[1]LILIANNA KONOPSKA, [2]JERZY SAWICKI
[1]Szczecin University, [2]West Pomeranian University of Technology in Szczecin

Studies of the VOT parameter in realizations of Polish voiceless and voiced plosive phonemes*

SUMMARY

The paper presents the results of phonetic and acoustic studies concerning the VOT. (*Voice Onset Time*) parameter in the word-initial realizations of voiceless and voiced plosive phonemes /p/, /t/, /k/, /b/, /d/, /g/ in 18 speakers with standard pronunciation, aged from 5 to 19 years.
Keywords: VOT, plosive consonants, children.

INTRODUCTION

Stop consonants or plosives are the only kind of consonants present in each of the world's known languages. They can be realized in all parts of the vocal tract, from the lips to the glottis (Ladefoged, Maddieson 2008, 13-15). Out of many possible places of articulation, some are more preferred than others. According to the UCLA (UPSID) data[1] of 1990 the occurrence of plosives in respect of the place of articulation in 393 investigated languages had a decreasing tendency in the following order: dental/alveolar[2], bilabial, velar, glottal, palatal, uvular, retroflex, postalveolar, epiglottal, labiodental, and linguolabial consonants. The universal boundaries of the articulation area of stop consonants are marked off by the lips

* The study financed from budgetary funds for science in 2010-2013 as a research project no. N N104 084639.

[1] University of California, Los Angeles, UCLA), *Phonological Segment Inventory Database,* UPSID). This database contains data about distribution of 919 different segments in 451 languages of the world.

[2] The aforecited study specified one common area for dental and alveolar plosives according to the calculation rules used in UPSID (Henton et al. 1992).

and the uvula[3] (Henton et al. 1992). Out of all stop consonants, the most frequently present are voiceless unaspirated ones, then voiced consonants, and then voiceless aspirated consonants. If a language has only one series of stop consonants, they are always pulmonic unaspirated voiceless stops. As a rule, the most common places of articulation are connected with the most common types of phonation and the mechanism of producing the articulatory airstream mechanisms; therefore the most frequent stop consonant in the investigated languages of the world is the unaspirated voiceless consonant [t] (dental/alveolar) (Henton et al. 1992).

In the articulation of stops with the pulmonic airstream mechanism, several essential phases are distinguished: the onset of the closure (occlusion), when one articulator is moving toward the other; the occlusion, when the closed articulators block the airflow, and the release of occlusion, when the articulators are moving away from each other and the occluded air is suddenly released (plosion). Changes in the glottal activity taking place in the phases of formation, duration, and release of occlusion can – as the separating activities – be comparably important in the process of auditory perception of stop consonants. The occlusion segment with the following burst and aspiration noise are the most important information carriers about a stop consonant (Henton et al. 1992). Speakers of particular languages use different states of glottal activity, and, as a result, also a different airstream passing through the glottis and the vocal tract, as well as combinations of these activities in order to achieve the phonological distinction of stops in a given language (Fig. 1.).

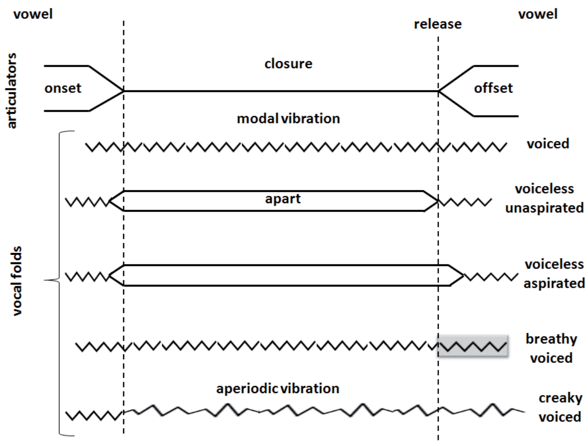

Fig. 1. The schematic time structure of individual phases of stop consonants in the intervocalic position with specification of different states of glottal activity and the passage of the articulatory airstream[4]

[3] Exluding labiodental and linguolabial articulations (Henton et al. 1992).

[4] Prepared and modified based on: Henton, Ladefoged, Maddieson (1992, 49, 66).

It is therefore indisputable that in the process of production of stop consonants in individual languages, also depending on the position of a consonant in the word and its articulation phase, different types of phonation may appear, e.g. preaspired[5] – Iceland, Gaelic; voiced – most languages; voiceless – most languages; creaky voiced – Hausa and Mazatec; stiff voiced – Jingpho and Korean; aspirated – Danish and Thai; breathy voiced – Hindi and Marathi; slack voiced – Javan and Wu. The airstream mechanism and its direction can also be different, e.g. pulmonic egressive (vast majority of languages) but also glottalic ingressive (e.g. Sandhi, and in the African language group – Hausa and Khoisan). Stop consonants can be also initiated differently in the phase of the occlusion formation (e.g. word-initial glottalization, and word-initial nasalization[6]), although, generally, these actions are less varied and regarded as less important than the occlusion and/or actions ending the occlusion since phonemically significant phenomena occur in these phases (Henton et al. 1992; Ladefoged, Maddieson 2008)[7].

VOT PARAMETER

Out of the possible phonetic cues that signal the phonological correlation of voicing within stop consonants, the prevailing manner of contrasting are different states of glottal activity. One of the most measurable phonetic-acoustic parameters of stops is the moment, expressed in the time domain, of the onset of voicing (vibration of vocal folds) relative to the moment of the occlusion release (Henton et al. 1992; Ladefoged, Maddieson 2008). It is assumed that the idea of *Voice Onset Time* or the temporal relations between the occlusion release in a plosive and the onset of periodic vibrations of vocal folds goes back to the late 19th century and was expressed by H. Adjarian in the description of realization of Armenian stop consonants[8]. The history of VOT as a parameter used in speech synthesis dates back to the 1950s and research work conducted in the Haskins Laboratories. The VOT parameter was applied in practice in a special pattern-playback machine (speech synthesizer), in which optical-mechanical methods were used to process information contained in the spectrogram into its corresponding acoustic sequence (Rottenberg 2009). The VOT parameter became widespread in studies on speech signals when in 1964 L. Lisker and A. Abramson published the pioneering results

[5] In the intervocalic or final position (Henton et al. 1992).

[6] Prenasalization usually occurs with voiced stop consonants and in many languages in which no voiced stop consonants are contrasted with word-initial nasal consonants; it may be a way to make it easier to maintain voicing of the stop consonant (Henton et al. 1992).

[7] A detailed discussion of the barely signalled phenomena would have to go beyond the scope of this study.

[8] Source: Wikipedia. http://en.wikipedia.org/wiki/Voice_onset_time (accessed: 25 August 2013).

of phonetic-acoustic investigations, in which the *Voice Onset Time* was defined as: "the interval between the release of occlusion and the onset of glottal vibration, that is, voicing" (Lisker, Abramson 1964, 389); or "the time interval between the burst that marks release and the onset of periodicity that reflects laryngeal vibration" (Lisker, Abramson 1964, 422). The goal of the research was to find acoustic features that would serve as cues for discriminating stop consonants by marking off temporal relations between the release of occlusion and the onset of voicing. Consequently, VOT duration values were measured in realizations of stops (labial, dental, alveolar, retroflex and velar) in the word-initial position in 11 selected languages divided into three groups depending on the number of categories of stops. The group that uses two categories of stops embraced Dutch, American English, Cantonese, Hungarian, Spanish (Puerto Rico) and Tamil; the group with three categories consisted of Korean, Eastern Armenian and Thai; and the third group with four categories comprised Hindi and Marathi. In the procedure (described by Lisker and Abramson) for marking off the interval between the release of occlusion of the consonant being realized and the onset of vocal folds vibration the point zero adopted in the time domain is the burst (explosion). The conducted measurements of VOT values show that in the 11 languages investigated the word-initial stop consonants were located along the VOT continuum[9] essentially in two out of the foregoing three categories of VOT[10] (Lisker, Abramson 1964), which, with time, became the grounds for dividing languages into those based on voicing and those based on aspirations (Keating 1984). The three ways of the realization of voicing specified by Lisker and Abramson, based on the values assumed by the VOT parameter in word-initial plosive consonants in the pre-consonantal position, are as follows:

1) negative values of VOT – voicing lead: periodicity appears before the consonantal burst in the segment of occlusion of the articulators, VOT values typically ranging from (-125) [ms] to (-75) [ms], the mean value being ca. (-100) [ms] (voiced unaspirated plosives);

2) small positive VOT values – short voicing lag: periodicity starts at the moment of or right after the release of occlusion of the articulators, typical VOT values ranging from 0 [ms] to 25 [ms], the mean value being ca. 10 [ms] (voiceless unaspirated plosives);

3) large positive VOT values – long voicing lag: periodicity starts after the release of occlusion of the articulators and consequent aspiration or silence after the release of the blockage, typical VOT values ranging from 60 [ms] to 100 [ms], the mean value being ca. 75 [ms] (voiceless aspirated plosives) (Lisker, Abramson 1964).

[9] In their interpretation, VOT should not be treated as an acoustic continuum but as an articulatory or physiological continuum (Abramson 1977, 296).

[10] Thai, East Armenian, and Korean use three VOT categories (Lisker, Abramson 1964).

In Lisker and Abramson's interpretation, VOT is a laryngeal dimension, i.e. the laryngeal time coordination of voicing with a complex set of intersecting, overlapping or even discretely expressed acoustic cues manifested inter alia in the energy of explosion, amplitude of aspiration noise, the F_1 formant cutback, lower frequency of the first formant in the initial phase of the following consonant, and changes in basic frequency (Lisker, Abramson 1964; Abramson, Lisker 1965; Abramson 1977). It follows from the research started by Lisker and Abramson, and from the extremely numerous studies following their publication, that VOT values (long/short duration in the time domain) depend to some extent on the place of articulation, the VOT parameter as a rule assuming higher values, with the change of articulation place shifting in the anteroposterior direction. Relations between the place of articulation, the mass of the articulators and the speed of articulatory movements, and VOT values, also termed universals, are as follows:

1) the more posterior the consonant occlusion, the higher the VOT values (longer VOT), velar consonants always having higher values than the other stops,

2) the more extensive the contact of the occluded articulators, the higher the VOT values (longer VOT),

3) the faster the movement of the articulators, the lower the VOT values (shorter VOT),

4) the less time the articulators take to reconfigure their arrangement (from the place of the closure to the position of the tongue mass characteristic of the following consonant), the smaller the VOT values are (Cho, Ladefoged 1999).

There are also other, generally known, qualitatively different factors influencing the length of VOT: contextual and those related to the speed of utterance. Many of the factors have been discussed in another study devoted to the duration of voiced fricatives (Konopska 2013). Studies on VOT universals show that the duration of VOT is longer before higher vowels than lower ones, in a stressed syllable than in an unstressed, in single words than in sentences, in slower utterances than in faster ones (Lisker, Abramson 1964; Lisker, Abramson 1967; Lehiste 1977, Flege, Brown 1982; Docherty 1992; Cho, Ladefoged 1999; Mortensen, Tøndering 2013).

It has just been half a century since the first study about VOT was published. During that time the characteristics of voiced and voiceless plosives in many languages were studied and described using VOT, inter alia in the aspects of development, categorial perception, foreign language teaching, and in bilingualism (Caramazza et al. 1973; Macken, Barton 1979; Macken, Barton 1980; Keating et al. 1980; Davis 1995; Kehoe et al. 2004; Hoonhorst et al. 2009; Ringen, Kulikov 2012)[11]. The VOT parameter was also applied in clinical studies on different

[11] To name the vast literature on investigations on VOT is beyond the scope of the present study.

speech disorders, inter alia as a reliable tool for measuring and assessing the perception and realization of ±voicing in realizations of plosives in the word-initial position (e.g. Itoh et al. 1982, Auzou et al. 2000, Özsancak et al. 2001, Arnaut, Ávila 2008, Collet et al. 2012). Out of the many studies using VOT, only several are concerned with Polish, barely three of them being the source of published data on VOT values in the speech signal of Polish native speakers. Polish plosives have their values in two VOT categories. For the sound realizations of phonologically voiced phonemes /b, d, g/ it is the voicing lead, and for the sound realizations of phonologically voiceless phonemes /p, t, k/ it is the short voicing lag. The first Polish study, in which the results of VOT measurements for one adult speaker were published, is A. Kopczyński's report dated 1971 (Kopczyński 1971)[12]. He presented the obtained data also in the contrastive studies on Polish and American English (Kopczyński 1977, 72). The next studies published between 1977 and 1981[13] by P. Keating et al. discuss the issues of VOT production and perception, thus being the second source of data about the values of VOT in the case of Polish-speaking adults (Moslin, Keating 1977; Mikoś et al. 1978; Keating 1980; Keating et al. 1981).

The measurements of VOT in the speech signal of Polish-language child speakers were taken by A. Trochymiuk (Trochymiuk 2005; 2007; 2008). Her study devoted to the auditory and acoustic analysis of articulation in deaf children using cued speech also discusses the issues connected with the realization of voicing in word-initial stop consonants. The need to solve the research problem in the group with disordered articulation contributed to starting identical tests in the control group. Data were thus obtained, including VOT values in the speech signal in child speakers with normal articulation, aged 8-12 years. In the context of studies on learning a second language and foreign language teaching methods the studies by W. Gonet (2001), A. Rojczyk (2009; 2010), and E. Szalkowskiej-Kim (2010) should be mentioned, and in the context of bilingualism – a note by L. Newlin-Łukowicz (2010).

OWN RESEARCH

In Konopska's own studies into desonorization dyslalia,[14] attempts were also made, using the acoustic phonetics methods, to solve the research problem concerning the quality of sound realizations of plosive, affricate and fricative oppo-

[12] Measuremenst taken at University of California, Los Angeles.

[13] For /t/ and /d/ – based on recordings of 24 speakers, and for all stop consonants in Polish – based on recorded speech signals of five adult speakers.

[14] To describe the type of dyslalia distinguished in Polish logopedics and conventionally termed "voiceless speech", own name of the disorder was adopted: desonorization dyslalia (Konopska, Tarnowska 2005).

sition phonemes. In the adopted 37-element phonological system of contemporary Polish, authored by B. Rocławski, this disorder applies to voiced non-tense phonemes in 10 pairs of opposition phonemes (/b/ : /p/, /d/ : /t/, /g/ : /k/, /v/ : /f/, /z/ : /s/, /ʑ/ : /ɕ/, /ʒ/ : /ʃ/, /dz/ : /ʦ/, /ʥ/ : /ʨ/, /ʤ/ : /ʧ/ (Rocławski 1986, 2005)[15]. The VOT parameter was used in those studies in order to assess the realization quality of the contrast of ±voicing in word-initial realizations of voiced and voiceless plosive phonemes – labial, dental, and velar – in the speech signal in children with desonorization dyslalia and in speech signals in children with correct pronunciation (control group). It was necessary to form a control group for the acoustic-phonetic analysis of speech signals because the interpretation of disordered speech signals requires identical analyses of speech signals obtained from speakers with correct pronunciation. This report presents the results of acoustic analyses concerning VOT measurements obtained from the control group, which is presented as the main group in this study. The subject of this report is the presentation of the investigation results, in which the answer was sought to the question: *What values does VOT assume in the word-initial realizations of opposing stop phonemes in speakers with normal articulation*?

TEST GROUP

The test group consists of 18 persons (9 females and 9 males) aged from 4.7 to 18 years (the mean age being 9.6 years) qualified to participate in the recording sessions out of the group of those who volunteered to take part in the recordings. All the subjects are brought up in a full family, go to public education institutions, and have learning successes. No one in the group in question has been certified as disabled or hearing impaired, or treated for a neurological disorder, or has had speech therapy treatment; however, logopedic tests were performed in most of them during the screening carried out in health care institutions. Interviews with parents provided data on the correct physical and mental development and on the correct development of speech. A logopedic test was carried out in all the subjects, on the basis of which abnormalities in articulation, speech fluency, articulatory organs, and in the occlusion and tongue frenulum were excluded. All the subjects had complete dentition in the anterior occlusion and had not been orthodontically treated. The participation in all the recording sections was unpaid.

[15] B. Rocławski's detailed justifications for the association of voicing with non-tenseness, and voicelessness with tenseness, and the use of the combination of both features in phonological description are included in his studies (Rocławski 1986; 2005).

LEXICAL MATERIAL AND RECORDING SESSIONS

The guiding principle in the selection of the lexical material was mainly the possibility of comparing sound realizations of obstruent opposition phonemes in the same vocalic context, and the need to record in each speaker several realizations of individual phonemes for the sake of example. The next determinants were connected with the requirement of uniformizing the lexical material for all the age groups and, for the youngest speakers, with an assumed at least passive knowledge of the words that were intended to be used. The need to uniformize the way of obtaining the lexical material from the speakers determined the choice: isolated words were elicited by showing pictures. The speech signal realized in acoustic tests was recorded during the recording sessions which were conducted as a "linguistic game" using the naming of a series of pictures in the test ("the Picture Naming Game") (Krajna 2008, 6 and 12). The task of the child was to say the missing word in a sentence uttered by the person conducting the test, e.g. "*Pani ma parasolkę, bo deszcz bardzo mocno* … (*pada*) [A lady has an umbrella because the rain is heavily … (pouring)]". In case of difficulty in eliciting the desired word used in a sentence, the next word on the list was to be elicited, and then the previous task was repeated (Łobacz 1996). This way of obtaining the lexical material is also termed "controlled spontaneous speech" (Kuijpers 1991, after: Łobacz 1996, 99). In order to facilitate the task situation and to guide the speakers into producing a specific verbal response in the picture-naming test, single genuine color photographs of people, animals, phenomena, and actions performed by people were used, these being graphic representations of the word lists.

Table 1 contains the list of words used to measure VOT in the word-initial realizations of plosive phonemes written down in orthographic notation and in broad – phonematic – transcription using the symbols of the IPA alphabet (Jassem 2003; Trochymiuk, Święcicki 2004). The phonematic transcription of the lexical material in *lento*, harmonized with the IPA symbols for Polish presented in W. Jassem's study (2003), was made in accordance with B. Rocławski's phonemic system of contemporary Polish adopted in investigations (2005).

All recording session were carried out in the soundproof studio of the Acoustics and Sound Recording Technology Laboratory at the Department of Systems Engineering, Signals, and Electronics, West Pomeranian University of Technology in Szczecin. The participants in each session were as follows: the sound engineer responsible for recording speech signals and supervising the technical side of recordings, and the tester staying with the child in the soundproof studio. All participants in the recordings were informed about the purpose and the way of conducting them; they were given the opportunity to become acquainted with the

Table 1. Word list of the picture-naming test taking into consideration the number of realizations (planned for acoustic analysis) of voiceless and voiced plosive phonemes in the word-initial position

<table>
<tr><th colspan="4">Plosive phonemes</th><th rowspan="2">Orthographic notation</th><th rowspan="2">Phonematic transcription</th></tr>
<tr><th colspan="2">Number of realizations (51)</th><th>normatively voiceless (24)</th><th>normatively voiced (27)</th></tr>
<tr><td>/p/</td><td rowspan="2">19</td><td>8</td><td>–</td><td>pada, pasek, pije, piłka, pokój, półka, puszka, pudełko</td><td>/pada/, /pasek/, /pije/, /piwka/, /pokuj/, /puwka/, /puʃka/, /pudewko/</td></tr>
<tr><td>/b/</td><td>–</td><td>11</td><td>bada, Basie, bazie, bije, bilet, bocian, bułka, buty, bucik, budy, budzik</td><td>/bada/, /baɕe/, /baʑe/, /bije/, /bilet/, /botɕan/, /buwka/, /butɨ/, /butɕik/, /budɨ/, /budʑik/</td></tr>
<tr><td>/t/</td><td rowspan="2">12</td><td>6</td><td>–</td><td>tata, talerz, Tomek, tory, tygrys, tyczka</td><td>/tata/, /taleʃ/, /tomek/, /torɨ/, /tɨgrɨs/, /tɨtʃka/</td></tr>
<tr><td>/d/</td><td>–</td><td>6</td><td>data, dach, domek, dom, dywan, dym</td><td>/data/, /dax/, /domek/, /dom/, /dɨvan/, /dɨm/</td></tr>
<tr><td>/k/</td><td rowspan="2">20</td><td>10</td><td>–</td><td>kapa, kasa, kawa, kije, kino, kozy, kość, koziołek, kury, kubek</td><td>/kapa/, /kasa/, /kava/, /kije/, /kino/, /kozɨ/, /koɕtɕ/, /koʑowek/, /kurɨ/, /kubek/</td></tr>
<tr><td>/g/</td><td>–</td><td>10</td><td>gapa, gazeta, garaż, gitara, gips, gość, godzina, gotuje, guma, góry</td><td>/gapa/, /gazeta/, /garaʃ/, /gitara/, /gips/, /goɕtɕ/, /godʑina/, /gotuje/, /guma/, /gurɨ/</td></tr>
</table>

persons carrying out recording sessions, the studio room, and the recording equipment. Recording sessions were carried out during single meetings, the session duration not exceeding 20 minutes.

RECORDING OF A SPEECH SIGNAL

To receive a speech signal two different joint electro-acoustic converters were used: a basic wideband headset microphone AKG C-555L and an additional LGF-24 Navcomm double throat microphone which recorded in detail the lower range of signal frequency, and which was placed at the throat level above the thyroid cartilage (Flege, Brown 1982; Acker-Mills et al. 2005)[16]. The speech signal from the two converters was recorded in two parallel channels at the sampling

[16] The concept and execution of the technical side of sound recording was devised by Jerzy Sawicki, PhD (Engineering) of the Department of Systems Engineering, Signals and Electronics, West Pomeranian University of Technologiy in Szczecin.

frequency of 44.1 kHz with 16-bit resolution per sample (standard CD-Audio) in a professional digital studio recorder Fostex FR-2 LE. After completion of recordings, the digital audio files were archived in CD-ROM in lossless WAV format.

The use of the first converter – the headset microphone – in registering a speech signal secures the constant distance of the microphone diaphragm from the speaker's mouth (approximately 3-5 cm), which matters in the case of child speakers because during the recordings the natural freedom of children's behaviors could have produced unexpected and significant changes in the position of the microphone relative to the mouth and, consequently, changes in the value of the sound level and the time of propagation of the acoustic signal from the mouth to the microphone. The use of the other converter was motivated by the fact that the throat microphone records in detail the range of low frequencies of a speech signal, including the signal of basic tone even at the low level of the amplitude of laryngeal vibrations: it thus provides complete data on the timing of laryngeal excitation, which was of particular importance in the tests conducted. The time shift of signals from both microphones, resulting from different times of signal propagation (the microphones were placed at different distances from the source), was analyzed in detail for each speaker using the *Praat* edition, analysis and speech synthesis program (Boersma, Weenink 2010). For each speaker, based on three to four recording excerpts, the mean time lag was determined, and then signals from the two microphones were synchronized and recorded in two-channel sound files. Independently of the sound file with synchronized signals from both microphones, an additional sound file in WAV format with the recording from the basic microphone was separated. All the results of VOT measurements obtained in the tests were determined during the manual segmentation of speech signals using the 5.2. *Praat* program (Boersma, Weenink 2010).

MEASUREMENT OF VOT IN REALIZATIONS OF VOICELESS AND VOICED PLOSIVE PHONEMES

The procedure for measuring VOT was carried out in accordance with the description in the study by L. Lisker and A. Abramson (1964, 389). VOT values were determined by marking off a time interval between the release of occlusion of the consonant being realized and the start of vocal folds vibrations. The moment of realization of the stop burst was adopted as the zero reference point in the time dimension. The VOT parameter takes on negative values if periodicity occurs before the moment of occlusion release, and positive values when periodicity occurs after occlusion release, and it assumes the zero value if periodicity occurs at the moment of the release of the articulators (Fig. 2.).

The start of burst realization (occlusion release) of voiceless and voiced plosives was assumed to be the moment when the burst (plosion) impulse is visible

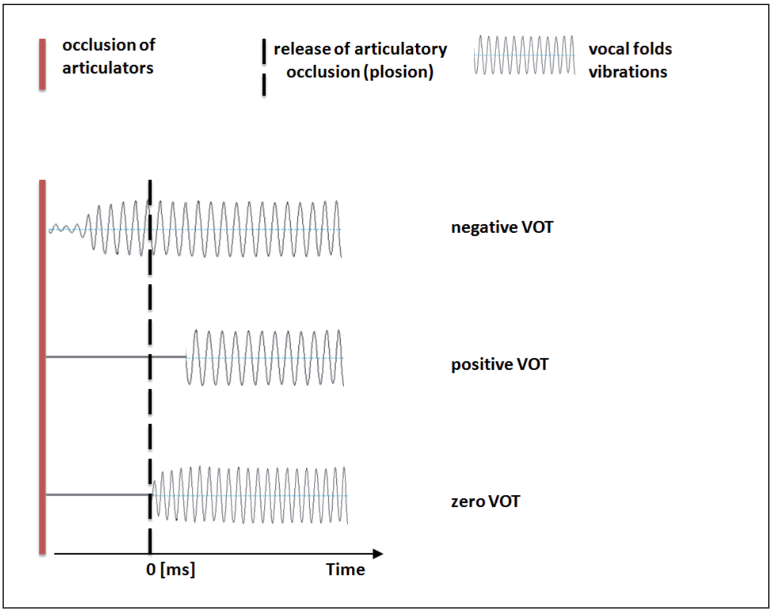

Fig. 2. Timings of three possible variants of VOT measurements[17]

in the oscillogram. Whenever it was possible the first measuring point was marked off at the moment of the zero crossing by the waveform of an acoustic speech signal impulse[18], when its deflection from equilibrium changed increasingly towards positive or negative values. In the case of voicing lead when the impulse overlapped with the periodicity of the recording and occurred e.g. at the peaks of the sound wave maximums, the transition point was defined as the point where the impulse started. In order to make as accurate measurements as possible, with the weak burst of the voiceless consonant [p], the graphic enlargement of the oscillogram was adjusted, and the range of spectrogram dynamics was changed to better image the details of the burst spectrum. The release of [p] occlusion is usually manifested in a smaller amplitude and in a more spectrally dispersed burst than in the other stop consonants (Stevens 1980, 190, [after]: Henton et al. 1992), which makes it sometimes difficult to determine VOT values[19]. If several impulses recurred in the course of burst, the test point was marked off with the first impulse (Davis 1995 after: Özsancak et al. 2001; Trochymiuk 2008). The second test of

[17] Prepared based on the figure at: http://www.phon.ucl.ac.uk/

[18] The oscillographic recording does not always cross the zero at the moment of burst release. The procedure for such case has been described below in the description of VOT determination.

[19] Some studies point out the impossibility of determining the VOT parameter, e.g because of the absence of a visible burst, which is explained by the insufficient occlusion of the articulators (usually during /p/ realizations, less often with /k/, least frequently with /t/). The available data show that the percentage of unmeasurable speech signal samples in determining VOT ranges from ca.

VOT was marked off in the oscillogram: a) in the voicing lead – at the moment of zero crossing towards positive values by the first period of the continuous series of periodic vocal folds vibrations until a burst occurred, b) in the short voicing lag – at the moment of zero crossing towards positive values by the first period of the series of periodic vocal folds vibrations in the vowel following a plosive consonant. The establishment of the transition points in the analyzed sequence made it possible to determine the sought-after duration of the interval between transition points, and its turn, and to read the value of VOT on this basis.

In marking off VOT, the leading role was played by the oscillographic recording from the basic microphone, whereas the three-dimensional, wideband spectrographic recording had a supporting role in the analysis. In ambiguous recordings of speech signals, e.g. in the case of lag in crossing of the oscillogram axis by an impulse or periodic sound wave[20], support was sought in the spectrogram of a signal from the basic microphone and/or in the oscillogram from the throat microphone, in which the time record of acoustic pressure often depicted the complex course of supralaryngeal-laryngeal positions and respiratory-phonatory-articulatory movements. In the studies presented here, a separate analysis of the realization of voicing in the word-initial realizations of the consonant phonemes /b/, /d/, /g/ covered all the realizations in which voicing occurred before the occlusion release in a consonant but a silence segment also occurred before the burst[21]. In Konopska's own research, the term "voiced before a voiceless segment before plosion" was adopted for this type of realizations. The described mode of voicing realization not fulfilling the criterion for measuring VOT is shown in Figure 3.

The number of all word-initial realizations of plosive consonant phonemes that were acoustically analyzable was 918 for 18 persons. Following the playbacks and already in the course of speech signal segmentation, 29 realizations were excluded from the acoustic analysis, which accounts for 3.16% of exclusions. The largest group of exclusions are voiced realizations with a voiceless segment before the burst (72.4%), the second largest being the realizations in which the vocalic segment after a plosive consonant was realized as voiceless (13.8%). In the remaining cases (consonant substitution 3.45%, impossible measurement 6.90%, prothetic sound 6.90%) the distribution of realizations of plosive

---

4% to 15%-20%. In French-speaking tests the percentage of unmeasurable samples in speakers with hypokinetic dysarthria was 20%, and for speakers in the control group - 6% (Özsancak et al. 2001; Auzou et al. 2000). On problems with marking off the VOT parameter in the initial realizations of the voiced consonant [d] in Polish-language speakers, see P. Kaeting (1980, 36-38).

[20] See for example P. Keating (1980, 37).

[21] P. Łobacz defines this type of voicing realization as "partially voiceless pronunciation" – "(…) at the beginning, a part of occlusion is voiced, then follows an at least 30 ms-long segment of silence before plosion and aspiration" (1996, 181). This type of realization of voicing of plosives is similarly described by A. Trochymiuk (2008, 151–152).

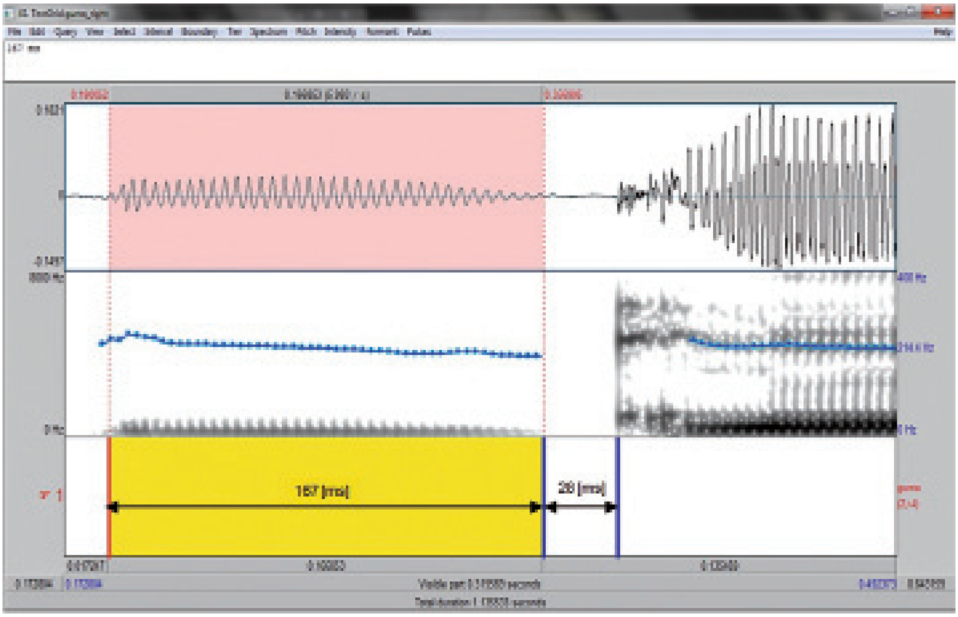

Fig. 3. The marked-off boundaries and durations of the course of +*periodicity* (167 miliseconds) and – *periodicity* (28 miliseconds) in realization of the phoneme /g/ in the initial sound of the word *guma* [rubber] uttered by a male speaker aged 10 years, and received as /guma/

phonemes excluded from acoustic analysis is similar. Acoustic analysis covered a total of 889 realizations of plosive phonemes in the word-initial position, of which 424 were sound realizations of the voiceless phonemes /p/, /t/, /k/ and 465 – sound realizations of the voiced phonemes /b/, /d/, /g/.

STATISTICAL METHODS

The following measures of statistical description were applied: arithmetic mean, to measure the degree of dispersion (variability) of results – the measure of variability degree (minimum and maximum values and standard deviation), standard error (SE) and 95% credibility interval (CI). The t-Student test and variance analysis were applied in statistical inference. Pearson's Chi-square test, the accurate Fisher test, and Spearman's rank correlation were used for analysis. Statistically significant results in all tests were the results for which $p \leq 0.05$ (boldfaced in the Tables); central tendency results were those for which probability p ranges from $p > 0.05$ to $p \leq 0.10$ (italics in the Tables), while the results for which $p > 0.10$ were regarded as statistically insignificant (Brzeziński 1996; Ferguson, Takane 2002).

Results of VOT measurements in realizations of voiceless plosive phonemes /p/, /t/, /k/

All the determined VOT values in the realizations of voiceless plosive phonemes are positive and range from 1 do 165 milliseconds: for the phoneme /p/ ranging from 1 [ms] to 67 [ms] with the mean VOT value – 24.4 [ms], for the phoneme /t/ ranging from 9 [ms] to 79 [ms] with the mean VOT value – 27.3 [ms], and for the phoneme /k/ ranging from 11 [ms] to 165 [ms] with the mean VOT value – 55.0 [ms] (Tab. 2.).

Table 2. Results of VOT measurements [ms] in realizations of voiceless phonemes /p/, /t/, /k/ taking into consideration (±SD), (SE), (CI), limit values, and statistical analysis

VOT values in realizations of phonemes /p/, /t/, /k/								
Phoneme	Number of realizations (424)	Arithmetic mean [ms]	Standard Deviation (±SD)	Standard Error (SE)	95% Credibility Intervals (CI)		Limit values [ms]	
					(-95% CI)	(+95% CI)	minimum	maximum
/p/	137	24.4	13.2	1.1	22.2	26.6	1.0	67.0
/t/	108	27.3	11.7	1.1	25.1	29.6	9.0	79.0
/k/	179	55.0	23.4	1.7	51.6	58.5	11.0	165.0
Statistical analysis								
Compared			/p/ : /t/		/p/ : /k/		/t/ : /k/	
t-Student test			*0.07087*		**0.00001**		**0.00001**	
ANOVA (dispersion of results)			0.18394		**0.00001**		**0.00001**	

The VOT value is significantly higher in sound realizations of the phoneme /k/ than in sound realizations of the phonemes /p/ and /t/ (p=0.00001). The difference between mean VOT values in realizations of the phonemes /p/ and /t/ is at the limit of statistical significance (p=0.07087). The analysis of the obtained VOT values in the realizations of the voiceless plosive phonemes /p/, /t/, /k/ shows that in the correct-pronunciation group in question:

– mean VOT values are positive and for Polish they are located along the VOT continuum in the short voicing lag,

– mean VOT values of the realization of the velar consonant are significantly higher than meat VOT values of plosives with the anterior place of articulation (i.e. labial, dental),

– the value of VOT changes as the place of articulation changes in the antero-posterior direction, and this is an increasing tendency from the labial to the dental to the velar consonant ([p] → [t] → [k]) (Fig. 4.).

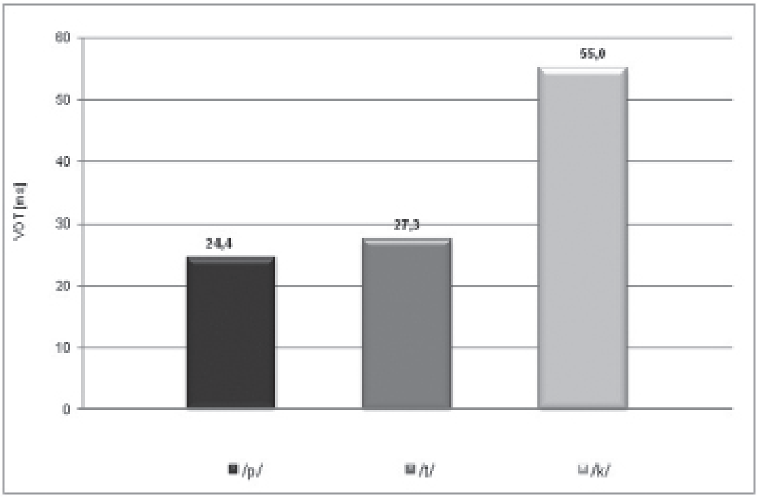

Fig. 4. Mean VOT values for normatively voiceless realizations of plosive phonemes /p/, /t/, /k/

Results of VOT measurements in realizations of voiced plosive phonemes /b/, /d/, /g/

The realizations of the voiced consonant phonemes /b/, /d/, /g/, recorded in the speech signal, have the same vocalic right-side context as the discussed realizations of the voiceless consonant phonemes /p/, /t/, /k/. The obtained results of mean VOT values in the word-initial realizations of voiced plosive phonemes are presented in Table 3.

The VOT minimum and maximum limit values in the tested group assume positive and negative values ranging from 38 [ms] to (-238) [ms]. Mean VOT values are negative and are: (-101.9) [ms] for sound realizations of the phoneme /b/, with limit values ranging from 17 [ms] to (-238) [ms]; (-109.5) [ms] for sound realizations of the phoneme /d/, with limit values ranging from 8 [ms] to (-210) [ms]; (-82.3) [ms] for sound realizations of the phoneme /g/, with minimum and maximum values ranging from 38 [ms] do (-192) [ms]. The mean negative VOT values for /g/ are significantly lower than the obtained mean negative VOT values for /b/ and /d/ ($p = 0.00001$). In contrast, the difference in mean VOT values in the realizations of /b/ and /d/ is at the limit of statistical significance ($p = 0.07347$) (Tab. 3, Fig. 5.).

Figure 6 shows an extremely high VOT value amounting to (-238) [ms] in the realization of the phoneme /b/ in the word *bada* [is examining] said by a seven-year-old boy.

The analysis of the obtained mean VOT values in the realizations of the normatively voiced plosive phonemes /b/, /d/, /g/ in the correct-pronunciation group shows that:

Table 3. Results of VOT measurements in realizations of voiced phonemes /b/, /d/, /g/ taking into account (±SD), (SE), (CI), limit values and statistical analysis

VOT values in realizations of phonemes /b/, /d/, /g/								
Phoneme	Number of realizations (465)	Arithmetic mean [ms]	Standard Deviation (±SD)	Standard Error (SE)	95% Credibility Intervals (CI)		Limit values [ms]	
					(-95% CI)	(+95% CI)	minimum	maximum
/b/	196	-101.9	34.7	2.5	-97.0	-106.8	17.0	-238.0
/d/	103	-109.5	34.7	3.4	-102.7	-116.3	8.0	-210.0
/g/	166	-82.3	46.2	3.6	-75.2	-89.4	38.0	-192.0
Statistical analysis								
Compared			/b/ : /d/		/b/ : /g/		/d/ : /g/	
t-Student test			*0.07347*		**0.00001**		**0.00001**	
ANOVA (dispersion of results)			0.97101		**0.00013**		**0.00198**	

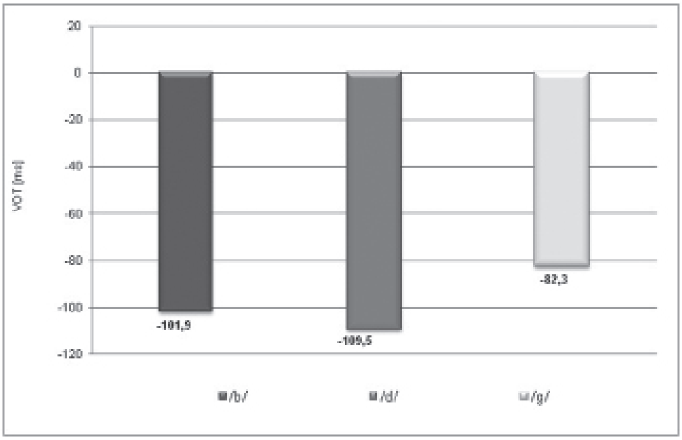

Fig. 5. Mean VOT values in realizations of normatively voiced plosive phonemes

– minimum and maximum VOT values are positive and negative,

– mean VOT values for sound realizations of the normatively voiced plosive phonemes /b/, /d/, /g/ are negative and are located along the VOT continuum in voicing lead,

– the mean absolute value of VOT decreases progressively as the place of articulations moves from the front to the back: from the dental to the labial to the velar consonant ([d] → [b] → [g]).

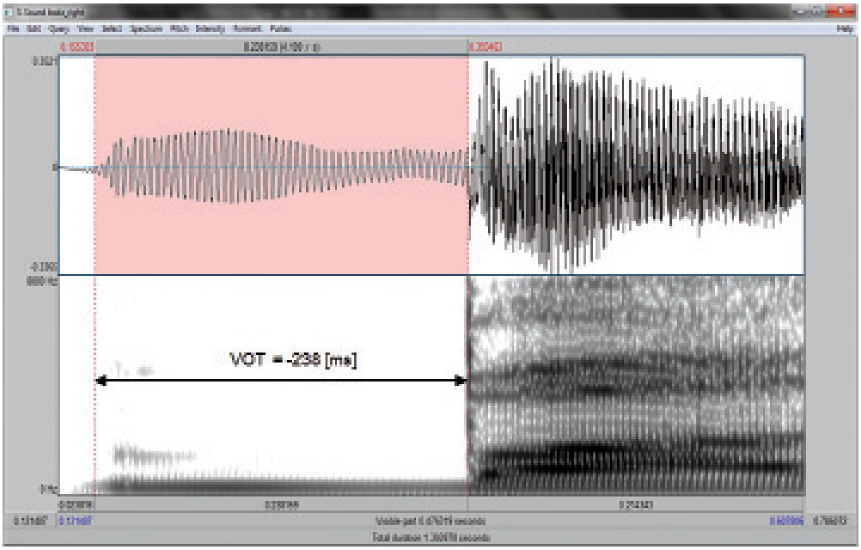

Fig. 6. The oscillogram and spectrogram of the realization of the consonant phoneme /b/ in the word *bada* said by a seven-year-old male speaker

DISCUSSION

The overall specification of mean VOT values obtained in the studies to date in the realizations of Polish minimal-pair obstruent phonemes in speech signals of adult Polish-language speakers and child speakers with correct pronunciation is presented in Table 4.

The analysis of the collected data shows that VOT values for sound realizations of the voiceless plosive phonemes obtained in Konopska's own studies are similar to those obtained by A. Trochymiuk and by P. Keating et al. (Keating et al. 1981; Trochymiuk 2008). The VOT values given by A. Kopczyński for /p/ and /t/ slightly differ from the other data, which is probably due to the fact that the analyzed speech signal comes from one speaker (Kopczyński 1971). Mean VOT values for the voiced labial consonant range from 21.5 [ms] do 37.5 [ms], for the dental – ranging from 27.3 [ms] do 33 [ms], and for the velar – ranging from 49 [ms] to 55 [ms]. In contrast, mean VOT values for sound realizations of the Polish voiceless plosive phonemes /p, t, k/ in the tests conducted by P. Keating et al., by A. Trochymiuk and in Konopska's own studies range from 34.1 [ms] to 35.6 [ms] (Tab. 4).

In the tests performed in Polish child speakers aged 8-12 years, the minimum and maximum VOT values in the realizations of the voiced phonemes /b, d, g/ assume positive and negative values like in Konopska's own studies, while the mean VOT values in the realizations of /b/ are (-88.9) [ms], in the realizations of

/d/ – (-107.2) [ms], and in the realizations of /g/ – (-76,2) [ms] (Trochymiuk 2008, 117-125, Tab. 4.). The obtained results of VOT measurements in her own investigations approximate to the measurement results obtained by L. Konopska in the case of the velar and the dental consonant whereas they differ in the mean values obtained for the labial consonant. These differences should be accounted for by the vowel neighborhood. In Konopska's own studies the VOT parameter in the realizations of /b/ was analyzed in the context of four vowels [a, o, u, i] (one low, one middle, and two high vowels) whereas the study cited analyzed VOT in the context of one low ([a]) and one high ([u]) vowel (Trochymiuk 2008, 117-121). Moreover, in Konopska's own investigations higher VOT values were obtained in the realizations of /b/ before the vowel [a][22] than in A. Trochymiuk's investigations, which might have also impacted the differences established.

The results of VOT measurements obtained in the presented tests make it possible to calculate and compare generalized mean VOT values. For voiced con-

Table 4. Comparative analysis of the hitherto obtained mean VOT values in realizations of the Polish plosive phonemes: voiceless /p/, /t/, /k/ and voiced /b/, /d/, /g/

Mean VOT values [ms] in realizations of the Polish plosive phonemes								
Author, year, test group	A. Kopczyński (1971) one adult (age unknown)		P. Keating et al. (1981) 5 adults (students)		A. Trochymiuk (2008) 10 children aged 8 to 12		L. Konopska, J. Sawicki (2013) 18 persons aged 5 to 18	
Phonemes	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless
/p, b/	-78	37.5	-88.2	21.5	-88.9	21.6	-101.9	24.4
/t, d/	-72	33	-89.9	27.9	-107.2	28.8	-109.5	27.3
/k, g/	-61	49	-66.1	52.7	-76.2	54.2	-82.3	55.0
Mean VOT	**-70.3**	**39.8**	**-81.4**	**34.1**	**-90.8**	**34.9**	**-97.9**	**35.6**

Sources: Kopczyński A. (1971). *Degree of voicing in initial stops in educated Polish and American English*. "Studia Anglica Posnaniensia", 3, 75-79; Keating, P. A., Mikoś, M. J., Ganong III, W. F. (1981). *A cross-language study of range of voice onset time in the perception of initial stop voicing*. "Journal of the Acoustical Society of America", 70: 1261-1271; Trochymiuk A. (2008). *Wymowa dzieci niesłyszących. Analiza audytywna i akustyczna*. Wyd. UMCS, Lublin; Konopska L.: own research.

[22] The analysis of VOT in the context of the vowel following the plosive consonant is the subject of a separate study (in preparation).

sonants these mean values range from (-70.3) [ms] to (-97.9) [ms] or – having excluded the data from one speaker – from (-81.4) [ms] to (-97.9) [ms], and for voiceless consonants they range from 34.1 [ms] to 39.8 [ms] or – having excluded the data from one speaker – from 34.1 [ms] do 35.6 [ms]. The presented data from the phonetic-acoustic analyses of the speech signals of adult and child speakers confirm that in Polish the voiceless stop consonants locate their values in the short voicing lag while the voiced consonants – in the voicing lead.

In three out of the presented four independently conducted VOT tests in Polish language users in the realizations of voiceless plosives the VOT parameter increases progressively as the place of articulation shifts in the anteroposterior direction: from the labial to the dental to the velar consonants ([p] → [t] → [k]). In the realizations of voiced plosives, VOT decreases progressively as the place of articulation moves in the anteroposterior direction: from the dental to the labial to the velar consonant ([d] → [b] → [g]). These figures fully confirm the cited universals about VOT, i.e. that the more anterior the consonant occlusion and the more extensive the contact of the occluded articulators is, the higher the VOT value; and that the faster the articulator movement, the lower the VOT value. Differences concerning the time aspect of VOT within the stop consonants are explained in terms of rules resulting from aerodynamics laws (pressure differences related to the volume of resonating cavities behind and before the place of the articulators closure), the speed of articulatory movements of individual articulators and the size of occlusion surfaces during the contact of the articulators, and the variable degree of opening of the rima glottitidis (Cho, Ladefoged 1999, Ohala 1997). At present, differences in the speeds of the anterior (as with [t]), the middle and posterior (as with [k]) part of the tongue – apart from the quantitatively different muscle mass and the varying participations of individual parts of the tongue – are explained by myological studies. Research work on the presence of individual types of muscle fibers (fast twitch fibers, slow twitch fibers and intermediate fibers between fast- and slow twitch fibers) within the internal tongue muscles (the vertical muscle, the transversal muscle, and the longitudinal muscle) demonstrates that the prevalent fibers in the internal tongue muscle are fast twitch fibers (60%), but their location is varied. In the front part of the tongue smaller fast twitch fibers predominate – 71%, in the posterior tongue part – large slow twitch fibers and intermediate fibers – 66%. In the authors' view this makeup of fibers probably reflects the genotypic and phenotypic functional specialization of the functions of the oral cavity. The predominance of type II (fast twitch) fibers enables fast and flexible activities in positioning and arranging the tongue movements in the process of chewing, swallowing, breathing, and speaking (Stål et al. 2003).

The present studies complement the existing knowledge on VOT values in Polish-language speakers[23]. However, accurate investigation of the process of improving the realization of voicing in children's productions requires cross-sectional studies on a large number of populations. The development of particular types of voicing realizations in plosives, in particular in the pronunciation of preschool children, should be the subject of separate phonetic and acoustic tests using VOT.

BIBLIOGRAPHY

Abramson A.,1977, *Laryngeal timing in consonant distinctions*, „Phonetica", 34, s. 295–303.

Abramson A., Lisker L., 1965, *Voice onset time in stop consonants. Acoustic analysis and synthesis*, [w:] *Proceedings of the 5th International Congress of Acoustics*, Liège, A51.

Acker-Mills B., Houtsma A., Ahroon W., 2005, *Speech Intelligibility with Acoustic and Contact Microphones*, [in:] *New Directions for Improving Audio Effectiveness* (pp. 7–1 – 7–14). Meeting Proceedings RTO-MP-HFM-123, Paper 7. Neuilly-sur-Seine, France: RTO. (Pobrano 19 września 2010 r. ze strony: http://ftp.rta.nato.int/Public/PubFullText/RTO/MP/RTO-MP-HFM-123/MP-HFM-123-07.pdf).

Arnaut M., Ávila C., 2008, *Devoicing of plosive phonemes in dysphonic children's speech*, „Rev. Soc. Bras. Fonoaudiol., 13 (1), s. 37–44.

Auzou P., Özsancak C., Morris R., Jan M., Eustache F., Hannequin D., 2000, *Voice onset time in aphasia, apraxia of speech and dysarthria. A review*, „Clinical Linguistics and Phonetics", 14, s. 131–150.

Boersma P., Weenink D., 2010, *Praat: doing phonetics by computer (Version 5.2)* [Computer program]. Pobrano 30 października 2010 r. ze strony: http://www.praat.org/.

Brzeziński J., 1996, *Metodologia badań psychologicznych*, PWN, Warszawa.

Caramazza A., Yeni-Komshian G., Zurif E., Carbone E., 1973, *The acquisition of a new phonological contrast. The case of stop consonants in French-English bilinguals*, „Journal of the Acoustical Society of America", 54, s. 421.

Cho T., Ladefoged P., 1999, *Variations and universals in VOT. Evidence from 18 languages*, „Journal of Phonetics", 27, s. 207–222.

Collet G., Colin C., Serniclaes W., Hoonhorst I., Markessis E., Deltenre P., Leybaert J., 2012, *Effect of phonological training in French children with SLI: perspectives on voicing identification, discrimination and categorical perception*, „Research in Developmental Disabilities", 33(6), s. 1805–1818.

Davis K., 1995, *Phonetic and phonological contrasts in the acquisition of voicing. Voice onset time production in Hindi and English*, „Journal of Child Language", 22, s. 275–305.

Docherty G., 1992, *The Timing of Voicing in British English Obstruents,* s. 9.

Ferguson G., Takane Y., 2002, *Analiza statystyczna w psychologii i pedagogice*, PWN, Warszawa.

Flege J., Brown W., 1982, *The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance*, „Journal of Phonetics", 10, s. 335–345.

---

[23] Lilianna Konopska would like to extend her gratitude to Jerzy Sawicki, PhD (Engineering) of the Department of Systems Engineering, Signals, and Electronics, West Pomeranian University of Technology in Szczecin for introducing her to the complex issues of the signal theory in the aspect of acoustic phonetics, and to Dr Anita Lorenz (Trochymiuk), UMCS Department of Logopedics and Applied Linguistics in Lublin for her assistance in deciding ambiguous cases in the recorded realization speech signal.

Gonet W., 2001, *Obstruent Voicing in English and Polish. A Pedagogical Perspective*, „International Journal of English Studies", 1 (1), s. 73–92.

Henton C., Ladefoged P., Maddieson I., 1992, *Stops in the World's Languages,* „Phonetica", s. 49.

Hoonhorst I., Colin C., Markessis E., Radeau M., Deltenre P., Serniclaes W., 2009, *French native speakers in the making. From language-general to language-specific voicing boundaries*, „Journal of Experimental Child Psychology", 104(4), s. 353–366.

Itoh M., Sasanuma S., Tatsumi I., Murakami S., Fukusako Y., Suzuki T., 1982, *Voice onset time characteristics in apraxia of speech*, „Brain and Language", 17(2), s. 193–210.

Jassem W., 2003, *Illustrations of the IPA. Polish*, „Journal of the International Phonetic Association", 33 (1), s. 103–107.

Keating P., 1980, A Phonetic Study of a Voicing Contrast in Polish. Unpublished Ph.D. dissertation, Brown University. (Pobrano 29 września 2010 r. ze strony: http://www.linguistics.ucla.edu/people/keating/thesis1-55.pdf

Keating P., 1984, *Phonetic and phonological representation of stop consonant voicing*, „Language", 60, s. 286–319.

Keating P., Mikoś M., Ganong W., 1981, *A cross-language study of range of voice onset time in the perception of initial stop voicing*, „Journal of the Acoustical Society of America", 70, s. 1261–1271.

Keating P., Westbury J., Stevens K., 1980, *Mechanisms of stop-consonant release for different places of articulation*, „Journal Acoust. Soc. Am.", 67, s. 93. Pobrano: 21 września 2010 r. ze strony: http://www.linguistics.ucla.edu/people/keating/KeatingWestburyStevensASA19 80.pdf

Kehoe M., Lleó C., Rakow M., 2004, *Voice onset time in bilingual German-Spanish children*, „Bilingualism. Language and Cognition", 7(01), s. 71–88.

Konopska L., 2013, *Iloczas głoskowych realizacji bezdźwięcznych fonemów trących* /f/, /s/, /ɕ/ i/ ʃ/, [w:] *Język. Człowiek. Społeczeństwo*, red. J. Panasiuk, T. Woźniak, Wyd. UMCS, Lublin, s. 441–477.

Konopska L., Tarnowska Cz., 2005, *Dyslalia desonoryzacyjno-fonacyjna, czyli o tzw. mowie bezdźwięcznej – inaczej*, [w:] *Logopedia. Teoria i praktyka*, red. M. Młynarska, T. Smereka, Ag. Wyd. A linea, Wrocław, s. 307–315.

Kopczyński A., 1971, *Degree of voicing in initial stops en educated Polish and American English*. „Studia Anglica Posnaniensia", 3, s. 75–79.

Kopczyński A., 1977, *Polish and American English Consonant Phonemes. A Contrastive Study*. PWN, Warszawa.

Krajna E., 2008, *100-wyrazowy test artykulacyjny*. *Założenia teoretyczne*. *Podręcznik*, Wyd. Komlogo, Gliwice.

Ladefoged P., Maddieson I., 2008, *The Sounds of the World's Languages*, Blackwell Publishing.

Lehiste I., 1977, *Suprasegmentals*, MIT Press, Cambridge–Massachusetts–London.

Lisker L., Abramson A., 1964, *A cross language study of voicing in initial stops. Acoustic measurements*, „Word", 20, s. 384–422.

Lisker L., Abramson A., 1967, *Some effects of context on voice onset time in English stops, „Language and Speech"*, 10, s. 1–28.

Łobacz P., 1996, *Polska fonologia dziecięca*. *Studia fonetyczno-akustyczne*, Wyd. Energeia, Warszawa.

Macken M., Barton D., 1979, *The acquisition of the voicing contrast in Spanish. A phonetic and phonological study of word-initial stop consonants*, „Journal of Child Language", 7(03), s. 433–458.

Macken M., Barton D., 1980, *A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time*, „Journal of Child Lang.", 7, s. 41–74.

Mikoś M., Keating P., Moslin B., 1978, *The perception of voice onset time in Polish*, „Journal of the Acoustical Society of America" (S1), 63, S19.

Mortensen J., Tøndering J., 2013, *The effect of vowel height on Voice Onset Time in stop consonants in CV sequences in spontaneous Danish*. *Proceedings of Fonetik 2013 The XXVIth Annual Phonetics Meeting 12–13 June 2013*, „Sweden Studies in Language and Culture", Linköping University, Linköping, s. 21.

Moslin B., Keating P., 1977, *Voicing distinction in Polish word-initial stop consonants*, „J. Acoust. Soc. Am.", 62, S1, S27–S27.

Newlin-Łukowicz L., 2010, *L1 phonetic interference in the VOTs produced by Polish-English bilinguals*. Dostęp na stronie: http://wa.amu.edu.pl/plm/2013/files/Abstracts/Newlin-Lukowicz_Luiza_PLM2013b.pdf

Ohala J., 1997, *Aerodynamics of phonology*. *Proceedings of the Seoul International Conference on Linguistics*, Linguistic Society of Korea, Seoul, s. 92–97.

Özsancak C., Auzou P., Jan M., Hannequin D., 2001, *Measurement of voice onset time in dysarthric patients. Methodological considerations*, „Folia Phoniatrica at Logopedica", 53, s. 48–57.

Ringen C., Kulikov V., 2012, *Voicing in Russian stops. Cross-linguistic implications*, „Journal of Slavic Linguistics".

Rocławski B., 1986, *Zarys fonologii, fonetyki, fonotaktyki i fonostatystyki współczesnego języka polskiego*, Wyd. Uniw. Gdańskiego, Gdańsk.

Rocławski B., 2005, *Podstawy wiedzy o języku polskim dla glottodydaktyków, pedagogów, psychologów i logopedów*, Glottispol, Gdańsk.

Rojczyk A., 2009, *Parametr VOT w języku polskim i angielskim. Badanie percepcji*, „LingVaria" 4, s. 29–47.

Rojczyk A., 2010, *Temporal and spectral parameters in perception of the voicing contrast in English and Polish*, Wyd. Uniw. Śląskiego, Katowice.

Rottenberg M., 2009, *Voice onset time vs. articulatory modeling for stop consonants,* „Logopedics Phoniatrics Vocology", 34, s. 171–180.

Stål P., Marklund S., Thornell L., De Paul R., Eriksson P., 2003, *Fibre composition of human intrinsic tongue muscles,* „Cells Tissues Organs", 173, s. 147–61.

Szalkowska-Kim E., 2010, *Charakterystyka fonetyczna kontrastu homoorganicznych spółgłosek zwartych w językach: polskim i koreańskim*, „Postscriptum Polonistyczne", (2 (6), s. 31–57.

Trochymiuk A., 2005, *Wymowa bezdźwięczna w świetle analizy akustycznej*, [w:] *Logopedia. Teoria i praktyka*, red. M. Młynarska, T. Smereka, Ag. Wyd. A linea, Wrocław, s. 499–519.

Trochymiuk A., 2007, *Wykorzystanie akustycznego parametru VOT w analizie mowy bezdźwięcznej*, [w:] *Język. Interakcja. Zaburzenia mowy. Metodologia badań*, t. 2, red. T. Woźniak, A. Domagała, Wyd. UMSC, Lublin, s. 212–240.

Trochymiuk A., 2008, *Wymowa dzieci niesłyszących. Analiza audytywna i akustyczna*, Wyd. UMCS, Lublin.

Trochymiuk A., Święcicki R., 2004, *Symbole podstawowej transkrypcji Międzynarodowego Towarzystwa Fonetycznego (IPA) i jej rozszerzenia (ExtIPA)*, „Audiofonologia", XXV, s. 97–114.

Internet sources

http://en.wikipedia.org/wiki/Voice_onset_time
http://www.phon.ucl.ac.uk/