The EMA Study on the Inter-individual Variability and Differences in Articulation between Polish Oral and Nasalised Vowels

Robert Wielgat[a,*], Anita Lorenc[b,c]

[a] *State Higher Vocational School in Tarnów, Mickiewicza 8, 33-100 Tarnów, Poland*
[b] *Maria Curie-Skłodowska University, Department of Speech Therapy and Applied Linguistics, Sowińskiego 17, 20-040 Lublin, Poland*
[c] *Warsaw University, Institute of Applied Polish Studies, Department of Speech and Language Therapy and Voice Production, Krakowskie Przedmieście 26/28, 00-927 Warszawa, Poland*

Article history:
Received 1 December 2017
Received in revised form
22 December 2017
Accepted 23 December 2017
Available online 27 December 2017

Abstract

The electromagnetic articulography (EMA) is a relatively exact and efficient method used in study on speech production physiology. It allows to precisely estimate movement trajectories of speech articulators like tongue, lips, jaw etc. by tracking position of sensors fixed to the articulators. This paper presents results of EMA research on Polish oral and nasalised vowels in orthography represented by the graphemes <e>, <o>, <ę>, <ą>. Inter-individual variability of tongue and lips position in X-axis direction during realization of the same phoneme has been estimated. Differences between oral and nasalised vowels in terms of movement of articulators in X-axis direction have been assessed too.

Key words: electromagnetic articulography, inter-individual variability, articulatory phonetics, oral vowels, nasalised vowels

Introduction

Electormagnetic articulography (EMA) is a relatively new research technique that emerged in the phonetic instrumental research at the end of the twentieth century [1, 2]. Due to applied technology, this method allows to record, store, visualise and assess articulator movements (tongue, lips, mandible, soft palate) in 3D space in real time when pronouncing various speech sounds including vowels. Vowel articulation for different world languages are broadly described in the phonetic literature. In the EMA research of vowels the most often assessed are tongue position during articulation [3], relationships between movements of tongue and lips [4] or tongue and mandible [5, 6], quantity, amplitude and velocity of speech articulator movements [7], dependency of articulation with context [3] as well as coordination between articulation gestures and vowel fundamental frequency [6, 8]. Polish vowels have not been described using EMA so far.

In this paper an analysis of Polish oral vowels [ɛ] and [ɔ] and their nasalised counterparts has been provided. According to phonetic articulatory classification [9, 10] vowel [ɛ] is described as front (in terms of tongue movements in the horizontal plane), open-mid (in terms of lips opening size), unrounded (in terms of shape of lips) and oral (in terms of position of soft palate). Vowel [ɔ] in turn according to the same classification is described as back (in terms of tongue movements in the horizontal plane), open-mid (in terms of lips opening size), rounded (in terms of shape of lips) and oral (in terms of position of soft palate).

Then the most important differences between the mentioned vowels are:

- the movements of the tongue in the horizontal plane dividing them into the front and back vowels,
- the shape of the lips – as a consequence the unrounded and rounded vowels are distinguished,
- the soft palate and the associated division into oral and nasalised vowels.

Polish nasalised vowel represented by the grapheme <ę> is associated with an open-mid front unrounded vowel while the vowel <ą> represents an open-mid back rounded vowel. Their realisation in word-internal position is phonotactically possible only before non-palatalized fricative consonants and usually is transcribed [11] using two phonetic symbols (the proper oral open-mid vowel [ɛ] or [ɔ] followed by nasalised labio-velar approximant [w̃]):

a) *męski* [mɛw̃sci] ‘male’/’manly’, *węże* [vɛw̃zɛ̡] ‘snakes’, *węch* [vɛw̃x] ‘smell’,
b) *wąsy* [vɔws̃ɨ] ‘moustache’, *mąż* [mɔw̃ʂ] ‘husband’, *wąchać* [vɔw̃xatɕ] ‘to smell’.

This article supplements the results of the articulation analysis of Polish oral vowels and their nasalised counterparts presented in monograph of A. Lorenc [12]. In the follow up analyses of the

*Corresponding author: rwielgat@poczta.onet.pl

aforementioned speech sounds, it was decided to examine the relationship between nasal or oral articulation and movements of the tongue or lower lip, and to present them in relation to the inter-individual variability of speech organs and differences between Polish oral and nasalised vowels. Thus, nasalised vowels [ɛw̃], [ɔw̃] and the corresponding oral vowels [ɛ] and [ɔ] are compared, assuming as the primary critical articulator the front part of the tongue in the case of front vowels and its back part in the case of back vowels. Secondary articulation, the movement of the lower lip, was also taken into account. It was analysed in the same time relationship as the primary articulation.

Materials and Methods

The research involved 20 adult Polish speakers (10 females and 10 males) who, in the opinion of the team of experts (phonetics and speech therapists), represented careful style of the standard variety of contemporary Polish and met the criteria of linguistic and biological norms [12]. Because of insufficient number of recorded vowels for some of the speakers (1 or 2 examples), statistical testing of hypotheses was possible for only 8 female and 7 male recordings, which were selected for further analysis.

The pronunciation of selected Polish vowels was analysed in two-syllable words in the stressed position in the preceding context of the bilabial consonant [p] and the following context of the fricative dental consonant [s]. Three examples of each examined vowel were provided. Nasalised vowels were evaluated with three repetitions of the same word, as no other examples were found to fulfill the conditions of the assumed phonetic structure (in some cases, additional words were also analysed from the complementary list – for example - words *węzeł* and *wąsy* – as vowels of interest appeared in the similar contexts). Nasalised vowels were compared to the corresponding oral vowels – [ɛ] or [ɔ]. Table 1 presents the words that are included in the list for the evaluation of proper vowel variants, which are considered in the further part of the analysis. The letters representing the vowels are in bold.

Table 1. The word list (in the orthographic transcription) used to assess oral and nasalised Polish vowels

Vowel	Word		
[ɛ]	*pesel*	*pestka*	*pesto*
[ɛw̃]	*Pęset*	*pęset*	*pęset*
[ɔ]	*Posąg*	*posag*	*postrzał*
[ɔw̃]	*Pąsy*	*pąsy*	*pąsy*

Articulation research was performed by synchronously registering several types of signals: articulographic, video and audio [2, 13, 14]. For this purpose the Carstens articulograph (AG500), a vision system consisting of three Point Gray high speed cameras (Gazelle GZL-CL-22C5M-C), and an audio recorder with 16-channel microphone array [15, 16] were used. The construction of the measuring system and its dependencies are shown in Figure 1, while its operation is described in detail in [17].

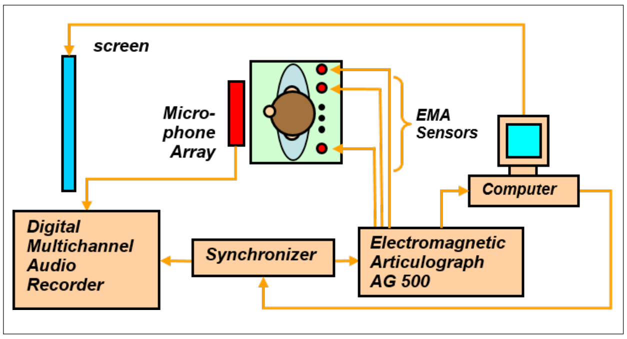

Figure 1. Block diagram of the measurement system

The recorded word material was subjected to acoustic segmentation to determine the boundaries of the analysed vowels. It was developed manually using the Praat program for acoustic signal analysis [18], using the principles described in detail in [12, 19]. The analysis of articulation gestures was then carried out, using the phoneEMAtool software developed specifically for the needs of the experiment [20]. The toolkit allows to dynamically visualize motion trajectories of all sensors (except the reference) in X-axis (front-back) and Z (top-down) and for the analysis and extraction of information related to the location of individual sensors in time across all axes, taking into account the angular tilt φ and θ. Moreover, the software allows to calculate the motion speed of sensors in time and to determine its minima and maxima. In addition, it is possible to set a 20% level of increasing and decreasing velocities. Adding this criterion helps to avoid typical problems in determining the beginnings and ends of movements. For example, it is stated that fixed-length vowels may not have an unequivocal point on the zero-velocity axis that separates the CV (consonant-vowel) and VC (vowel-consonant) movements (see [5, 7]). An extended amount of almost zero speed makes it difficult to precisely determine the time of the beginning of the articulation gesture on this type of continuum. Adopting the 20 percent velocity threshold criterion helps to establish such a segmentation point on a relatively steep, hence well-defined in time, velocity curve. A comparison of different velocity thresholds [25] found that using 20% value, unlike other evaluations, provided the best and most stable results. The arguments in question have also been used to determine the velocity of the sensors, with the possibility of calculating its minimum and maximum values, as well as the 20% acceleration and deceleration levels on the rising and falling slopes. In order to eliminate unwanted distortions, the Savitzky-Golay filter was used. For more precise smoothing, the algorithm was applied twice.

PhoneEMAtool [20] allows simultaneous and synchronous processing of three types of data: (1) audio (* .wav), (2) EMA

(* .txt) and (3) acoustic segmentation (* .TextGrid). Figure 2 shows the *phoneEMAtool* working window during exemplary data analysis.

g) GOFF (gesture offset) – end of articulation gesture (final 20% of PVEL2 threshold on the descending edge).

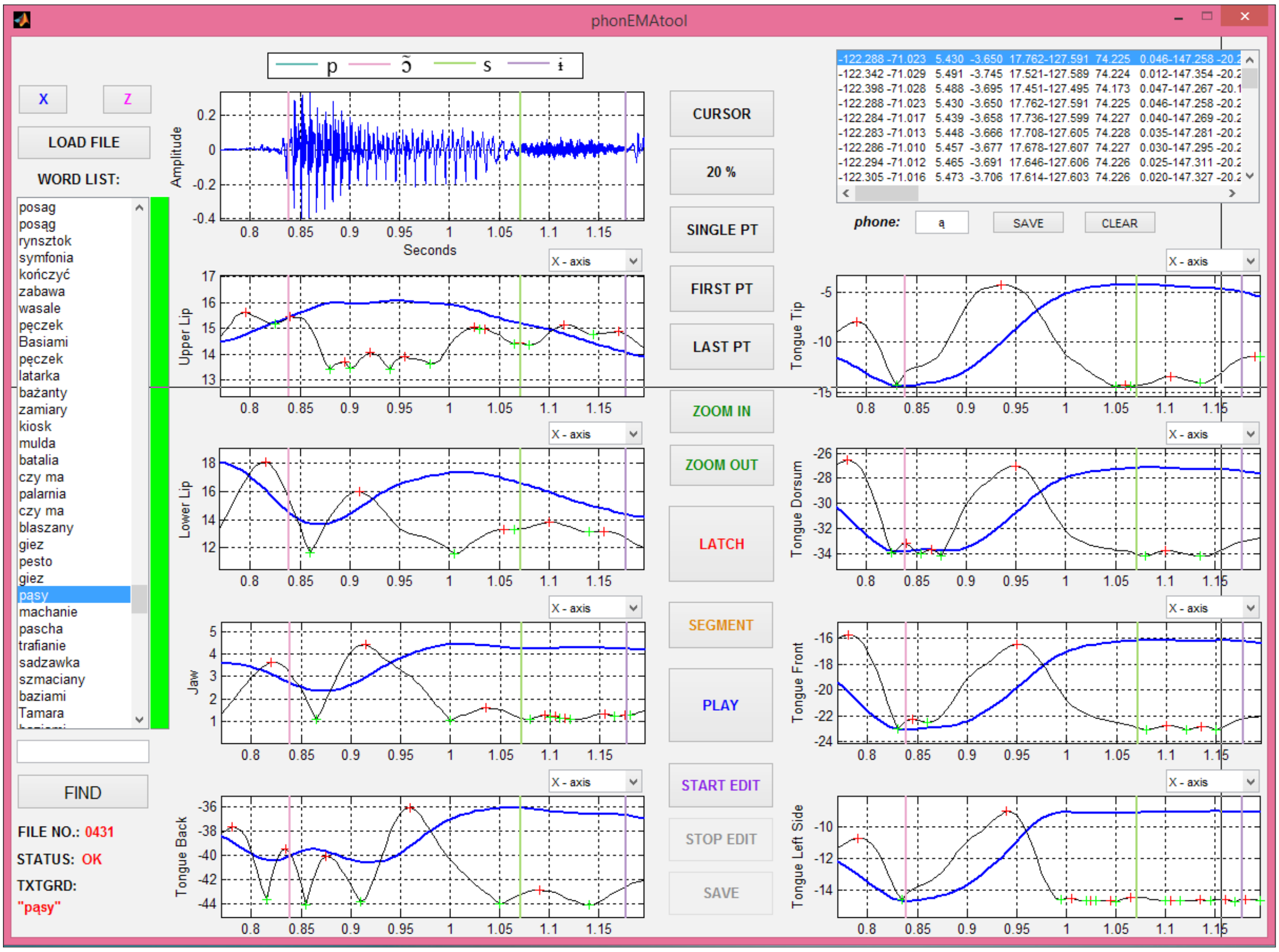

Figure 2. The *phoneEMAtool* working window during the analysis of articulation gestures along the X axis of the Polish word ‘pąsy’ (speaker PT_m, file 431)

Following the approach ofa Best et al. [23] an analysis of vowel articulatory gestures and their landmarks was carried out as illustrated in Figure 3.

The symbols shown at the bottom of the Figure 3 are used to indicate the following gesture boundaries:

a) GONS (gesture onset) – the beginning of the articulation gesture (the initial 20% of PVEL1 threshold on the rising edge);

b) PVEL1 (peak velocity) – the peak of velocity at the beginning of the gesture;

c) NONS (nucleus onset) – the beginning of the articulation gesture core (final 20% of PVEL1 threshold on the descending edge: the beginning of the constriction);

d) NOFF (nucleus offset) – the end of the articulation gesture core (initial 20% of PVEL2 threshold on the rising edge: end of the constriction);

e) MinVEL (minimum velocity) – minimum velocity within the core of the articulation gesture (between NONS and NOFF);

f) PVEL2 (peak velocity) – peak velocity at the end of the gesture;

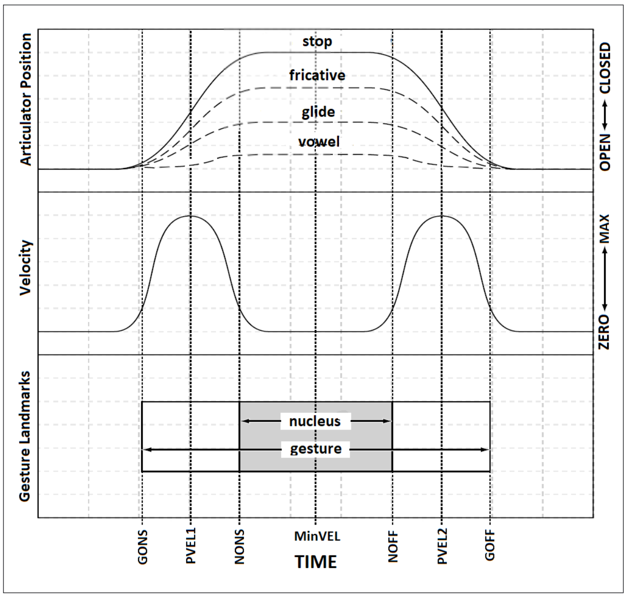

Figure 3. Constriction gesture landmarks in articulator movement path as identified from the velocity profile of the EMA sensor [23] using *phoneEMAtool* program [20].

Apart from defining the components of a dynamic articulation gesture analysis, the assessment also requires determining the relevant elements of their complex structure and indicating the articulator whose movement plays a central role in creating the articulation. In the publications addressing this problem, it is often referred to as the so-called 'critical articulator or 'crucial articulator' (see [24, 25]). Although the number of key articulators for each phoneme may be greater than one, usually one of them is considered primary. Relevant time points for other secondary critical articulators may be determined by the time position of the primary articulator.

In the analysis of Polish vowels, it was decided to check the relationship between nasalised or oral articulation and the movements of the tongue or lower lip. Consequently, back vowels, oral [ɔ] and nasalised [ɔw̃] were compared in the analysis, assuming back part of the tongue as critical articulator. As they are rounded vowels, secondary articulation (lower lip movement) is also taken into account – by analysing it in the same time relationship as the primary articulation. Both gestures were evaluated by observing changes of the tongue back (TB) and lower lip (LL) sensor position in the horizontal axis (X), binding them with reversing or foregoing the tongue mass and rounding and flattening the lips.

Similarly, the non-rounded front vowels, [ɛ] and [ɛw̃] were compared, except that the primary articulator was considered the front part of the tongue, and the second articulation was associated with the lower lip movement, as in the case of the back vowel. Measurements of the primary articulation gesture relied on tracking the position of the tongue front (TF) sensor in the horizontal axis (X). In the same temporal relation, the secondary articulation gesture of the lower lip (LL) was also evaluated, also across the X axis. The initial measurement point was set at the place where the first minimal velocity (MinVEL1) was received corresponding to the reception of the first extreme position by the sensor of the frontal (TF) or back (TB) part of the tongue. The articulation gestures of the mentioned parts of the tongue were still present during the segment of the short before the consonant [p] or just after its explosion. The last measurement point was determined at the end of the articulation gesture (GOFF), where the critical articulator sensor reached a 20% velocity threshold on the decreasing slope. The end of the articulation gesture usually appears in the final part of the evaluated vowel. The scheme of measuring the trajectory of the movement of the back of the tongue during the implementation of the rear vowels [ɔ] and [ɔw̃] is illustrated in Figure 4 (for the example [ɔw̃]).

The scheme of trajectory measurement of the frontal part of the tongue during the articulation of front vowels [ɛ] and [ɛw̃] is illustrated in Figure 5 (for [ɛw̃] example).

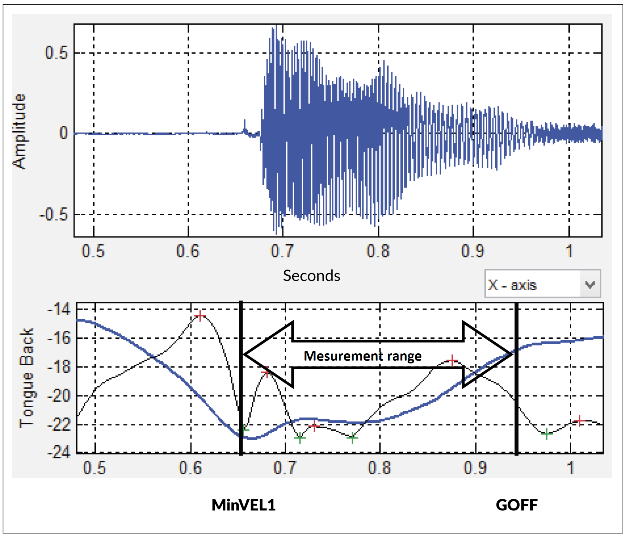

Figure 4. Oscillogram and trajectory and velocity of movement of the back part of the tongue (TB) in the X axis (front-back) during nasalised vowel articulation [ɔw̃] of the Polish word 'pąsy' (speaker MK_f, file 234)

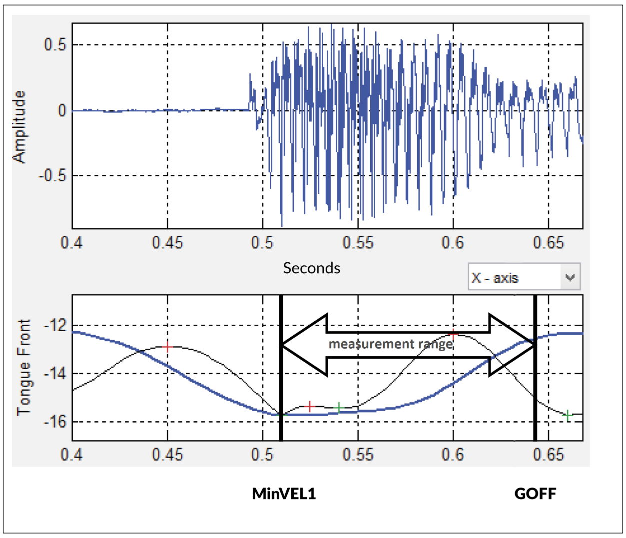

Figure 5. Oscillogram and trajectory and velocity of movement of the front part of the tongue (TF) in the X axis (front-back) during nasalised vowel articulation [ɛw̃] of the Polish word 'pęset' (speaker RD_m, file 212)

Results and Discussion

As mentioned in the introduction , there is an intrinsic variation in the location of selected sensors placed on the tongue when pronouncing nasalised vowels [ɛw̃], [ɔw̃] and corresponding oral vowels [ɛ], [ɔ]. The sensor positions were measured starting with the first minimal velocity MinVel, achieved when the tongue is in the correct vowel position, until the end of a typical articulation gesture (GOFF). The maximum and minimum swings of the sensors analysed in the X-axis (front-to-back direction) were also taken into account. Based on the these sensor positions for each uttered vowel , the values of two parameters were obtained:

- difference between values of EMA sensor X coordinate in points MinVEL1 and GOFF for movements of lower lip (LL); tongue tip (TF) – in the case of vowels [ɛw̃] and [ɛ];

tongue back (TB) – in the case of vowels [ɔw̃] and [ɔ]:
D = X_GOFF – X_MinVEL1
Where:
X_GOFF – X coordinate of the sensor placed on LL, TF and TB respectively in point GOFF,
X_MinVEL1 – X coordinate of the sensor placed on LL, TF and TB respectively in point MinVEL1,

- Amplitude of movements in X-axis being the difference between maximal and minimal values of X coordinates achieved by sensors placed on lower lip (LL); tongue tip (TF) – in the case of vowels [εw̃] and [ε]; tongue back (TB) – in the case of vowels [ɔw̃] and [ɔ]:
Amp = max_X – min_X
Where:
max_X – maximal value of X coordinate achieved by a sensor placed on the LL, TF and TB respectively
min_X – minimal value of X coordinate achieved by a sensor placed on the LL, TF and TB respectively

Variation in values of these two parameters were assessed in terms of inter-individual variability as well as differences between vowels.

Inter-individual Variability

Inter-individual variability was first evaluated in terms of differences between male and female speakers. In order to assess those variability ANOVA and MANOVA tests were carried out for D and Amp parameters. Values of test probabilities p of the tests are summarized in Table 2.

Whenever possible, the parametric ANOVA variance test was used, in which case test probabilities are indicated in Table 2 with a normal font. In case of failure to fulfil the assumptions of parametric test, the nonparametric Kruskal-Wallis test was used to perform the ANOVA analysis and the test probabilities are marked in bold. Adopted level of statistical significance was set to α = 0.05. If the test showed a significant difference in the assumed level of statistical significance, the probability value is indicated in red. Wherever the conditions allowed, a multivariate analysis of variance (MANOVA) was performed taking into account the parameters D and Amp for TF (TB) and LL simultaneously. Mean and median values are marked in blue.

As can be seen from the Table 2, statistically significant differences between men and women occurred at the level α = 0.05 only for oral back vowel [ɔ] in the case of D parameter calculated for the lower lip (LL). In the rest of cases differences were not statistically significant at the level α = 0.05.

The second analysis of the inter-individual variability was carried out on the group of men and women taken together. Results of the analysis are presented in Table 3.

Table 3. Number of pairs with statistically significant differences in mean values

Vowel	D			Amp		
	LLx	TFx	TBx	LLx	TFx	TBx
[ε]	10	**0** (0)	-	**1** (9)	**0** (0)	-
[εw̃]	**0** (12)	31	-	**3** (19)	30	-
[ɔ]	**1** (11)	-	10	12	-	10
[ɔw̃]	**4** (20)	-	23	**3** (15)	-	27

Results in the Table 3 reports the numbers of pairs of speakers for which statistically significant difference occurred. These pairs were determined by the HSD Tukey's test (numbers marked by a normal font in the Table 3) and by Kruskal-Wallis test when initial assumptions for the Tukey's test were not met (numbers marked in bold). Numbers in brackets indicate results of Tukey's test in the case when the initial assumptions for Tukey's test are not fulfilled. However violation of the mentioned assumptions was in most cases not significant. More information is given in the section 3.3. Numbers of sample vowels coming from particular speakers were rather low and not even. Usually there were between 3 and 5 examples of vowels per speaker. Speakers with 2 or 1 samples per speaker were not included into analyses.

Table 2. Values of test probabilities for ANOVA and MANOVA statistical test for detecting statistically significant difference between groups of male and female speakers

Vowel	D			Amp			mean	median
	ANOVA for LL	ANOVA for TF (TB)	MANOVA for LL vs. TF (TB)	ANOVA for LL	ANOVA for TF (TB)	MANOVA for LL vs. TF (TB)		
[ε]	**0,0514**	0,801		0,883	0,912	0,783	0,686	0,801
[εw̃]	0,992	0,7	0,78	0,454	**0,796**		0,744	0,780
[ɔ]	**0,0492**	(0,53)		**0,856**	(0,657)		0,523	0,594
[ɔw̃]	**0,158**	**(0,128)**		**0,176**	**(0,139)**		0,150	0,149
Mean	0,313	0,540	0,780	0,592	0,626	0,783		
Median	0,105	0,615	0,780	0,655	0,727	0,783		

Results from the Table 3 show that opposite to the variability caused by sex, the variability caused by individual features of a given speaker is much higher. Only in 3 cases out of total 16 analysed, tests did not show any variability between pairs of speakers.

In the analysis of inter-individual variability among speakers the MANOVA test was not carried out because the assumptions for the test were not met. Therefore for every speaker values of each parameter (D or Amp) for lower lip and tongue front (back) were visualised on diagram plane for better assessment of inter-individual variability. An exemplary diagram for the case with strongest inter-individual variability is shown in Figure 6.

The situation presented in Figure 6 is consistent with results of HSD Tukey's test presented in Figure 7.

Differences between oral and nasalised vowels

Differences between oral and nasalised vowels in terms of mean and median values of D and Amp parameters have been shown in Figure 8 and Figure 9. Mentioned values were taken from our former research described in [26]. It is evident that mean and median values of the movement amplitude and D parameter for the lower lip in nasalised vowel group are significantly higher than the same values for oral vowel group. This dependency is true in each case when groups of men and women are taken separately.

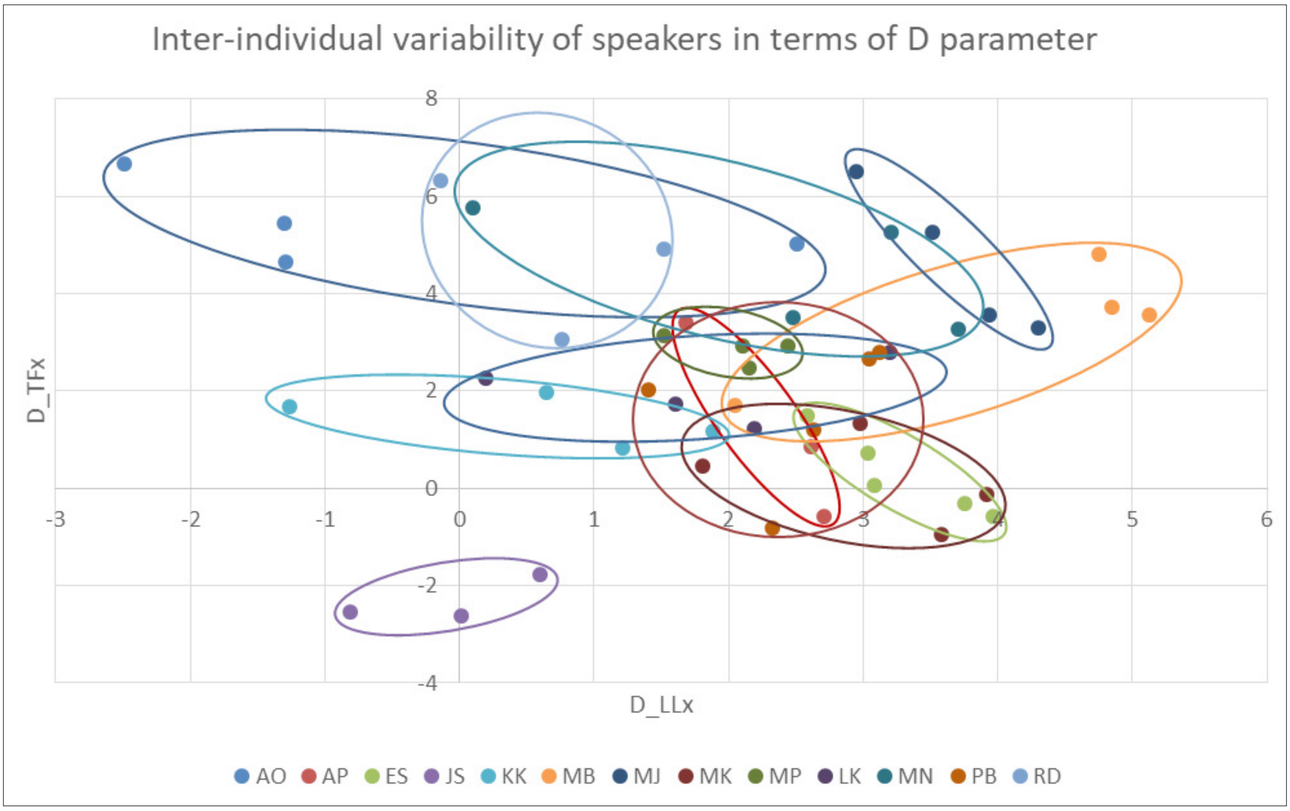

Figure 6. Diagram of inter-individual variability in values of D parameter for tongue front (D_TFx) vs. lower lip (D_LLx) for Polish nasalised vowel [ɛw̃]

	HSD (nierówne N); zmn.: D_TFx (ew_(M)ANOVA_D) Zaznaczone róznice sa istotne z p < ,05000												
mówca	{1} M=5,4450	{2} M=1,2200	{3} M=,27000	{4} M=-2,320	{5} M=1,4175	{6} M=3,4475	{7} M=4,6550	{8} M=,17750	{9} M=2,8675	{10} M=2,0050	{11} M=4,4475	{12} M=1,5720	{13} M=4,7633
AO_f {1}		0,003285	0,000137	0,000131	0,000920	0,413992	0,998452	0,000135	0,107545	0,007063	0,987611	0,001565	0,999917
AP_f {2}	0,003285		0,997729	0,024751	1,000000	0,467432	0,033093	0,994703	0,850657	0,999646	0,057548	1,000000	0,024521
ES_f {3}	0,000137	0,997729		0,249728	0,963241	0,017178	0,000317	1,000000	0,101775	0,626515	0,000571	0,833382	0,001466
JS_f {4}	0,000131	0,024751	0,249728		0,014101	0,000145	0,000131	0,298001	0,000250	0,002431	0,000131	0,008974	0,000131
KK_f {5}	0,000920	1,000000	0,963241	0,014101		0,389613	0,014067	0,936811	0,836856	0,999921	0,027766	1,000000	0,042125
MB_f {6}	0,413992	0,467432	0,017178	0,000145	0,389613		0,947253	0,012616	0,999931	0,841384	0,987354	0,510641	0,965061
MJ_f {7}	0,998452	0,033093	0,000317	0,000131	0,014067	0,947253		0,000258	0,583173	0,087856	1,000000	0,023406	1,000000
MK_f {8}	0,000135	0,994703	1,000000	0,298001	0,936811	0,012616	0,000258		0,078376	0,550076	0,000433	0,868764	0,001108
MP_f {9}	0,107545	0,850657	0,101775	0,000250	0,836856	0,999931	0,583173	0,078376		0,996486	0,748712	0,915921	0,701529
LK_m {10}	0,007063	0,999646	0,626515	0,002431	0,999921	0,841384	0,087856	0,550076	0,996486		0,154060	0,999997	0,176468
MN_m {11}	0,987611	0,057548	0,000571	0,000131	0,027766	0,987354	1,000000	0,000433	0,748712	0,154060		0,045091	1,000000
PB_m {12}	0,001565	1,000000	0,833382	0,008974	1,000000	0,510641	0,023406	0,868764	0,915921	0,999997	0,045091		0,063206
RD_m {13}	0,999917	0,024521	0,001466	0,000131	0,042125	0,965061	1,000000	0,001108	0,701529	0,176468	1,000000	0,063206	

mówca	HSD (nierówne N); zmn.: D_LLx (ew_(M)ANOVA_D) Zaznaczone róznice sa istotne z p < ,05000												
	{1} M=-,6425	{2} M=2,3333	{3} M=3,2800	{4} M=-,0667	{5} M=,62000	{6} M=4,1950	{7} M=3,6750	{8} M=3,0725	{9} M=2,0525	{10} M=1,8000	{11} M=2,3750	{12} M=2,5020	{13} M=,71333
AO_f {1}		0,102374	0,001190	0,999985	0,925461	0,000154	0,000347	0,002467	0,073111	0,147278	0,026988	0,017799	0,954421
AP_f {2}	0,102374		0,997654	0,344354	0,808798	0,715673	0,957691	0,999793	1,000000	0,999994	1,000000	1,000000	0,858599
ES_f {3}	0,001190	0,997654		0,039443	0,080932	0,993669	0,999999	1,000000	0,938103	0,811303	0,994246	0,995961	0,252277
JS_f {4}	0,999985	0,344354	0,039443		0,999903	0,002686	0,012903	0,068217	0,532986	0,712300	0,319700	0,251284	0,999644
KK_f {5}	0,925461	0,808798	0,080932	0,999903		0,004039	0,023896	0,143474	0,841303	0,952806	0,599791	0,494538	1,000000
MB_f {6}	0,000154	0,715673	0,993669	0,002686	0,004039		0,999977	0,967061	0,301346	0,166515	0,545639	0,651054	0,027188
MJ_f {7}	0,000347	0,957691	0,999999	0,012903	0,023896	0,999977		0,999889	0,707669	0,500252	0,910159	0,954744	0,105925
MK_f {8}	0,002467	0,999793	1,000000	0,068217	0,143474	0,967061	0,999889		0,984271	0,921562	0,999509	0,999938	0,369510
MP_f {9}	0,073111	1,000000	0,938103	0,532986	0,841303	0,301346	0,707669	0,984271		1,000000	1,000000	0,999995	0,958250
LK_m {10}	0,147278	0,999994	0,811303	0,712300	0,952806	0,166515	0,500252	0,921562	1,000000		0,999932	0,999477	0,991930
MN_m {11}	0,026988	1,000000	0,994246	0,319700	0,599791	0,545639	0,910159	0,999509	1,000000	0,999932		1,000000	0,837326
PB_m {12}	0,017799	1,000000	0,995961	0,251284	0,494538	0,651054	0,954744	0,999938	0,999995	0,999477	1,000000		0,763356
RD_m {13}	0,954421	0,858599	0,252277	0,999644	1,000000	0,027188	0,105925	0,369510	0,958250	0,991930	0,837326	0,763356	

Figure 7. Results of two HSD Tukey's tests checking statistically significant differences between speakers in values of D parameter for lower lip (upper table) and tongue front (lower table) for Polish nasalised vowel [ɛw̃]. Results obtained by STATISTICA software

One can also notice another relation of inter-individual variability. Namely mean and median values of the movement amplitude (Amp) and D parameter for the tongue front (in the case of front vowels [ɛ] and [ɛw̃]) and for the tongue back (in the case of back vowels [ɔ] and [ɔw̃]) are usually higher for women than for men. These relations, however, in some cases like mean and median values of D parameter for the oral vowel [ɛ] and mean values of amplitude for the oral vowel [ɔ] are not true.

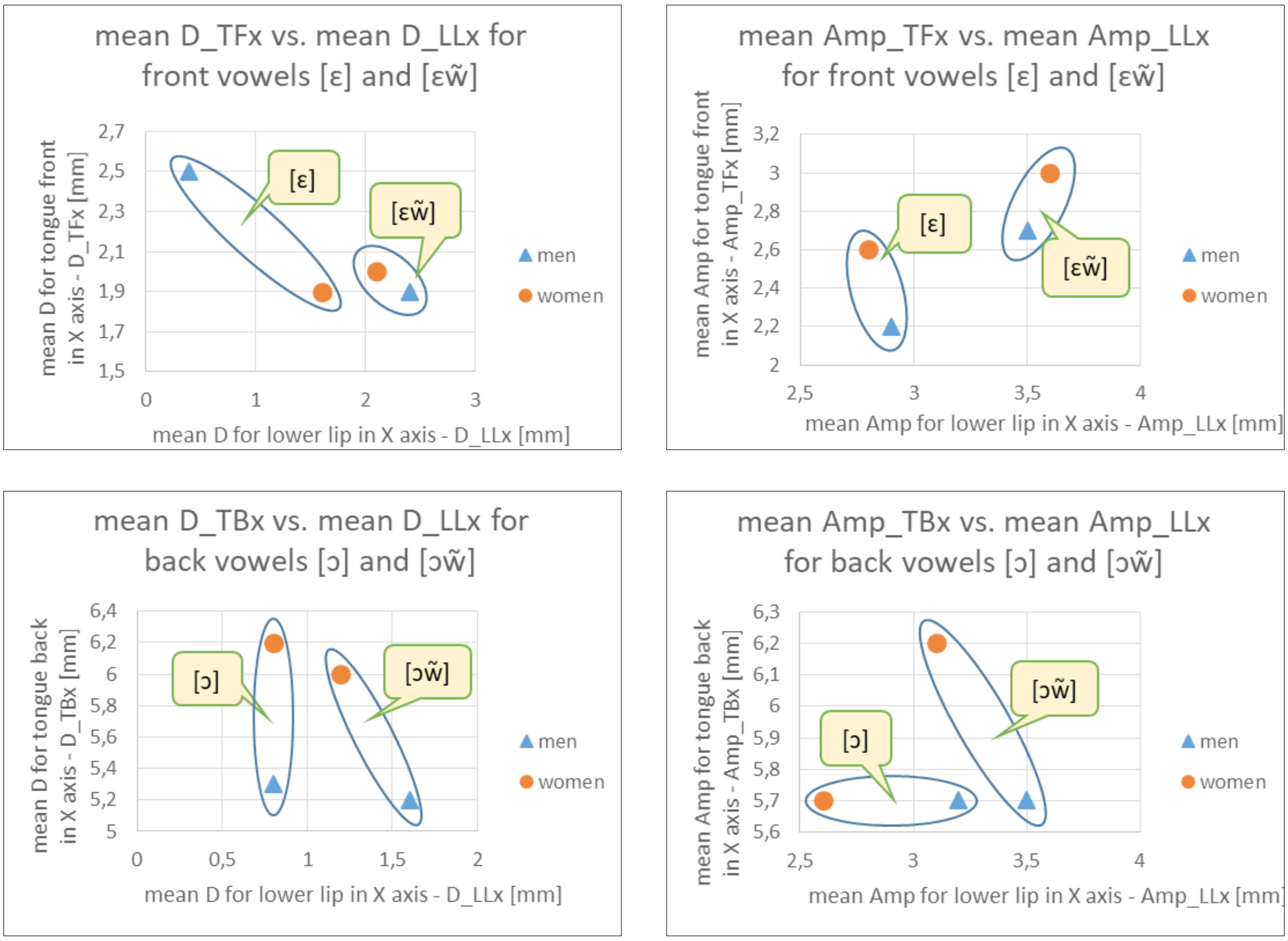

Figure 8. Diagrams of mean values for D and Amp parameters for different Polish oral and nasalised vowels

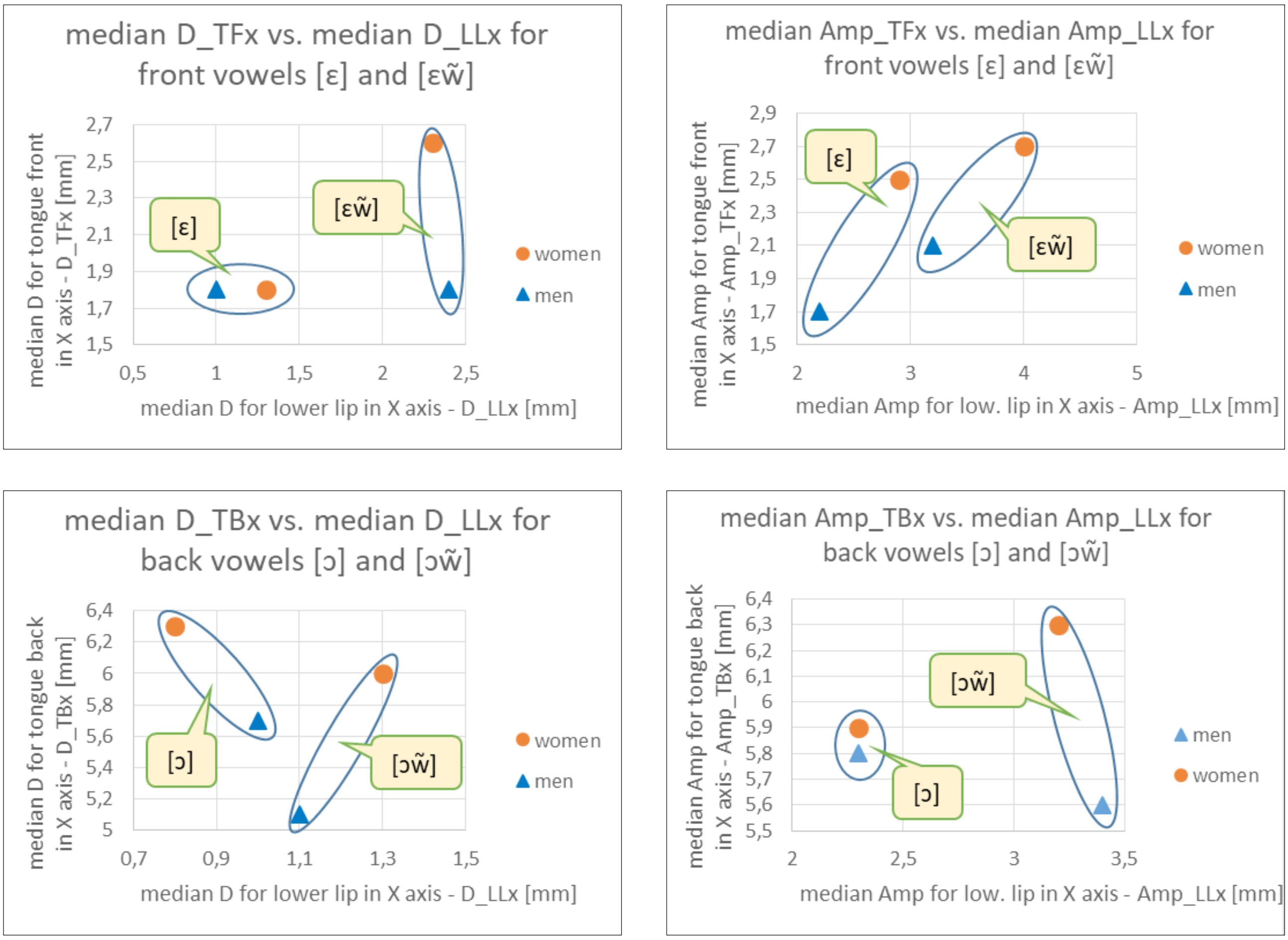

Figure 9. Diagrams of median values for D and Amp parameters for different Polish oral and nasalised vowels

Discussion

Research presented in this paper supplements another works described in [12] and [26]. Here more attention was paid to the inter-individual variability and differences between oral and nasalised vowels.

Inter-individual variability was assessed taking into account two factors: sex of the speaker as well as individual speaker features.

Formally, results of the statistical tests indicates rather low variability caused by sex of the speaker at the level of statistical significance $\alpha = 0.05$. However results of the tests from Table 2 for oral vowels and for lower lip movements along X axis proved such a variability. It should be also noted that results for nasalised vowel [ɔw̃] are significantly closer to meet criteria of statistically significant difference than the rest of the obtained results. Indeed for the vowel [ɔw̃] there is statistically significant difference between male and female speakers for D and Amp parameters at the level $\alpha = 0.15$ for tongue back and at the level $\alpha = 0.2$ for lower lip movements in X-axis. Table 2 shows also another relatively distinct relations. The first one concerns front vowels and indicates lower variability than back vowels. The second is that parameter D shows higher variability than parameter Amp.

Much higher variability exists between speakers in comparison with variability in terms of sex. Results in Table 3 show, that variability indicated by Kruskall-Wallis test is much lower than variability indicated by HSD Tukey's test with slightly violated assumptions. In order to decide which test is more reliable, variability was also assessed visually by observation of diagrams like that from figure 6. Observation proved that more reliable are results of the HSD Tukey's test what can be also easily checked on the example illustrated on figure 6 by comparing it with results from upper table in the Figure 7 where the assumptions of Tukey's test were slightly violated. For instance points which belong to the speaker JS are very well separated from the points belonged to speakers ES, MB, MJ in categories of D_LLx parameter what is also indicated in the upper table in the Figure 7.

The last analysis presented in the paper concerns differences between oral and nasalised vowels. As mentioned in the section 3.2 mean and median values of Amp and D parameter for the lower lip in nasalised vowel group are distinctly higher than the same values for oral vowel group. These differences range from 0.1 mm up to 2 mm if vowels are considered separately in the group of men and women. In practical terms, this means that the range of movement of the lower lip is higher than the range of the movement of the tongue front or tongue back.

As a side result of the analysis of differences between oral and nasalised vowels, yet another interesting relation resulted for the inter-individual variability in the sex category. Namely mean and median values of the movement amplitude and D parameter for the tongue are usually higher for women than for men. This relation is more pronounced for median than for mean values. Results obtained by median calculation should be considered as more reliable because median is more robust for outliers than average value. However it should be also taken into account that these differences between male and female speakers are rather not statistically significant as shown in the Table 2.

Conclusion

Research described in this paper proved rather low variability in the basic articulation of nasalised and oral vowels between male and female speaker. The only exceptions are the movements of lower lip for nasalised vowels. In spite of low variability in the movement of articulators in terms of speaker's gender, slightly higher range of the tongue movement occurred in the group of women.

Much higher variability is observed when each speaker is taken into consideration separately. According to the HSD Tukey's statistical test, number of pairs of speakers with statistically significant difference ranged from 0 to 31. Mean percentage of pairs with statistically significant difference equals about 20%.

The results indicated also higher lower lip movements for nasalised vowels than for oral ones however this relation should be proved statistically in the future analysis.

Possible further research directions may include:

- analysis of articulator movements in Z axis
- analysis of time dependencies during vowel articulations
- correlational and regressive analysis of EMA sensor movements.

Acknowledgements

Work described in this paper was supported by grant Nr 2012/05/E/HS2/03770 titled Polish Language Pronunciation. Analysis Using 3-dimensional Articulography with A. Lorenc as the principal investigator. The project is financed by The Polish National Science Centre on the basis of the decision Nr DEC-2012/05/E/HS2/03770.

References

1. J. S. Perkell, M. H. Cohen, M.A. Svirsky, M.L. Matthies, I. Garabieta, and M.T. Jackson, *The Journal of the Acoustic Society of America,* 1992, **92(6)**, 3078–3096.
2. J. Beskow, O. Engwall and B. Granström, *Simultaneous measurements of facial and intraoral articulation*, Proceedings of Fonetik 2003, Dept. of Linguistics, Stockholm University, pp. 57–60, 2003.
3. P. Hoole, B. Kühnert, *Proceedings XIIIth Int. Conf. Phon. Sci*., 1995, **2**, 442–445.
4. J. Frid, S. Schötz, A. Löfqvist, *Proceedings of Fonetik*, 2012, 85–88.
5. P. Hoole, B. Kühnert, *Tongue-jaw coordination in German vowel production*, Proceedings of the 1st ESCA tutorial and research workshop on Speech Production Modelling/4th Speech Production Seminar, Autrans, 97–100, 1996.
6. J.A. Shaw, W. Chen, M.I. Proctor, D. Derrick, E. Dakhoul, *On the inter-dependence of tonal and vocalic production goals in Chinese*. 10th ISSP, Cologne, 5–8 May, pp. 395–398, 2014.
7. P. Hoole, C. Mooshammer, H. G. Tillmann, *Proc. ICSLP 94*, Yokohama, 1994, **1**:53–56.
8. A. Stella, M. Vanrell, M. Iraci, P. Prieto, B.G. Fivela, *Inter-gestural coordination between tonal and oral gestures in Catalan*, Italian and Spanish. 10th ISSP, Cologne, 5–8 May, pp. 417–420, 2014.
9. IPA (International Phonetic Association), Handbook of the International Phonetic Association. A Guide to the Use of the International Phonetic Alphabet, Cambridge University Press, Cambridge, 1999.
10. W. Jassem, *Journal of the International Phonetic Association*, 2003, **33(1)**, 103–107.
11. L. Madelska, *Słownik wariantywności fonetycznej współczesnej polszczyzny*, Collegium Columbinum, Kraków, 2005.
12. A. Lorenc, *Wymowa normatywna polskich samogłosek nosowych i spółgłoski bocznej*, Dom Wydawniczy ELIPSA, Warszawa 2016, ISBN 978-83-8017-090-2.
13. H. Kjellström, O. Engwall, *Speech Communication*, 2009, **51(3)**, 95–09.
14. D. Schabus, M. Pucher, P. Hoole, *The MMASCS multi-modal annotated synchronous corpus of audio, video, facial motion and tongue motion data of normal, fast and slow speech*, Proceedings of the 9th Language Resources and Evaluation Conference, 26–31 May, Reykjavik, Iceland, pp. 3411–3416, 2014.
15. D. Król, A. Lorenc and R. Święciński, *Detecting Laterality and Nasality in Speech with the Use of a Multi-Channel Recorder*, Proceedings of the 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5147–5151, 2015.
16. A. Lorenc, D. Król, R. Święciński, *Assessment of Sound Laterality with the Use of a Multi-Channel Recorder*, 18th International Congress of Phonetic Sciences (ICPhS) 2015, 10–15 August, Glasgow, Scotland, http://www.icphs2015.info/pdfs/Papers/ICPHS0895.pdf, 2015.
17. Ł. Mik, R. Wielgat, A. Lorenc, D. Król, R. Święciński, R. Jędryka, *Multimodal Speech Data Acquisition with the Use of EMA Fast-speed Video Cameras and a Dedicated Micro-*

phone Array, 23rd International Conference Mixed Design of Integrated Circuits and Systems (MIXDES), Łódź, Poland, June 2016.
18. P. Boersma, D. Weenink, *Praat: doing phonetics by computer* [computer program, version 5.3.57], http://www.praat.org/, 2014.
19. P. Machač, R. Skarnitzl, *Principles of phonetic segmentation*, Epocha Publishing House, Praha 2009.
20. Ł. Mik, A. Lorenc, *phoneEMAtool* [computer program] 2015.
21. C. Kroos, *Eingiplige und zweigipflige Vokale des Deutchen? Kinematische Analyse der Gespanntheitsopposition im Standarddeutchen*, Hausarbeit zur Erlangung des Magistergrades, München: Ludwig-Maximilians Universität, 1996.
22. V. Bukmaier, J. Harrington, U. Reubold, F. Kleber, *Synchronic variation in the articulation and the acoustics of the Polish threeway place distinction in sibilants and its implications for diachronic change*, 15th Annual Conference of the International Speech Communication Association (INTERSPEECH) 2014, 14–18 September, Singapore, pp. 203–207, 2014.
23. C.T. Best, C. Kroos, R.L. Bundgaard-Nielsen, B. Baker, M. Harvey, M. Tiede, L. Goldstein, *Articulatory basis of the apical/laminal distinction: tongue tip/body coordination in the Wubuy 4-way coronal stop contrast* [w:] Fuchs S., Grice M., Hermes A., Lancia L., Mü cke D. (ed.), *Proceedings of the 10th International Seminar on Speech Production (ISSP)*, 2014, 5–8 May, Cologne, Germany, pp. 33–36, 2014.
24. J. Kim, S. Lee, S. Narayanan, *Estimation of the movement trajectories of non-crucial articulators based on the detection of crucial moments and physiological constraints*, 15th Annual Conference of the International Speech Communication Association (INTERSPEECH) 2014, 14–18 September, Singapore, pp. 163–168, 2014.
25. J. Kim, A. Toutios, S. Lee, S. Narayanan, *Journal of the Acoustical Society of America*, 2015, **137(3)**, 1411–1429.
26. A. Lorenc, R. Wielgat, *Wewnątrzosobnicza zmienność położenia narządów artykulacyjnych podczas wymowy wybranych samogłosek polskich*, Tarnow Scientific Colloquia – Humanities, (2) 1/2017, Tarnów, 2017, ISSN 2451-2990, pp. 129–157.