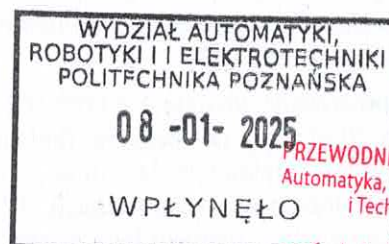


RECENZJA



PRZEWODNICZĄCY RADY DYSCYPLINY
Automatyka, Elektronika, Elektrotechnika
i Technologie Kosmiczne
prof. dr hab. inż. Wojciech Szeląg

naukowego osiągnięcia habilitacyjnego Pana doktora inż. Szymona Drgasa pt.

„Interpretowalne i niskokosztowe metody poprawy jakości mowy”

oraz ocena aktywności naukowej (dorobku naukowo-badawczego) Habilitanta
wykonana na zlecenie Rady Dyscypliny AEEiTK Politechniki Poznańskiej (uchwała z dnia 8.10.2024)

- A. Ocena osiągnięcia naukowego
- B. Ocena aktywności zawodowej
- C. Wniosek końcowy

A. Ocena osiągnięcia naukowego

A.1. Ogólna charakterystyka osiągnięcia habilitacyjnego

Osiągnięciem habilitacyjnym Pana dr inż. Szymon Drgasa jest **monotematyczny cykl jego 13 publikacji**, w których zajmował się on zagadnieniem poprawy jakości cyfrowego sygnału mowy (głównie jej zrozumiałości) w trzech oddzielnych zastosowaniach/systemach: automatycznego rozpoznawania mowy (ASR), automatycznego rozpoznawania mówców oraz aparatach/implantach dla osób z ubytkiem słuchu. Swoje prace kontynuował na przestrzeni ostatnich 12 lat od obrony doktoratu w 2013 roku na Politechnice Poznańskiej. Publikacje cyklu można podzielić na dwie główne grupy, w których wykorzystywane są odmienne narzędzia matematyczne/algorytmiczne do rozwiązywania dwóch różnych problemów:

- 1) nieujemna faktoryzacja oraz nieujemny rozplot macierzy spektrogramu sygnału mowy, czyli *Nonnegative Matrix Factorization* (NMF) i *Nonnegative Matrix Deconvolution* (NMD) macierzy wartości bezwzględnych współczynników krótko-czasowej transformacji Fouriera – *Short-Time Fourier Transform* (STFT) sygnału mowy, czyli generatywne metody uczenia maszynowego, są używane do wyznaczania wzorców charakteryzujących mówcę;
- 2) dedykowane, zaproponowane przez Habilitanta nisko-kosztowe, różniczkowalne modyfikacje metod uczenia maszynowego, w szczególności sieci neuronowych, są wykorzystywane do poprawy zrozumiałości sygnału mowy w systemach ASR i w implantach słuchu.

W tym miejscu recenzent potwierdza spełnienie przez Habilitanta formalnego wymogu, dotyczącego posiadania udokumentowanego osiągnięcia naukowego, które może być uznane za osiągnięcie habilitacyjne.

A.2. Omówienie publikacji cyklu habilitacyjnego

Na recenzowane osiągnięcie naukowe składa się jednotematyczny cykl **13 publikacji naukowych**, w tym **9 artykułów** w czasopismach z listy JCR (sumaryczny IF=27.61, łącznie 805 punktów MNiSzW) oraz 4 referaty konferencyjne. Prace te są scharakteryzowane poniżej.

- [A1] **Szymon Drgas (100%)**, “Speech intelligibility prediction using generalized ESTOI with fine-tuned parameters”, *Speech Communication*, vol. 159 (2024), str. 1-10, doi: 10.1016/j.specom.2024.103068. (140 pkt, IF=2,4).

Opracowano metodę FT-GESTOI (*Fine-Tuned Generalized Extended Spectro-Temporal Objective Intelligibility*) do uczenia funkcji strat podczas oszacowywania zrozumiałości sygnału mowy. Metoda wykorzystuje, m.in., podpasmową dekompozycję sygnału mowy (zespoły filtrów krótkoczasowej transformacji Fouriera STFT), filtrację 2D w dziedzinie czas-częstotliwość, progowanie i regresję logistyczną. Eksperymentalnie wykazano użyteczność metody. Praca zawiera 42 odnośniki.

- [A2] **Szymon Drgas (70%)**, Lars Bramslov, Archontis Politis, Gaurav Naithani i Tuomas Virtanen. “Dynamic processing neural network architecture for hearing loss compensation”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 32 (2023), str. 203-214, doi: 10.1109/TASLP.2023.3328285 (140 pkt, IF=4,1).

Opracowano metodę DPNN (*Dynamic Processing Neural Network*) do kompensacji ubytków słuchu z użyciem splotowych sieci neuronowych. Przetwarzanie sygnału zrealizowano w dziedzinie czas-częstotliwość z użyciem współczynników transformacji STFT. Stosowano filtrację 2D i wzmacniano obwiednie zmiennych składowych TF sygnału. Porównano wyniki uzyskane dla różnych konfiguracji sieci z użyciem dwóch metryk jakości zrozumiałości mowy: STOI (*Spectro-Temporal Objective Index*) oraz HASPI (*Hearing Aid Speech Intelligibility*). Wykazano skuteczność metody. Praca zawiera 42 odnośniki.

- [A3] **Szymon Drgas (100%)**, “A survey on low-latency DNN-based speech enhancement”, *Sensors* vol. 23 (2023), 1380, str. 1-26, doi: 10.3390/s23031380 (100 pkt, IF=3,4).

W publikacji dokonano bardzo obszernego przeglądu rozwiązań stosowanych do uzdatniania sygnału mowy z użyciem głębokich sieci neuronowych DNN (*Deep Neural Networks*). Jako miary poprawy jakości mowy użyto dwóch metryk: STOI, tak jak poprzednio, oraz PESQ (*Perceptual Evaluation of Speech Quality*). Praca zawiera 117 odnośników.

- [A4] **Szymon Drgas (85%)** i Tuomas Virtanen, “Joint speaker separation and recognition using non-negative matrix deconvolution with adaptive dictionary”, *Computer Speech & Language*, vol. 70 (2021), s. 101223, str. 1-14, doi: 10.1060/j.csl.2021.101223 (100 pkt, IF=3,252).

W pracy zaproponowano modyfikację nieujemnego rozplotu macierzy NMD, którą nazwano NMD z adaptacyjnym słownikiem, czyli DANMD (*Adaptive-Dictionary NMD*). Modyfikację zastosowano do separacji mówców. W metodzie NMD macierz transformacji STFT jest przedstawiana jako liniowa superpozycja macierzy ze słownika. W modyfikacji metody macierze słownika należą do dwóch grup: macierzy niezależnych od mówców oraz zależnych od mówców. Wykazano skuteczność metody. Praca zawiera 38 odnośników.

- [A5] **Szymon Drgas (70%)**, Magdalena Blaszkiewicz i Anna Przekoracka-Krawczyk, “The Combination of neural tracking and alpha power lateralization for auditory attention detection”. *Journal of Speech, Language, and Hearing Research*, vol. 64 (2021), no. 9, str. 3603–3616, doi: 10.1044/2021_JSLHR-20-00608. (100 pkt, IF=2,674).

W pracy przedstawiono wyniki badań dotyczących wykrywania uwagi słuchowej AAD (*Auditory Attention Detection*), dokładniej – kierunku z którego pochodzi dźwięk interesujący słuchacza. Dokonywano rejestracji sygnałów encefalografu EEG (fal mózgowych) oraz stosowano metody NT (*Neural Tracking*) oraz APL (*Alpha Power Lateralization*). Stosowano analizę sygnału w podpasmach częstotliwościowych. Wykazano, że oddzielne zastosowanie NT i APL daje podobne wyniki, natomiast łączne – zdecydowanie lepsze. Praca zawiera 35 odnośników.

- [A6] **Szymon Drgas (70%)**, Tuomas Virtanen, Jorg Lucke i Antti Hurmalainen, “Binary non-negative matrix deconvolution for audio dictionary learning”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25 (2017), no. 8, str. 1644–1656, doi: 10.1109/TASLP.2017.2709909 (25 pkt, IF=2,95).

W pracy zaproponowano wersję binarną metody nieujemnego rozplotu macierzy NMD i zastosowaną ją do macierzy STFT sygnału mowy. Macierz STFT jest przedstawiana jako suma opóźnionych macierzy wzorcowych (należących do tzw. słownika, który trzeba zbudować), wziętych z binarnymi wagami. Metodę zastosowano do separacji i identyfikacji mówców. Otrzymano lepsze wyniki niż w metodach GMM (*Gaussian Mixture Model*) i I-vector. Praca zawiera 48 odnośników.

- [A7] **Szymon Drgas (80%)** i Adam Dabrowski, “Speaker recognition based on multilevel speech signal analysis on Polish corpus”, *Multimedia Tools and Applications*, vol. 74 (2015), str. 4195–4211, doi: 10.1007/s11042-013-1502-0 (**30 pkt, IF=1,331**).

Artykuł dotyczy zagadnienia identyfikacji mówcy niezależnie od wypowiedzianej przez niego treści. Zaproponowano i przetestowano skuteczność łącznego użycia cech widmowych mowy z cechami wysokopoziomowymi (prozodycznymi, artykulacyjnymi i leksykalnymi). Zbudowano i użyto korpus nagrań rozmów telefonicznych w języku polskim. Użyto metod *fast scoring* and *AdaBoost* do fuzji wielu cech w procesie rozpoznawania. Przetestowano różne metody normalizacji. Zwrócono uwagę na korzystne właściwości normalizacji z-norm2. Praca zawiera 23 odnośniki.

- [A8] Tomasz Grzywalski i **Szymon Drgas (50%)**, “Speech enhancement using U-nets with wide-context units”, *Multimedia Tools and Applications*, vol. 81 (2022), str. 18617–18639, doi: 10.1007/s11042-022-12632-6 (**70 pkt, IF=3,6**).

Zaproponowano nową metodę poprawy zrozumiałości i jakości sygnału mowy z użyciem sieci neuronowej. Użyto architektury U-net, składającej się z bloku kodera i dekodera. W celu zwiększenia kontekstu przetwarzania w warstwach ukrytych zastosowano rekurencyjne komórki sieci oraz rozszerzone elementy splotowe. Uzyskano lepsze wskaźniki zrozumiałości mowy: SI-SDR (*Scale-Independent Speech-to-Distortion Ratio*), STOI oraz PESQ, niż w innych metodach. Praca zawiera 47 odnośników.

- [A9] Tomasz Grzywalski i **Szymon Drgas (45%)**, “Speech enhancement by multiple propagation through the same neural network”, *Sensors*, vol. 22 (2022), s. 2440, str. 1-14, doi: 10.3390/s22072440 (**100 pkt, IF=3,9**).

Użyto trzech struktur sieci neuronowych (U-net, ResBLSTM oraz Transformer-Net) do wielokrotnego (maksymalnie 5-krotnego) przetwarzania sygnału mowy w celu poprawy jej zrozumiałości. Jako miar jakości zrozumiałości mowy użyto metryk I-SDR, STOI, and PESQ. Stwierdzono, że poprawa zrozumiałości jest coraz mniejsza z każdą kolejną iteracją. Wykazano, że architektury U-Net and Transformer-Net bardziej zyskują na operacji powtarzania przetwarzania niż architektura ResBLSTM. Praca zawiera 36 odnośników.

- [A10] **Szymon Drgas (85%)** i Tuomas Virtanen, “Speaker verification using adaptive dictionaries in non-negative spectrogram deconvolution”, *12-th Int. Conf. on Latent Variable Analysis and Signal Separation (LVA/ICA-2015)*, Liberec 2015, Springer, LNCS 9237, str. 462–469, doi: 10.1007/978-3-319-22482-4_54 (**15 pkt WoS**).

Praca stanowi pierwszą, konferencyjną publikację Habilitanta, dotyczącą zaproponowanej modyfikacji metody nieujemnego rozplotu macierzy NMD, tzn. dodania do niej adaptacyjnego słownika i otrzymania DANMD (*Adaptive-Dictionary NMD*). Później metodę tę szeroko opisano w [A4]. Praca zawiera 11 odnośników.

- [A11] Tomasz Grzywalski i **Szymon Drgas**, “Speech enhancement by iterating forward pass through U-net”, *2020 Conf. on Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, Poznań 2020, IEEE, str. 157–162, doi: 10.23919/SPA50552.2020.9241307 (**20 pkt**).

Praca stanowi konferencyjną, wczesną wersję publikacji [A8, A9]. Jest tu tylko rozpatrywana architektura U-net sieci neuronowej z rozszerzonym kontekstem, tak jak w [A8]. Wielokrotne przetwarzanie jest przeprowadzane tak jak w [A9]. Użyto metryk I-SDR, STOI, and PESQ. Liczne wyniki symulacji potwierdzają wartościowość propozycji Habilitanta. Praca zawiera 11 odnośników.

- [A12] Tomasz Grzywalski i **Szymon Drgas (50%)**, “Using recurrences in time and frequency within U-net architecture for speech enhancement”, *2019 IEEE Int. Conf. on Acoustics, Speech and Signal*

Processing (ICASSP), Brighton 2019, str. 6970–6974, doi: 10.1109/ICASSP.2019. 8682830 (**70 pkt CORE**).

W pracy przeprowadzono analizę doboru wielkości elementów spłotowych oraz liczby elementów rekurencyjnych w różnych warstwach sieci neuronowej o architekturze U-net, w sytuacji kiedy sieć ta jest stosowana w zadaniach uzdatniania (poprawy jakości/zrozumiałości) sygnału mowy w obecności szumu. Rozpatrywano problem rozszerzania kontekstu analizy danych w warstwach ukrytych. Rozpatrywano sieci: FCLN (*Fully Connected Layers Network*), RNN (*Recursive Neural Network*) oraz U-net bez oraz z rekurencją. Najlepsze wyniki otrzymano dla tej ostatniej. **Praca ta jest najczęściej cytowaną pracą z cyklu habilitacyjnego (15 cytowań)**. Praca zawiera 11 odnośników.

- [A13] Tomasz Grzywalski i **Szymon Drgas**, “Application of recurrent U-net architecture to speech enhancement”, *2018 Conf. on Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, IEEE, Poznań 2018, str. 82–87, doi: 10.23919/SPA.2018.8563364 (**20 pkt**).

W pracy zaproponowano rekurencyjną architekturę sieci U-net, odpowiednią dla zadania poprawy jakości mowy. Sieć pracuje na spektrogramach (STFT) mowy. Stosuje nad-próbkowanie. Praca zawiera 11 odnośników.

Habilitant podzielił publikacje cyklu na dwie grupy:

- 1) publikacje [A4, A5, A6, A7, A10] – pod wspólnym tytułem: „opracowanie generatywnych metod uczenia maszynowego do wyznaczania wzorców charakteryzujących mówcę”, głównie dotyczących modyfikacji metody NMD [A4, A6, A10];
- 2) publikacje [A1, A2, A3, A8, A9, A11, A12, A12] – pod wspólnym tytułem: „niskokosztowe, różniczkowalne metody uczenia maszynowego do poprawy zrozumiałości mowy”.

A.3. Ocena osiągnięcia habilitacyjnego – wkład w rozwój dyscypliny

Wkład w rozwój dyscypliny można ocenić poprzez miejsce publikacji osiągnięć naukowych (rangę czasopism i konferencji) oraz poprzez poziom ich cytowania na poziomie światowym.

Uwzględniając pierwsze kryterium, wkład Habilitanta w rozwój dyscypliny można określić jako znaczący. Swoje prace opublikował on w renomowanych czasopismach światowych, wymienionych w tabeli I. Na szczególną uwagę zasługują dwa pierwsze czasopisma, bardzo blisko związane z tematyką badań. Sumaryczny, indywidualny IF osiągnięcia habilitacyjnego (uwzględniający udział procentowy w publikacjach) wynosi $IF_{\text{suma}}^{\%} = 20$ oraz odpowiada mu **615 punktów MNiSzW $_{\text{suma}}^{\%}$** (także z uwzględnieniem udziału procentowego). W przeliczeniu na jedną publikację otrzymujemy $IF_{\text{mean}}^{\%} = 2.2$ oraz **68 punktów MNiSzW $_{\text{mean}}^{\%}$** , czyli duże wartości.

Tabela I. Porównanie wskaźników biblio-metrycznych publikacji Habilitanta w czasopismach JCR.

Czasopismo	Rok	Punkty MNiSzW	IF
<i>IEEE/ACM Trans. on Audio, Speech, and Language Processing</i>	2024, 2017	140, 25	4.1
<i>Speech Communication</i>	2024	140	2.4
<i>Sensors</i>	2023, 2022	100, 100	3.4, 3.9
<i>Computer Speech & Language</i>	2021	100	3.3
<i>Journal of Speech, Language, and Hearing Research</i>	2021	100	2.7
<i>Multimedia Tools and Applications</i>	2022	70	1.3

W tabeli II porównano wszystkie publikacje cyklu habilitacyjnego pod względem liczby cytowań na podstawie bazy SCOPUS. Uwzględniono w nim tylko cytowania obce, tzn. nie uwzględniono cytowań pochodzących od samych autorów prac. Zestawienie to jest bardzo interesujące. Wynika z niego, że indeks Hirscha prac cyklu bez autocytowań wynosi **H=4** i że decydują o tym prace

[A3, A10, A12, A13], z których trzy ostatnie to starsze publikacje konferencyjne, a praca pierwsza ma charakter przeglądowy. Dziwi tylko jedno cytowanie pracy [A6], opublikowanej w uznanym czasopiśmie w 2017 roku. Nie dziwi brak cytowań dwóch pierwszych, najważniejszych prac cyklu z powodu ich publikacji w latach 2023-2024. Zdaniem recenzenta indeks Hirscha **H=4 bez autocytowań** oraz łącznie **34 cytowania 4 najbardziej popularnych prac** Habilitanta z ocenianego cyklu to bardzo dobry wynik, także potwierdzający istotny wkład Habilitanta w rozwój dyscypliny.

Tabela II. Cytowania obce (poza autorami prac) publikacji Habilitanta w bazie SCOPUS.

Praca	Rok	IF	PKT	Cytowania
A1	2024	2.4	140	--
A2	2023	4.1	140	--
A3	2023	3.4	100	7
A4	2021	3.3	100	--
A5	2021	2.7	100	--
A6	2017	3.0	25	1
A7	2015	1.3	30	3
A8	2022	3.6	70	2
A9	2022	3.9	100	2
A10	2015	--	15	5
A11	2020	--	70	--
A12	2019	--	20	15
A13	2018	--	20	7

Podsumowując, znaczenie otrzymanych wyników można ocenić za pomocą wartości wskaźników naukometrycznych osiągnięcia. **Indywidualny** sumaryczny **IF=20 i 615 punktów MNiSW** oraz **indeks Hirscha H=4** (bez autocytowań wszystkich autorów) samego osiągnięcia habilitacyjnego są moim zdaniem wystarczającym dowodem na poparcie tezy o istotności naukowej opiniowanego cyklu artykułów.

Moim zdaniem recenzowane osiągnięcie habilitacyjne ma następujące **zalety**:

- 1) finezję rozwiązań matematycznych, zastosowanych podczas modyfikacji nieujemnego rozplotu macierzy spektrogramu sygnału mowy oraz podczas użycia nowej metody;
- 2) bardzo duże wyczucie algorytmiczne dotyczące sieci neuronowych, prowadzące do zaproponowania ich nowej architektury, która okazała się niezwykle skuteczna w zadaniu odsumiania i poprawy zrozumiałości sygnału mowy (dodanie elementów rekurencyjnych oraz splotowych, rozszerzających kontekst w warstwach ukrytych);
- 3) umiejętne dokonanie fuzji cech niskopoziomowych i wysokopoziomowych (prozodycznych, artykulacyjnych i leksykalnych) w zadaniu rozpoznawania mówcy;
- 4) prowadzenie badań we współpracy z największymi specjalistami światowymi: Prof. Tuomasem Virtanenem z Tampere University of Technology (Finlandia) oraz Prof. Jorgiem Lucke z University of Oldenburg (Niemcy).

A.4. Podsumowanie

Uwzględniając powyższe uwagi **uważam, że opiniowany, jednotematyczny cykl publikacji Pana dr inż. Szymona Drgasa może być uznany za osiągnięcie o znaczącym wkładzie w rozwój dyscypliny automatyka, elektronika, elektrotechnika i technologie kosmiczne**, w rozumieniu obowiązującej ustawy „Prawo o szkolnictwie wyższym i nauce” z dnia 20 lipca 2018 roku. Osiągnięcie jest osobiste, oryginalne i wartościowe, o czym świadczy wysoki sumaryczny indywidualny **IF=20** (z

uwzględnieniem udziału procentowego) oraz indeks Hirscha $H=4$ tylko prac cyklu bez autocytowań. Osiągnięcie habilitacyjne jest doceniane na świecie, o czym przekonuje kilkadziesiąt jego cytowań (dokładnie 42), pomimo opublikowania najważniejszych dwóch prac cyklu w latach 2023-2024. Cztery prace cyklu są łącznie cytowane aż 34 razy.

B. Ocena aktywności naukowej

B.1. Życiorys zawodowy

Wykształcenie. Pan Szymon Drgas uzyskał w 2005 roku stopień naukowy magistra fizyki na Wydziale Fizyki Uniwersytetu im. Adama Mickiewicza w Poznaniu w specjalizacji akustyka, broniąc pracę dyplomową p.t. „*Dyskryminacja częstotliwości modulacji amplitudowej*” (promotor prof. dr hab. A. Sęk). Z kolei w roku 2006 otrzymał na tym samym wydziale i uczelni stopień naukowy magistra informatyki stosowanej, po obronie pracy dyplomowej pt. „*Automatyczne rozpoznawanie mowy przy użyciu niejawnych modeli Markowa*” (promotor prof. dr hab. A. Dąbrowski). 7 lat później, czyli w 2013 roku, obronił pracę doktorską pt. „*Automatic speaker recognition based on multilevel analysis of speech signal*” na Wydziale Informatyki Politechniki Poznańskiej (promotor prof. dr hab. A. Dąbrowski) i uzyskał tytuł doktora nauk technicznych z zakresu automatyki i robotyki.

Zatrudnienie. Od ukończenia studiów wyższych w 2006 roku Habilitant pracuje na Politechnice Poznańskiej, w latach 2006-2013 na stanowisku asystenta, a w latach 2013-2024 – na stanowisku adiunkta, cały czas na Wydziale Automatyki, Robotyki i Elektrotechniki.

B.2. Charakterystyka i ocena aktywności publikacyjnej

Zdecydowana większość dorobku naukowego Pana Doktora Szymona Drgasa dotyczy zagadnienia przetwarzania sygnałów w systemach rozpoznawania mowy, rozpoznawania mówców oraz uzdatniania sygnału mowy na potrzeby implantów słuchu. W szczególności, Habilitant zajmował się także następującymi zagadnieniami i posiada publikacje z tego zakresu: separacja dźwięków, rozpoznawanie mowy szeptanej, wykrywania patologii głosu, wykrywania nieprawidłowych dźwięków osłuchowych płuc i oskrzeli, a także – filtry cząsteczkowe oraz optymalizacja parametrów uogólnionego podobieństwa kosinusowego.

Pan dr inż. Szymon Drgas jest autorem łącznie **54 publikacji**, napisanych w ciągu niespełna 20 lat swojej pracy zawodowej, w tym:

- 1) **26 publikacji przed doktoratem** (12 artykułów w czasopismach i 14 referatów konferencyjnych),
- 2) **28 publikacji po doktoracie** (7 rozdziałów w monografiach, 11 artykułów w czasopismach i 10 referatów konferencyjnych).

Tabela III. Liczba publikacji Habilitanta

	Rozdz. w monografiach	Artykuły w czasopismach	Wystąpienia konferencyjne	Razem
Przed dr	--	12	14	26
Po dr	7	11	10	28
Razem	7	23	24	54

Artykuły zostały opublikowane w następujących czasopismach:

- 1) **po doktoracie:** *IEEE/ACM Trans. on Audio, Speech and Language Processing* (2024, IF=4.1, 2017, IF=2.95), *Speech Communication* (2024, IF=2.4), *Sensors* (2023, 2022, IF=3.4), *Multimedia Tools and Applications* (2022, IF=3.6, 2015, IF=1.33), *Computer Speech &*

Language (2021, IF=3.3), *Journal of Speech, Language and Hearing Research* (2021, IF=2.7), *Przegląd Elektrotechniczny* (2016), *Elektronika: Konstrukcje Technologie Zastosowania* (2013);

- 2) **przed doktoratem:** *Hearing Research* (2010, IF=2.4), *Journal of Speech, Language and Hearing Research* (2009, IF=2.4), *Applied Acoustics* (2012, IF=1.1), *Archives of Acoustics* (2008, 2007 2x), *Przegląd Elektrotechniczny* (2012), *Elektronika: Konstrukcje Technologie Zastosowania* (2011 3x, 2008, 2007).

Referaty zostały wygłoszone na następujących konferencjach:

- 1) **po doktoratem:** *IEEE Int. Conf. on Acoustics, Speech and Signal Processing ICASSP* (2019, 70 pkt.), *IEEE Int. Conf. on Big Data* (2018, 70 pkt), *Federated Conf. on Computer Science and Information Systems* (2018, 70 pkt), *Int. Conf. on Methods and Models in Automation and Robotics* (MMAR), *Międzyzdroje* (2016 2x, 15 pkt), *Signal Processing: Algorithms, Architectures, Arrangements and Applications* (SPA). Poznań (2016, 2013, 15 pkt), *Int. Conf. on Latent Variable Analysis and Signal Separation LVA/ICA* (2015, 15 pkt);
- 2) **przed doktoratem:** *INTERSPEECH* (2011), *Int. Conf. on Signals and Electronic Systems ICSES* (2012 2x, 2008), *Signal Processing: Algorithms, Architectures, Arrangements and Applications* (SPA). Poznań (2012, 2011, 2010 2x, 2009, 2008, 2007), *Int. Conf. on Multimedia, Communications, Services and Security MCSS*, Kraków, (2011).

Summaryczny *Impact Factor* całego dorobku naukowego Pana dr inż. Szymona Drgasa jest równy **IF=33.5**, zaś tylko po uzyskaniu stopnia doktora - **IF=27.7**. W bazie SCOPUS posiada on łącznie **114 cytowań obcych** i indeks Hirscha **H=5** (bez autocytowań). W bazie WoS posiada łącznie **135 cytowań obcych** oraz indeks Hirscha **H=6** (bez autocytowań). Łącznie 22 jego publikacje niższej rangi (rozdziały w monografiach oraz referaty w materiałach konferencyjnych WoS/CORE) zebrały **595 punktów** MNiSzW, natomiast 12 publikacji wyższej rangi (artykuły w czasopiśmie JCR) ocenionych zostało na **894 punktów** MNiSzW. Są to **liczby potwierdzające dobrą jakość całego dorobku naukowego Habilitanta**.

Podsumowując tę część recenzji można stwierdzić, że tak pod względem liczby publikacji jak i ich jakości **oceniana aktywność naukowa jest bardzo dobra**. Należy podkreślić wysoką rangę czasopism, w których Habilitant publikował swoje prace po doktoracie. Trzeba jednak zwrócić uwagę na udział Habilitanta w **bardzo małej liczbie konferencji (4) o najwyższej randze światowej**. Najprawdopodobniej jest to spowodowane bardzo wysokimi kosztami udziału w tych konferencjach.

B.3. Charakterystyka i ocena wpływu publikacyjnego na rozwój dyscypliny

Za jedno z kryteriów oceny wpływu Pana dr inż. Szymona Drgasa na rozwój dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne można przyjąć poziom cytowania jego publikacji. A zdaniem recenzenta jest on wysoki. Jak już zaznaczono powyżej, prace **[A12, A3, A13]** cyklu habilitacyjnego zostały zauważone i docenione w światowym środowisku naukowym i posiadają, odpowiednio, aż **15, 7 i 7 cytowań obcych**. Dodatkowo dwie prace Habilitanta, nie wchodzące w skład cyklu, posiadają łącznie aż **31 cytowań**:

- 1) Kociński, J., **Drgas, S.**, Ozimek, E., "Evaluation of Blind Source Separation for Different Algorithms Based on Second Order Statistics and Different Spatial Configurations of Directional Microphones", *Applied Acoustics* 73.2 (2012), str. 109-116, IF=1,097 (**16 cytowań**);
- 2) Grzywalski, T., Maciaszek, A., Biniakowski, A., Orwat, J., **Drgas, S.**, Piecuch, M., ... ,Szaryński, K. (2018), „Parameterization of Sequence of MFCCs for DNN-based Voice Disorder Detection”, *2018 IEEE Int. Conf on Big Data*, 70 pkt (CORE) (**15 cytowań**).

Na tej podstawie można jednoznacznie stwierdzić, że dr inż. Szymon Drgas jest rozpoznawalnym w świecie naukowcem z zakresu zaawansowanych metod przetwarzania i analizy sygnałów akustycznych w systemach automatycznej separacji i identyfikacji akustycznej.

B.4. Charakterystyka i ocena istotnej współpracy z sektorem gospodarczym

W 2016 roku Habilitant podjął współpracę z firmą StethoMe (<https://www.stethome.com>), która stworzyła stetoskop działający w oparciu o metody uczenia maszynowego. Umożliwia on analizę i wykrywanie nieprawidłowych dźwięków osłuchowych w układzie oddechowym w warunkach domowych, oraz przesyłanie wyników badań na odległość do lekarza. W okresie od kwietnia 2017 do listopada 2018 Habilitant pracował w firmie StethoMe na stanowisku inżyniera przy doborze i opracowaniu algorytmów do przetwarzania i ekstrakcji informacji ze stetoskopu. Zatrudnienie to było współfinansowane ze środków Europejskiego Funduszu Rozwoju Regionalnego w ramach projektu pt. "Unikalna w skali świata technologia identyfikacji i klasyfikacji dźwięków z badania osłuchowego przeprowadzonego w warunkach domowych, z wykorzystaniem inteligentnych algorytmów, jako wsparcie zdalnej diagnostyki i monitorowania chorób układu oddechowego." W trakcie wspólnych prac badawczych powstały publikacje [E11, E12], których Habilitant jest współautorem. Firma StethoMe istnieje i oferuje swoje produkty w chwili obecnej, co bardzo dobrze świadczy pomocy naukowej udzielonej w przeszłości ww. firmie przez Habilitanta.

B.5. Charakterystyka i ocena innych rodzajów aktywności i osiągnięć naukowych

Poniżej skrótowo scharakteryzowano pozostałe rodzaje aktywności i osiągnięcia Habilitanta.

1) Wizyty naukowe po obronie doktoratu

- 5.08.2013 - 25.08.2013, **Finlandia**, Tampere University of Technology (obecnie: Tampere University), Audio Research Group, prof. Tuomas Virtanen, badania dotyczące przetwarzania sygnału mowy.
- 23.02.2015 - 18.03.2015, maj 2015, **Niemcy**, Carl von Ossietzky Universitat, Machine Learning Group, prof. Jorg Lucke, badania dotyczące uczenia reprezentacji sygnałów.
- 3.06.2015 - 24.06.2015, **Finlandia**, Tampere University of Technology (obecnie: Tampere University), Audio Research Group, prof. Tuomas Virtanen, badania dotyczące uczenia reprezentacji sygnałów.

2) Udział w projektach badawczych/konkursowych po obronie doktoratu

- **POIR.01.01.01-00-0528/16-02**: "Unikalna w skali świata technologia identyfikacji i klasyfikacji dźwięków z badania osłuchowego przeprowadzonego w warunkach domowych, z wykorzystaniem inteligentnych algorytmów, jako wsparcie zdalnej diagnostyki i monitorowania chorób układu oddechowego", w ramach programu Szybka Ścieżka, **jako pracownik firmy StethoMe**.
- **Pro-IDEAS-2013** nr 09/93/DSMK/4106, Wydział Informatyki Politechniki Poznańskiej, Konkurs badawczy dla młodych pracowników nauki, "Cyfrowe metody analizy sceny dźwiękowej (CMASD)" – **kierownik**.
- **Inżynier Przyszłości** nr 012/2/2014/MD - Wzmocnienie potencjału dydaktycznego Politechniki Poznańskiej, "Ekstrakcja informacji o mówcy z nagrań jednokanałowych" – **stypendysta**.

3) Udział w projektach badawczych/konkursowych przed obroną doktoratu

- **PPBW (Polska Platforma Bezpieczeństwa Wewnętrznego)**, podprojekt nr 0020/1/R/T00/2010/05: „System analizy, klasyfikacji i wspomagania rozpoznawania osób na podstawie rozmów z telefonów alarmowych oraz kontroli operacyjnej”, projekt finansowany przez NCBR, realizowany w latach: 2008 – 2011 - **wykonawca**.
- **EU FP7 INDECT** („Inteligentny system informacyjny wspierający obserwację, wyszukiwanie i detekcję dla celów bezpieczeństwa obywateli w środowisku miejskim”), projekt nr 218086 7-ego Programu Ramowego Unii Europejskiej, realizowany w latach 2008 – 2013. – **wykonawca**.

- **Projekt KBN nr 3 T11 C 040 30** – „Rozwój parametryczno-lingwistycznych metod analizy mowy w celu identyfikacji osoby mówiącej” (zakończony w 2009 r.) - **wykonawca**.

4) Recenzje artykułów i referatów

Habilitant 15 razy (15) recenzował prace zgłaszane do renomowanych czasopism światowych: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2), *Signal Processing Letters* (1), *Computer Speech and Language* (1), *Archives of Acoustics* (4), *Multimedia Tools and Applications* (6), *PlosONE* (1).

Jest on także wieloletnim recenzentem referatów zgłaszanych do dwóch największych konferencji światowych z zakresu analizy i przetwarzania sygnału mowy, a mianowicie *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* oraz *INTERSPEECH*. Po obronie doktoratu w 2013 roku jest on stałym recenzentem artykułów zgłaszanych na *IEEE SPA Conference*, odbywającej się corocznie w Poznaniu.

5) Członkostwo w komitetach redakcyjnych czasopism

Habilitant jest obecnie redaktorem działu „*Speech, computational acoustics, and signal processing*” czasopisma *Archives of Acoustics*, wydawanego przez Polską Akademię Nauk.

6) Członkostwo w komitetach konferencji

Habilitant jest od 2006 roku członkiem Komitetu Organizacyjnego międzynarodowej konferencji *IEEE Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)* odbywającej się od prawie 30 lat w Poznaniu pod auspicjami dwóch Polskich Sekcji IEEE: *Signal Processing* oraz *Circuits and Systems*. Ponad to brał udział w pracach Komitetu Naukowego Ogólnopolskiej Studenckiej Konferencji Akustyków (OSKA).

7) Członkostwo towarzystw naukowych

W roku 2019 Habilitant należał do organizacji IEEE.

8) Promocja kadry naukowej

Promotor 14 prac dyplomowych inżynierskich i 6 prac magisterskich.

9) Dodatkowo – działalność dydaktyczna, popularyzatorska i organizacyjna

Wykładowca akademicki odpowiedzialny za przedmioty: Akustyka techniczna, Algebra z geometrią, Przetwarzanie sygnałów i informacji, Teoria i metody optymalizacji. Opiekun Koła Naukowego „Decybel”, działającego na Politechnice Poznańskiej. Od 2021 roku zaangażowany w przygotowanie projektu „*Alzheimer Prediction Project*”, mającego na celu wczesne wykrywanie choroby Alzheimera w fazie prodromalnej (instytucje współpracujące: UAM-Poznań, PAN, Uniwersytet w Oldenburgu/Niemcy) - opiekun strony internetowej alz.put.poznan.pl. Osoba zaangażowana w budowę oraz ciągły rozwój laboratorium akustyki technicznej na Politechnice Poznańskiej. Organizator popularyzacyjnych demonstracji w ww. laboratorium. Członek *The EEG Research Hub* – grupy prowadzonej przez prof. Roba van Lubbe.

B.3. Podsumowanie

Moim zadaniem całkowita aktywność naukowo-badawcza Pana dr inż. Szymona Drgasa zasługuje na jednoznacznie pozytywną ocenę. Przekonuje o tym:

- 1) bardzo aktywna, międzynarodowa współpraca badawcza i publikacyjna (szczególnie z Prof. T. Virtanenem z University of Tampere),
- 2) bardzo owocna współpraca z zapleczem/sektorem gospodarczym (firma StethoMe),
- 3) bardzo wyraźne poprawienie wskaźników nauko-metrycznych w ostatnim okresie (prace bardziej samodzielne w czasopiśmie wysokiej rangi).

Niedosyt stanowi jedynie brak udziału Habilitanta w projektach NCN i NCBiR po obronie doktoratu, tak w charakterze wykonawcy jak i kierownika. Co w pewnym stopniu kompensuje współpraca z firmą StethoMe.

C. Wniosek końcowy

Uwzględniając pozytywną ocenę, wystawioną osiągnięciu habilitacyjnemu (monotematycznemu cyklowi 13 publikacji) chciałbym jednoznacznie poprzeć wniosek o nadanie Panu dr inż. Szymonowi Drgasowi stopnia naukowego doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych, w dyscyplinie automatyka, elektronika, elektrotechnika i technologie kosmiczne. Przekonuje mnie o tym dodatkowo pozytywna ocena wystawiona także całej aktywności naukowo-badawczej Habilitanta (choć z punktu formalnego nie jest ona wymagana).

Tomasz Zieliński

Kraków, 7.01.2025