

# Phonak Field Study News.

**AutoSense OS™ 7.0 poprawia rozumienie mowy dzięki dostarczaniu wysokiej jakości dźwięku w trudnych warunkach słuchowych**

**Badanie kliniczne przeprowadzone w Phonak Audiology Research Center (PARC) wykazało statystycznie istotną poprawę rozumienia mowy w hałasie rzeczywistego otoczenia w przypadku urządzeń nowej generacji Infinio R i Sphere™ Infinio, które mają udoskonaloną dokładność klasyfikacji środowiskowej.**

Wright, A., Cox, M., Keller, M., Galster, J. sierpień 2025.

## Wprowadzenie

Phonak od 25 lat wykorzystuje sztuczną inteligencję (AI) w aparatach słuchowych. Pierwszą komercyjną aplikację wprowadzono w aparacie Phonak Savia z automatycznym systemem AutoSelect, a następnie zastąpiono ją systemem SoundFlow w aparacie Phonak Exélia. W 2014 roku na rynku pojawiła się pierwsza wersja systemu operacyjnego AutoSense OS (ASOS), który wykorzystywał analizę otoczenia dźwiękowego w czasie rzeczywistym do sterowania przetwarzaniem sygnału. Badania kliniczne nad skutecznością oryginalnego programu ASOS wykazały, że zapewnia on poprawę stosunku sygnału do szumu (SNR) o 1,3 dB w zakresie rozumienia mowy w hałasie w porównaniu z programami manualnymi.<sup>1</sup>

Od momentu wprowadzenia na rynek system ASOS był stale udoskonalany, a jego możliwości były jeszcze większe dzięki zwiększeniu dokładności klasyfikacji dźwięku i zapewnieniu dostępu do jeszcze większej liczby funkcji, aby zmaksymalizować korzyści i satysfakcję użytkownika, niezależnie od środowiska akustycznego. Obecnie ASOS obejmuje sześć programów do sygnałów akustycznych na wejściu: Spokojna sytuacja, Mowa w hałasie, Mowa (sferyczna) w dużym hałasie, Mowa w samochodzie, Komfort w hałasie i Muzyka. Cztery z tych sześciu programów bazują na obecności mowy na różnych poziomach i różnych rodzajach hałasu tła (w tym na jego braku), co sprawia, że

komunikacja jest najważniejszym elementem celów projektowych ASOS.

ASOS 7.0 został przeszkolony na 18 razy większej liczbie środowisk dźwiękowych niż poprzednia generacja, co czyni go jak dotąd najdokładniejszą implementacją technologii AutoSense – ze wzrostem dokładności klasyfikacji o 24%, zgodnie z pomiarami technicznymi.<sup>2</sup> Aby podkreślić znaczenie dokładności klasyfikacji, w badaniu klinicznym oceniono wpływ, jaki miałyby błędne zaklasyfikowanie mowy w hałasie jako zwykłego hałasu, na rozumienie mowy i jakość dźwięku. Porównanie to odzwierciedla dwa najbardziej prawdopodobne wybory automatycznego układu kierowniczego w hałaśliwym otoczeniu akustycznym. Dobrze wyszkolony system, który odczytuje dużą ilość zróżnicowanej mowy w nagraniach z hałasem, będzie prawidłowo rozpoznawał scenę, stosując klasę dźwięków, która redukuje hałas i poprawia rozumienie mowy. Z kolei mniej zaawansowany system może mieć trudności z rozróżnianiem wzorców mowy w hałasie, co może skutkować zastosowaniem klasy dźwięków zaprojektowanej pod kątem komfortu słuchania, która redukuje hałas, lecz nie poprawia rozumienia mowy.

## Metodologia

### Uczestnicy

Do badania włączono siedemnaście dorosłych użytkowników aparatów słuchowych ze średnim do średnio ciężkim obustronnym ubytkiem słuchu w wieku od 62 do 92 lat (*średnia wieku* = 76 ±8). Przed badaniem została przeprowadzona otoskopia, a w razie potrzeby przeprowadzano różne formy postępowania z woskowiną.

### Aparaty słuchowe

Para aparatów słuchowych typu RIC (z zewnętrzną słuchawką) Audéo Sphere™ Infinio 90 została zaprogramowana zgodnie z zaleceniami programu do dopasowywania Target, z uwzględnieniem formuły dopasowania aparatów słuchowych Phonak (APD 3.0) i proponowanego typu nasadki. Siedmiu uczestnikom założono obustronnie nakładki typu Power, sześciu nakładki wentylowane, a czterem uczestnikom jedną kopułę Power i jedną wentylowaną. U wszystkich uczestników przeprowadzono badanie sprzężenia zwrotnego i pomiary w uchu rzeczywistym, stosując protokół obejmujący pomiary reakcji ucha rzeczywistego bez aparatu słuchowego (REUR), reakcji ucha rzeczywistego z zamkniętym uchem (REOR), reakcji ucha rzeczywistego wspomaganego aparatem słuchowym (REAR) i maksymalnym poziomie wyjściowym (MPO). Nie wykonano żadnego dokładnego strojenia.

## Porównania

Porównano programy manualne dla Mowy w dużym hałasie (MwDH) i Sferycznej mowy w dużym Hałasie (SMwDH) z programem manualnym Komfort w Hałasie (KwH).

SiLN i SSiLN to klasy dźwięków wybrane przez ASOS 7.0 dla testowanej sceny dźwiękowej w pomocy nowej generacji Infinio 90R i Sphere™ Infinio. Poprzednie wersje ASOS były trenowane na mniejszych zestawach nagrań z rzeczywistych warunków, w związku z czym mogły nadmiernie klasyfikować mowę w hałaśliwym otoczeniu jako KwH. Zatem KwH oznacza klasę dźwięku wybraną przez system, który nie został przeszkolony przy użyciu tak wielu rzeczywistych nagrań mowy w hałasie, jak ASOS 7.0.

Programy te różnią się kształtowaniem wzmocnienia i zastosowaniem zaawansowanych funkcji oczyszczania dźwięku, takich jak mikrofony kierunkowe i redukcja szumu (RSz). Tabela 1 podsumowuje różnice w funkcjach między programami.

	KwH	MwDH	SMwDH
Kierunkowość Mik	Real-Ear Sound	StereoZoom 2.0	Stała kierunkowość
Dynamiczna Redukcja Szumu	--	Tak	--
Sferyczna Wyrazistość Mowy	--	--	Tak
NoiseBlock	12	8	--

Tabela 1. Podsumowanie zestawów funkcji dostępnych w trzech programach testowych.

Technologia KwH stosuje globalną redukcję wzmocnienia w całym zakresie częstotliwości dla wszystkich poziomów wejściowych i wykorzystuje konfigurację mikrofonu kierunkowego, która symuluje efekty działania małżowiny usznej (Real-Ear Sound). Wykorzystuje także umiarkowane ustawienie funkcji NoiseBlock – funkcji redukcji szumów, która jest najskuteczniejsza w przypadku stacjonarnych źródeł hałasu.

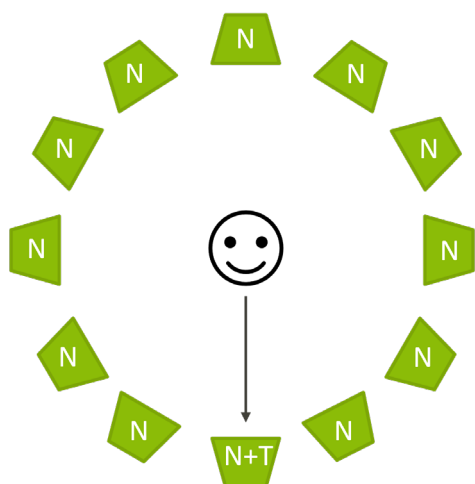
Technologia MwDH również wykorzystuje technologię NoiseBlock, ale wzbogacono ją o technologię Dynamicznej Redukcji Szumu, która współpracuje z mikrofonem kierunkowym. W tym przypadku trybem mikrofonu jest adaptacyjna obuuszna charakterystyka kierunkowa – StereoZoom 2.0 – zdolna do poprawy współczynnika SNR nawet o 6,4 dB.<sup>3</sup>

Technologia SMwDH łączy stałą kierunkowość mikrofonu z redukcją szumu (Rsz) sterowaną przez głęboką sieć neuronową o nazwie Sferyczna Wyrazistość Mowy, która może poprawić współczynnik SNR nawet o 10,2 dB.<sup>3</sup>

Wszystkie te programy mają na celu redukcję hałasu i zapewnienie komfortu słuchania. Jednakże tylko MwdH i SMwDH zostały zaprojektowane specjalnie z myślą o jednoczesnej poprawie rozumienia mowy – MwdH poprzez kierunkowość, a SMwDH poprzez redukcję szumu opartą na sieciach neuronowych (DNN).

### Rozumienie mowy

Rozumienie mowy w hałasie mierzono przy użyciu szumu ambisonicznego wyższego rzędu o natężeniu 70 dBA (sumarycznie) emitowanego przez wszystkie głośniki w układzie 12 głośników, przy czym zdania AzBio<sup>4</sup> były prezentowane pod kątem 0°. Hałas symulował typowe i trudne warunki akustyczne – rozmowę w samolocie na wysokości przelotowej. Nagranie zostało wykonane z prawdziwego samolotu i zawierało typowe dźwięki przejściowe, a także dominujący dźwięk silnika.



Rysunek 1. Schemat konfiguracji testowej. W kabine dźwiękochłonnej zastosowano zestaw 12 głośników. Szum ambisoniczny (Sz) odtwarzany ze wszystkich 12 głośników. Mowa docelowa (D) została umieszczona w szumie pod kątem 0°.

Zdania przedstawiono na jednym z pięciu ustalonych poziomów SNR, ustalonych na podstawie indywidualnych wyników w adaptacyjnym zadaniu słownym spółgłoska-jądro-spółgłoska (SJS) przy użyciu najlepiej wspomaganym warunków – w tym przypadku SMwDH. Zadanie SJS miało na celu zidentyfikowanie SRT-50 każdego uczestnika, czyli progu SNR, przy którym osiągnął 50% poprawnych odpowiedzi. Wartość SRT-50 każdej osoby dopasowano do najbliższego stałego poziomu SNR spośród pięciu dostępnych opcji: -5, -3, 0, +3 i +5 dB. Czternastu z siedemnastu uczestników poddano testowi przy -5 dB, a pozostałych trzech przy -3 dB.

Po zakończeniu adaptacyjnego zadania SJS i wybraniu stałego stosunku sygnału do szumu (SNR) do testowania, każdy uczestnik wypełnił jedną ćwiczebną listę zdań AzBio oraz dwie listy na każdy warunek testowy, co łącznie dało siedem list na uczestnika. Dla wszystkich uczestników zastosowano te same siedem list, równoważąc warunki i równoważąc listy pomiędzy warunkami w celu zminimalizowania kolejności i efektów listy.

Postawiono hipotezę, że aktywacja MwdH lub SMwDH będzie prowadzić do lepszego rozumienia mowy w porównaniu do sytuacji, gdy uczestnicy korzystali z programu Kwh. Hipotezę tę sprawdzono przy użyciu liniowego modelu efektów mieszanych. Warunki testu były kontrolowane poprzez połączenie z aplikacją myPhonak. Uczestnik i badacz oceniający test nie znali warunków.

Oprócz pomiaru zrozumienia mowy za pomocą testu behawioralnego, uczestników poproszono o subiektywną ocenę zrozumienia mowy po każdym warunku testu. Oceny przyznawano w 5-stopniowej skali, gdzie 1 oznaczało, że *wypowiedź była zupełnie niezrozumiała*, a 5, że *wypowiedź była tak zrozumiała, jak to możliwe*.

### Jakość dźwięku

Oceny jakości dźwięku zebrano przy użyciu narzędzia porównawczego A/B. Uczestnicy korzystali z tabletu, aby przełączać się między programami aparatów słuchowych w każdym porównaniu parzystym (Kwh vs. MwdH oraz Kwh vs. SMwDH). Następnie oceniali programy według następującej 5-punktowej skali:

- A jest o wiele lepszy niż B
- A jest nieco lepszy niż B
- A jest taki sam jak B
- B jest nieco lepszy niż A
- B jest o wiele lepszy niż A

Uczestnicy wystawili oceny obu parom porównawczym, uwzględniając siedem wymiarów jakości dźwięku: zrozumiałość mowy, naturalność mowy, głośność mowy, zakłócanie hałasu, wysiłek słuchowy, ogólną jakość dźwięku i preferencje. Warunki równoważono w ramach porównań parami, a każdy wymiar mierzono dwukrotnie.

Ta sama scena dźwiękowa na tym samym poziomie głośności została użyta zarówno do oceny jakości dźwięku, jak i do zadania rozumienia mowy. Współczynnik SNR był również taki sam pomiędzy zadaniami. Mowa docelowa w ocenie jakości dźwięku to naturalny fragment wypowiedzi wygłoszony przez kobietę, która wykorzystuje realistyczny wysiłek wokalny. Fragment trwał mniej niż dwie minuty i

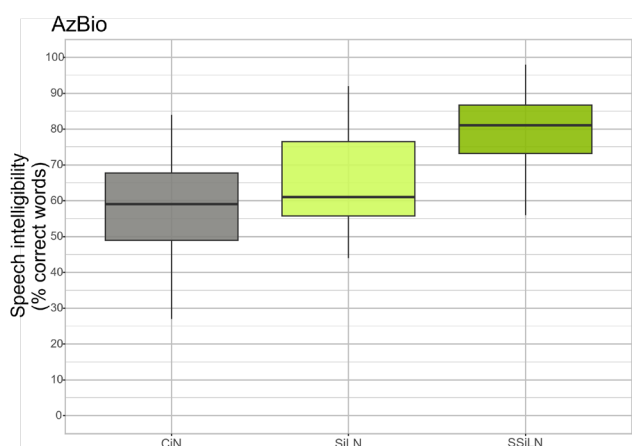
można go było powtarzać, jeśli uczestnik potrzebował więcej czasu.

## Wyniki

### Rozumienie mowy

MwDH uzyskała lepsze wyniki niż KWH w zadaniu AzBio, znacząco zwiększając poprawność identyfikacji słów średnio o 8 punktów procentowych ( $p < 0,001$ ).

Aktywacja SMwDH znacząco zwiększyła także poprawną identyfikację słów, średnio o 22 punkty procentowe ( $p < 0,001$ ).

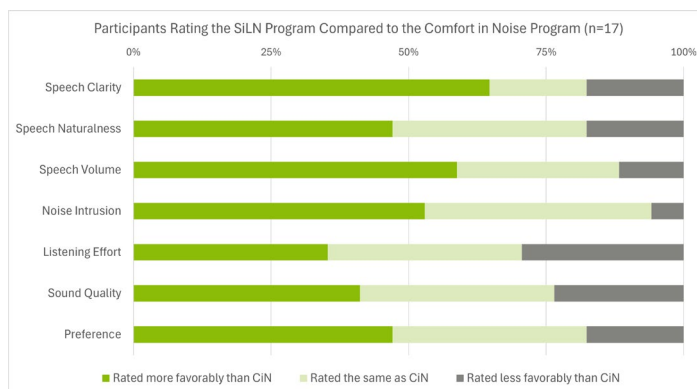


Rysunek 2. Wyniki rozumienia mowy dla wszystkich trzech warunków u 17 uczestników. KWH uzyskał średnią na poziomie 58% ( $\pm 15$ ). MwDH uzyskała średnią na poziomie 66% ( $\pm 14$ ). SMwDH uzyskała średnią na poziomie 80% ( $\pm 11$ ).

W ocenie subiektywnego rozumienia mowy 88% uczestników uznało MwDH za tak samo zrozumiałą lub bardziej zrozumiałą niż KWH. W przypadku SMwDH odsetek ten wzrósł do 94%.

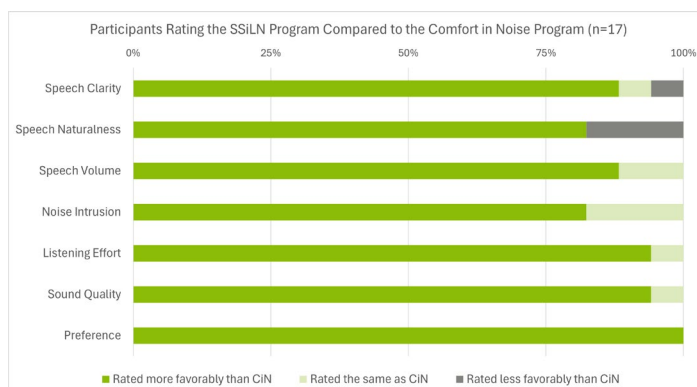
### Jakość dźwięku

Większość uczestników oceniła MwDH tak samo lub korzystniej niż KWH we wszystkich wymiarach jakości dźwięku. Najwyżej oceniono głośność mowy i przeszkadzający hałas, co wskazuje na odczuwalną poprawę współczynnika SNR, potwierdzoną wynikami badań behawioralnych.



Rysunek 3. Oceny jakości dźwięku dla programu MwDH w porównaniu z programem KWH.

SMwDH również został oceniony wysoko – pięć z siedmiu wymiarów zostało jednomyślnie ocenionych tak samo lub lepiej niż KWH: głośność mowy, przeszkadzający hałas, wysiłek związany ze słuchaniem, ogólna jakość dźwięku i preferencje. Co jeszcze bardziej imponujące, 94% uczestników oceniło SMwDH lepiej niż KWH pod względem wysiłku związanego ze słuchaniem i jakości dźwięku. Zgodnie z oczekiwaniami, w przypadku tych ocen, preferencja dla programu SMwDH była duża – 100% uczestników preferowało program SMwDH względem programu KWH.



Rysunek 4. Oceny jakości dźwięku dla programu SMwDH w porównaniu z programem KWH.

## Podsumowanie

Technologia MwDH spełniła oczekiwania, zapewniając jakość dźwięku równą lub lepszą niż technologia KWH, a jednocześnie zapewniając dodatkowe korzyści w zakresie rozumienia mowy dzięki funkcjom StereoZoom 2.0 i dynamicznej redukcji szumu. Ponieważ technologia StereoZoom poprawia stosunek sygnału do szumu dzięki kierunkowości, można się spodziewać, że przewaga MwDH nad KWH w rozumieniu mowy będzie jeszcze większa, gdy mowa będzie oddzielona przestrzennie od szumu.

W związku z tym SMwDH uzyskała lepsze wyniki niż KWH, a nawet lepsze niż MwDH. Technologia SMwDH została

zaprojektowana tak, aby umożliwiała rozumienie mowy bez względu na konfigurację przestrzenną. Intensywne szkolenie głębokiej sieci neuronowej pozwala również na zachowanie większej ilości docelowej mowy podczas oczyszczania dźwięku, czego dowodem są niemal idealne wyniki jakości dźwięku uzyskane przy SMwDH.

Wyniki te potwierdzają tezę, że nie wszystkie hałaśliwe środowiska należy traktować tak samo. W przypadku obecności mowy globalna redukcja głośności bez aktywacji zaawansowanych funkcji oczyszczania dźwięku może skutkować utratą możliwości komunikacji i obniżeniem wydajności. Dlatego ważne jest, aby algorytmy oparte na sztucznej inteligencji, takie jak AutoSense OS, były nadal udoskonalane, dzięki czemu użytkownicy końcowi będą mogli cieszyć się pełnią korzyści płynących z używania aparatów słuchowych.

## Bibliografia

1. Übelacker, E., Tchorz, J., Latzel, M. (2015). AutoSense OS: Benefit of the next generation of technology automation. Artykuł Phonak Field Study News, dane pozyskane ze strony <https://www.phonak.com/evidence>
2. Sanchez, C., & Giurda, R. (2025). System operacyjny AutoSense OS™ 7.0 dostosowuje się o 24% precyzyjniej do warunków akustycznych, dzięki czemu użytkownicy mogą cieszyć się wyjątkową jakością dźwięku w każdej chwili. Phonak Insight, dostępne na stronie <https://www.phonak.com/en-int/professionals/audiology-hub/evidence-library>
3. Rauffer, S., Kohlhauer, P., Jehle, F., Kühnel, V., Preuss, M., Hobi, S. (2024). Spheric Speech Clarity proven to outperform three key competitors for clear speech in noise. Artykuł Phonak Field Study News, dane pozyskane ze strony: <https://www.phonak.com/evidence>
4. Spahr, A. J., Dorman, M. F., Litvak, L. M., Van Wie, S., Gifford, R. H., Loizou, P. C., Loiseau, L. M., Oakes, T., & Cook, S. (2012). Development and validation of the AzBio sentence lists. *Ear and hearing*, 33(1), 112–117. <https://doi.org/10.1097/AUD.0b013e31822c2549>

## Autorzy i badacze

należy zarządzanie wewnętrznymi badaniami klinicznymi z udziałem osób dorosłych.

### Wewnętrzny badacz/Autor

**Ashley Wright, dr audiologii, starsza audiolog**



**badawcza** Ashley jest starszym audiologiem badawczym w Phonak Audiology Research Center (PARC), Aurora, Illinois. Ukończyła studia doktoranckie na Uniwersytecie Rush w Chicago i dołączyła do PARC w 2018 roku. Do jej głównych obowiązków

### Wewnętrzny badacz/Autor

**Marina Cox, dr audiologii, audiolog badawcza**



Marina jest audiologiem badawczym w Phonak Audiology Research Center (PARC), Aurora, Illinois. Ukończyła studia doktoranckie na Northwestern University w Chicago i dołączyła do PARC w 2024 roku.

### Wewnętrzny badacz/Autor

**Matthias Keller, doktor, badania audiologiczne**



Matthias Keller uzyskał doktorat z psychologii na Uniwersytecie w Zurychu w Szwajcarii, badając związane z wiekiem różnice w przetwarzaniu języka mówionego przez neurony. Po pracy w firmie Phonak/Sonova w Szwajcarii od 2019

roku, przeniósł się do Stanów Zjednoczonych i dołączył do zespołu PARC w 2023 roku.

### Główny badacz

**Jason Galster, doktor, wiceprezes ds. strategii badań klinicznych**



Jason współpracuje z globalną siecią zespołów zajmujących się badaniami klinicznymi, zajmując się rozwojem i badaniami naukowymi nad technologią słuchową jutra. Jest uznanym na świecie autorem,

wykładowcą i profesorem nadzwyczajnym audiologii, którego zainteresowania badawcze obejmują metody statystyczne służące charakteryzowaniu zmienności indywidualnej oraz pomiar wyników słuchowych w życiu codziennym.

# Phonak Field Study News●

## Jednostronne podsumowanie

**AutoSense OS™ 7.0 poprawia rozumienie mowy dzięki wysokiej jakości dźwięku w trudnych warunkach słuchowych**

**Badanie kliniczne przeprowadzone w Phonak Audiology Research Center (PARC) wykazało statystycznie istotną poprawę rozumienia mowy w hałasie otoczenia w przypadku urządzeń nowej generacji Infinio R i Sphere™ Infinio dzięki ulepszonej dokładności klasyfikacji środowiskowej.**

Wright, A., Cox, M., Keller, M., Galster, J. July 2025.

### Główne wnioski

- System operacyjny AutoSense OS 7.0 dostosowuje się o 24% precyzyjniej do warunków akustycznych, dzięki czemu użytkownicy mogą cieszyć się wyjątkową jakością dźwięku w każdej chwili.
- 100% użytkowników Sphere™ Infinio preferuje system AutoSense OS 7.0 od poprzedniej generacji.
- Prawidłowa klasyfikacja sytuacji słuchowej poprawiła rozumienie mowy o 8% w przypadku Infinio R i o 22% w przypadku Sphere™ Infinio.

### Wnioski praktyczne

- Na podstawie ustaleń niniejszego badania oczekuje się, że klienci, którym założono aparaty słuchowe nowej generacji Infinio R i Sphere™ Infinio, doświadczą lepszego rozumienia mowy w hałasie bez pogorszenia jakości dźwięku.
- Protetycy słuchu mogą być pewni, że ich pacjenci będą zadowoleni z AutoSense OS 7.0 jako programu startowego.

- Marka Phonak wierzy w model ciągłego doskonalenia, co oznacza, że funkcje sztucznej inteligencji, takie jak AutoSense OS, zawsze będą najnowocześniejsze.