JACEK JĘDRYCZKOWSKI [iD]

Didactic Film in Computer-Aided Education of People with Visual and Hearing Impairments

ORCID: 0000-0001-6707-263X, Ph.D., University of Zielona Góra, Faculty of Social Sciences, Institute of Pedagogy, Department of Media and Information Technology, Poland

Abstract

Artificial intelligence mechanisms enable fast speech transcription and printed text recognition (Optical Character Recognition). Speech synthesizers for reading any digital documents have also become popular. In most cases, these are free solutions that can significantly support the education and functioning of people with visual and hearing impairments. Unfortunately, the knowledge of parents and caregivers on this subject is insufficient. Bearing this in mind, a suitable series of educational videos was prepared and published on YouTube. Subsequently, research was conducted to determine whether and to what extent there is a demand for video materials on this topic. For this purpose, the method of implicit observation was used, carried out using YouTube Analytics mechanisms. Its results are available in the form of numerous reports and summaries. Based on the collected materials, a considerable interest in the study was found. It was possible to determine which issues are of particular interest to the recipients. The obtained results will allow for further development of the educational offer.

Keywords: didactic film, YouTube Analytics, e-learning, educational process, deaf people, blind people

Introduction

The dynamic development of artificial intelligence tools has made it possible to improve speech recognition and transcription solutions in a way that was not possible before. This is possible both for dictating to a computer and for processing existing audio-video recordings. Similarly, free tools for recognizing text directly from photos or scans and programs that allow the computer to read it aloud have developed. Unfortunately, few teachers, parents and caregivers have the IT skills to use these solutions. Bearing in mind the need to popularize this type of information, the author of the article prepared and posted a series of

educational videos on YouTube. The aim of the study was to determine whether there is a demand for this type of topic and to find out which issues interest the recipients. The information on this was obtained through YouTube Analytics reports.

The subject matter of the study

The article discusses the procedure and results of the evaluation of the didactic usefulness of the author's video materials. The subjects of the study are the recipients of educational videos who help people with visual and hearing impairments. It was assumed that these are teachers preparing teaching aids and parents and caregivers supporting their wards in the learning process and in using cultural goods available in electronic form. The condition for conducting the research was to develop a series of videos and publish them on a playlist on YouTube. This solution made it possible to evaluate the viewership of the published materials and the covert observation of the viewers' interaction with the video material (watching time of individual parts; scrolling and re-watching and pausing). Based on the research material collected in the form of YouTube Analytics reports, it can be stated which parts of the material aroused the most interest. The conclusions from the research allow for modifying educational activities and improving or modifying the films made.

Research methodologies and tools

The ways of using the videos were analyzed using the research tools offered by YouTube Analytics. This is a reporting system that provides information on the viewership and all interactions of the recipients with the published video materials. YouTube Analytics does not allow assigning the observational material to a specific viewer. However, the quantitative nature of this type of research is their advantage. If the research problem is general, allowing for the development of video material on a topic of interest to a wider audience, then the research sample can include thousands of people. This makes it possible to generalize the results with a high degree of probability. Łobocki (2000, p. 46; 1999, pp. 48–51) defines observation as a method of scientific research and distinguishes its types: direct and indirect; standardized and non-standardized; overt and covert. YouTube Analytics reports can be considered as an equivalent of documentation prepared during indirect observation. However, it should be noted that the monitoring system operates covertly as in direct observation.

When planning research using YouTube Analytics tools, you can specify the variables very precisely. We then have a typical standardized observation. However, the ability to combine the data collected by YouTube Analytics with the observed variable means that their review can lead to a number of unexpected conclusions, similar to the case of non-standardized observation.

YouTube Analytics reports provide information, among others, on: number of views, watch time, average watch time; average percentage viewed and traffic sources, i.e. the ways in which Internet users found a given video. All data can be assigned to a specific time period and to the source of visits. The most valuable from the point of view of analyzing the learning process are the audience retention reports. They inform about the interactions of the recipients with the video material. These are two charts: "Absolute audience retention" and "Relative audience retention". In both cases, the tool allows you to simultaneously track the content of the video and the vertical axis moving along the chart. This allows you to correlate specific content with the range of interactions.

The parameter defined as "Absolute audience retention" informs how many percent of the total number of views of the video are the playback of individual parts. This means that values greater than 100% may appear. If the chart is flat and indicates a low level of engagement of the recipients, you should use the second chart. It informs about the level of attention determined on the basis of the activity of the recipients of a given video in comparison with all videos from YouTube of a similar length. This solution allows you to determine whether there were content that engaged the attention of the recipients. The cognitive aim of the research was to determine the scale of interest in educational materials illustrating the use of information and communication technologies supporting the education process of people with visual and hearing impairments. The contents that aroused the most interest of the recipients were also sought, taking into account the practical purpose, which was to develop the concept of subsequent videos.

The formulated research problems were questions about: number of views; watch time; average watch time; average percentage viewed and parts of the videos that aroused the most interest. The researcher was also interested in the sources of visits corresponding to the highest viewership.

Educational YouTube channel in the education process

Based on the research on the educational impact of multimedia messages and the research on the difficulties related to understanding laboratory instructions, on June 7, 2015, the author of the study launched the YouTube channel "JJ Kursy" offering educational videos related to broadly understood information and communication technologies (Jędryczkowski, 2015). The "JJ Kursy" channel currently offers 257 original videos, supplementing the content of laboratory classes and lectures. These videos are also available on websites dedicated to individual laboratory groups at the University of Zielona Góra. This solution allows for collecting detailed statistics informing about the learning process. To date, the video materials have been viewed 1 465 901 times and watched for 50 825 hours (5 years and 292 days). Only 21 468 views (2.9%) correspond to

the students of the university, which corresponds to 1615 hours of watching (over 67 days). This means that the YouTube channel currently serves as a Massive Open Online Course (MOOC), which is used by Internet users in the country and abroad (Jędryczkowski, 2019a, pp. 167–180).

Results of analyzes of the possibilities of using didactic films to support the education of people with visual and hearing impairments

The educational videos were published on the "JJ Kursy" channel on YouTube on the playlist "Supporting the education of people with visual and hearing impairments" (Jędryczkowski, 2023). The data on the viewership are from November 23, 2023. The following materials were posted:

1. Converting speech to text in Windows 10 and 11.
2. How to use free cloud-based office suites?
3. Clipchamp – automatic creation of subtitles for videos for the deaf.
4. Converting text to speech in Windows.
5. OCR in Google Docs. Recognizing text from photos and locked PDF files.

The first three videos present solutions useful in the education process of people with hearing impairments. The mechanisms that convert speech to text in real time and the programs for transcription, obtaining text from audio and video files, including for automatic generation of subtitles for videos, are discussed. The fourth and fifth videos focus on tools designed for people with visual impairments. Numerous solutions that convert text to speech were discussed. Free Google Docs mechanisms for recognizing text from photos and scans were also analyzed. This information may be useful for teachers and parents preparing educational materials for blind people. The text obtained in this way can be automatically converted to speech.

The video "Converting speech to text in Windows 10 and 11" was published on December 14, 2023. The material lasting 06:33 min was viewed 4234 times and watched for 133.8 hours. The average watch time of the video is 1:53 min, and the average percentage viewed is 28.9%. The video received 29 likes and contributed to gaining 17 subscribers. The most popular sources of visits are: Google search (2250 views with 71.1 hours of watching) and YouTube search (823 views with 24.3 hours of watching).

Analyzing the absolute audience retention of the recipients (Figure 1), it was found that the most interest was aroused by the fragment from 0:16 to 0:30. It was repeatedly scrolled to re-analyze, which made it longer watched than the rest of the content. The highest value of audience retention in this case reached 96.5%. It was found that at this point the keyboard shortcut for dictating in Windows 11 is discussed. On this chart, it is difficult to indicate other fragments that attract attention, so the values of the relative audience retention (Figure 2) were

analyzed. A value above the average was observed in the ranges from: 00:00 to 01:08 and from 05:06 to 05:26. The first fragment is the full procedure for configuring dictation in Windows 11, and the second is the way of dictating to Google Docs. It should be noted that almost all the time (except for the last 20 seconds) the level of attention does not fall below the average value, which means the involvement of the viewers expressed by numerous interactions.
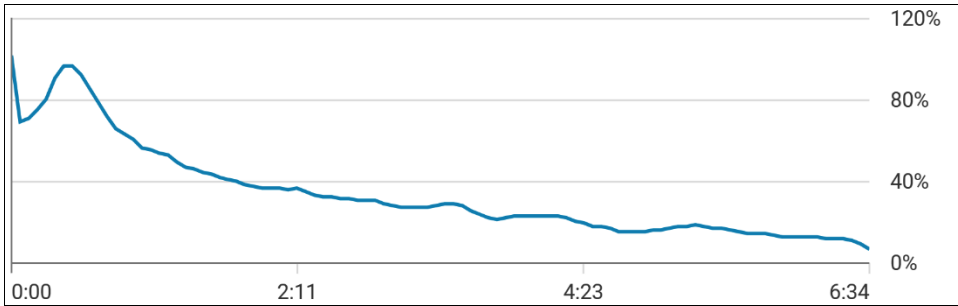

Figure 1. Absolute audience retention while watching the video "Converting speech to text in Windows 10 and 11"
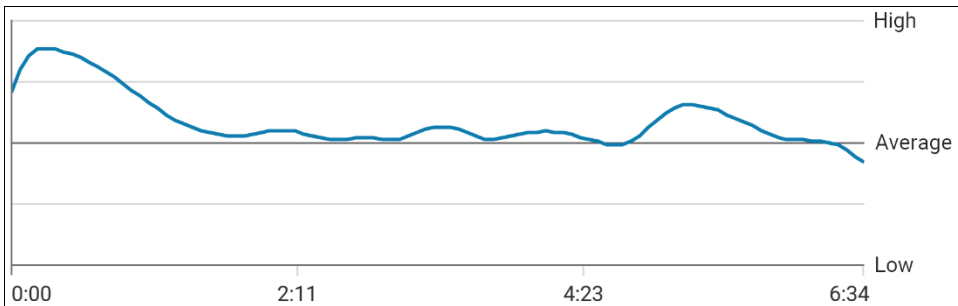
Fogure 2. Relative audience retention while watching the video "Converting speech to text in Windows 10 and 11"

The video "How to use free cloud-based office suites?" was published on August 25, 2023. The material lasting 08:55 min was viewed 122 times and watched for 6.1 hours. In this case, only the part dedicated to text transcription in Microsoft Word 365 (online version), which is in the range from 01:17 to 02:20, was analyzed. The video received 7 likes and did not contribute to gaining subscribers. The most important sources of visits are: playlists (35 views with 1.4 hours of watching); suggested videos (24 views with 1.4 hours of watching) and browsing features (23 views with 0.8 hours of watching). Based on the analysis of the absolute audience retention referring to the analyzed fragment, attention engagement at the level of 52.6% was found (Figure 3). However, in the case of

the relative indicator, values above the average were observed, the highest in relation to the rest of the video (Figure 4). This suggests that some viewers watched only this part of the video. This could happen because the link leading to the discussed fragment was also posted in the description directly under the video.
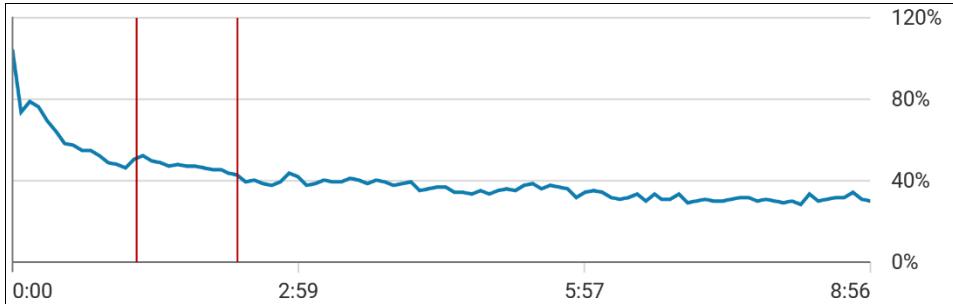

**Figure 3. Absolute audience retention while watching the video
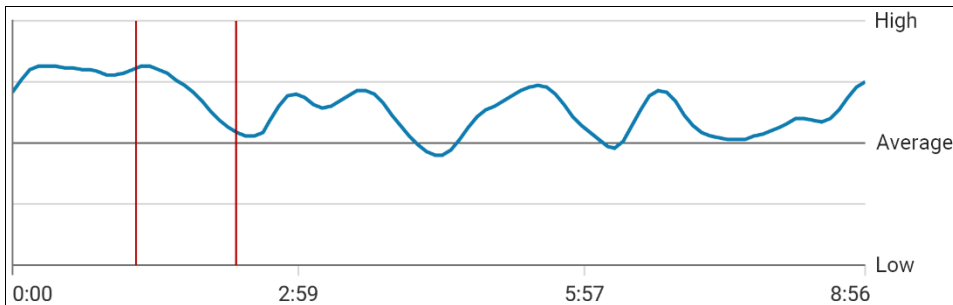"How to use free cloud-based office suites?"**

**Figure 4. Relative audience retention while watching the video
"How to use free cloud-based office suites?"**

The video "Clipchamp – automatic creation of subtitles for videos for the deaf" was published on July 12, 2023. The material lasting 03:03 min was viewed 227 times and watched for 4.7 hours. The average watch time of the video is 1:14 min, and the average percentage viewed is 40.6% (a very high indicator on YouTube). The video received 5 likes and contributed to gaining 1 subscriber. The most popular sources of visits are: Google search (80 views with 1.5 hours of watching) and YouTube search (36 views with 0.8 hours of watching). Analyzing the absolute audience retention of the recipients (Figure 5), it was found that the most interest was aroused by the fragment from 00:00 to 01:24, taking values above 50%. In the case of the relative indicator (Figure 6), values above the average were found in the ranges from 00:17 to 01:40 – laun-

ching the program and automatically generating subtitles (transcription) and from 02:25 to 02:45 – starting to watch the video from its end. This is a cognitive strategy based on scrolling the video to the end and checking what the effect of the actions will be. If it is satisfactory, the viewer returns to the moment where the important procedures are demonstrated, usually skipping the introduction and theory (Jędryczkowski, 2019b, pp. 41–50).
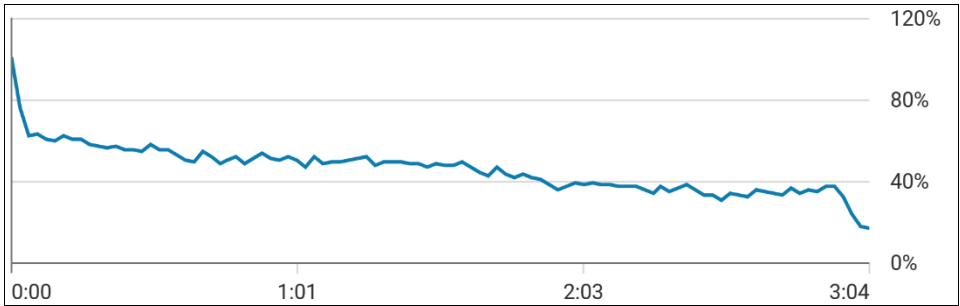

Figure 5. Absolute audience retention while watching the video "Clipchamp – automatic creation of subtitles for videos for the deaf"
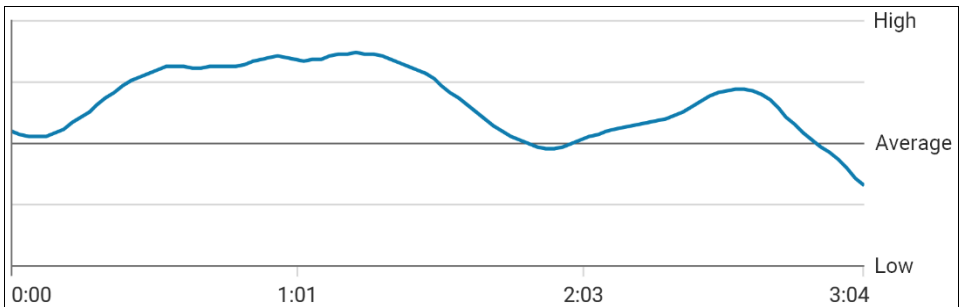

Figure 6. Relative audience retention while watching the video "Clipchamp – automatic creation of subtitles for videos for the deaf"

The video "Converting text to speech in Windows" was published on March 28, 2023. The material lasting 19:19 min was viewed 1526 times and watched for 58.6 hours. The average watch time of the video is 2:17 min, and the average percentage viewed is 11.9%. The video received 18 likes and contributed to gaining 4 subscribers. The most popular sources of visits are: Google search (629 views with 25.8 hours of watching); suggested videos (344 views with 17.2 hours of watching) and YouTube search (258 views with 9.2 hours of watching). The absolute audience retention chart does not clearly show the content that particularly interested the recipients (Figure 7). However, in the range from 00:11 to 01:10, values above 50% were recorded. At this time, the work with the

178

"Narrator" of Windows 11 was discussed. The interest in this topic is also confirmed by the relative audience retention (Figure 8). On this chart, you can also observe that other issues caught the attention of the viewers. An increase in interest, to the average level, was found in the range from 06:50 to 09:46 min. This is a fragment dedicated to e-book readers reading aloud the text of books. The range from 10:33 to 13:28 min corresponds to an increase in interest to average values. Then, the configuration and possibilities of the QTranslate program were discussed. In the range from 11:54 to 17:46, interactions were observed at the average level up to the level above the average. This fragment concerned the ways of installing extensions for loud reading of text in the Google Chrome browser.
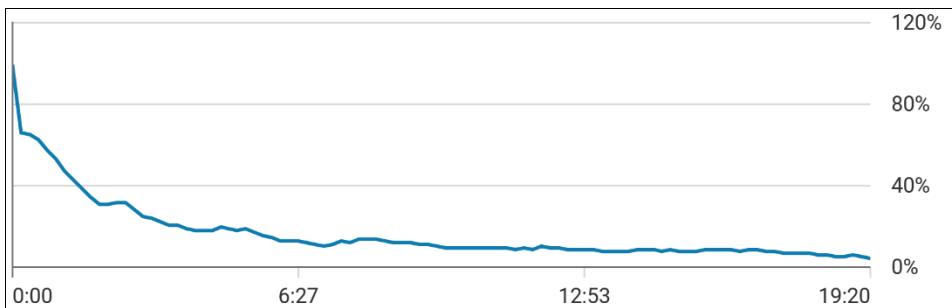

Figure 7. Absolute audience retention while watching the video "Converting text to speech in Windows"
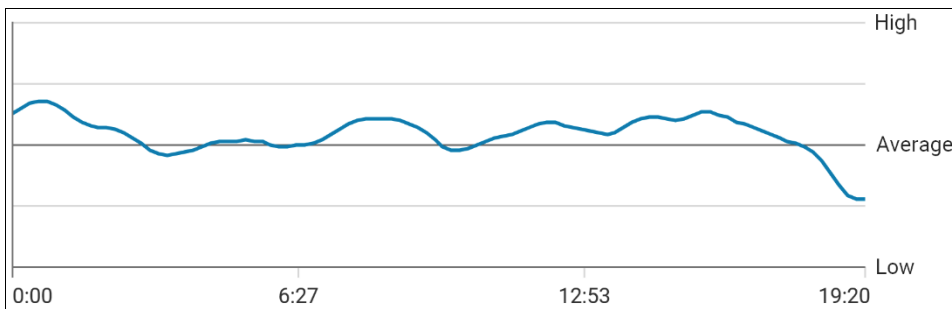

Figure 8. Relative audience retention while watching the video "Converting text to speech in Windows"

The video "OCR in Google Docs. Recognizing text from photos and locked PDF files" was published on November 26, 2022. The material lasting 07:33 min was viewed 1596 times and watched for 65.8 hours. The average watch time of the video is 2:27 min, and the average percentage viewed is 32.4%. The material received 22 likes and contributed to gaining 6 subscribers. The most popular

sources of visits are: Google search (662 views with 25.7 hours of watching) and YouTube search (354 views with 15.7 hours of watching). The absolute audience retention chart showed only one very clear increase in audience retention in the range from 02:50 to 3:20 min (Figure 9). It reached a value of 69.4%. At this time, the procedure for obtaining text from a photo using Google Docs was discussed. The relative audience retention indicator (Figure 10) showed interest above the average in the range from: 00:05 to 03:13 min (photographing and uploading photos to Google Drive and the entire OCR procedure), i.e. obtaining text from an image. Another increase in interest was noted in the range from 05:53 to 6:48 min (the procedure for obtaining text from a secured PDF document).
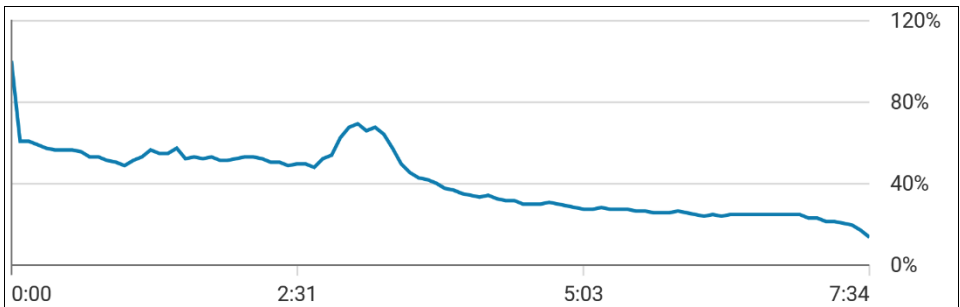

Figure 9. Absolute audience retention while watching the video "OCR in Google Docs. Recognizing text from photos and locked PDF files"
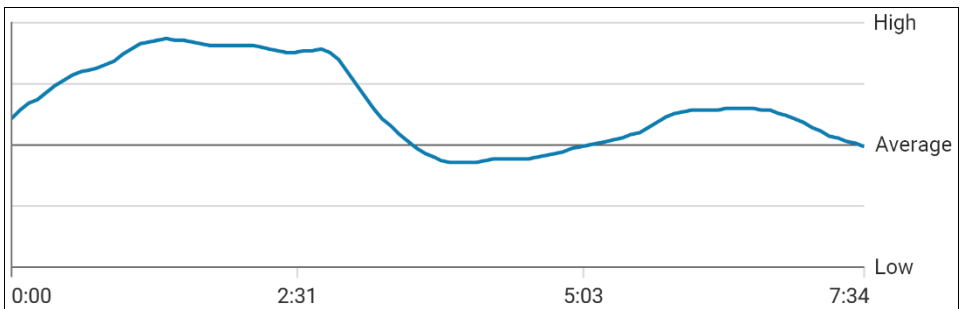

Figure 10. Relative audience retention while watching the video "OCR in Google Docs. Recognizing text from photos and locked PDF files"

Conclusion

The analyzed video materials with a total time of 35 min 25 sec. They were viewed 7705 times and watched for 269 hours. It was found that the videos were used by people interested in specific topics. Such a conclusion could be reached by analyzing the sources of visits. Most often it was searching in Google and

directly on YouTube. This is also confirmed by the likes (81) and the number of subscriptions obtained (28). Content that particularly interested Internet users was also found. These were most often technical solutions that facilitate the conversion of voice to text, transcription of text from audio and video files, and obtaining text from photos and locked PDF documents. The scale of interest in the published videos indicates that there is a demand for this type of content. The ease with which modern digital technologies transform speech into text has made it possible to meet the expectations of people with hearing impairments without incurring financial costs. Adding subtitles to videos becomes not only an expression of care for diversity and accessibility of content, ensuring equal access to culture and education. In the case of educational impacts, subtitles will facilitate full understanding of the content, which in turn translates into efficiency and equalization of educational opportunities (Domagała-Zyśk, 2017, pp. 41–53).

From the point of view of digital typhlopedagogy, speech recognition tools are an indispensable element of functioning of people with visual impairments in modern society (Pawłowska-Cyprysiak, Hildt-Ciupińska, 2022, pp. 18–22). There are a number of free solutions that allow parents and caregivers to independently process paper publications into digital form, and then read them by speech synthesizers. The ability of a blind person to communicate with a computer using speech is not insignificant. Copilot, an artificial intelligence that, in addition to answering very complex questions, is intended to enable computer control, is currently reaching computers with Windows. In a similar way, you can now control devices in the "digital home". Due to the great interest in this topic, it will be included in the implementation of subsequent video materials.

References

Domagała-Zyśk, E. (2017). *Napisy jako forma wspierania edukacji i funkcjonowania społecznego osób niesłyszących i słabosłyszących*. W: W. Otrębski, K. Mariańczyk (esd.) *Przepis na rehabilitację. Dzieci i młodzież z niepełnosprawnością* (pp. 41–53). Gdańsk: Harmonia Universalis.

Jędryczkowski, J. (2015). *JJ Kursy – Kanał YouTube*. Retrieved from: https://www.youtube.com/ @JJKursy (23.11.2023).

Jędryczkowski, J. (2019a). Learning strategies when using online multimedia courses. ICT in Educational Design Processes. *Materials, Resources – KEGA Edition*, *14*, 41–50.

Jędryczkowski, J. (2019b). Masowy otwarty kurs online w procesie samokształcenia. *Lubelski Rocznik Pedagogiczny*, 38(4). doi:10.17951/lrp.2019.38.4.167-180.

Jędryczkowski, J. (2023). *Playlista: Wspomaganie kształcenia osób z wadami wzroku i słuchu*. Retrieved from: https://www.youtube.com/playlist?list=PL_JBpYyA8qn0ZDkj38XUv3Hby Kkasgwei (23.11.2023).

Łobocki, M. (1999). *Wprowadzenie do metodologii badań pedagogicznych*. Kraków: Impuls.

Łobocki, M. (2000). *Metody i techniki badań pedagogicznych*. Kraków: Impuls.

Pawłowska-Cyprysiak, K., Hildt-Ciupińska K. (2022). Nowe technologie i e-kompetencje w ocenie osób z niepełnosprawnością narządu wzroku. *Bezpieczeństwo Pracy – Nauka i Praktyka*, *10*, 18–22.