

Łukasz Dutka

Uniwersytet Warszawski

Wydział Lingwistyki Stosowanej

Instytut Lingwistyki Stosowanej

Zakład Badań nad Przekładem Ustnym i Audiowizualnym

Laboratorium przekładu audiowizualnego AVT Lab

## **Autoreferat rozprawy doktorskiej**

### **Live subtitling with respeaking in Polish: technology, user expectations and quality assessment**

tytuł polski: „Napisy na żywo metodą respeakingu w języku polskim:  
technologia, oczekiwania użytkowników i ocena jakości”

przygotowanej pod kierunkiem dr hab. Agnieszki Szarkowskiej

#### **1. Wybór i znaczenie tematu**

Dostęp do mediów jest ważnym warunkiem pełnego uczestnictwa w życiu społecznym (Greco, 2018; Jankowska, 2020; Neves, 2005). Od wielu lat podejmowane są liczne wysiłki, aby uczynić media takie jak telewizja bardziej dostępnymi dla odbiorców, w tym osób z niepełnosprawnością wzroku czy słuchu. W obszarze dostępności mediów (ang. *media accessibility*) upowszechnił się obecnie szereg usług dostępnościowych (ang. *access services*). Jedną z nich są napisy dla niesłyszących, które ułatwiają osobom głuchym i słabosłyszącym odbiór treści audiowizualnych. Napisy odgrywają obecnie ważną rolę w zagwarantowaniu odbiorcom z niepełnosprawnością słuchu dostępu do edukacji i rozrywki (Greco, Jankowska, 2020; Zarate, 2021).

W wielu krajach liczba programów telewizyjnych dostępnych z napisami dla niesłyszących wzrosła znacząco na przestrzeni ostatnich lat, zarówno jeśli chodzi o treści telewizyjne nagrywane wcześniej, jak i nadawane na żywo, w których to przypadku konieczne jest tworzenie napisów na żywo. W kontekście polskim,

ustawodawstwo unijne i krajowe również doprowadziło do wzrostu liczby programów telewizyjnych opatrzonych napisami, a także do zastosowania napisów na żywo w programach informacyjnych, czego od wielu lat domagali się widzowie z niepełnosprawnością słuchu (Sacha, Kasperkowiak, 2012; Szarkowska, 2010; Szczygielska, 2019). Polskie przepisy dotyczące dostępności mediów nałożyły na nadawców obowiązek, by część programów telewizyjnych zawierała napisy dla niesłyszących. Wymóg ten początkowo dotyczył 10% audycji telewizyjnych. W 2020 roku kwota ta została podniesiona do 25% wszystkich programów telewizyjnych, a od 2022 roku wzrosła do 35%. Do 2024 roku ma wynieść 50% (KRRiT, 2018; Szczygielska, 2019). Od 2020 roku wszyscy trzej główni nadawcy w Polsce (TVP, Polsat i TVN) zapewniają napisy na żywo do niektórych swoich programów informacyjnych i debat politycznych. Są to zarówno napisy w całości przygotowane na żywo, jak i napisy częściowo przygotowane z wyprzedzeniem (tzw. napisy semi-live), które powstają z wykorzystaniem tekstów dostępnych w systemach telewizyjnych, np. tekstów odczytywanych z promptera przez prezenterów serwisów informacyjnych.

Choć coraz więcej programów jest opatrywanych napisami dla niesłyszących, jakość tych napisów nie zawsze jest zadowalająca (Romero Fresco, 2021). Konieczne jest zagwarantowanie wysokiej jakości napisów dla niesłyszących, w tym napisów na żywo, tak by napisy służyły nie tylko spełnieniu przez nadawców wymogów formalnych, ale także zapewniały widzom z niepełnosprawnością słuchu realny dostęp do treści telewizyjnych na żywo i możliwość komfortowego odbioru audycji telewizyjnych. Osiągnięcie tego celu wymaga m.in. badań nad jakością napisów na żywo oraz metodami ich tworzenia.

W związku z powyższym, w prezentowanej rozprawie doktorskiej dokonano przeglądu dostępnych metod, technik i narzędzi umożliwiających tworzenie napisów na żywo w języku polskim, a także przeanalizowano szereg miar jakości napisów na żywo i zastosowano je do oceny zebranego korpusu napisów na żywo w języku polskim w celu oceny aktualnej jakości napisów na żywo w telewizyjnych kanałach informacyjnych w Polsce.

## **2. Przedmiot rozprawy i pytania badawcze**

Celem niniejszej rozprawy było zbadanie jakości napisów na żywo w polskich telewizjach informacyjnych. W związku z tym postawiono następujące pytania badawcze:

(1) Czy dzięki napisom na żywo widzowie głusi i słabosłyszący mają realny dostęp do programów informacyjnych w telewizji; to znaczy, czy napisy dokładnie odzwierciedlają treść i są zrozumiałe?

(2) Czy jakość napisów na żywo różni się w zależności od nadawcy?

(3) Czy jakość napisów na żywo różni się od jakości napisów przygotowywanych częściowo z wyprzedzeniem (tzw. napisów semi-live)?

(4) Czy jakość napisów na żywo rośnie czy maleje wraz z upływem czasu?

Postawiono hipotezę, że jakość napisów semi-live jest wyższa niż jakość napisów na żywo oraz że jakość napisów na żywo wzrasta z czasem, w miarę jak nadawcy nabierają doświadczenia w tworzeniu napisów na żywo.

## **3. Struktura rozprawy**

Rozprawa składa się z sześciu rozdziałów. Rozdział pierwszy przedstawia napisy na żywo w ujęciu teoretyczno-konceptualnym i podsumowuje rozwój przekładu audiowizualnego jako dyscypliny naukowej oraz teoretyczne aspekty dostępności mediów. Analizie poddane zostały relacje pomiędzy przekładem audiowizualnym a dostępnością mediów.

Rozdział drugi przyjmuje perspektywę diachroniczną i opisuje historię rozwoju napisów na żywo jako usługi dostępnościowej. Szczególną uwagę poświęcono początkom i rozwojowi respeakingu jako metody tworzenia napisów, a także przedstawiono historię rozwoju napisów dla niesłyszących w ogólności, a napisów na żywo w Polsce w szczególności, zarówno w przypadku telewizji, wydarzeń na żywo, jak i streamingu w Internecie. Jako że proces upowszechniania się usług dostępnościowych jest silnie związany ze zmieniającymi się wymogami prawnymi oraz działaniami regulacyjnymi, w rozdziale podsumowano uregulowania prawne dotyczące

dostępności mediów w Polsce na tle prawodawstwa obowiązującego w Unii Europejskiej i Stanach Zjednoczonych.

W rozdziale trzecim omówione zostały metody, techniki i narzędzia wykorzystywane do tworzenia napisów na żywo. Poddano je krytycznej analizie pod kątem możliwości ich zastosowania w przypadku tworzenia napisów na żywo w języku polskim. Następnie przedstawiono bardziej szczegółowo metodą respeakingu oraz wykorzystywaną przez tę metodę technologię komputerowego rozpoznawanie mowy, a także jeden z wariantów metody respeakingu: respeaking z równoległą moderacją (Saerens i in., 2020), gdy tekst dyktowany przez respikera jest na żywo korygowany przez moderatora, zanim zostanie wyświetlony widzom.

Rozdział czwarty poświęcony jest odbiorcom napisów i omawia ich potrzeby oraz oczekiwania, a także ich rolę jako jednego z interesariuszy napisów na żywo. Jako że wszyscy odbiorcy zgodnie oczekują wysokiej jakości napisów, w piątym rozdziale omówiono pojęcie jakości w kontekście napisów, a także to, jak można ją skonceptualizować i zmierzyć. Po zidentyfikowaniu i przeanalizowaniu różnych wymiarów jakości napisów na żywo – zarówno lingwistycznych, jak i technicznych – przedstawiono istniejące miary jakości napisów na żywo oraz zaproponowano nowe, niestosowane wcześniej miary. Następnie omówiono różnice między subiektywnym postrzeganiem jakości przez odbiorców a obiektywnymi pomiarami. Rozdział zamyka podsumowanie obecnej wiedzy na temat wpływu różnych sposobów tworzenia napisów na żywo na jakość napisów.

W rozdziale szóstym przedstawiono szczegółowo badanie różnych wymiarów jakości napisów na żywo i napisów przygotowywanych częściowo z wyprzedzeniem (tzw. semi-live) w trzech polskich kanałach informacyjnych: TVP Info, Polsat News i TVN24. Po omówieniu metody, materiałów i procedury, przedstawiono i poddano krytycznej analizie wyniki dotyczące różnych wymiarów jakości. Omówiono także metody statystyczne, które zostały użyte do zbadania danych i potwierdzenia statystycznej istotności wyników. W rozdziale tym zidentyfikowano także główne różnice pomiędzy napisami na żywo dostarczonymi przez trzech polskich nadawców programów informacyjnych, a także omówiono interakcje, która zachodzą pomiędzy różnymi wymiarami jakości.

Rozprawę zamyka podsumowanie wyników badań, a także zestaw zaleceń dla nadawców telewizyjnych oraz Krajowej Rady Radiofonii i Telewizji. W wersji elektronicznej rozprawy dołączono do niej także dwa aneksy, które zawierają dane badawcze analizowane w rozprawie.

#### **4. Metodologia i materiał badawczy**

Jakość napisów na żywo i napisów semi-live przeanalizowano na podstawie próbek z 96 programów telewizyjnych z trzech polskich kanałów informacyjnych (TVP Info, Polsat News i TVN24) pochodzących z trzech kwartałów (II kw. 2021, III kw. 2021 i I kw. 2022). Każda z próbek powstała poprzez nagranie około 10-minutowego fragmentu programu. Łączny czas trwania wszystkich przebadanych próbek wyniósł 970 minut; zawierały one łącznie 13 620 napisów. Próbki zebrano z programów będących przykładami dwóch różnych gatunków audycji telewizyjnych: serwisów informacyjnych (*Wiadomości, Fakty, Wydarzenia*) oraz debat politycznych (*Gość Wiadomości, Kropka nad i, Gość Wydarzeń*).

W badaniu wykorzystano uznane miary jakości napisów na żywo: model NER (jako miarę tego, jak dokładnie napisy oddają treść wypowiedzi) (Romero Fresco, 2021; Romero Fresco, Pöchhacker, 2018), opóźnienie napisów (ang. *latency*), prędkość czytania i stopień redukcji tekstu. Wykorzystano je zgodnie ze standardami ustalonymi w dotychczasowych badaniach (Fresno, 2019; Fresno, Sepielak, 2020; Romero Fresco, 2016; Romero Fresco, Martínez Pérez, 2015; Szarkowska, Barreto Silva, Orrego-Carmona, 2021; Szarkowska, Gerber-Morón, 2018). Zaproponowano też dwa nowe wskaźniki: miarę segmentacji tekstu oraz pomiar odstępów między napisami. Uważa się, że poprawna segmentacja tekstu (podział tekstu na dwulinijkowe napisy oraz właściwy układ tekstu w ramach napisu, w tym podział tekstu pomiędzy liniami) ułatwiają czytanie napisów. Natomiast dłuższe przerwy między następującymi po sobie napisami z zapisem tej samej wypowiedzi mogą utrudniać czytanie i skracają czas na odczytanie napisów. Zaproponowane miary stanowią nowatorski wkład pracy w badania nad napisami na żywo. Dokonano również pomiaru tempa wypowiedzi w programach telewizyjnych. Wszystkie wyniki poddano analizie z wykorzystaniem

metod statystycznych, w tym analizy wariancji, analizy kowariancji, korelacji i liniowych modeli mieszanych.

## 5. Wyniki badania

Badanie wykazało, że jakość napisów na żywo jest dość zróżnicowana, w tym różni się w zależności od nadawcy. Wyniki pokazały, że respeaking z moderacją jest skuteczną metodą tworzenia napisów na żywo w języku polskim; może być stosowany w telewizji oraz pozwala na uzyskanie zadowalającej jakości.

Wyniki pokazują również, że respeaking z moderacją jest skuteczną metodą tworzenia napisów na żywo w języku polskim i może być stosowany w stacjach telewizyjnych w celu uzyskania dobrej jakości, porównywalnej z wynikami uzyskiwanymi w języku angielskim przez nadawców w Wielkiej Brytanii lub USA. (Należy poczynić zastrzeżenie, że wyjątkiem jest tu opóźnienie napisów, które jest zwykle nieco większe przy tej metodzie. Można temu zaradzić poprzez zastosowanie opóźnienia sygnału telewizyjnego, co niweluje opóźnienie napisów). Jednocześnie należy podkreślić, że nie zidentyfikowano żadnej innej metody, która dawałaby lepsze rezultaty w przypadku tworzenia napisów na żywo w języku polskim.

Analiza korpusu napisów pokazuje, że jakość napisów do programów na żywo w polskich stacjach informacyjnych jest dość zróżnicowana. W świetle wszystkich analizowanych wymiarów jakości, programy informacyjne posiadały napisy, które można uznać za doskonałe, bardzo dobre, dobre, złe lub bardzo złe w zależności od nadawcy, rodzaju napisów i kwartału, z którego pochodziły próbki. Jakość napisów na żywo do niektórych programów telewizyjnych była nieakceptowalna. Napisy wymagały od widzów zbyt dużej prędkości czytania, były bardzo opóźnione, niedokładne, a czasami niezrozumiałe.

Badanie potwierdziło, że napisy przygotowywane częściowo z wyprzedzeniem mają lepszą jakość niż napisy przygotowywane w pełni na żywo, co potwierdza zalecenie, by nie tworzyć napisów na żywo w przypadku programów, które nie są w pełni spontaniczne i powstają z wykorzystaniem napisanych wcześniej tekstów, co umożliwia zastosowanie napisów semi-live. Zarówno TVP Info, jak i Polsat News

stosują napisy semi-live z dość dużym powodzeniem i zwykle osiągają bardzo dobrą dokładność, niskie opóźnienie i dobrą segmentację tekstu.

Należy jednak zaznaczyć, że napisy semi-live osiągały w badaniu dość wysokie wartości, jeśli chodzi o prędkość wyświetlania napisów, co może uniemożliwiać widzom odczytanie całości tekstu w dostępnym czasie. Podczas gdy średnia prędkość napisów wynosiła 16,81 znaków na sekundę, aż 36,9% wszystkich napisów semi-live osiągnęło prędkości wyższe niż 18 znaków na sekundę, co wskazuje na potrzebę większej kondensacji (dla porównania prędkość czytania zalecana dla napisów dla niesłyszących w Polsce wynosi 12 znaków na sekundę). Niestety w przypadku wszystkich trzech telewizji informacyjnych jakość zarówno napisów na żywo, jak i napisów semi-live pogorszyła się w badanym okresie. W skali 10-punktowej, w ciągu trzech kwartałów objętych badaniem, dokładność napisów na żywo spadła z akceptowalnego poziomu 5/10 do poziomu nieakceptowalnego 2/10, zarówno w przypadku TVP Info, jak i Polsat News. Stacja TVN24 pozostała poza skalą z wynikiem 0/10. Zwiększyło się opóźnienie napisów, prędkość czytania oraz liczba błędów w podziale tekstu.

Wyniki badania wskazują, że szczególnie problematycznymi obszarami w kwestii jakości napisów na żywo w Polsce są prędkość wyświetlania napisów (w przypadku napisów semi-live) oraz dokładność i opóźnienie (w przypadku napisów na żywo). Należy podkreślić, że jednym z głównych wniosków jest statystycznie istotny i głęboko niepokojący trend obniżania jakości (przede wszystkim spadek dokładności w połączeniu z rosnącym opóźnieniem) obserwowany u wszystkich trzech nadawców w obu rodzajach napisów. Wyniki badania pokazują zatem, że o ile obecne ramy regulacyjne odniosły sukces w zapewnieniu widzom większej liczby napisów do treści telewizyjnych na żywo (co jest ważnym krokiem naprzód i sukcesem regulacyjnym, który należy docenić), o tyle nie udało się uzyskać odpowiedniej jakości napisów na żywo dostępnych dla polskich widzów. W 2022 roku, podczas gdy napisy na żywo w polskich kanałach informacyjnych formalnie służą wypełnieniu prawnego obowiązku zapewnienia napisów, przynajmniej część z nich nie spełnia potrzeb użytkowników. Niska dokładność i wysokie opóźnienia oznaczają, że wiele programów telewizyjnych na żywo opatrzone napisami, które albo przekazują niedokładne lub wręcz fałszywe

informacje, albo są bardzo trudne do zrozumienia. W najgorszym przypadku napisy niskiej jakości nie oferują odbiorcom żadnej wartości dodanej (w porównaniu z brakiem napisów) i nie zapewniają widzom głuchym i słabosłyszącym realnego dostępu do treści telewizyjnych.

## **6. Znaczenie teoretyczne i praktyczne zastosowania**

Badanie potwierdziło, że możliwe jest dokonanie kompleksowego przeglądu wielu aspektów jakości napisów na żywo, w tym obiektywny pomiar takich aspektów jakości jak segmentacja tekstu czy odstępy pomiędzy napisami, które to nie podlegały ocenie w dotychczasowych badaniach. Przedmiotowe badanie może stanowić wzorzec w zakresie pomiaru różnych aspektów jakości napisów na żywo i może stanowić pierwszy krok w kierunku regularnego monitorowania jakości napisów na żywo w polskiej telewizji.

W przypadku każdego z analizowanych nadawców telewizyjnych zidentyfikowano główne problemy w zakresie jakości napisów, a zamieszczone na końcu rozprawy zalecenia dla nadawców telewizyjnych zawierają wskazówki w zakresie konkretnych działań, które każdy z nadawców może podjąć, by polepszyć jakość napisów na żywo. Zidentyfikowano też działania, które mogłaby podjąć Krajowa Rada Radiofonii i Telewizji na rzecz lepszej jakości napisów na żywo, takie jak opublikowanie wytycznych w zakresie jakości napisów na żywo, wydanie rekomendacji w zakresie maksymalnego dopuszczalnego opóźnienia napisów na żywo oraz prowadzenie regularnego monitoringu jakości napisów.

## **7. Ograniczenia badania i implikacje dla dalszych badań**

Ograniczenia badania wynikają głównie z dostępności danych. Ponieważ nadawcy telewizyjni nie udostępniają napisów do analizy i nie jest też możliwe ich pozyskanie w sposób automatyczny, konieczna jest pracochłonna i czasochłonna rekonstrukcja plików napisowych na podstawie nagrań oraz przygotowanie transkrypcji wypowiedzi. O ile przeanalizowany korpus napisów można uznać za reprezentatywny dla programów newsowych nadawanych w porze najwyższej



oglądalności (tzw. prime-time), nie obejmuje on nagrań z programów nadawanych w innych godzinach oraz został on ograniczony do dwóch gatunków: serwisu informacyjnego i debaty politycznej. Badanie można by rozszerzyć o inne rodzaje programów telewizyjnych opatrywanych napisami na żywo takie jak magazyn reporterski, talk show o tematyce niepolitycznej, transmisja nabożeństwa religijnego czy transmisja rozgrywek sportowych.

Ponadto badanie skupia się na obiektywnych miarach jakości. Nie ma pewności co do tego, jak te miary przekładają się na subiektywne postrzeganie jakości przez odbiorców napisów. By to ustalić, należałoby przeprowadzić badania recepcji napisów przez widzów oraz dokonać triangulacji danych uzyskanych poprzez zastosowanie miar obiektywnych z oceną dokonaną przez odbiorców napisów.

## 8. Bibliografia

- Fresno, N. (2019). Of bad hombres and nasty women; the quality of the live closed captioning in the 2016 US final presidential debate. *Perspectives*, 27(3), 350–366. DOI: <https://doi.org/10.1080/0907676X.2018.1526960>.
- Fresno, N., Sepielak, K. (2020). Subtitling speed in Media Accessibility research: some methodological considerations. *Perspectives*, 1–17. DOI: <https://doi.org/10.1080/0907676X.2020.1761841>.
- Greco, G. M. (2018). The nature of accessibility studies. *Journal of Audiovisual Translation*. (1), 205–232. Pobrane z: <https://www.jatjournal.org/index.php/jat/article/view/51>.
- Greco, G. M., Jankowska, A. (2020). Media Accessibility Within and Beyond Audiovisual Translation. W: Ł. Bogucki & M. Deckert (red.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (57–81). Cham: Springer International Publishing.
- Jankowska, A. (2020). Audiovisual media accessibility. W: E. Angelone, M. Ehrensberger-Dow, G. Massey (red.), *Bloomsbury companions: 2020: 1. The Bloomsbury companion to language industry studies*. London: Bloomsbury Academic.
- Neves, J. (2005). *Audiovisual Translation: Subtitling for the Deaf and Hard-of-Hearing* (PhD). Roehampton University, London. Pobrane z: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.1405&rep=rep1&type=pdf>.

- Romero Fresco, P. (2016). Accessing communication: The quality of live subtitles in the UK. *Language & Communication*, 49, 56–69.
- Romero Fresco, P. (2021). Negotiating quality assessment in media accessibility: the case of live subtitling. *Universal Access in the Information Society*, 20(4), 741–751. DOI: <https://doi.org/10.1007/s10209-020-00735-6>.
- Romero Fresco, P., Martínez Pérez, J. (2015). Accuracy Rate in Live Subtitling: The NER Model. W: R. Baños Piñero & J. D. Cintas (red.), *Palgrave studies in translating and interpreting. Audiovisual translation in a global context. Mapping an ever-changing landscape* (nr. 3, s. 28–50). Basingstoke: Palgrave Macmillan. DOI: [https://doi.org/10.1057/9781137552891\\_3](https://doi.org/10.1057/9781137552891_3).
- Romero Fresco, P., Pöchhacker, F. (2018). Quality assessment in interlingual live subtitling: The NTR Model. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 16(o). Pobrane z: <https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/download/438/402>.
- Sacha, A., Kasperkowiak, M. (2012). Napisy dla osób głuchych oraz słabosłyszących i tłumaczenia w języku migowym w telewizji – standardy, dobre praktyki, stan obecny. Pobrane z: [http://www.krrit.gov.pl/Data/Files/\\_public/Portals/o/konsultacje/2012/prawo\\_medi\\_alne/19-niepelnosprawni\\_fdc\\_zalacznik-2.pdf](http://www.krrit.gov.pl/Data/Files/_public/Portals/o/konsultacje/2012/prawo_medi_alne/19-niepelnosprawni_fdc_zalacznik-2.pdf).
- Saerens, G., Tampir, M., Dutka, Ł., Szczygielska, M., Szarkowska, A., Romero-Fresco, P., . . . Robert, I. (2020). How to implement live subtitling in TV settings: Guidelines on making television broadcasts accessible to hard of hearing and deaf people as well as foreigners. Pobrane z: <http://ka2-ilsa.webs.uvigo.es/guidelines/>.
- Szarkowska, A. (2010). Accessibility to the media by hearing impaired audiences in Poland: problems, paradoxes, perspectives. W: J. Neves, A. Matamala, J. Díaz Cintas (red.), *Approaches to Translation Studies. New Insights Into Audiovisual Translation and Media Accessibility* (s. 139–158). Amsterdam: Brill | Rodopi.
- Szarkowska, A., Gerber-Morón, O. (2018). Viewers can keep up with fast subtitles: Evidence from eye movements. *PloS one*, 13(6), e0199331. DOI: <https://doi.org/10.1371/journal.pone.0199331>.
- Szarkowska, A., Barreto Silva, B., Orrego-Carmona, D. (2021). Effect of subtitle speed on proportional reading time: Re-analysing subtitle reading data with mixed effects models. *Translation, Cognition & Behavior*. Artykuł w czasopiśmie internetowym. DOI: <https://doi.org/10.1075/tcb.00057.sza>.
- Szczygielska, M. (2019). Napisy dla niesłyszących w Polsce. Historia, problemy, wyzwania. *Półrocznik Językoznawczy Tertium*, 4(193-210). DOI: <https://doi.org/10.7592/Tertium2019.4.1.Szczygielska>.
- Zarate, S. (2021). *Captioning and Subtitling for d/Deaf and Hard of Hearing Audiences*. UCL Press.