized labial consonants [pʲ] and [bʲ] (Pompino-Marschall & Żygis 2003; Rochoń & Pompino-Marschall 1999). Since 2009 articulographic studies of Polish have been conducted using the AG500 equipment at the UMCS Department of Speech Pathology and Applied Linguistics, eg.: detailed testing of two speakers – with normative and defective pronunciation (Trochymiuk & Świącieński 2009; Lorenc & Świącieński 2014), comparative tests describing the articulation basis in the Poles while they speak Polish and English (Świącieński 2013) and dynamic visualizations of Polish pronunciation obtained owing to the electromagnetic articulography, a circular microphone array and high-speed video (Lorenc [forthcoming], Król & Lorenc [forthcoming]).

In recent years there has been great interest in articulographic experiments serving to develop the animation of movements of individual speech organs during speaking (cf. Gibert et al. 2012; Badin & Serrurier 2006; Engwall 2003; Revéret et al. 2000). Studies in this area have also been made in cooperation between Polish scientists⁴ and business⁵ – the results of this cooperation are described below.

3. The experiment

3.1. The research tool

Pronunciation tests were carried out using an articulograph of type AG500.⁶ This device enables recording, presentation, and assessment of the movements of the articulators (lips, tongue, mandible, and palate) in a real-time three-dimensional image during speech. Its basic function consists in measuring the position of points marked in space at a specified time. The standard version of the Articulograph AG500 comprises the following components:

⁴ P. Łobacz became the head of the team, with A. Lorenc and R. Świącieński as members.

⁵ Experiments were carried out under the project UDA-POIG.08.01.00-24-228/09-00 “Portal edukacyjny polskiej fonetyki stosowanej w zakresie normy i patologii mowy” [Educational portal of Polish applied phonetics in normal and pathological speech], Innovative Economy Operational Program, implemented by the company Fonem.eu.

⁶ The AG500 articulograph belongs to the Maria Curie-Skłodowska University, Department of Speech Pathology and Applied Linguistics.

1. The principal unit with an IDA computer with software that controls the flow of magnetic waves emitted by DTC6 transmitters and records the data from the PSR12 receiver. This internal computer is connected to an external computer by an internet cable.

2. The IDA computer-controlled DTC6 transmitters (coils) generate a magnetic field at specified frequencies, thus forming a spherical measurement bandwidth within the AG500 EMA cube to which they are attached.

3. The PSR12 receiver, which intercepts a signal from 12 sensors, passing it to the IDA computer. When more than 12 sensors are needed, another twelve-channel receiver can be connected to the system.

4. The HS220-AG500 sensors (HQ220-L165-S indicators) consist of receiving coils (see Figure 4).



Figure 4: A HQ220-L165-S sensor. Source: own work

Before a test, they are placed on the speaker's articulators. The alternating magnetic field in the system's cube induces the alternating current flow in the sensors: its amplitude is a function of the distance between the receiving coil and particular transmitter coils. The coils functioning as receivers enable measurement taking account of five variables: the spatial coordinates on the X, Y and Z axes, and two angle directions (phi and theta). These sensors are repeatedly used during measurement of articulator positions in different speakers. Because of equipment depreciation and for hygienic reasons, they should be disinfected and replaced.

1. The AG500 EMA cube constitutes the space within which the tested subject's head is placed during measurement (cf. Figure 5). The position of the EMA can be adjusted. The EMA cube has six built-in, appropriately positioned transmitters, which generate an alternating magnetic field.

2. The AG500 calibrator – a calibration device used for calibrating the data from the sensors.

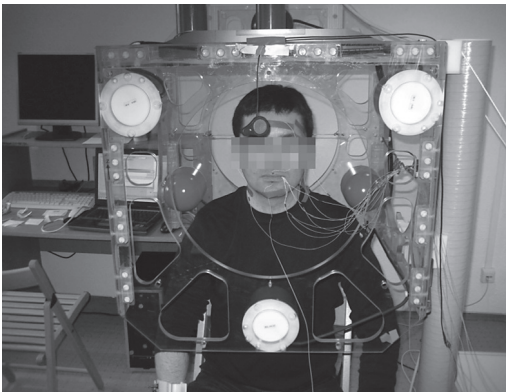


Figure 5: A subject being tested in the AG500 EMA cube. Source: own research

3.2. The recording procedure

The system was calibrated before the planned testing. For this purpose, a calibration device was placed within the cube, to which device three magazines with previously inserted sensors were attached (four sensors in each magazine). Sensor cables were then connected to the central unit, and the program calibrating the apparatus was activated. After completion of the calibration process, the report was verified to make sure that the parameters obtained showed the correct functioning of the system.

After verifying the correctness of calibration, the sensors were disconnected from the central unit and removed from the magazines for further preparation. They were first placed in a sterilizing solution for one hour, and after drying were coated with latex milk.

The prepared sensors were positioned in the previously selected points of the speaker's mobile articulators, using the non-toxic tissue glue to fix them. Three sensors had the function of enabling subsequent correction of undesirable head movements that occurred during testing: they were placed on the mastoid processes behind the ears and in the depression between the forehead and the nose. These places were selected because the sensors could not move relative to one another during testing. Then the sensors controlling lip movements were attached. They were positioned on the central facial axis, but not on the lips themselves – because the highly sensitive tissue might be torn off if the sensors came unstuck – but on the skin right on the margins of the lips. Another four sensors were placed on the medial part of the tongue: one on the tip, one in the postdorsal area and two at equal intervals between the outermost sensors (in the subject with normative speech three sensors were used, placed on the tip and on the medial and rear parts of the tongue). The next two sensors were glued onto the sides of the tongue's upper surface between the first (tip) and second sensor (cf. Figure 6 and 7). The last sensor was glued at the border between the lower incisors and gums, to monitor mandibular movements.

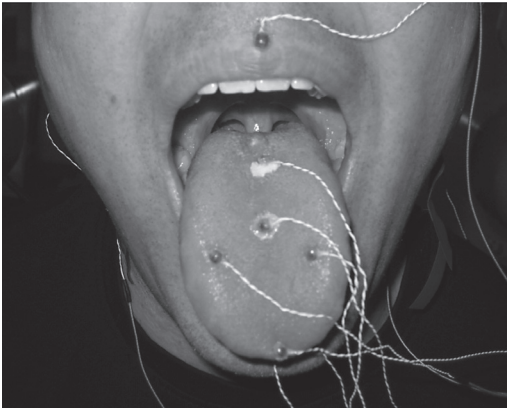


Figure 6: The positioning of selected sensors recording the action of mobile articulators in the subject with standard articulation. Source: own research

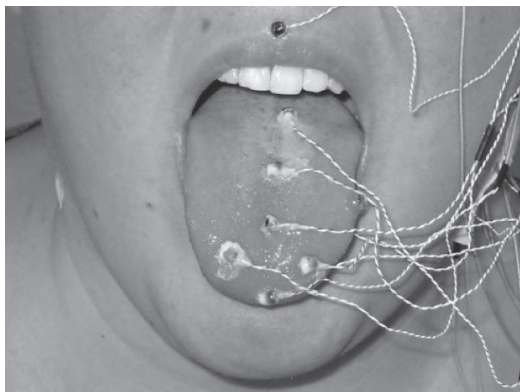


Figure 7: The positioning of selected sensors recording the action of mobile articulators in the female subject with non-standard articulation. Source: own research

One sensor served to outline the shape of the alveolar ridge, the hard palate and partly the short palate. This enabled later representation of the diagram of the abovementioned speech organs. The use of glue to attach all sensors in the mouth involved drying the surfaces to which they were affixed. Sterile gauze swabs were used for this purpose. The cables placed in the facial area were attached with adhesive plaster to the cheeks, neck and forehead so that they would not limit articulator movements with their weight.

The subjects prepared in this way were seated at the main unit, sitting in a chair in such a way that the head was in the center of the cube. After the sensors were connected to the system, and the grounding clip was clasped on the hand, the subjects were ready for testing.

After several trial recordings, the tester made the outline of the palate along the central axis, using an additional sensor attached to a wooden spatula. The subjects were then asked to read out words which appeared on the screen placed in front of them at eye level. The words appeared every three seconds.

3.3. Word list

Word lists were devised in such a way that the target consonants occupied the intra-word (medial) position, if possible between non-high and non-rounded vowels [a] or [ɛ] (to minimize coarticulatory influence), whose neutral articulation would emphasize the consonants' features, while changing their pronunciation as little as possible. Oral vowels were also evaluated in the medial position in the vicinity of plosive consonants. The contrast between the maximum consonantal plosiveness and maximum vowel openness would emphasize the articulation features of vowels. All the test words are shown in Table 1 below. The symbols corresponding to the target sounds are given in bold type in the phonemic transcription.

After the recordings were completed, the cables were disconnected from the main unit and the sensors were removed from the subject's body surfaces. After each session, the data obtained were imported from the main articulograph unit to the controlling computer. Using the manufacturer-supplied CalcPos software, the computer calculated the amplitude values obtained from individual sensors into spatial data for each sensor, taking account of the previously obtained calibration data. Then the correction for head movements was applied on

Table 1: The word list used for testing

No.	Orthographic transcription	IPA phonemic transcription	No.	Orthographic transcription	IPA phonemic transcription
1.	bity	/biti/	24.	beza	/beza/
2.	byty	/biti/	25.	Leszek	/leʂek/
3.	bety	/beti/	26.	leżak	/leʒak/
4.	baty	/bati/	27.	pasie	/paʂe/
5.	boty	/boti/	28.	bazie	/baze/
6.	buty	/buti/	29.	echo	/exɔ/
7.	węszy	/vɛʂi/	30.	heca	/xetsa/
8.	wąsy	/vɔʂi/	31.	sadze	/sadʒe/
9.	kapa	/kapa/	32.	leczy	/letʂi/
10.	żaba	/ʒaba/	33.	gadżet	/gadʒet/
11.	papier	/papʲer/	34.	zaciek	/zatsɛk/
12.	obiekt	/ɔbʲekt/	35.	podział	/podʒaw/
13.	etap	/etap/	36.	temat	/temat/
14.	pedał	/pedaw/	37.	zamiar	/zamʲar/
15.	zakiet	/zacet/	38.	denat	/denat/
16.	ogier	/ɔjer/	39.	taniec	/tanɛts/
17.	rzeka	/ʒɛka/	40.	tango	/tango/
18.	zegar	/zegar/	41.	Ela	/ela/
19.	efekt	/efekt/	42.	era	/era/
20.	kawa	/kava/	43.	zadra	/zadra/
21.	mafia	/mafʲa/	44.	daje	/daje/
22.	pawie	/pavʲe/	45.	cała	/tsawa/
23.	deser	/deser/			

the basis of the data obtained from the control sensors, again using software supplied by the articulograph manufacturer. The data thus prepared were further analyzed.

3.4. Speakers

For the study the authors investigated the articulation of a 32-year-old woman, who displayed abnormally convex articulation of the phonemes /ʂ/, /ʒ/, /tʂ/ i /dʒ/⁷ and realized

⁷ Jassem (2003) proposes the symbols [ʃ] and [ʒ] to transcribe Polish post-alveolars; Hamann (2002; 2003; 2004) and Żygis (2004) point to the retroflex character of Polish post-alveolar sibilants and suggest that appropriate symbols for retroflex sounds, i.e. [ʂ] and [ʒ] need to be used. The present investigation confirms the retroflex status of the standard pronunciation of the sounds in question and the symbols [ʂ] and [ʒ] are used throughout the text as more adequate for presenting the actual standard Polish pronunciation. Moreover, the symbols [ʃ] and [ʒ] appear more suitable for representing articulations with the convex tongue setting since, in the IPA system of notation, they indicate articulations with this type of tongue curvature (see IPA 1999; Laver 1994).

The use of the symbols [ʃ] and [ʒ] for standard pronunciation, as suggested by Jassem (2003), requires that dorsal articulations should be represented by different graphic forms. Krajna (2008) suggests [ʃ], in which the diacritic [̠] denotes advanced tongue root. It does not, however, reflect the relevant for this investigation differences in the overall shape of the tongue body.

the alveolo-palatal consonants [ɕ], [ʐ], [tɕ], [dʐ]⁸ as dento-prepalatal [ɕʲ], [ʐʲ], [tɕʲ], [dʐʲ].⁹ Moreover, the woman displayed nonstandard tongue body setting in the articulation of most Polish sounds, which consisted in pronouncing them with the tongue excessively arched towards the palate and advancing the place of articulation of most of the coronal consonants. The examination revealed that the speaker had no impediments of physical and phonemic hearing as well as no anatomical abnormalities other than the one described below.

During the interview, it was established that the disordered articulation in the examined speaker was caused by adenoid hypertrophy diagnosed when the speaker was at a kindergarten age. At the age of 13, the speaker underwent adenoidectomy and reports that the size of the removed tissue was large enough to fit in her palm. Such an overgrown organ caused difficulty with breathing through the nose and chronic infections of pulmonary airways, due to which the speaker kept the mouth constantly open. Because of the fact that the pathogenic factor acted during the period of speech development, the speaker acquired a non-standard articulatory lingual setting with the mass of the tongue shifted to the front of the oral cavity. As a consequence, the articulation of particular phonemes became non-standard. The present paper focuses on the description of dorsal articulation displayed by the speaker. For comparative purposes, the same type of data were acquired with the same methodology from a 40-year-old male speaker who was judged by three independent phoneticians to display standard Polish pronunciation.

4. Results

The model of the vocal tract subsequently used in the animation of the speech organs during speech was developed by a multi-specialist team which, apart from graphic artists and programmers, also included phoneticians, speech therapists, orthodontists, and laryngologists. The design took into consideration the actual anatomical conditions of the speakers' facial skeletons taken from photographs and video recordings. The shape of the palate was designed on the basis of the contour of passive articulators made using one sensor during the articulographic recording (cf. Figures 8 and 9).

The final shape of the speakers' vocal tracts was determined by plotting the above-described contours of passive articulators onto the schematic cross section of their facial skeletons. It also takes into account the positioning of sensors on the active articulators (the tongue, lips and mandible). Subsequently, the X and Y axes were marked as the points of reference for vertical and horizontal movements of the articulators (cf. Figures 10 and 11).

⁸ These consonants are most frequently described in Polish speech therapy and phonetic sources as 'mediodorsal' or 'palatal'.

⁹ During the articulation of these consonants the mass of the tongue was significantly advanced and the raised blade of the tongue created a narrowing in the post-dental/alveolar area, the greatest constriction being in the alveolar region, rather than in the front part of the hard palate as in normatively realized sounds. The obstruction extended to the end of the front surface of the hard palate and was created by a significantly greater part of the tongue muscle than in the norm. The centre and back of the tongue were elevated higher and were more advanced, which caused that the apex was located behind the teeth rather than in the alveolar area. The post-alveolar constriction was formed with the centre of the tongue, not its front.

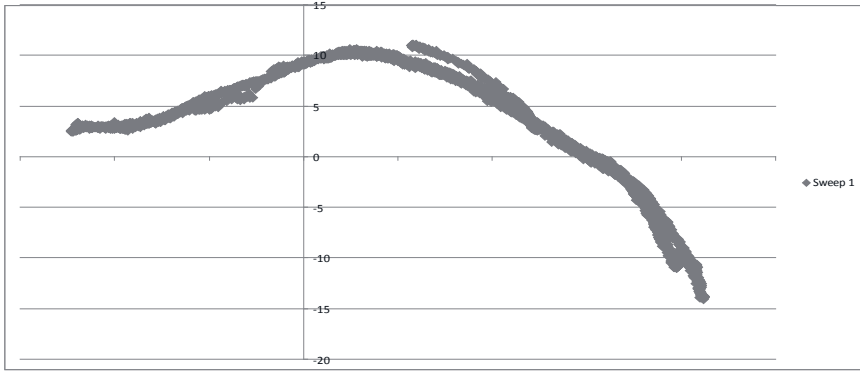


Figure 8: The contour of passive articulators of the subject with standard pronunciation. Source: own research

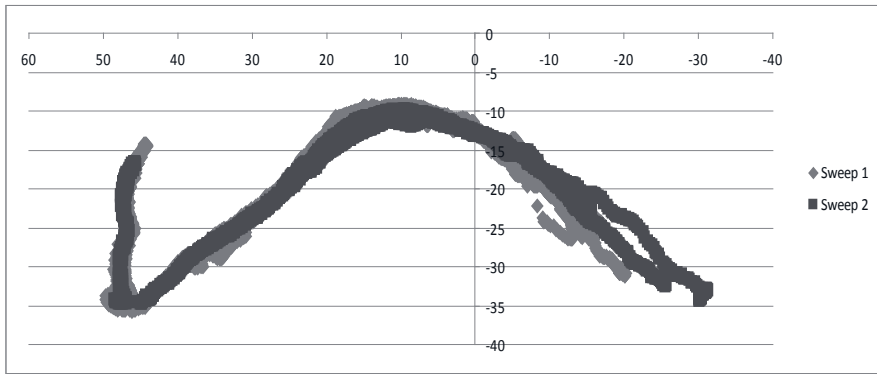


Figure 9: The contour of passive articulators of the female subject with non-standard pronunciation. Source: own research

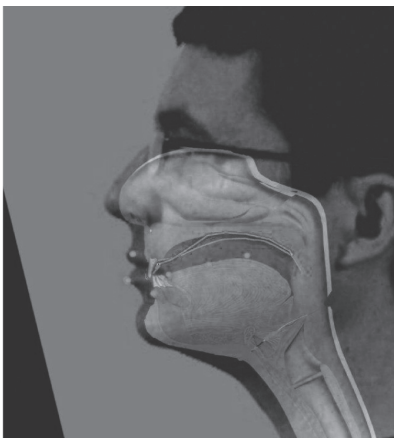


Figure 10: The model of the vocal tract taking into account the speaker's anatomical conditions, the contours of passive articulators and sensor positions. Source: own research

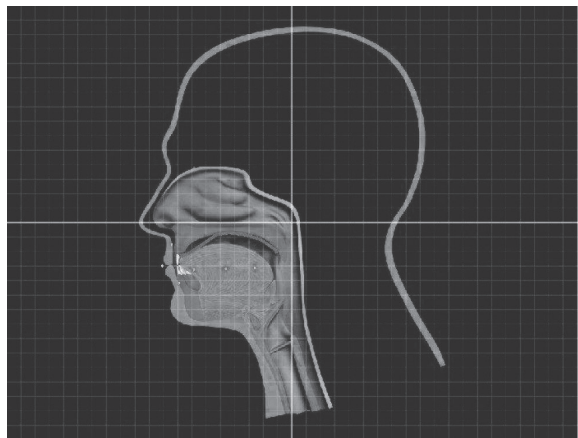


Figure 11: Schematic model of the vocal tract plotted onto the coordinate axes. Source: own research

Because the articulograph records the position of the sensors in space with the sampling frequency of 200 Hz (every 5 milliseconds) with an accuracy of 0.5 mm, the same units were adopted in designing the animations. On the basis of articulographic testing, a total of 45 animations were made for standard articulation, and 8 animations for non-standard articulation. According to the structure of the aforementioned word list, the sounds tested are always in the medial position in the word. The main advantage of the created animations is that one can observe the work of individual speech organs in their dynamic movement, stop the film at any point and replay it.

Figure 12 presents an illustration of the non-standard realization of the word *Leszek*. The vertical axis indicates the position of the sensors in the top-down dimension and the horizontal axis shows their positioning in the front-back plain (the front is in the left part of the figure). The units are millimetres and the axes intersect in the centre of the articulograph cube.¹⁰ Looking at the traces, one can notice that the sensor attached to the back of the tongue crosses the contour of the soft palate. At this point it needs to be explained that during the tracing of the roof of the mouth, the speakers were breathing through the nose. Therefore, the velum was lowered below the position in which it appears in no-nasal articulations.

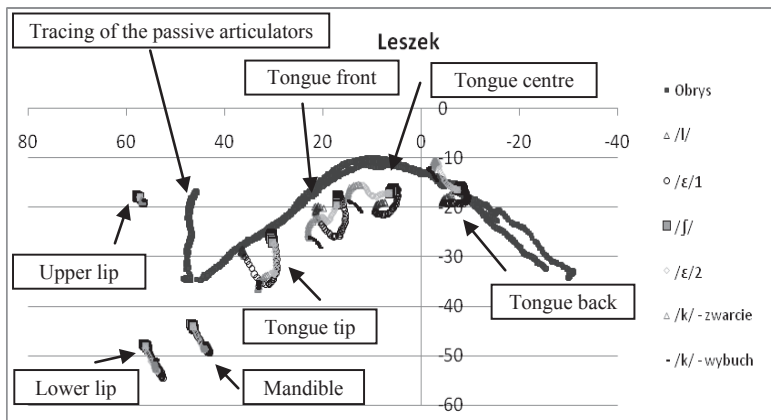


Figure 12: Articulator positions and movements during the production of the word *Leszek* by the speaker with non-standard pronunciation. The axes show horizontal and vertical position values in millimetres.

During the standard articulation of /s/ the lips become protruded, which indicates that they also become rounded. The tip of the tongue is raised and creates a narrowing in the post-alveolar region. It is an apical articulation. The mass of the tongue behind the constriction is lower than the main articulator – the apex. It can be observed that behind the apex the tongue is concave, which is confirmed by isolated trace of the consonant presented in Figure 13. It is apparent that the sensors placed on the central part of the tongue are below the line connecting the tip of the tongue and the back of it.¹¹ It is also evident that the longitudinal muscles of the tongue are contracted to a greater degree in the production of the

¹⁰ It ought to be remembered that the recorded people were not static during the recordings and their heads moved inside the articulograph cube. Therefore, the values reported here are not absolute, but result from normalizing computations that extracted the movements of the articulators from head movements.

¹¹ The acquired data show the great extent of anticipatory coarticulation. In the word *Leszek*, for instance,

sibilants than in the neighbouring vowels, which is visible when one compares the distance between the tongue sensors in the last phase of the articulation of the following vocoid. The tongue is much longer before making the velar closure.

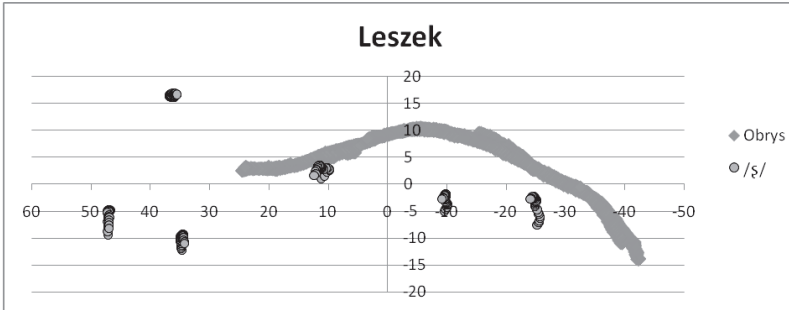


Figure 13: The segment /s/ extracted from the recording of the word *Leszek* (standard pronunciation)

Figure 14 shows the disordered articulation of [ʃ]. The trajectories of the sensors glued in the vicinity of the lips show that this sound, similarly to the normative articulation, is realized with lip protrusion and jaw raising, which also indicates lip rounding. The tip and the front of the tongue form an extensive narrowing with the alveolar ridge. The surface of the centre of the tongue approximates the post-alveolar area and the hard palate. The articulation is laminal with the mass of the tongue elevated behind the constriction. The shape of the surface of the tongue is convex – the two middle tongue sensors are above the line linking the other two sensors.

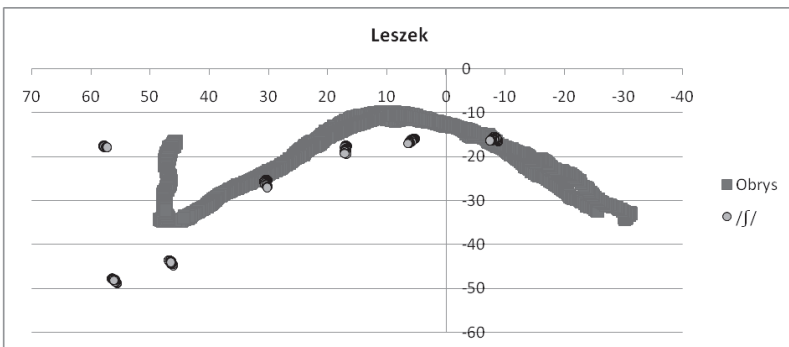


Figure 14: The segment /ʃ/ extracted from the recording of the word *Leszek* (non-standard pronunciation)

The figures below present one of the film frames, in which the position of articulators is seen during the realization of the phoneme /s/ by speakers with standard and with non-standard articulation (cf. Figures 15 and 16).

the back of the tongue displays a rising movement throughout the whole duration of the post-alveolar fricative so as to create the velar closure after realizing the vowel /ɛ/.



Figure 15: Middle part of the segment /s/ extracted from the animation of the word *Leszek* (standard pronunciation). Source: <http1>

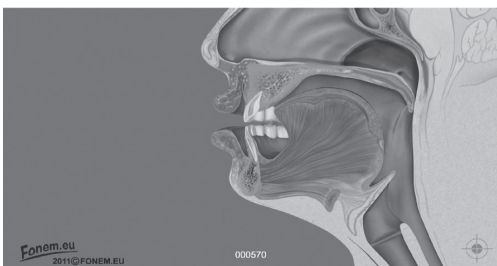


Figure 16: Middle part of the segment /ʃ/ extracted from the animation of the word *Leszek* (non-standard pronunciation). Source: <http1>

The investigation of articulation using the articulograph led to the finding that abnormally convex tongue settings are characteristic not only of the realization of the consonants [ʃ] [ʒ] [tʃ] [dʒ] (for more on this subject see the article Lorenc & Świąciński 2014). It turned out that this type of non-normative position of the tongue's mass is also found in the articulation of other consonants and vowels. The articulation in question can be observed in the lateral projections of the realization of the phonemes /l/ and /ɛ/ in the word *Leszek* (cf. Figures 17, 18, 19, 20) as well as in the selected animation frames (cf. Figures 21, 22, 23, 24).

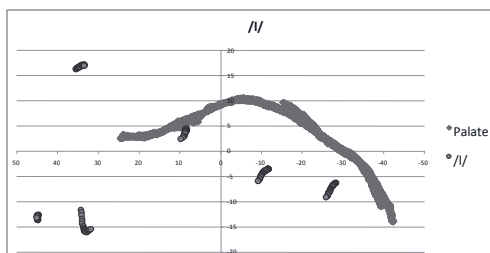


Figure 17: The segment /l/ extracted from the recording of the word *Leszek* (standard pronunciation). Source: own research

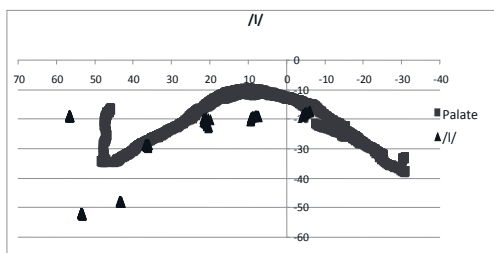


Figure 18: The segment /l/ extracted from the recording of the word *Leszek* (non-standard pronunciation). Source: own research

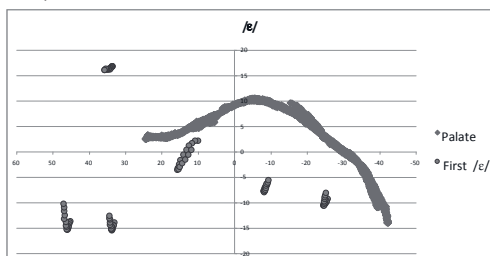


Figure 19: The segment /ɛ/ extracted from the recording of the word *Leszek* (standard pronunciation). Source: own research.

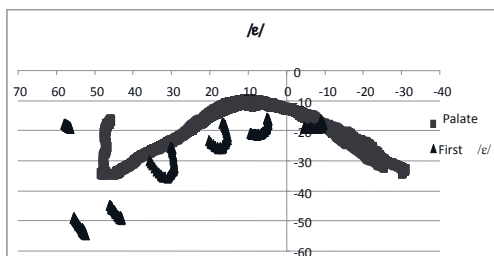


Figure 20: The segment /ɛ/ extracted from the recording of the word *Leszek* (non-standard pronunciation). Source: own research

All of the animations are available on the Polish phonetics portal (cf. <http1>). The project “Portal edukacyjny polskiej fonetyki stosowanej w zakresie normy i patologii mowy”

[Education portal of Polish applied phonetics in normal and pathological speech] was one of the winners of the 11th European Language Label (ELL) Competition, which is a European quality certificate in language education.

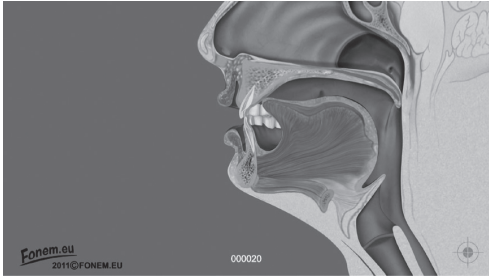


Figure 21: Middle part of the segment /l/ extracted from the animation of the word *Leszek* (standard pronunciation). Source: [http1](http://1)



Figure 22: Middle part of the segment /l/ extracted from the animation of the word *Leszek* (non-standard pronunciation). Source: [http1](http://1)

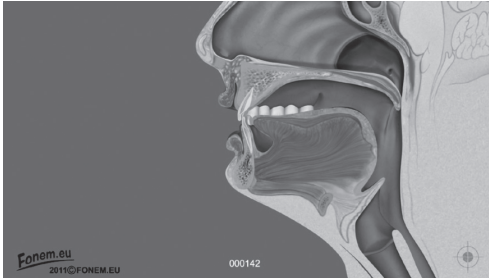


Figure 23: Middle part of the segment /e/ extracted from the animation of the word *Leszek* (standard pronunciation). Source: [http.1](http://1)



Figure 24: Middle part of the segment /e/ extracted from the animation of the word *Leszek* (non-standard pronunciation). Source: [http1](http://1)

5. Future research

The imperative goal of the new project „Współczesna wymowa polska. Badanie z wykorzystaniem trójwymiarowej artykulografii elektromagnetycznej” [Polish Language Pronunciation. Analysis Using Three-Dimensional Electromagnetic Articulatory] ¹² is to create the first objective, quantitative and qualitative description of the standard pronunciation of Polish sounds. This goal may be achieved through creating a state-of-the-art articulation laboratory, bringing together a team of researchers specializing in this type of studies. The proposed experiment will focus on gaining knowledge and data related to the characteristics of standard Polish pronunciation, creating a detailed description of Polish speech dynamics, the coordination of articulators, their positioning in respective articulatory stages as well as presenting research findings in accessible graphic forms.

¹² The project nr 2012/05/E/HS2/03770 managed and coordinated by A. Lorenc is financed by the Polish National Science Center.

The study will consist in recording the pronunciation of a selected group of speakers with the use of the articulograph. The participants in the experiment will be 10 females and 10 males displaying standard pronunciation of Polish who will have been subjected to speech and language diagnostic tests aiming at excluding speakers with impaired hearing, anatomical abnormalities of the speech apparatus or impaired motor skills of the speech organs.

The experiment will be carried out with the use of AG500 articulograph. Audio and video signals will be recorded simultaneously with the positional coordinates of the articulograph sensors placed on the tongue, lips and several additional locations. The audio signals are going to be registered with a broadband 16-channel microphone array (cf. Król & Lorenc [forthcoming])¹³ in order to enhance the quality of sound data. The device will allow for a more precise synchronization of the articulographic, audio and video signals. Camcorders will register the speakers' face images. Due to white markers specifically positioned on the speakers' faces, it will be possible to obtain exact motion estimations of outer articulators, such as the jaw and lips. Three high frame-rate camcorders will be used for the registration of the face front as well as the left and right profile. Audio and video recordings as well as data from articulograph sensors will be processed in order to obtain selected parameters in the temporal and spectral domain. The parameters will be used for dynamic Bayes network (DBN) training. DBNs will represent stochastic models of Polish utterances examined in the project. Evaluation of the precision of the tongue motion estimation by DBNs, where reference motion will be obtained from the articulograph, is also planned. All data concerning the speakers, articulograph signals, sounds, camcorder images, signal parameters and articulation models in the form of DBNs will be stored in a dedicated relational database.

References

- Akdemir, Eren & Çiloğlu, Tolga. 2008. The use of articulator motion information in automatic speech segmentation. *Speech Communication* 50(7). 594–604.
- Badin, Pierre & Serrurier, Antoine. 2006. Three-dimensional modeling of speech organs: Articulatory data and models. *IEICE Technical Report* Vol. 106, No 177, SP2006-26. Kanazawa, Japan, 20-21 July 2006. The Institute of Electronics, Information and Communication Engineers. 29–234.
- Benni, Tytus 1931. *Palatogramy polskie*. Kraków: Wydawnictwo Towarzystwa Miłośników Języka Polskiego.
- Byrd, Dani & Lee, Sungbok & Riggs, Daylen & Adams, Jason. 2005. Interacting effects of syllable and phrase position on consonant articulation. *Journal of the Acoustical Society of America* 118(6). 3860–3873.
- Code, Christopher & Ball, Martin. 1984. *Experimental Clinical Phonetics. Investigatory Techniques in Speech Pathology and Therapeutics*. Worcester: Billing & Sons Limited.
- Dłuska, Maria. 1950. *Fonetyka polska. Artykulacja głosek polskich*. Kraków: Wydawnictwo Studium Słowińskiego.
- Engwall, Olov. 2003. Combining MRI, EMA and EPG measurements in a three-dimensional tongue model. *Speech Communication* 41. 303–329.
- Gibert, Guillaume & Attina, Virginie & Tiede, Mark & Bundgaard-Nielsen, Rikke & Kroos, Christian & Kasisopa, Benjawan & Vatikiotis-Bateson, Eric & Best, Catherine T. 2012. Multimodal Speech Animation from Electromagnetic Articulography Data. In *20th European Signal Processing Conference*, Bucharest, Romania, 2807–2811.

¹³ In previous articulographic research (e.g. Akdemir & Çiloğlu 2008, Byrd et al. 2005) standard microphones were used. Such an approach has a major drawback in the form of variable time delays between the recorded audio and articulographic signals. These synchronization problems appear due to the unavoidable head movements of the recorded person.

- Gick, Bryan. 2002. The use of ultrasound for linguistic phonetic fieldwork. *Journal of the International Phonetic Association* 32. 113–122.
- Hamann, Silke. 2002. Postalveolar fricatives in Slavic languages as retroflexes. In Baauw, S. & Huiskes, M. & Schoorlemmer, M. (eds.), *OTS Yearbook 2002*, 105–127. Utrecht: Utrecht Institute of Linguistics.
- Hamann, Silke. 2003. *The Phonetics and Phonology of Retroflexes*. Utrecht: LOT Press.
- Hamann, Silke. 2004. Retroflex fricatives in Slavic Languages. *Journal of the International Phonetic Association* 34(1). 53–67.
- Hardcastle, William J. 1972. The use of electropalatography in phonetic research. *Phonetica* 25. 197–215.
- Hardcastle, William & Gibbon, Fiona & Jones, Wilf. 1991. Visual display of tongue-palate contact: electropalatography in the assessment and remediation of speech disorders. *British Journal of Disorders of Communication* 26. 41–74.
- IPA (International Phonetic Association). 1999. *Handbook of the International Phonetic Association. A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Jassem, Wiktor. 2003. Illustration of the IPA: Polish. *Journal of the international Phonetic Association* 33(1). 103–107.
- Kiritani, Shigeru. 1986. X-Ray microbeam method for measurement of articulatory dynamics – techniques and results. *Speech Communication* 5. 119–140.
- Koneczna, Halina & Zawadowski, Witold. 1951. *Przekroje rentgenograficzne głosek polskich*. Warszawa: PWN.
- Krajna, Ewa. 2008. *100-wyrazowy test artykulacyjny*. Gliwice: Wydawnictwo Komlogo.
- Król, Daniel & Lorenc, Anita (forthcoming). Rozkład pola akustycznego w badaniach artykulacji nosowych i bocznych. *Prace Filologiczne*.
- Laver, John. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Lorenc, Anita (forthcoming). Dynamic visualisations in studies on contemporary Polish pronunciation. *Odpowiedzialność w edukacji niesłyszących*. Lublin: Wydawnictwo KUL.
- Lorenc, Anita & Święciński, R. 2014. Application of phonetics in speech therapy: A case of abnormal convex tongue setting in Polish. In Szpyra-Kozłowska, J. & Guz, E. & Steinbrich, P. & Święciński, R. (eds.). *Recent Developments in Applied Phonetics. Studies in Linguistics and Methodology* 6, 287–324. Lublin: Wydawnictwo KUL.
- Pompino-Marschall, Bernd & Żygis, Marzena. 2003. Surface palatalization of Polish bilabial stops: Articulation and acoustics. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona 3–9 August, 1751–1754.
- Revêret, Lionel & Bailly, Gérard & Badin, Pierre 2000. MOTHER: a new generation of talking heads providing a flexible articulatory control for video-realistic speech animation. In *6th International Conference of Spoken Language Processing*, Beijing, China, 755–758.
- Rochoń, Marzena & Pompino-Marschall, Bernd 1999. The articulation of secondarily palatalized coronals in Polish. In *Proceedings of XIVth International Congress of Phonetic Sciences*, San Francisco, 1897–1900.
- Stone, M. & Stock, G. & Bunin, K. & Kumar, K. & Epstein, M. A. & Kambhamettu, Chandra & Li, M. & Parthasarathy, Vijay & Prince, Jerry. 2007. Comparison of speech production in upright and supine position. *The Journal of the Acoustical Society of America* 122(1). 532–541.
- Święciński, Radosław. 2013. An EMA study of articulatory settings in Polish speakers of English. In Waniek-Klimczak, E. & Shockey, L.R. (eds.), *Teaching and researching English accents in native and non-native speakers*, 73–82. Heidelberg: Springer Publications.
- Trochymiuk, Anita, & Święciński, Radosław. 2009. Artykulograficzne badanie wymowy grzbietowej. Studium przypadku. *Logopedia* 38. 173–201.
- Wierzchowska, B. 1967. *Opis fonetyczny języka polskiego*. Warszawa: PWN.
- Wierzchowska, B. 1980. *Fonetyka i fonologia języka polskiego*. Wrocław: Zakład Narodowy im. Ossolińskich & Wydawnictwo PAN.
- Żygis, Marzena. 2004. Dlaczego polskie sybilanty š i ž są retrofleksami? *Logopedia* 33. 119–132.

Internet sources

<http://www.fonem.eu>