

Express Yourself with Whispp

Real-time On-Device
Voice Reconstruction AI

July 2026



Qualcomm

Executive Summary

For millions of people living with voice conditions, everyday communication can be an ongoing challenge. Conditions such as spasmodic dysphonia, vocal cord paralysis, total laryngectomy after throat cancer make speech unclear or strained, limiting participation in professional, social, and educational settings. Also severe stuttering can have a huge impact on someone’s life.

To address this challenge, Qualcomm for Good collaborated with Whispp, a startup specializing in real-time voice AI technology, to pilot an innovative, on-device solution that reconstructs natural speech from whispered or impaired input. As part of this initiative, 30 participants across multiple countries were equipped with Snapdragon® X Elite-powered laptops running Whispp’s software, enabling real-time, audio-to-audio voice transformation.

By enabling more seamless and inclusive communication, the solution helped participants re-engage in conversations with greater confidence, independence, and connection to others.

Findings at a Glance

74%

Of participants would use video calling more if they did not have a voice condition.

84%

Of participants rated their AI laptop an 8 or higher out of 10 when asked how likely they were to recommend it to a friend or family member.

96%

Of participants said they would use it when it is commercial product.



Letitia

“Thank you for the opportunity to participate and rediscover what it is like to be a part of a conversation, instead of sitting back or trying to type a response and missing the moment because the conversation moved on. So thank you from the bottom of my heart.”

The Challenge: Understanding Voice Conditions at Scale

Voice conditions are more common than often recognized and can have a significant impact on individuals' ability to communicate and participate fully in daily life.

29%

Of individuals experience a voice condition in their lifetime. ¹

\$2 Billion

In lost productivity annually due to voice conditions. ²

3.5 billion

People will need at least one assistive technology product by 2050. ³

The Variability of Voice Conditions

Voice disorders affect the vocal cords and change how the voice sounds, causing hoarseness, weakness, or pitch changes.

They may be temporary, such as during a flu or long-lasting

These conditions may be temporary, such as during an illness, or long-lasting.

Voice disorders generally fall into three categories:

Functional

Result from voice misuse or tension without structural damage (e.g., muscle tension dysphonia).

Organic

Arise from physical changes or damage to the vocal folds or surrounding structures, such as nodules, polyps, cysts, or laryngeal cancer.

Neurological

Occur when nerve problems affect vocal fold movement, as seen in vocal fold paralysis or spasmodic dysphonia.

“

Globally, 300 million people struggle to make themselves heard due to voice conditions. This leads to social withdrawal, professional exclusion, and mental health challenges, including heightened rates of anxiety and depression.⁴

The Technology Gap in Assistive Speech

Assistive technology is any item, piece of equipment, software program, or product system that is used to increase, maintain, or improve the functional capabilities of individuals with disabilities.³

For people with voice conditions, assistive technology can help support or replace natural speech. Examples include voice amplifiers, which make speech louder and reduce strain or text-to-speech apps, which allow users to communicate when speaking is difficult or not possible.³

While these technologies help people with voice conditions neither are a complete solution. Voice amplifiers simply make affected speech louder and are often incompatible with technology.

Meanwhile, text-to-speech to speech-to-text options have a higher latency and lack ease of use which creates barriers to conversations with a natural 'flow'.

Most speech enhancement, noise-cancelling and Automatic Speech Recognition systems are designed for loud, spoken voices, not for whispers, low-volume, atypical or affected speech. In fact, sometimes affected speech is not recognized by speech recognition systems. This leaves technology inaccessible for people with voice conditions.⁵⁻⁶

AI Meets Assistive Technology

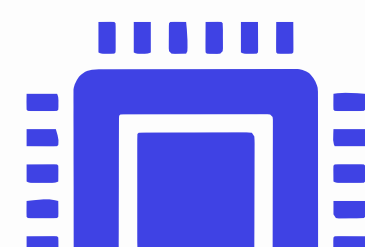
Why Whispp is Different

Unlike conventional assistive speech tools, which often rely on text-to-speech conversion or simple voice amplification, Whispp employs proprietary on-device Voice AI to transform an individual's impaired speech into a natural, expressive voice in real time.

By performing inference directly on the device rather than relying on continuous cloud connectivity, Whispp delivers low-latency, reliable speech transformation that works even in bandwidth-constrained or privacy-sensitive environments. This edge-first approach enables more fluid, natural conversations and reduces the friction that often disrupts assistive communication tools.

Edge AI

Mobile and PC intergration



What is Voice AI

Traditional voice processing tools such as noise suppression, spectral masking, or voice isolation are effective at removing unwanted sound, but they cannot restore what is missing from affected speech. Whispp's advanced Voice AI goes further.

Using real-time, audio-to-audio inference at the edge, Whispp reconstructs and restores speech by reviving lost pitch, tone, and natural intonation directly on the device. Whether dealing with whispers or affected voices, the technology intelligently rebuilds missing vocal characteristics to deliver clear, lifelike sound.

Because processing happens locally, users experience immediate feedback without the delays associated with cloud-based speech pipelines. The result is speech that not only sounds clear, but feels authentic, expressive, and true to the speaker's original voice.

Real time conversation

50 ms



Private, Personal and Responsible

Ethically-Collected Dataset

Whispp trains its models on an ethically collected proprietary dataset and never trains on user data.

This on-device AI architecture further strengthens privacy and trust by ensuring sensitive voice data remains on the user's device, rather than being transmitted or stored externally. By combining ethical data practices with local inference, Whispp aligns responsible AI development with real-world accessibility needs.

Language independent

Audio-to-Audio based AI



Personal voice & accent

Preserved speaker identity



Voice input → On-device processing → Reconstructed voice output

Powered by Snapdragon X Elite

Whispp's AI-powered voice technology was optimized to run efficiently on Snapdragon X Elite platform, leveraging the dedicated Neural Processing Unit (NPU) to deliver high performance with significantly lower power consumption.

Key Finding: Extended Battery Life with On-Device AI

Running Whispp on the NPU more than doubles battery life compared to CPU-based processing:

~310 minutes on NPU

~130 minutes on CPU

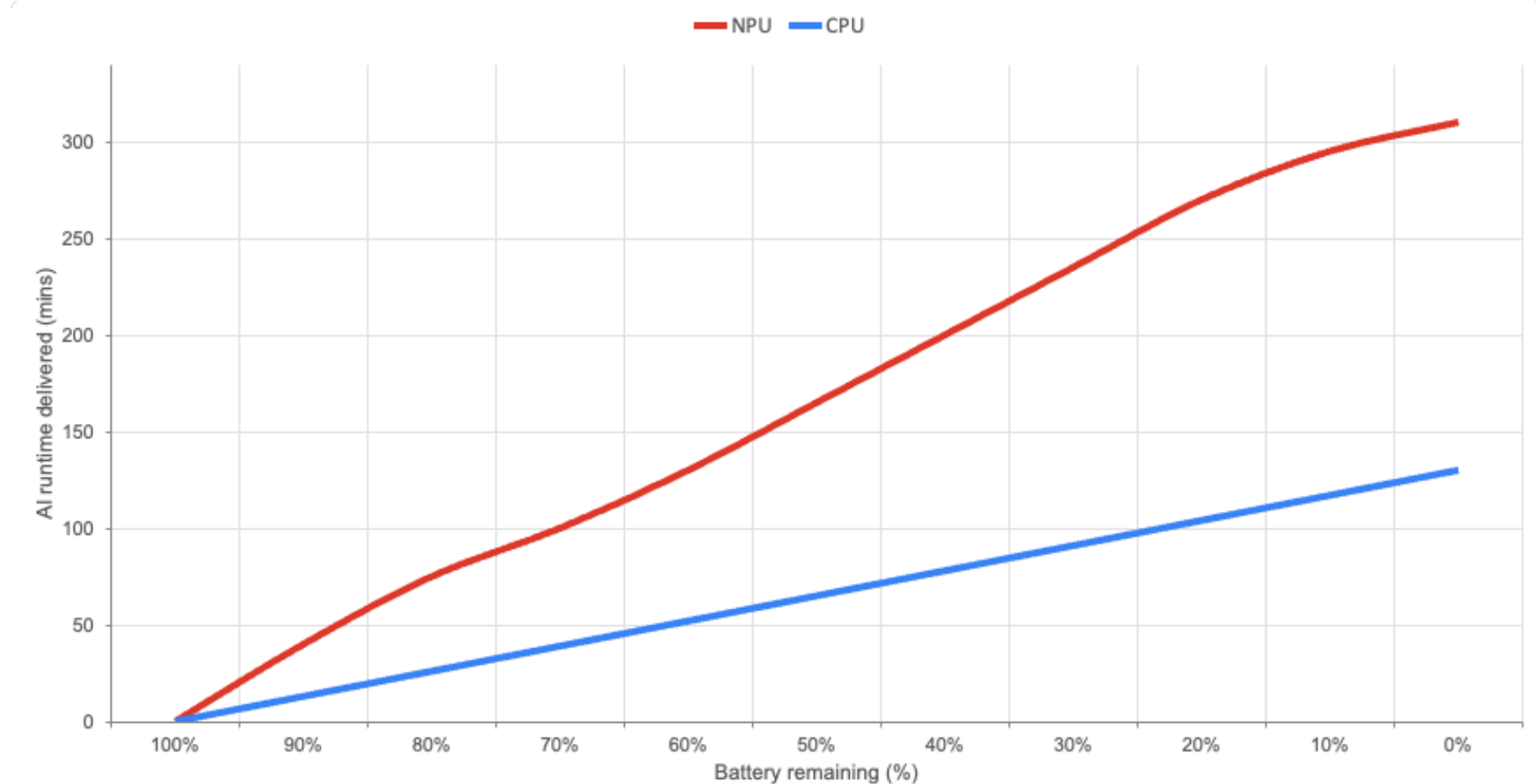
This demonstrates that offloading AI workloads to the NPU significantly reduces energy consumption, enabling longer, uninterrupted use on battery-powered devices.

Why This Matters

Efficient on-device AI processing is critical for real-time communication tools. By utilizing the NPU:

- Users benefit from longer device usage without frequent charging
- Performance remains consistent during extended use
- The solution becomes more practical for daily, continuous

Whispp Battery Consumption over Time: NPU vs. CPU



Comparison of battery percentage over time when running an AI application on the CPU and the NPU, highlighting significantly improved energy efficiency with NPU processing

All measurements were conducted under consistent conditions on the same device:

- Device fully recharged between tests.
- Battery percentage recorded at fixed intervals.
- Tests conducted in high-performance mode.
- Final readings at 0% battery are approximate.



Qualcomm® AI optimized NPU

Enabling Reliable, Real-Time Speech Processing

Real-time communication depends on consistent, low-latency performance. Testing shows that running Whispp on Qualcomm NPU delivers significantly more stable and reliable processing compared to CPU-based execution.

Key Finding: Lower and More Stable Latency

- The NPU maintains consistent latency (~48–50 ms) with minimal variation
- The CPU shows significant fluctuations, with latency spikes up to ~200 ms
- CPU-based processing also introduced streaming disruptions, where real-time performance could not be maintained

These disruptions were not observed when using the NPU.

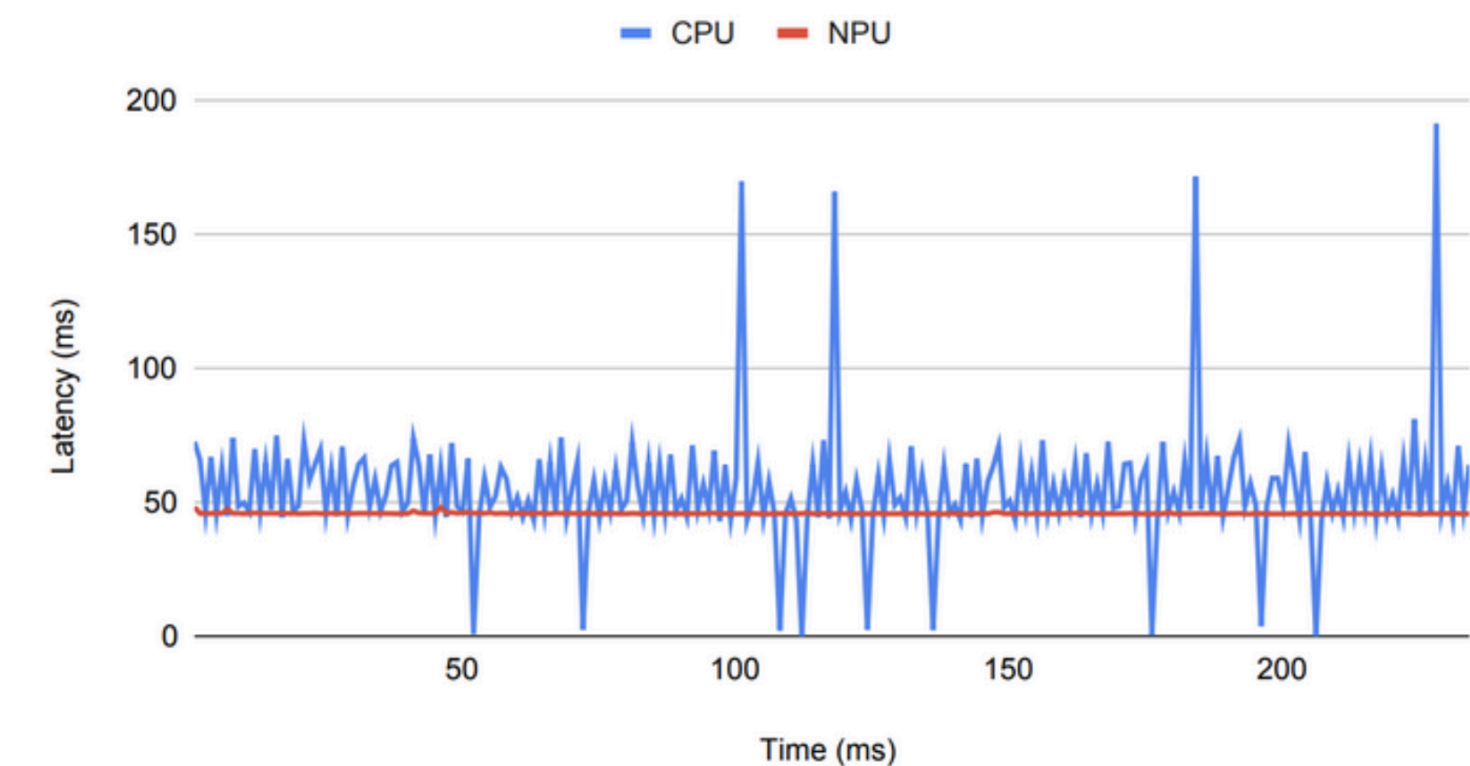
Performance Benefits of NPU Processing

Using the NPU enables:

- Lower latency per frame
- Consistent processing speed over time
- Ability to support more complex or multiple AI models simultaneously
- Improved overall audio quality and streaming performance

All measurements were taken with the laptop in high-performance mode.

Whispp Latency when Powered by CPU vs NPU

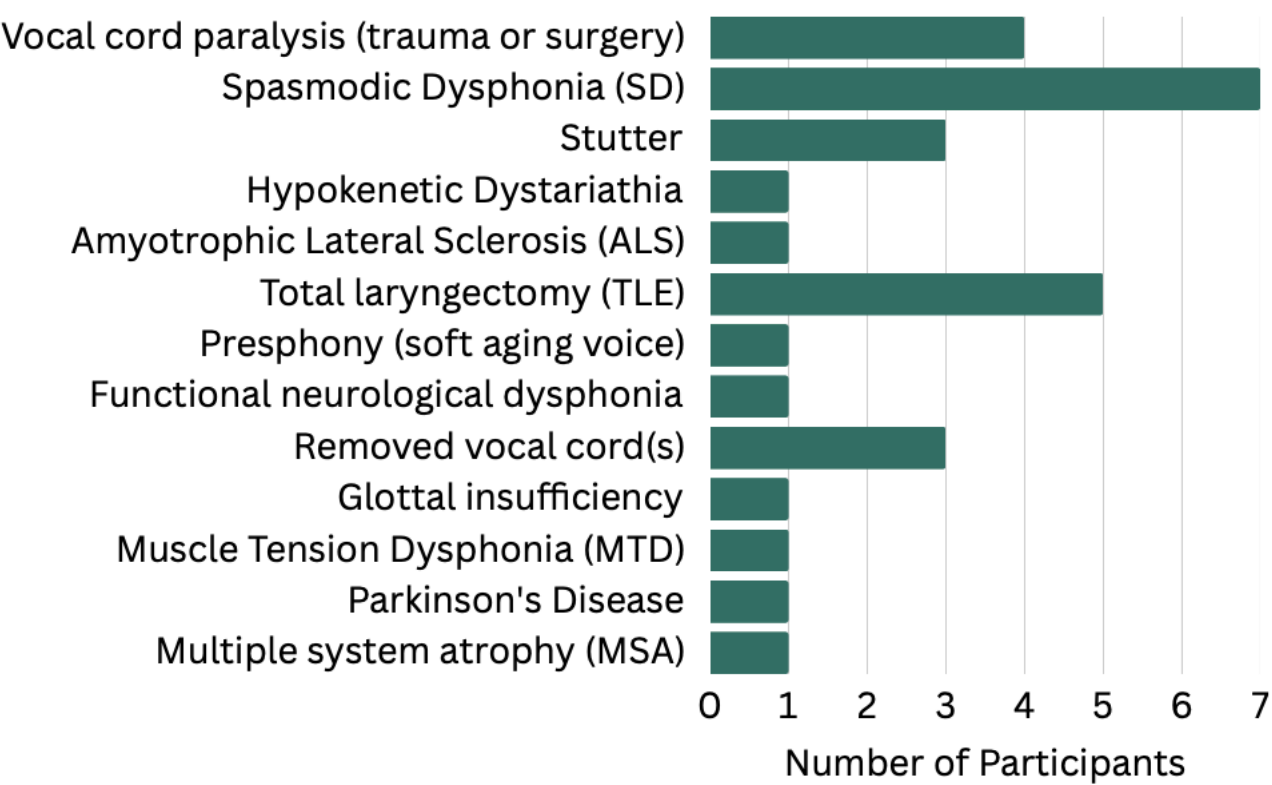


CPU vs NPU latency during sustained AI processing, showing significantly greater stability with NPU execution.

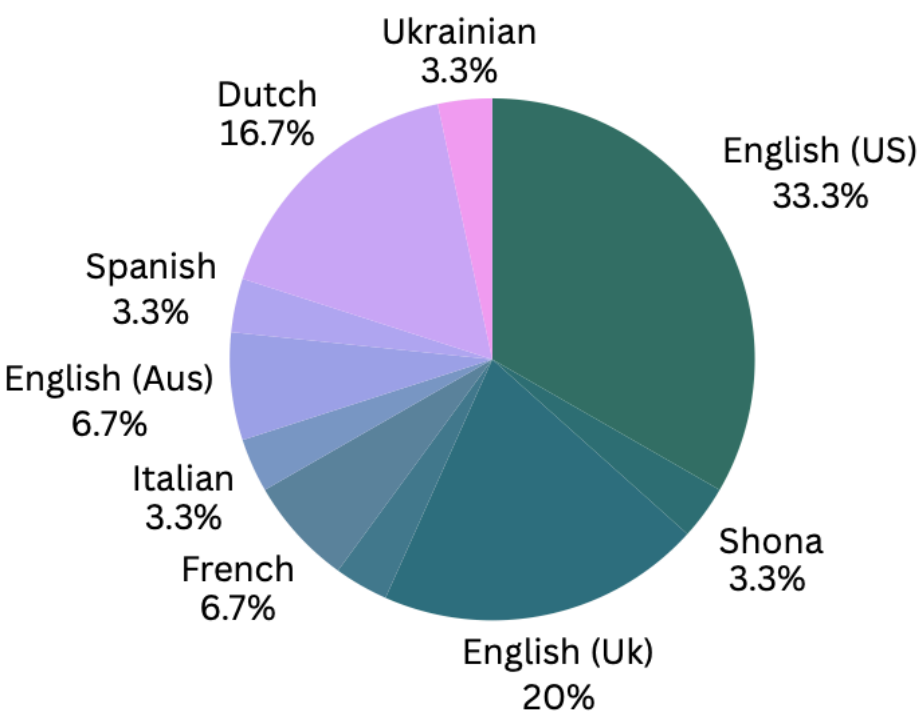
Project Overview: Methodology & Participants

A total of 30 participants across nine countries were selected, representing a wide range of voice conditions, languages, and communication needs. Participants were chosen to reflect:

13 different voice conditions



10 languages



Scope and Diversity of Use

Thirteen different voice conditions and ten languages were represented across participants. This breadth reinforced the scalability of the solution and its relevance for multilingual, inclusive communication environments.

Technology Development

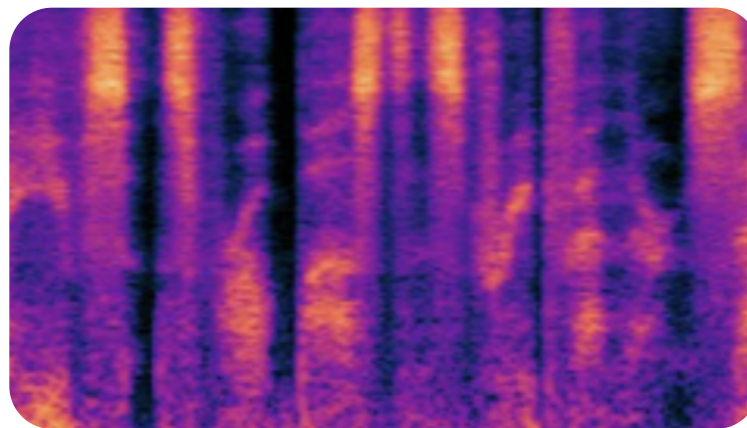
Each participant received a laptop powered by Snapdragon X Elite, with Whispp pre-installed. Participants integrated the system into their daily routines and used it across professional, social, and personal communication scenarios. By running speech reconstruction directly on the device, the deployment delivered consistent, low-latency performance across regions with varying connectivity conditions, demonstrating the practical advantages of edge-based AI for assistive communication.

Feedback and Support

The pilot ran for six weeks, with weekly check-ins conducted through surveys and one-to-one online video sessions. Technical support was provided on demand through remote assistance throughout the duration of the project. Structured feedback loops enabled ongoing assessment of reliability, user experience, and real-time performance in everyday use, rather than isolated testing scenarios.

How Whispp's AI Reconstructs the Human Voice

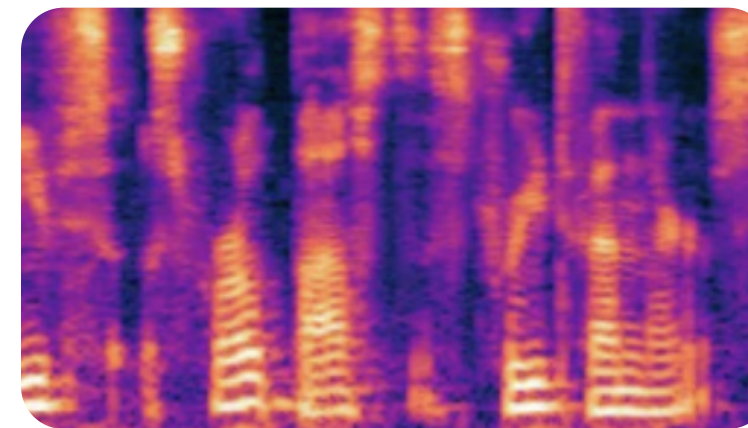
Whispered Speech



Whispered speech retains some formant information, allowing words to remain intelligible, but loses key features of speech:

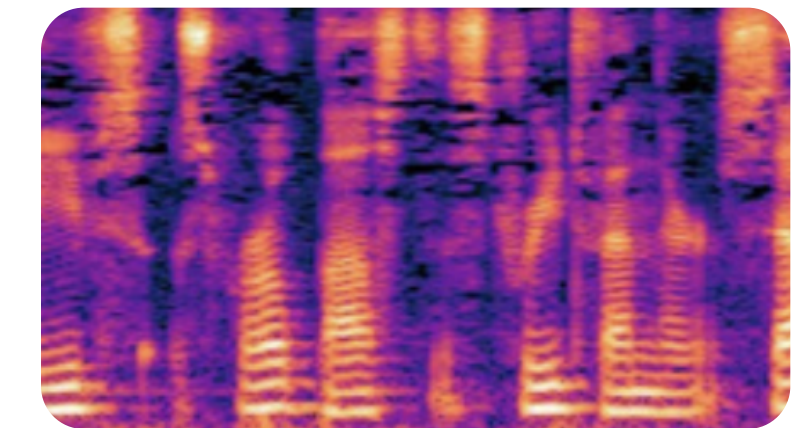
- Fundamental frequency (F_0): no visible pitch band because the vocal folds do not vibrate.
- Harmonic structure: replaced by diffuse, noisy energy.
- Natural intonation: speech sounds flat and monotone since there's no pitch variation.

Whispp Reconstructed Speech



VS

Voiced Speech



Whispp's proprietary voice AI restores the key features of natural spoken speech from a whispered or affected voice sample:

- Reconstructs fundamental frequency (F_0): visible as the lowest horizontal band representing vocal fold vibrations (pitch).
- Reconstructs harmonic structure: visible as multiple frequency bands stacked above the F_0 , giving richness and timbre to the speech.
- Enhances formant clarity, improving intelligibility.
- Reintroduces prosody and emotional expressiveness, allowing speech to sound authentically human again.

Describe A Moment Whispp Made a Real Difference

In addition to quantitative results, participants shared their reflections of how Whispp transformed their ability to communicate in everyday situations, ranging from professional interactions to personal connections.

"Wow factor for friends to hear my converted voice."



Roger

"I had to make a call to our internet provider about something and I used Whispp as a phone call (so no video) I was clearly heard and understood. They didn't call me sir. "



José

"When I spoke with a professional during 30 min when this professional was driving in a noisy environment."



David

"The first video call with my colleagues, they were really happy for me."



Philippe

Pilot Results: Reducing Communication Barriers and Improving Usability

The pilot demonstrated strong adoption, high usability, and meaningful improvements in participants' ability to communicate effectively in everyday digital interactions.

Before the pilot

Prior to using Whispp, participants experienced significant challenges in real-time communication - particularly during video calls and virtual meetings.

Participants rated the difficulty of participating in video calls at an average of 4.2 out of 5, indicating that communication was often very difficult.

Many relied on alternative methods, such as typing responses or asking others to speak on their behalf, which limited participation and disrupted conversational flow.

High Usability and Lasting Impact

84%

of participants rated Whispp an 8 or higher out of 10 when asked how likely they were to recommend it to someone with a voice condition.

79%

of Whispp participants rated the software as easy or very easy to use across a variety of tasks.

83%

of participants indicated that on-device processing and data privacy are important factors in their experience.

During the pilot

After gaining access to Whispp's AI-enabled voice reconstruction technology, participants reported noticeable improvements in communication ease and quality.

100% of participants used Whispp at least once per week.

33% used it daily, integrating the technology into their regular communication routines.

Users reported being able to engage in longer, more natural conversations with reduced vocal strain.

Listeners described Whispp-enhanced voices as clearer, more natural, and easier to understand.

Some listeners described their Whispp-enhanced voices as "clearer, louder, and more natural." Many users reported being able to hold longer, more relaxed conversations without vocal strain.

Case study: Enabling Independent Communication

Michael is a self-employed pharmaceutical naming consultant and CEO who has lived with a severe stutter since childhood. While video conferencing is a critical part of his daily work, speaking in real time, particularly in high-pressure situations such as meeting new clients, has historically been a significant challenge.

As a result, Michael often relied on colleagues to communicate on his behalf during meetings. Whispering reduces stuttering by up to 85%, the Whispp product allows him to clearly share his thoughts and expertise without depending on others.

Michael can speak for himself again in online meetings. Because whispering reduces stuttering by up to 85%⁷, the Whispp product will allow him to clearly share his thoughts and expertise without depending on others, even in situations that once felt overwhelming.



"It was amazing to use the update with my personal voice and being able to just whisper and talk with people."

Michael

After the Pilot: Lasting Impact

Following the pilot, participants reported sustained improvements in communication confidence and independence.

Key outcomes included:

Increased participation in professional and social conversations

Reduced reliance on others to communicate on their behalf

Greater comfort using voice-based communication tools

Participants also highlighted the emotional and social impact of the technology, particularly the ability to reconnect with others and communicate in their own voice.

"It's an application that could become a must-have for video calls."



Philippe

"Finally I can make a call and I don't need to ask other people to make a call for me."



David

"Wow factor for friends to hear my converted voice."



Roger

"I loved to use it for your pilot and I experienced it was fun for the people I called. They were really impressed. All my friends and family can't imagine me without my disabled voice so it was a real improvement."



José

"I absolutely love that I can connect with people with calls - it's been 3 years since I could last have a phone call and be heard. Whispp is opening doors of reconnection for me. It is bringing inclusion back to my life."



Letitia

Appendix



Whispp AI uniquely transforms whispered and affected speech into a natural voice in real-time, enabling millions of people with voice problems and severe stuttering to hear their voices again. By converting whispered speech and affected speech into clear, natural voice during phone and video calls and through voice and video messages, Whispp enables people with voice problems to communicate in various situations confidently. Unlike conventional speech aids that rely on speech-to-text, Whispp's innovative audio-to-audio AI technology circumvents bypasses delay problems, enabling instant, natural-sounding speech conversion.



Qualcomm relentlessly innovates to deliver intelligent computing everywhere, helping the world tackle some of its most important challenges. Our proven solutions drive transformation across major industries, and our Snapdragon X Elite platform power extraordinary consumer experiences. Building on our nearly 40-year leadership in setting industry standards and creating era-defining technology breakthroughs, we deliver leading edge AI, high-performance, low-power computing, and unrivaled connectivity. Together with our ecosystem partners, we enable next-generation digital transformation to enrich lives, improve businesses, and advance societies. At Qualcomm, we are engineering human progress.



Qualcomm

1. Cohen SM: Self-reported impact of dysphonia in a primary care population: An epidemiological study. The Laryngoscope 2010, 120:2022-2032.
2. Schwartz SR, Cohen SM, Dailey SH, Rosenfeld RM, Deutsch ES, Gillespie MB, Granieri E, Hapner ER, Kimball CE, Krouse HJ: Clinical practice guideline: hoarseness (dysphonia). Otolaryngology–Head and Neck Surgery 2009, 141:1-31.
3. Organization WH, Fund UNCs: Global report on assistive technology: World Health Organization; 2022.
4. Dietrich M, Abbott KV, Gartner-Schmidt J, Rosen CA: The frequency of perceived stress, anxiety, and depression in patients with common pathologies affecting voice. Journal of voice 2008, 22:472-488.
5. Jefferson, M., 2019. Usability of automatic speech recognition systems for individuals with speech disorders: Past, present, future, and a proposed model.
6. Moore, M., Venkateswara, H. and Panchanathan, S., 2018. Whistle-blowing ASRs: Evaluating the need for more inclusive speech recognition systems. Interspeech 2018.
7. van de Vorst, R., & van Vugt, F. T. (2023). Advances in Cognitive Psychology, 19(1), 92–103.