Reading rate in filmic audio description*

ANNA JANKOWSKA
Jagiellonian University in Kraków
Universitat Autònoma de Barcelona
anna.m.jankowska@uj.edu.pl

BARTOSZ ZIÓŁKO
AGH University of Science and Technology
Techmo
bziolko@agh.edu.pl

MAGDALENA IGRAS-CYBULSKA
AGH University of Science and Technology
Techmo
migras@agh.edu.pl

AGATA PSIUK
Seventh Sense Foundation
agata.psiuk@siodmyzmysl.org

ABSTRACT

The study discussed in this article was carried out as a pilot study to assess the process, resources and data management scheme (Thabane et al., 2010) to be used in a large-scale experiment on filmic audio description (AD) reading rate. As part of this study we defined the reading rate in filmic AD context. We described the characteristic features of Polish filmic AD scripts and recordings and examined the reading rate of Polish AD for three Polish fiction films: a comedy, a

drama, and an action film. We calculated the average length of breath pauses and the maximum, minimum and average reading rate measured in characters per second (CPS) and words per minute (WPM) – two measures commonly used in audiovisual translation. The main finding of this study is the validation of the research procedure for testing the AD reading rate. We also computed the average reading rate for Polish filmic AD (179 WPM) and discovered that it changes depending on the film genre (167 WP for drama, 182 for comedy and 189 for action). When it comes to breath pauses in Polish AD, we calculated their average length at 190 ms – a value much lower than expected for breath pauses in Polish. The results of our study are discussed in the context of research on the speech tempo.

Key words

Audio description, speech rate, delivery rate, reading speed, voicing.

1.   Introduction

The study discussed in this article was carried out within the "AudioMovie - Cinema for All" project (2015-2018), funded by the Polish National Centre for Research and Development under the Social Innovations II programme. The project is carried out by a consortium of six partners: Seventh Sense Foundation (leader), Foundation for Audio Description Progress "Katarynka", Jagiellonian University in Kraków, Institute of Innovative Technologies EMAG, Kino Pod Baranami (cinema) and Centre of Technology Transfer EMAG.

The project is carried out in Poland and responds to the current challenges of mainstreaming AD in Poland (Jankowska & Walczak, forthcoming). The project has four main goals: (1) elaborating a code of good practice for the production of legal AD; (2) developing a cloud-based system for storing and sharing alternative audio tracks: AD and audio subtitles (AST); (3) developing a mobile application that will enable the reproduction of alternative audio tracks in cinemas and television, and (4) producing software for AD/AS editing and recording. The study we report on in this article was carried out within the last of the above-mentioned objectives.

One of the features of AD editing software is an algorithm that calculates the ideal duration of the description entered into the available time slot as well as the minimum and maximum tolerated deviation from the ideal reading rate. In order to do that, the algorithm needs to be given reading rate values it can calculate upon. Unfortunately, there is no reliable data regarding AD reading rate values. At present there are no standards regarding this aspect of AD in Poland, apart from the internal guidelines of some recording studios which require the audio describers to prepare the script with the reading rate set at 15 characters per second (CPS) (Walczak, personal communication, March 1, 2017). A similar situation can be observed in other countries (see Subchapter 3 of this article for a detailed description).

In order to obtain reliable data, we set up a study within the AudioMovie project to assess the current AD reading rate in Polish. Since to the best of our knowledge this is one of the first studies carried out on the AD reading rate, a pilot study to assess the methods, resources and data management scheme (Thabane *et al.*, 2010) was needed. This paper presents and discusses the pilot study in question.

The paper is organized as follows: in Section 2 we introduce the notions of speech tempo, speech rate and reading rate. In Section 3 we present the state of the art in research and professional practice regarding the AD reading rate. In Section 4 we discuss the peculiarities of the materials used in the study as the way in which AD is drafted and recorded influences research design. Sections 5-7 are devoted to the corpora and methodology adopted in the study. The results are presented in Section 8 and discussed in Section 9. The conclusions are presented in Section 10.

2.   Speech tempo, speech rate and reading rate

Speech tempo or simply *tempo* is "the pace at which a stretch of connected discourse is delivered by the speaker" (Jacewicz *et al.*, 2009, p. 234). The measurement of tempo can be strongly affected by hesitations and pausing (Jacewicz *et al.*, 2010). For this reason, by general agreement, it is usual to distinguish between *speech rate* for speech tempo that includes pauses and *articulation rate* for speech tempo that excludes them (Jacewicz *et al.*, 2009; Trouvain, 2003). Depending on the scope of research, the more general umbrella term s*peech rate* is often given a more precise form, i.e. *speaking rate* for studies on discourse delivered during spontaneous speech and *reading rate* for studies on discourse produced while reading a structured text.

Pauses play a central role in speech rate measurements (Trouvain et al., 2016). Research on speech rate shows that there are three types of pauses: unfilled (UFPs), filled (FPs) and breath pauses (BPs) (Rodero, 2012; Igras-Cybulska et al., 2016; Trouvain et al., 2016). The UFPs or silent pauses are defined as regions of signal where no voice activity is recorded. The FPs also called fillers or pseudo-words are defined as "sound produced during speech that represents a silent period filled by a vocalization" (Rodero, 2012, p. 394), e.g. "yyy", "eee", "hmmm". The appearance of FPs does not affect the meaning of a sentence but perturbates discourse fluency. The third type of pauses that we consider are BPs – pauses correlated with audible inhales or exhales.

In AD-related literature *pronunciation speed* (Díaz-Cintas, 2006), *reading speed* (McGonigle, 2007), *speed of description* (Fels & Udo, 2009), *description speed/presentation rate* (Udo & Fels, 2010), *narration speed* (Cabeza-Cáceres, 2013; Vercauteren, 2016) or *speech rate* (Fryer, 2016) are used interchangeably to indicate the tempo in which AD is delivered by a voice talent. However, it has to be noted that none of the above-mentioned researchers provides a definition of the terms they used, thus it is not clear what they encompas.

In this article we will use the term *reading rate*. We believe that it is the most suitable term given the unique nature of AD scripts and recordings (see 4.1 and 4.2 for a detailed explanation), the aim of our study and the scope of this article. Reading rate in this article is understood as the rate at which recorded AD is presented in its final shape – after sound editing. It includes breath pauses but excludes pauses imposed by interpretation, important sounds or synchronisation with images.

3. AUDIO DESCRIPTION AND READING RATE

Filmic AD is a service for the blind and visually impaired that renders the visual cues of an image. AD is composed of a verbal description (script) that is delivered as an additional audio track (recording) between the dialogues. In recent years, AD has received growing interest from researchers. However, most of the research concentrates on the textual aspect of AD, while the audio component has been largely neglected and so far studied only by a limited number of researchers (Fryer, 2010, 2016; Remael, 2012; Szarkowska & Orero, 2014; van der Heijden, 2009, 2011) especially in the context of text-to-speech AD (Fernández-Torné & Matamala, 2015; Fryer & Freeman, 2014; Linder & Martínez, 2011; Szarkowska & Jankowska, 2012; Walczak, 2011).

Researchers have also shown some interest in the AD reading rate (Díaz-Cintas, 2006; McGonigle, 2007; Fels & Udo, 2009; Udo et al., 2010; Cabeza-Cáceres, 2013; Fryer, 2016; Vercauteren, 2016). Most of them notice that reading rate is a factor to be considered in AD creation and/or reception process (Díaz-Cintas, 2006; Fryer, 2016; McGonigle, 2007), however so far only three studies have taken it into consideration: two in Canada (Fels & Udo, 2009; Udo et al., 2010) and one in Spain (Cabeza-Cáceres, 2013). In Canada, reading rate was included alongside with other questions about description quality (i.e. understanding, enjoyment of the event and AD, language level, style, errors, technical quality, volume, etc.) in order to assess the impact of emotional style of AD of a fashion show (Fels & Udo, 2009) and a theatre play (Udo *et al.*, 2010). Although the authors of both experiments do not provide any information about the reading rate used, they conclude that the speed of the description was accepted. In Spain, Cabeza-Cáceres (2013) tested how different reading rates (slow: 14 CPS; medium: 17 CPS; fast: 20 CPS) influenced the comprehension and enjoyment of film clips. Based on the results, Cabeza-Cáceres concluded that although the reading rate does not influence enjoyment, the lowest reading rate (14 CPS) has a positive influence on comprehension.

The situation in professional practice resembles the state of the art in research. While all of them tackle the issue of creating AD, only some guidelines provide general instructions regarding reading rate, e.g. "speak clearly at a rate that can be understood" (Snyder, 2010), "description must not be hurried" (Independent Television Comission, 2000), "conform the rate of the description to the pace of the program" (DCMP Description Key, n.d.), "voice must be adapted to

the emotional content of the scene and the pace of action but must nevertheless remain fairly neutral" (Gonant & Morisset, 2008), "tone, style and rhythm of the event must be respected" (Federation suisse des aveugles et malvoyants, 2017). On the whole, very little attention is given in the worldwide guidelines to vocal delivery in general and the reading rate in particular.

When it comes to the recommended AD reading rate, only some guidelines give specific instructions about the advised/required reading rate, which is set at 160 WPM (Snyder, 2010; Netflix 2016 and SDI MEDIA, 2017) and seems to be applied regardless of language or country. Recommendations on the AD reading rate are also provided by some researchers. Ballester (2007), based on professional practice in Spain, recommends 180 WPM or 150 WPM depending on the scene type and space available. McGonigle (2007), relying on research on reading rates in blind children, advocates using 175 WPM in case of AD for children. So far the only recommendations regarding the AD reading rate, based on research evidence, have been made by Cabeza-Cáceres (2013) who computed the average AD reading rate in Catalan television (17 CPS), but based on his research he recommeds that reading rate in Catalan should be set at 14 CPS.

As shown above, two measures are used to describe the AD reading rate: CPS and WPM. While both CPS and WPM are commonly used in subtitling, WPM is also used in research on speech tempo alongside with syllables per second. The usage of CPS and WPM in AD is most probably due to the fact that these parameters are available in AD editing software and in subtitling software that is commonly used to create AD.

4. Characteristics of Polish audio description scripts and recordings

Before we move on to discussing the experiment and its results, it is necessary to present the distinctive features of both AD scripts and recordings, since the way in which they are prepared influences the analysis and its results. Below we present a review based on the examination of twelve AD scripts and recordings[1] gathered from Polish NGOs[2] as well as Polish AD guidelines and manuals[3].

1 *The Hexer* (dir. M. Brodzki, 2001, action), *Day of the Wacko* (dir. M. Koterski, 2002, comedy), *The Welts* (dir. M. Piekorz, 2004, drama), *Jasminum* (dir. J.J. Kolski, 2006, comedy), *Time to die* (dir. Dorota Kędzierzawska, 2007, drama), *Courage* (dir. G. Zglinski, 2011, action), *Yuma* (dir. P. Mularuk, 2012, action), *Ida* (dir. P. Pawlikowski, 2013, drama), *Planet Single* (dir. Mitja Okorn, 2011, comedy), *Raiders of the Lost Ark* (dir. S. Spielberg, 1981, action), *Matrix (dir.* The Wachowski Brothers, 1999, action) and *Star Trek* (dir. J.J. Abrams, 2009, action).

2   Seventh Sense Foundation, Foundation for Audio Description Progress "Katarynka" and Fundacja Kultury Bez Barier (Foundation for Culture Without Barriers).

3 *Standardy tworzenia audiodeskrypcji do produkcji audiowizualnych* [Standards for creating audio description for audoivisual productions] (Szymańska & Strzymiński, 2010), *Audiodeskrypcja – standardy tworzenia* [Audio description – principles of creation] (Żórawska et al., 2011), *Tłumaczenia filmowe w praktyce* [Film translation in practice] (Adamowicz-Grzyb,

Unlike subtitles, AD scripts in Poland may come in different formats that vary from describer to describer and from client to client. They can be also prepared using different software – currently AD scripts in Poland are created either in text editors or in subtitling software (e.g. EZ Titles and Subtitle Workshop, to name the most popular ones). Although there is not one standardized way of editing the scripts, some common features – for filmic AD scripts that are to be recorded in a studio – can be found. The way in which AD scripts are edited is directly influenced by the recording procedure. We explain it below in more detail.

It is very rare for Polish describers to read and record their own scripts. This is usually explained by the long-standing voice-over tradition in Poland – highly skilled and trained voice talents are widely available and the audience is used to professional voice performance. It is even claimed that three professions regarding AD creation process in Poland should be distinguished: (1) describer – a person who writes AD scripts; (2) lector – a person who voices AD scripts and (3) audio describer – a person who writes and voices AD scripts (Szymańska & Strzymiński, 2010). As a result, AD is mostly recorded in professional studios by trained voice talents. The usual practice is that the voice talents see neither the film nor the script before the recording. As the voice talents do not rehearse prior to the recording, the describers are asked to edit the script in such a way that the voice talents can read it as smoothly as possible, without unnecessary breaks and retakes. The editing includes segmenting and time coding the script as well as including reading tips regarding vocal delivery and pauses.

4.1.1.   SEGMENTATION

Segmentation is used in many audiovisual texts. In subtitling, the text is divided into subtitles. In dubbing, it is divided into takes or loops (Chaume Varela, 2012). When it comes to text segmentation in AD, so far it has been described only by Piety (2004) who, based on an analysis of four audio-described American films, suggests the following units or structural components:

   a) Insertions – continuous stretches of description that is uninterrupted by other significant audio content, such as dialogues.
   b) Utterances – short snapshots of language that describe some visible features, strung together to fill the space between the dialogue that can be as long as an insertion or much shorter (Piety, 2004:11).

2013), *Audiodeskrypcja w teorii i praktyce, czyli jak mówić o tym, czego nie można zobaczyć* [Audio description in theory and practice, or how to talk about something that cannot be seen] (Trzeciakiewicz, 2014), Audiodeskrypcja [Audio description] (Chmiel & Mazur, 2014).

However, it has to be stressed that Piety worked on transcripts he prepared himself based on audiovisual material. As Fryer (2016) aptly notices His proposal misses divisions that ears cannot spot and become visible only when looking at written AD scripts. Drawing from his proposal and the scrips analysed, we would like to suggest the following categorization of AD segmentation:

a) Insertions - stretches of AD uninterrupted by dialogue. Insertions are formed by smaller units called utterances. Insertions can be formed by one or more utterances (see Example 1 and Example 2). In AD scripts, insertions are usually distinguished by fragments of dialogue (in bold) entered into the script. This serves as signposting to the voice talents who can refer to both time codes (see 4.1.2) and the dialogues for calculating when to start reading.

Jest bardzo bystra. Inteligentna.

01:36:32   [szybko] Małpka zeskakuje z ramienia Marion. Ucieka.

To daktyl. Zjedz.

Example 1 Insertion formed by one utterance.

To daktyl. Zjedz.

01:36:52   Małpka biegnie opustoszałą uliczką. Na progu jednego z domów jednooki Egipcjanin z opaską na oku. Małpka podbiega do niego. Jednooki z małpką na ramieniu przemyka między kramami. Wpada na dwóch mężczyzn w beżowych garniturach. Odciągają go na bok.

01:37:12   Żywo gestykulują. Jednooki wskazuje na zatłoczoną uliczkę. Rozchodzą się.

Tam są!

Example 2 Insertion formed by two utterances.

b) Utterances - stretches of continuous AD that may be interrupted by short pauses and sound effects. Every new utterance is introduced with a new time code (see 4.1.2). Utterances may also contain additional information regarding vocal delivery (see 4.1.3) and pauses (see 4.1.4).

Time code (TC) is used in AD scripts to indicate when the voice talent is supposed to start reading (Remael, Reviers, & Vercauteren, 2014). Every new utterance in AD is assigned an in-TC. It is assumed that AD is written in such a way that, when reading at a steady pace, the voice talent will synchronise with the image and finish before a dialogue or other significant audio content appears. However, we observed two uses of TCs that further facilitate the task of synchronic reading. The first one consists in assigning out-TCs to utterances (see Example 3).

01:37:15 – 01:37:25	
	Marion i Indi idą wąską uliczką. Wypełnia ją tłum kupujących w tunikach, turbanach i burkach. Jednooki pędzi w ich kierunku. Wpada niemal prosto na nich. Chowa się za koszami. Obserwuje.

Example 3 Utterance with in-TC and out-TC.

The second use involves introducing an additional TC within longer utterances, i.e. when an utterance is longer than 20 seconds, TCs are introduced within the text ca. every 20 seconds (see Example 4).

01:38:10	Indi i Marion przystają przy kramie. Nadbiegają zamaskowani najemnicy. Rzucają się na Indiego. Indi wali pięściami na oślep.
01:38:30	[dość szybko] Marion naprzeciwko Indiego. Najemnik za plecami Marion! Indi odsuwa Marion. Wali go pięścią w nos. Kopniakiem powala drugiego. Trzeci obrywa z pięści, Czwartego uderza z łokcia.
01:38:42	Chwyta Marion za rękę. Odciąga na bok. Wrzuca na wóz pełen słomy. Chwyta bicz. Strzela nim na lewo! Na prawo. Najemnicy odskakują jak oparzeni. Wóz z słomą rusza. Mija jednookiego z małpką na ramieniu i mężczyznę w beżowym garniturze. Mężczyzna gestem dłoni przywołuje najemnika.
01:39:00	Wóz ze słomą zatrzymuje się. Marion zeskakuje. Biegnie najemnik. Marion podbiega do straganu. Chwyta patelnię. Najemnik jest tuż przy niej. Wyciąga długi ostro zakończony nóż. Uśmiecha się głupkowato.
01:39:17	Marion rzuca się do ucieczki. Wpada do otwartych drzwi obdrapanego domu. Najemnik za nią!. Bezwładne ciało napastnika pada na bruk.

Example 4 Utterance with additional TCs

Both of these practices are clearly borrowed from voice-over scripts and allow the voice talents to estimate if they are reading at the required pace.

Tips for vocal delivery are inserted within utterances and include guidance on reading rate, interpretation and pronunciation of foreign words. They are usually introduced according to the style guide used in voice-over scripts – as short instructions in square brackets inserted within the text. In Table 1 we present the most commonly used symbols, however it has to be noted that this is not a definite list.

No.	Symbol	Example
1	[tempo]	[fast] Rose runs to the stern.
2	[interpretation]	[romantic] Kisses his daughter on the forehead.
3	[phonetic transcription]	Jack [dżak] walks into the cabin.

Table 1 Vocal delivery symbols used in AD in Poland

As far as tempo is concerned, the describers are advised to write AD in such a way that it can be read at a steady pace. Whenever a given utterance needs to be read faster or slower, they are required to introduce an appropriate instruction within the script (see Table 1, Example 1). When it comes to voice modulation, on the whole the voice talents are advised to read AD with a neutral tone that does not distract the viewers (Szymańska & Strzymiński, 2010; Trzeciakiewicz, 2014; Żórawska, Więckowski, Künstler, & Butkowska, 2011). However, it is also recognized that the voice should not be monotonous and intonation can create dramaturgy (Szymańska & Strzymiński, 2010). In case a more interpretative voicing is required, the describers can signal it within the script, by providing the voice talent with interpreting suggestions (see Table 1, Example 2). When AD is provided for foreign films, the describers usually provide orthographic transcriptions of foreign words (see Table 1, Example 3), since voice talents often do not speak the language of the production. Transcription is usually provided in square brackets inserted after the foreign word.

4.1.4.   Signposting pauses

Marking pauses is one of the most important aspects of AD script editing. When analysing AD scripts for this study, we have identified two types of pauses: (a) *between-insertion pauses* and (b) *inner-insertion pauses*. We discuss them below in more detail.

a) Between-insertion pauses

Pauses to be made between insertions. They are easily spotted since dialogue comes between the insertions. When it comes to AD script, between-insertion pauses are usually signposted to the voice talent by pieces of dialogue (in bold) inserted into AD script (see Example 1 and 2 in 4.1.1). These pauses always render in the audio file.

b) Inner-insertion pauses

Pauses that occur within insertions. They can belong to one of the two categories: *between-utterance pauses* and *inner-utterance pauses*. Inner-insertion pauses do not always render in the audio file because of the sound editing performed after the recording (see 4.2 for more details).

Between-utterance pauses are pauses made by the voice talent between utterances. They are usually signposted in two ways – either by inserting an extra blank line between utterances (see Example 5 below) or by inserting a pause symbol at the end of an insertion (see Example 6 below).

0:55:00	[szybko] Indi pospiesznie oddala się. Wchodzi do namiotu, W środku związana Marion. W ustach knebel. Indi staje jak wryty. Wzruszony doskakuje do Marion. Przytula ją.
0:55:19	Indi namiętnie całuje Marion.
0:55:38	Indi rozluźnia sznury. Nieruchomieje. Chowa scyzoryk.
0:56:03	[szybko] Indi knebluje Marion. Wybiega z namiotu.

Example 5 Between-utterance pauses signposted by extra blank lines

01:04:07	Indi ściąga z pleców Satipo kosz. Zarzuca go przed wejściem do groty. Wchodzą do ciemnego skalnego tunelu. Indi pierwszy. Za nim Satipo z zapaloną pochodnią. Idą bardzo powoli. Jest coraz ciemniej. Ze ścian spływają strużki wody. Zwisają gęste zasłony pajęczyn i mchu. //
01:04:28	Mężczyźni przedzierają się przez gęsto utkane, gigantyczne pajęczyny. W oddali majaczy jasne światło.
01:04:39	Satipo przystaje, wyraźnie przestraszony.

Example 6 Between-utterance pauses signposted by the pause symbol „//'

The symbols used to introduce pauses are the same as in voice-over. The duration of the pauses is fixed only approximately. For symbols and their designated durations see Table 2 below.

Symbol	Example	Explanation
^	A sunny day. ^ Marion and Indi walk down a narrow alley.	Short pause – ca. one breath
/	Indi puts the bottle on the table. / He sits down.	Medium pause – ca. two breaths
//	Indi in the temple. He leans over a stone tablet. //	Long pause – ca. three breaths

Table 2 Symbols used to signpost pauses in AD

Inner-utterance pauses, i.e. pauses made by the voice talent within utterances, can be signposted it two ways: by symbols and/or by the script layout. The symbols borrowed from voice-over (see Table 2 above) are inserted into the body of the utterance (see Example 7).

01:03:02	Za jego plecami Baraka. ^ Wyciąga rewolwer. ^ Mierzy. Indi odwraca się! Strzela z bicza. Rewolwer ląduje w rzece. Przerażony Baraka ucieka. Indi patrzy za nim surowym wzrokiem. / Powoli zwija bicz. Sonduje wzrokiem okolicę. Satipo przygląda mu się z rozdziawionymi ustami. // Ruszają w górę wzniesienia. Przedzierają się między wiszącymi lianami. Jest coraz ciemniej.

Example 7 Inner-utterance pauses signposted by pause symbols.

Quite often pauses to be made within a given utterance are signposted by placing a new line of text within the utterance, but without a new TC or symbol (see Example 8).

01:37:27	[szybko] Na balkonie jednego z domów mężczyźni w beżowych garniturach. W dole w tłumie Indi i Marion. Mężczyźni wchodzą do mieszkania. W środku sześciu egipskich najemników w tunikach i turbanach. Zatłoczona uliczka. Marion.

Example 8 Inner-utterance pauses signposted by a new line of text.

Quite often pauses to be made within a given utterance are signposted by placing a new line of text within the utterance, but without a new TC or symbol (see Example 8).

It is also very common to signpost pauses that need to be made within the utterance (but also between utterances) in order to expose an important sound. In that case, the most common option is to introduce the sound into the script. Once again, the practices borrowed from voiceover are followed (see Table 3).

No.	Symbol	Example
1	[paralinguistic sound]	They carry a figure of Christ [hens].
2	(paralinguistic sound)	A glass doors. Behind them the face of Aniela – a petite old woman with straight grey hair. (knocking)
3	... PARALINGUISTIC SOUND	Novices leave the building DOORS.

Table 3 Symbols used to signpost important sounds in AD

4.2. CHARACTERISTICS OF POLISH AUDIO DESCRIPTION RECORDINGS

For our study we used a corpus based on real life materials: scripts and recordings supplied by different Polish AD providers. Recordings of AD constitute a very particular corpus. As opposed to corpora used usually in speech tempo research, e.g. recorded spontaneous or read speech, AD recordings are highly edited. As we have already explained, in Poland ADs are usually recorded in sound studios that specialize in audiovisual translation and sound editing. This means that they undergo the same processes as other audio tracks: once recorded, all ADs go through a mastering process. The entire procedure has implications that need to be considered in this study.

First of all, both during the recording and the mastering process AD script can be spontaneously changed by the voice talent who may decide to omit or rephrase certain passages of the script as well as add new utterances. Sometimes changes are also made by the sound editor or producer who may cut out parts of the recordings. Such changes are quite common and very often foreseen in the describer's contract. From the point of view of our study, this means that scripts to be analysed had to be rigorously checked against the recordings (see 7.1 for more details) in order for us to be able to carry out the time alignment procedure.

Another aspect of the script that is influenced during the recording and the mastering process are the pauses. As we have already suggested above (see 4.1.4 b), not all inner-insertion pauses marked in AD script render in the audio file. Their presence and length depend on the performance of the voice talent and the sound editor. While writing, the describer aims to achieve the "ideal reading rate" – which would allow the voice talent to read at a steady pace and synchro-

nise with the image. It is very difficult for the voice talent to reach this ideal. If the voice talent reads slightly slower, he/she may omit some inner-insertion pauses in order to be able to fit in the allocated space. Pauses are also omitted when the script is too dense and the voice talent needs to speed up to fit in. On the other hand, if the voice talent reads slightly faster, he/she might need to lengthen or even add pauses to follow the image. On the top of that, some pauses might be added due to breathing. Then, in the mastering process, pauses made during the recording can be further manipulated since the sound editor can adjust the audio position, i.e. move a block of recording along the sound timeline, shortening or lengthening the pauses between them. This is often done to better synchronise with the image or to expose important sound effects. The implications for our study are that in fact it is impossible to measure the AD reading rate unless we work with the original non-mastered recordings or define filmic AD reading rate as rate at which recorded AD is presented. We opt for the second option.

Last, but not least are the BPs. The main aim of our study was to calculate the AD reading rate understood as the rate at which recorded AD is presented. This definition of the AD reading rate includes BPs but excludes pauses imposed by interpretation, important sounds or synchronization with images. In recordings used in speech tempo studies breath pauses can be easily distinguished in the recordings, since they are audible in the audio track and often visible in the audio histogram. The situation with AD recordings is different. During the mastering process any possible noise is cleared. This means that breaths are cut out and all BPs render in the audio file as silent pauses. It is also impossible to spot them on the audio histogram. Given all that, for the purpose of our study, we had to design a procedure in order to estimate the duration of BPs (see 7.3 for more details).

5.   The present study

The aim of the study presented below is to set up and assess the research procedure for calculating the Polish filmic AD reading rate. Below we discuss the corpora used in the study and methods applied in order to calculate the reading rate: time-alignment procedure, and calculation of word and pause duration. This is followed by results and their discussion.

6.   Materials

In this study we used a corpus consisting of three ADs (scripts and recordings) prepared for Polish fiction films: *Ida* (dir. P. Pawlikowski, 2013, drama), *Yuma* (dir. P. Mularuk, 2012, action) and *Day of the Wacko* (dir. M. Koterski, 2002, comedy).

The materials were provided by Polish NGOs[4] and were prepared prior to and independently from this study. All ADs were drafted and voiced in Polish and were written by professional describers. They were voiced and recorded by professional voice talents in recording studios. Since we used real life materials, every description was drafted by a different describer and voiced by a different voice talent: *Ida* and *Day of the Wacko* by male voice talents and *Yuma* by a female voice talent.

7. METHODS

In order to compute the reading rate, we needed to calculate the duration of words and pauses as well as to distinguish between breath and non-breath pauses as the latter are not included in the reading rate. Below we present the methodological procedures.

7.1. TIME-SEGMENTATION PROCEDURE

In order to measure the reading rate, it is necessary to compute the duration of each word and each pause. The data can be obtained through a type of annotation also known as time-segmentation. It can be performed either manually or through automatic time alignment. In our study, time alignment was conducted automatically using SARMATA 2.0 Automatic Speech Recognition system (Ziółko *et al.*, 2015).

In order to carry out the procedure, the following steps were performed. First, the original AD scripts (see Example 8) were manually cleared of time codes, punctuation marks, capital letters and indications for the voice talents such as pause signposting or vocalisation tips. Numbers were transcribed as numerals. Foreign and phonetically ambiguous words were transcribed orthographically.

[00:24]	Pomieszczenie przyklasztorne.
	Anna stoi przy dużej figurze Chrystusa. W skupieniu pokrywa ją farbą.
	Anna to młoda, wątła nowicjuszka o niewinnym wyrazie twarzy. Jest
	ubrana w habit, na nim ma sweter, jej włosy przykrywa zakonny welon.
[00:43]	(na poprzedniej scenie) Cztery nowicjuszki przechodzą przez kurnik.
	Niosą rzeźbę Chrystusa. KURY
[00:55]	Nowicjuszki wychodzą na zewnątrz. DRZWI
[01:02]	Płatki śniegu wirują w powietrzu.

Example 8 AD script before cleaning.

---

4   Seventh Sense Foundation, Foundation for Audio Description Progress "Katarynka" and Fundacja Kultury Bez Barier (Foundation for Culture Without Barriers).

Then, the scripts were checked against the audio files as we discovered that discrepancies between the scripts and the audio files resulting from changes made during the recording or mastering were a common issue. After that, the audio files were cut into clips of approximately 20 minutes of length. This was necessary as SARMATA 2.0 was unable to process longer clips.

The AD scripts were cut accordingly and saved as text files (see Example 9).

```
pomieszczenie przyklasztorne anna stoi przy dużej
figurze chrystusa w skupieniu pokrywa ją farbą anna
to młoda wątła nowicjuszka o niewinnym wyrazie
twarzy jest ubrana w habit na nim ma sweter jej
włosy przykrywa zakonny welon cztery nowicjuszki
przechodzą przez kurnik niosą rzeźbę chrystusa
nowicjuszki wychodzą na zewnątrz w powietrzu wirują
płatki śniegu
```

Example 9 AD script prepared for procedure.

Finally, the AD scripts were fed into SARMATA 2.0 as the only possible output regarding speech content. Then, the system added time alignments to the scripts by matching the recognition hypothesis from the audio to the scripts.

7.2. CALCULATION OF WORD DURATION AND PAUSES

The time alignment procedure produced a segmentation of the audio into individual word segments with assigned time where each word begins and ends. This segmentation was generated in CSV files. The data from the CSV files of each film was imported into dedicated Excel sheets.

Given the information on each word's in and out time, we computed the duration of the words and pauses between them. Following that, the words were grouped into insertions. Then, the data was manually annotated in order to correlate the pauses detected in the audio files with commas, full stops and pause signposting in the AD scripts (pause symbols and/or new lines of text). On the one hand, not all pauses marked in the scripts were rendered in the audio files. On the other hand, not all pauses detected in the audio files could be correlated with punctuation marks or signposting. Those pauses were assigned to a category we called *other pauses*.

The main aim of our study was to calculate the AD reading rate understood as the rate at which recorded AD is presented. This definition of the AD reading rate includes BPs but excludes UFPs imposed by interpretation, important sounds or synchronization with images. This is why – after calculating pause duration and correlating them with commas, full stops and pause signposting - we proceeded to distil breath pauses from all the other pauses in the corpus.

In speech tempo studies, it is relatively easy to annotate BPs in a recording since they are usually accompanied by an audible sound of inhalation. Unfortunately, in AD recordings – due to the peculiarity of these recordings (see 4.2 for details) – all types of pauses render as silent. The unique nature of AD recordings guided the development of a new analytic procedure that would enable us to calculate the duration of BPs. The duration of all segmental phenomena in speech, including phonemes or pauses, should be modelled with a gamma-shaped or lognormal rather than normal distribution (Rosen, 2005; Wang *et al.*, 1994). Therefore, the corresponding expected values should be used as proper measures to describe the phenomena instead of mean and standard deviation. In order to investigate the distribution of punctuation-related pauses duration, histograms were constructed for each film separately (Figure 1). For selected parts of the histograms, the specific lognormal shape can be observed (marked with the red line). Breath pauses probably occurred in these regions. Pauses much longer than the average breath could probably mark moments when a breath was accompanied by another activity, for example pauses made deliberately to consult the image or to synchronize with it.
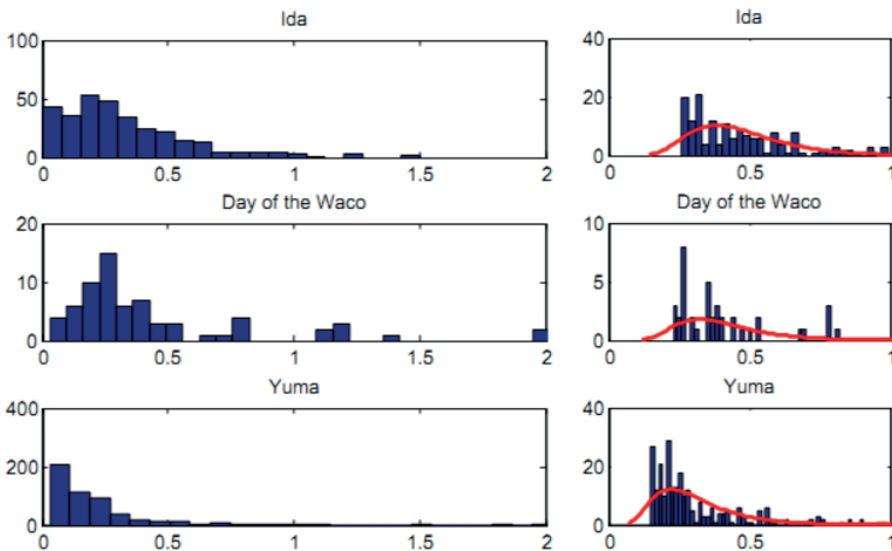

Figure 1 Distribution of full stop related pauses and their regions demonstrating a lognormal distribution.

After visual inspection of the histograms of pauses, we adopted a threshold value for each of the analysed films (see Table 3 below). The pauses longer than the values given in Table 3 below were considered non-breath pauses and the remaining ones were excluded from the analysis.

Ida (drama)	Day of the Wacko (comedy)	Yuma (action)
>0,25	>0,20	>0,14

Table 4 Length of breath pauses

7.4. CALCULATING READING RATE

Once the word duration was calculated, we proceeded to calculating the reading rate in each insertion. The reading rate was calculated based on the duration of words and breath pauses. We based the calculation on CPS and WPM – the measures used frequently in AVT and also in AD research and practice. Since there is no reliable data on the average word length in Polish, we calculated the average word length in our corpus, based on the number of words and characters without spaces. The average word length was estimated at 5 characters per word.

8. RESULTS

The overall results for the three films were analysed. The data corresponding to the maximum, minimum and average reading rate is presented in Table 5 below.

Film	Average reading rate		Maximum reading rate		Minimum reading rate	
	WPM	CPS	WPM	CPS	WPM	CPS
Ida	167	13.95	221	18	55	5
Day of the Wacko	182	15.17	221	18	97	8
Yuma	189	15.75	264	22	44	4

Table 5 AD reading rate

The average AD reading rate was calculated at 13.95 CPS for *Ida* (drama), 15.7 CPS for *Day of the Wacko* (comedy) and 15.75 CPS for *Yuma* (action). The maximum reading rate observed within insertions varied from 18 CPS in *Ida* and *Day of the Wacko* to 22 CPS in *Yuma*. The minimum reading rate within insertions was computed at 5 CPS for *Ida*, 8 CPS for *Day of the Wacko* and 4 CPS for *Yuma*.

When it comes to pauses, the results of the present experiment show that the expected value of a BP in our corpus is ~ 190 ms. Another finding is that the BPs constitute almost 60% of all inner-insertion pauses. The rest are silent pauses. Silent pauses could be correlated either with breaths and/or – given the way AD is recorded – pauses made deliberately by the voice talents to consult the image and synchronize with it. Another possibility is that silent pauses are introduced during the audio-editing stage. This can occur when AD does not fill the entire space available between the dialogues. In such cases it is a common practice to cut the recording into smaller chunks and redistribute them along the timeline for better synchronisation with the image and/or the soundtrack.

Finally, while investigating the correlation between the types of pauses and punctuation marks in AD scripts, we found out that 93.5% of breaths occur in the place of punctuation marks (91.5% for full stops and 2% for commas). Among all pauses correlated with full stops almost 60% are breath pauses. The rest are silent pauses. In the places where commas come up, breath pauses can be observed in 75% of cases. And in the places where signposting is used - in 72% of cases.

9.   DISCUSSION

The results of our study show that the average reading rate in our corpus exceeds the 150-160 WPM – the value currently advised for the AD reading rate – by 11-26%. However, the reading rate is highly dependent on the speaker and the language and what is valid for Polish might not apply to other languages. Another thing to consider is how this higher than the advised reading rate influences understanding and enjoyment, since Cabeza-Cáceres found out that although the reading rate does not influence enjoyment, it can influence comprehension (Cabeza-Cáceres, 2013).

The results of our research also show that the average AD reading rate changes depending on the film genre. This, however, requires further investigation and a larger corpus.

The expected value of breath pauses in our corpus (~190 ms) was significantly shorter than that in spontaneous and read speech. The previous research in this area showed that the expected value of breath pauses in spontaneous and read speech was ~400 ms (Igras & Ziółko, 2013). We hypothesize that the low value of these pauses could be attributed to changes performed by the sound engineer (e.g. cutting the audio track and moving selected segments to better match the action on the screen).

When it comes to the correlation between pauses and punctuation marks, our results also differed from spontaneous and read speech. In the previous research it was found out that although 90% of breaths occur in the places where punctuation marks come up, 50% can be attributed to full stops and 40% to commas. Among all full stops, almost 60% are correlated with breath pauses and the

rest of them with silent pauses. For commas, breath pauses can be observed only in 30% of cases (Igras-Cybulska *et al.*, 2016). This shows that the characteristics of Polish recorded filmic AD are different than those of read and spontaneous speech. Once again, the peculiarities of recorded AD can be due to the process in which they are created and require further studies.

For the purpose of this study we have used WPM and CPS – the measures most commonly used when referring to the AD reading rate. Both measures seem to be adequate to calculate the reading rate of the existing recordings, but they are not reliable measures when it comes to the estimated time needed to read a given text. From the technological point of view, in order to estimate the time needed, the following steps should be taken: 1) translating the orthographic notation into phonetic transcriptions, 2) matching the estimated time duration of each phoneme on the basis of statistical values for the given language, 3) calculating the duration of the utterance, taking into account both segmental and suprasegmental impact, e.g. preboundary lengthening (Igras & Ziółko, 2014), on the properties of the phones. For the needs of comparative analysis, both CPS and PPS (phonemes per second) could be accepted as equally informative measures, since the inventory structure of Polish phonemes is rather constant and independent from spoken or written language (cf. the frequency of phoneme appearance rankings based on the results obtained by Jassem, Łobacz, Rocławski, and Steffen (Sambor, 2004). Another measure of the AD reading rate could be the number of syllables per second. However, this approach would require a language-specific algorithm to automatically divide a written text into acoustic syllables.

An important limitation of this study is the fact that it was carried out on real life material – i.e. mastered ADs where all pauses were silent. This means that all of the results presented and discussed above are an estimation based on histograms created for the duration of punctuation-related pauses rather than on observation.

Another limitation of this study is the fact that its findings – in terms of reading rate and pause duration – can only be applied to Polish, since every language has different characteristics. It is also important to draw attention to the fact that, although AD speech rate is largely conditioned by the script as well as the film, the speech rate of each voice talent varies as it is part of an individual style. Additionally, the final shape of the recording is coined during the mastering process. As a consequence, a larger sample or a sample chosen from different voice talents or languages may bring about variations in the findings.

Conclusions

The present study has hopefully contributed to gaining some basic knowledge about the AD reading rate and the methodological issues in this area of research. The main findings of this study include validating research design for testing

AD reading rate and calculating the value of the average reading rate and breath pause for Polish AD. However, since it was a pilot study, all findings should be confirmed on a larger corpus.

Little previous research has been conducted on reading rate in AD, so we hope that this study inspires future research avenues in this area of AD research. Future studies could look into testing a more reliable measure of calculating the AD reading rate than CPS or WPM such as syllables per second or phonemes per second. It would also be interesting to verify whether AD perception, presence and engagement could be affected by the reading rate and how this related to AD quality. Establishing an AD audio profile (i.e. reading rate, pauses and the relationship between them) in comparison with other types of recorded speech is yet another area of research to be pursued in the future.

Bibliography

Adamowicz-Grzyb G. (2013) *Tłumaczenia filmowe w praktyce,* Warszawa: Fortima.

Ballester A. (2007) "La audiodescripción: Apuntes sobre el estado de la cuestión y las perspectivas de investigación", *TradTerm,* (13), pp.151-169.

Cabeza-Cáceres C. (2013) "Audiodescripció i recepció : efecte de la velocitat de narració, l'entonació i l'explicitació en la comprensió fílmica", Universitat Autònoma de Barcelona, https://www.academia.edu/23388431/Audiodescripció__i__recepció__efecte__de__la__velocitat__de__narració__lentonació__i__lexplicitació__en__la__comprensió__fílmica (last accessed 26/05/2017).

Chaume Varela F. (2012) *Audiovisual Translation: Dubbing,* Manchester: St Jerome.

Chmiel, A. & Mazur, I. (2014). *Audiodeskrypcja,* Poznań: Wydział Anglistyki im. Adama Mickiewicza w Poznaniu, https://repozytorium.amu.edu.pl/bitstream/10593/12861/1/Audiodeskrypcja__Chmiel__Mazur__2014.pdf (last accessed 20/05/2017).

DCMP Description Key (n.d.). "DCMP Description Key", http://descriptionkey.org/about__d.html#1 (last accessed 05/05/2017).

Díaz-Cintas J. (2006) "Competencias profesionales del subtitulador y el audiodescriptor", http://www.cesya.es/sites/default/files/documentos/informe__formacion.pdf (last accessed 21/04/2017).

Federation suisse des aveugles et malvoyants (2017) "Charte suisse de l'audiodescription", http://www.sbv-fsa.ch/sites/default/files/sbv/news/downloads/SBV__Schweizer__Charta__der__Audiodeskription__F__1.pdf (last accessed 26/05/2017).

Fels D. I. & Udo J. P. (2009) "Re-fashioning fashion: an exploratory study of live audio described fashion show" (No. 17), http://digitalcommons.ryerson.ca/trsitm/17 (last accessed 28/05/2017).

Fernández-Torné A. & Matamala A. (2015) "Text-to-speech vs. human voiced audio descriptions: a reception study in films dubbed into Catalan", *The Journal of Specialised Translation,* 24, pp. 61-88.

Fryer L. (2010) "Audio description as audio drama – a practitioner's point of view", *Perspectives,* 18 (3), pp. 205-213.

Fryer L. (2016) *An introduction to Audio Description A Practical Guide.* London and New York: Routledge.

Fryer L., & Freeman J. (2014) "Can you feel what I'm saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment", in *Proceedings of the International Society for Presence Research,* pp. 99-107.

Gonant F., & Morisset L. (2008) "La Charte de Qualité de l'Audiodescription", http://www.enaparte.org/audiodescription/pdf-doc/Charte-de-laudiodescription.pdf (last accessed 10/04/2017).

Igras-Cybulska M. Ziółko B. Żelasko P. & Witkowski M. (2016) "Structure of pauses in speech in the context of speaker verification and classification of speech type", *EURASIP Journal on Audio, Speech, and Music Processing,* 1:18, https://link.springer.com/article/10.1186/s13636-016-0096-7 (last accessed 21/04/2017).

Igras M., & Ziółko B. (2013) "Wavelet method for breath detection in audio signals", in *IEEE International Conference on Multimedia and Expo,* San Jose, CA, USA, pp 1-6.

Igras M. & Ziółko B. (2014) "The Role of Acoustic Features in

Marking Accent and Delimiting Sentence Boundaries in Spoken Polish", *Acta Physica Polonica*, 126(6), pp. 1246-1257.

Independent Television Comission (2000) "ITC Guidance On Standards for Audio Description", http://audiodescription.co.uk/uploads/general/itcguide__sds__audio__desc__word3.pdf (last accessed 05/04/2017).

Jacewicz E. Fox R. A. O'Neill C. & Salmons J. (2009) "Articulation rate across dialect, age, and gender", *Language Variation and Change*, 21(2), pp. 233–256.

Jacewicz E. Fox R. A. & Wei L. (2010) "Between-speaker and within-speaker variation in speech tempo of American English", *The Journal of the Acoustical Society of America*, 128(2), pp.839-850.

Jankowska A. & Walczak A. (forthcoming). Filmic audio description in Poland - present state and future challenges.

Linder G. & Martínez J. (2011) "Audio Description with a TTS application", Barcelona: Advanced Research Seminar on Audio Description, http://grupsderecerca.uab.cat/arsad/content/2011-seminar-0 (last accessed 18/08/2017).

McGonigle F. (2007) *The Audio Description of Children's Films*. University of Surrey.

Netflix (no date) "Audio description style guide", https://backlothelp.netflix.com/hc/en-us/articles/215510667-Audio-Description-Style-Guide-v1-0 (last accessed 18/08/2017).

Piety P. J. (2004) "The Language System of Audio Description: An Investigation as a Discursive Process -Audio Description - The Language System of Audio Description: An Investigation as a Discursive Process", *Journal of Visual Impairment and Blindness*, 98(8), pp. 453–469.

Remael A. (2012) "Audio description with audio subtitling for Dutch multilingual films: Manipulating textual cohesion on different levels", *Meta*, 57(2), pp.385-407.

Remael A. Reviers N., & Vercauteren G. (2014) "ADLAB Audio Description guideline", http://www.adlabproject.eu/Docs/adlab book/index.html (last accessed 01/04/2017).

Rodero E. (2012) "A comparative analysis of speech rate and perception in radio bulletins", *An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 32(3), pp.391-411.

Rosen K. M. (2005) "Analysis of speech segment duration with the lognormal distribution: A basis for unification and comparison", *Journal of Phonetics*, 33(4), pp.411-426.

Sambor J. (2004) *Współczesna Polszczyzna*, Lublin: Wydawnictwo UMCS.

Snyder J. (2010) *Audio Description Guidelines and Best Practices*, https://webcache.googleusercontent.com/search?q=cache:KMjSjhJa2vkJ:https://www.acb.org/adp/docs/AD-ACB-ADP%2520Guidelines%25203.1.doc+&cd=2&hl=ple&ct=clnk&gl=pl (last accessed 18/08/2017).

Szarkowska A. & Jankowska,A. (2012) "Text-to-speech audio description of voiced-over films. A case study of audio described Volver in Polish", in *Emerging topics in translation: Audio description*, Dir. par E. Perego, Trieste, EUT Edizioni Università di Trieste, pp. 81-98.

Szarkowska A. & Jankowska A. (2015) "Audio describing foreign films", *The Journal of Specialised Translation*, 23, pp.243-269.

Szarkowska A. & Orero P. (2014) "The importance of sound for audio description", in *Audio Description: New perspectives illustrated*. Dir. par A. Maszerowska, A. Matamala, & P. Orero, Amsterdam/Philadelphia, Benjamins, 121-140.

Szymańska B. & Strzymiński T. (2010) „*Standardy tworzenia audiodeskrypcji doprodukcji audiowizualnych*", http://avt.ils.uw.edu.pl/files/2010/12/AD-_standardy__tworzenia.pdf (last accessed 01/06/2017).

Thabane L. Ma J. Chu R. Cheng J. Ismaila A. Rios, L. P. Reid R. Marron T. Giangregorio L. Goldsmith, C. H. (2010) "A tutorial on pilot studies: the what, why and how", *BMC Medical Research Methodology*, 10(1), https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-10-1 (last accessed 01/04/2017).

Trouvain J. (2003) "Tempo variation in speech production implications for speech synthesis", http://www.coli.uni-saarland.de/groups/WB/Phonetics/contents/phonus-pdf/phonus8/Trouvain__Phonus8.pdf (last accessed 01/05/2017).

Trouvain J. Fauth C. & Möbius B. (2016) "Breath and Non-breath Pauses in Fluent and Disfluent Phases of German and French L1 and L2 Read Speech", *Speech Prosody*, http://www.ifcasl.org/docs/Trouvain__Fauth__Moebius__2016.pdf (last accessed 01/03/2017).

Trzeciakiewicz M. (Ed.). (2014) „Audiodeskrypcja w teorii i praktyce, czyli jak mówić o tym, czego nie można zobaczyć", http://avt.ils.uw.edu.pl/files/2014/01/Szarkowska__Kuenstler__AD-w-kinie-teatrze-i-muzeum.pdf (last accessed 18/08/2017).

Udo J. P., Acevedo B., & Fels D. I. (2010) "Horatio audio-describes Shakespeare's Hamlet", *British Journal of Visual Impairment*, 28(2), pp.139-156.

Udo J. P. & Fels D. I. (2010) "Universal design on stage: live audio description for theatrical

performances", *Perspectives*, 18(3), pp.189-203.

van der Heijden M. (2009a) "AD in The Netherlands: about sound-quality and internet accessibility", Barcelona: Advanced Research Seminar on Audio Description, http://grupsderecerca.uab.cat/arsad/content/2009-program (last accessed 18/08/2017).

van der Heijden M. (2009b) "How to combine subtitles with audiodescription in multilingual feature films?", Barcelona: Advanced Research Seminar on Audio Description, http://grupsderecerca.uab.cat/arsad/content/2009-program (last accessed 18/08/2017).

van der Heijden M. (2011) "Mixing and delivering AD for 3D films", Barcelona: Advanced Research Seminar on Audio Description, http://grupsderecerca.uab.cat/arsad/content/2011-seminar-0 (last accessed 18/08/2017).

Vercauteren G. (2016) "Enough is enough! But... just how much is enough?", Łódź: Intermedia Conference on Audiovisual Translation, http://intermedia.uni.lodz.pl/wp-content/uploads/2013/09/Intermedia-2016-programme.pdf (last accessed 18/08/2017).

Walczak A. (2011) "Text-to-speech audio description for visually impaired children", Barcelona: Advanced Research Seminar on Audio Description, http://grupsderecerca.uab.cat/arsad/content/2011-seminar-0 (last accessed 18/08/2017).

Wang X. Pols L. W. C. & ten Bosh L. F. M. (1994) "Analysis of context-dependent segmental duration for automatic speech rcognition", in *Proceedings ICSLP '94*, http://www.fon.hum.uva.nl/IFA-publications/ICSLP96/Xue_Wang/a461.html (last accessed 01/05/2017).

Ziółko B. Jadczyk T. Skurzok D. Żelasko P. Gałka J. Pędzimąż T. Gawlik I. Pałka, S. (2015) „SARMATA 2 . 0 Automatic Polish Language Speech Recognition System" https://www.researchgate.net/profile/Piotr_Zelasko/publication/281774738_SARMATA_20_Automatic_Polish_Language_Speech_Recognition_System/links/55f7d7b908aec948c4745dab.pdf (last accessed 01/03/2017).

Żórawska A. Więckowski R. Künstler I. & Butkowska U. (2011) „Audiodeskrypcja – zasady tworzenia", http://dzieciom.pl/wp-content/uploads/2012/09/Audiodeskrypcja-zasady-tworzenia.pdf (last accessed 01/04/2017).

FILMOGRAPHY

Ida (2013) dir. P. Pawlikowski, Poland.

Dzień Świra (Day of the Wacko) (2002) dir. M. Koterski, Poland.

Yuma (2012) dir. P. Mularuk, Poland.