



The effect of internet telephony and a cochlear implant accessory on mobile phone speech comprehension in cochlear implant users

Markus E. Huth¹ · Regula L. Boschung¹ · Marco D. Caversaccio¹ · Wilhelm Wimmer^{1,2} · Mantokoudis Georgios¹

Received: 9 January 2022 / Accepted: 30 March 2022 / Published online: 24 April 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Purpose In individuals with severe hearing loss, mobile phone communication is limited despite treatment with a cochlear implant (CI). The goal of this study is to identify the best communication practice for CI users by comparing speech comprehension of conventional mobile phone (GSM) calls, Voice over Internet Protocol (VoIP) calls, and the application of a wireless phone clip (WPC) accessory.

Methods This study included 13 individuals (mean age 47.1 ± 17.3 years) with at least one CI. Frequency response and objective voice quality were tested for each device, transmission mode and the WPC. We measured speech comprehension using a smartphone for a GSM call with and without WPC as well as VoIP-calls with and without WPC at different levels of white background noise.

Results Frequency responses of the WPC were limited (< 4 kHz); however, speech comprehension in a noisy environment was significantly improved compared to GSM. Speech comprehension was improved by 9–27% utilizing VoIP or WPC compared to GSM. WPC was superior in noisy environments (80 dB SPL broadband noise) compared to GSM. At lower background noise levels (50, 60, 70 dB SPL broadband noise), VoIP resulted in improved speech comprehension with and without WPC. Speech comprehension scores did not correlate with objective voice quality measurements.

Conclusion Speech comprehension was best with VoIP alone; however, accessories such as a WPC provide additional improvement in the presence of background noise. Mobile phone calls utilizing VoIP technology, with or without a WPC accessory, result in superior speech comprehension compared to GSM.

Keywords Cochlear implants · Speech comprehension · Communication aids for disabled · VoIP

Purpose

With the advent of mobile phones, mobile communication and availability for private and professional reasons have become ubiquitous. In hearing-impaired individuals, however, having mobile phone conversations is challenging despite treatment with hearing aids or cochlear implants (CI) [1]. Telephone speech comprehension in cochlear implant users (CIUs) is often compromised by reduced speech signal quality [2–5]. During the COVID-19 (coronavirus disease

2019) pandemic and the subsequent requisite social distancing, the importance of telephone speech comprehension during both private and professional telephone conversations and conferences has become paramount.

Telephone speech comprehension is often compromised due to a limited frequency bandwidth of 300–3400 Hz and digital signal compression [6, 7]. Voice over Internet Protocol (VoIP), however, transmits a wider frequency bandwidth (200–12,000 Hz), which results in improved speech comprehension independent of additional visual cues [6–9]. Mantokoudis et al. measured speech comprehension during a conventional telephone call with background noise at different signal-to-noise ratios (SNR) [10]. During quiet conditions, telephone speech comprehension in CIUs was 73%. Speech comprehension with background noise (SNR 10 dB), however, decreased significantly to only 12% in CIUs compared to 97% in normal-hearing individuals (NHI) [10, 11]. Using VoIP, speech comprehension with background noise

✉ Markus E. Huth
markus.huth@spital.so.ch

¹ Department of Otorhinolaryngology, Head and Neck Surgery, Inselspital, Bern University Hospital, University of Bern, 3010 Bern, Switzerland

² Hearing Research Laboratory, ARTORG Center, University of Bern, Bern, Switzerland

(SNR 10 dB) improved to 48% in CIUs; this resulted from higher speech signal quality due to transmission of a wider frequency spectrum and application of noise reduction algorithms in VoIP compared to conventional telephony [10, 11]. Using different mobile phone models, speech comprehension in CIUs improved by 13–15% for VoIP compared to Global System for Mobile Communication (GSM) transmission [12]. The ability to lip-read using the video signal during VoIP conversations further contributes to improved speech comprehension [13].

Furthermore, placing the phone receiver or the mobile phone next to the ear, as commonly held during telephone conversations, can produce interfering noise from physical contact of the phone receiver with the microphone of the speech processor [14, 15]. On the other hand, using the speakerphone can reduce the speech signal quality due to interference from ambient noise.

Notably, environmental background noise during mobile phone calls is an omnipresent predicament that comes with permanent availability [10] as surrounding traffic and people require a certain awareness. This predicament has already been recognized by the hearing aid industry.

While the newest CIs can link directly to a smartphone, previous CI generations require a coupling device. Currently, all three major CI manufacturers offer such coupling devices. MED-EL (MED-EL Elektromedizinische Geräte Gesellschaft m.b.H., Fürstenweg 77a, 6020 Innsbruck, Austria) offers the AudioLink. In addition to smartphone call transmission to the CI, the AudioLink can also transmit music from the smartphone and link to a TV. Advanced Bionics (Advanced Bionics LLC 28515 Westinghouse Place, Valencia, CA 91355, USA) offers the Naída CI Connect, which is directly attached to the CI speech processor. Cochlear (Cochlear Ltd., 1 University Avenue, Macquarie University, NSW, 2109, Australia) offers the Wireless Phone Clip (WPC). The WPC allows for direct, wireless routing of the speech signal from a smartphone to the hearing aid or cochlear implant, thereby reducing or optionally eliminating surrounding background noise. The WPC has a built-in microphone which enables optional acoustic transmission of environmental sounds with the phone signal, to increase surrounding awareness. Furthermore, the WPC has a button that allows for answering and terminating phone calls directly, effectively rendering the WPC, with its built-in microphone, a hands-free speakerphone.

Multiple studies performed by Cochlear Inc. demonstrate improved speech comprehension with WPC for conventional mobile (GSM) telephony in cochlear implant users [16–18]. The effect of the WPC on telephone speech comprehension in CIUs utilizing VoIP telephony, however, remains unclear. This is of particular interest, as VoIP telephony has become an integral part of the increasing amount of distant communication throughout the COVID-19 pandemic. Therefore,

this study investigates CIU telephone speech comprehension for GSM and VoIP telephony both with and without the WPC.

Methods

Ethical standards

The study protocol was approved by the local ethical review board prior to the start. The study was conducted in full accordance with the study protocol as approved and with the ethical standards as stated in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Written informed consent including consent for publication was obtained from all participants.

Measurement scenarios

Based on the two different transmission modes (GSM or VoIP) and the two coupling modes (acoustically with a mobile phone on the ear [A] or via WPC (Bluetooth connection to the speech processor [B])), four measurement scenarios resulted: GSM telephony with acoustic signal transmission (A-GSM); GSM telephony with Bluetooth signal transmission (B-GSM); VoIP telephony with acoustic signal transmission (A-VoIP); VoIP telephony with Bluetooth signal transmission (B-VoIP). GSM phone calls were routed through the mobile service provider. A mobile VoIP App was used for receiving VoIP calls. To prevent learning effects in participants due to repeated measurements in this study, the sequence of measurement scenarios was randomized.

Study equipment measurements

Prior to measuring human study participants, the characteristics of the study equipment were measured. An integrating sound pressure level (SPL) meter (Norsonic, Type 116, IEC 651/804 Type II, Norsonic AS, Tranby, Norway) was first calibrated with a 1 kHz signal at 114 dB SPL (Nor-1253 sound calibrator, Norsonic AS, Tranby, Norway) in a sound-attenuated audio booth.

A mobile phone (iPhone SE, iOS 10.2, Apple Inc., Cupertino, CA, USA) with the Skype software application (Version 7.35, Skype Communications S.à.r.l., 23–29 Rives de Clausen, L-2165 Luxemburg) was mounted to the audio booth. The Hochmair-Moser-Schulz (HSM) speech comprehension test [19] was played from a CD on a laptop computer using VLC media player software (Version 2.2.4, VideoLAN, 18 rue Charcot, 75013 Paris, France). The HSM-test is a validated speech comprehension test for cochlear implant users. The standardized

HSM-test sentences were transmitted directly via Skype software (Version 7.35, Skype Communications S.à.r.l., 23–29 Rives de Clausen, L-2165 Luxemburg) to the mobile phone in the audio booth either as a mobile GSM phone call or as a VoIP phone call to the mobile Skype software application (Fig. 1). The SPL meter measured sound levels at 1 cm distance to the mobile phone loudspeaker and the phone volume was set to 65 dB SPL.

A CI speech processor type Nucleus 7 was mounted and connected to a Head & Torso Simulator (Type 4128, Brüel & Kjær A/S, Skodsborgvej 307, DK-2850 Nærum, Denmark) in the audio booth. The mobile phone was mounted to the ear of the Head & Torso simulator at 1 cm distance to simulate a typical phone call position. The HSM test sentences were either played acoustically from the mobile phone receiver or transmitted directly to the CI speech processor via Bluetooth by the WPC. Every possible combination was subsequently measured: A-GSM, B-GSM, A-VoIP, and B-VoIP.

Frequency responses

The frequency responses of the GSM or VoIP transmission were measured with an audio analyzer (UPV Audio Analyzer, Rohde & Schwarz GmbH & Co. KG, Mühlhofstrasse 15, 81671 Munich, Germany) as described in a previous publication from our group [12]. The frequency responses were measured either acoustically (A, for the mobile phone receiver) or electronically (B, for the WPC) from the CI speech processor mounted on the Head & Torso simulator. Reproducible measurements within a 1% margin were considered valid.

Objective speech quality

To quantify the speech signal quality of the chosen speech comprehension test, a Perceptual Evaluation of Speech

Quality (PESQ) [20] was performed. The audio analyzer utilized the PESQ test, which is a validated objective measure of speech quality in telecommunication [21, 22] as well as in CIs [23]. The PESQ score ranges from 1 to 5, in which 1 is the lowest and 5 the highest speech signal quality. The Head & Torso simulator recorded the signal from the CI speech processor for the four measured scenarios.

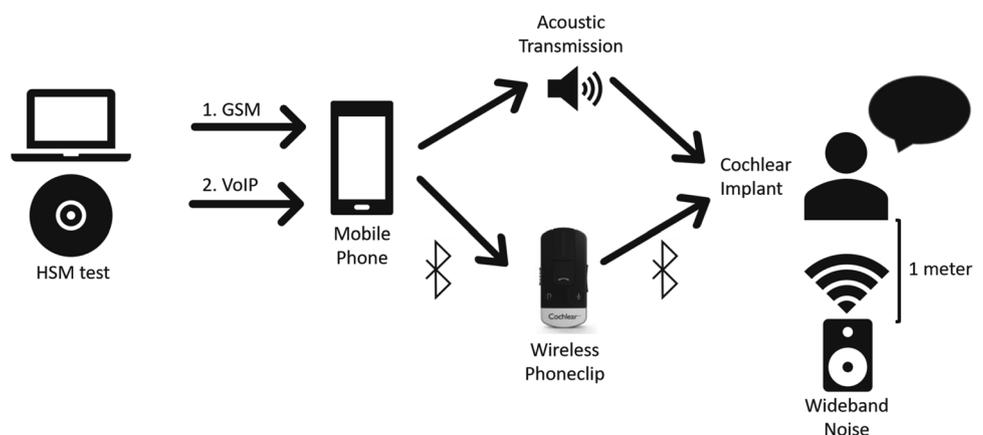
Study participants

To test speech comprehension, participating individuals were required to be native German speakers. Participants were recruited from our institutional CI patient database. Inclusion criteria were age ≥ 18 years, CI use for ≥ 3 months and a CI speech processor compatible with WPC. Exclusion criteria were minors and patients mentally or physically unable to participate.

Speech comprehension tests

All participants utilized their CI speech processor in their everyday user settings or a commonly-used setting for phoning. The four measurement scenarios (A-GSM, B-GSM, A-VoIP, B-VoIP) were measured in a randomized order. To minimize confounding effects of residual hearing of the contralateral, non-implanted ear, a foam earplug was inserted in the respective ear canal. Speech comprehension was quantified using the HSM-sentence test [19] (Fig. 1). The standardized HSM-test sentences were played from a CD on a laptop computer using VLC media player software. The speech signal was directly transmitted via VoIP software to the mobile phone (iPhone SE, iOS 10.2, Apple Inc., Cupertino, CA, USA) either as a GSM or VoIP call. The mobile phone and the transmission volume of the HSM-sentences were set at 65 dB SPL as measured during the calibration. A white background noise (WBN) signal was played at 50, 60, 70, and 80 dB SPL from a loudspeaker in the sound-attenuated audio booth at 1 m distance in front of

Fig. 1 Measurement setup. The Hochmair-Schulz-Moser (HSM)-test sentences were transmitted to a mobile phone via a conventional phone call (GSM) or via Voice over Internet Protocol (VoIP) and further transmitted either acoustically or via Bluetooth and Wireless Phoneclip to the cochlear implant while the participant was exposed to wideband noise at 50, 60, 70, or 80 dB SPL from a loudspeaker at 1 m distance



the participant. The HSM-sentences then were transmitted either acoustically or via WPC to the CI speech processor. The participants verbalized the HSM-sentences, and every correctly verbalized word contributed to the overall HSM-score (106 words = 106 points maximum score). After each measurement, participants were asked to quantify subjective comprehension of speech sound quality for all four scenarios by the mean opinion score (MOS, 5-point rating scale with 1 as the lowest and 5 as the highest score), a subjective speech quality test widely used in telecommunication [9, 24].

Statistics

We analyzed individual patient speech comprehension scores on the HSM-sentence test for all four scenarios. To eliminate confounding, the measurements of 6 participants were excluded at 80 dB SPL WBN due to zero comprehension of the HSM test sentences. Speech comprehension scores as directly measured by correctly verbalized words were compared by means of a Friedman-ANOVA test. To correct for multiple testing, a Wilcoxon-Nemenyi-McDonald-Thompson posthoc test was performed with a 5% level of significance. Statistical analysis was performed with Origin Pro (Version 8.6; OriginLab Corporation, One Roundhouse Plaza, Suite 303, Northampton, MA 01060, USA). We reported the subjective comprehension of speech sound quality according to the MOS as a secondary outcome.

Results

Fifteen individuals were identified as eligible. Two participants, however, demonstrated 0% speech comprehension in all tested scenarios and were thus excluded from the study.

We included 13 participants (mean age 47.1 ± 17.3 years; male:female 7:6, Table 1).

Frequency response

The frequency response was measured for all four scenarios from 20 to 10,000 Hz in dB V (Fig. 2). Acoustic signal transmission (A-GSM and A-VoIP) demonstrated similar frequency responses from 20 to 3200 Hz. In the low-frequency range from 70 to 200 Hz, A-GSM and A-VoIP showed a decreased frequency response compared to Bluetooth signal

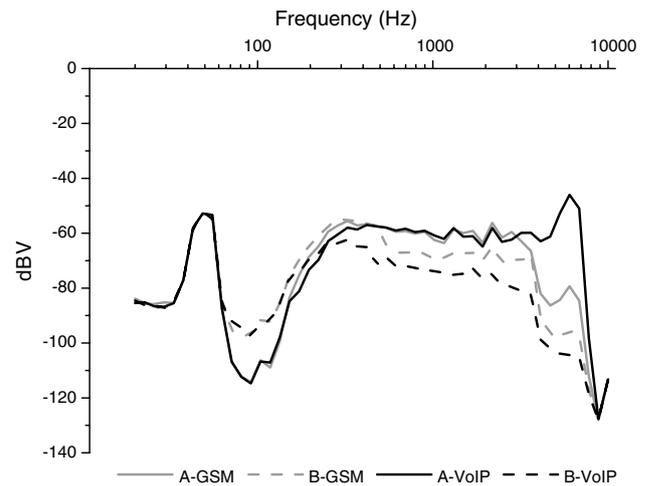
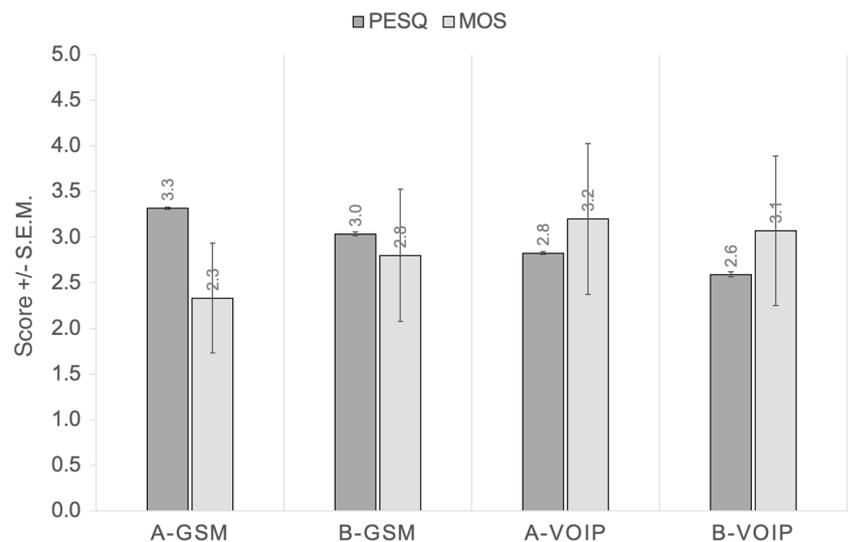


Fig. 2 Frequency response from 20 to 10,000 Hz in dB V for mobile telephony (GSM) with acoustic transmission (A-GSM), GSM with Bluetooth transmission via Wireless Phone Clip (WPC) (B-GSM), Voice over Internet Protocol (VoIP) telephony with acoustic transmission (A-VoIP), and VoIP with Bluetooth transmission via WPC (B-VoIP)

Table 1 Individual participant data regarding sex, age, cochlear implant processor and side of implantation. The male:female ratio of participants is 1.2 while the mean age is 47.1 years

ID	Sex	Age in years	CI-Processor	Side of CI	Time of deafness	Reason for deafness	Non-implanted ear
1	Male	23.6	CP800/CP900	Right	4 years	Congenital	No residual hearing
2	Female	68.5	CP910	Left	2 years	Progressive	No residual hearing
3	Female	31.4	CP910	Right	Unknown	Congenital	No residual hearing
4	Male	50.7	CP910	Bilateral	Unknown	Post-Meningoencephalitis	Not applicable
5	Male	45	CP910	Right	1 year	Congenital + Sudden hearing loss	Residual hearing, no hearing aid
6	Male	58.8	CP910	Right	Unknown	Ménière's Disease	No residual hearing
7	Male	77.9	CP910	Right	Unknown	Progressive	Residual hearing, hearing aid
8	Female	55.3	CP910	Right	Unknown	Ménière's Disease	No residual hearing
9	Male	55.8	Kanso	Left	Unknown	Sudden hearing loss	Residual hearing, hearing aid
10	Female	57	CP910	Right	Unknown	Congenital	No residual hearing
11	Female	34	CP910	Right	Unknown	Congenital	No residual hearing
12	Male	24.2	CP910	Left	Unknown	Unknown	No residual hearing
13	Female	30	CP910	Right	9 years	Congenital	No residual hearing

Fig. 3 Mean Perceptual Evaluation of Speech Quality (PESQ) scores (objective voice quality scores, scale 1–5) and mean opinion score (MOS, scale 1–5) of subjective audio signal quality ± 95% CI rated by the participants for mobile telephony (GSM) with acoustic signal transmission (A-GSM), GSM with Bluetooth transmission via Wireless Phone Clip (WPC) (B-GSM), Voice-over-Internet-Protocol (VoIP) telephony with acoustic transmission (A-VoIP), and VoIP with Bluetooth transmission via WPC (B-VoIP)



transmission (B-GSM and B-VoIP). In the mid-frequency range from 540 to 3200 Hz, however, A-GSM and A-VoIP showed increased frequency responses compared to B-GSM and B-VoIP. From 250 to 9000 Hz, B-VoIP demonstrated the lowest response compared to all other modalities. In the high frequencies above 4000 Hz, A-VoIP exhibited increased signal strength up to 9000 Hz compared to A-GSM, B-GSM, and B-VoIP, which all showed decreasing signal strength according to sequence (Fig. 2).

Objective speech quality

To quantify audio signal quality, the PESQ score was measured for the HSM-sentences in all four scenarios. A-GSM and B-GSM resulted in a PESQ-score of 3.3 and 3.0, respectively. A-VoIP and B-VoIP had PESQ-Scores of 2.8 and 2.6, respectively. Comparing acoustic to Bluetooth transmission, the application of the WPC resulted in a decreased PESQ-Score of 0.3 for B-GSM and 0.2 for B-VoIP (Fig. 3).

Speech comprehension

Figure 4 illustrates speech comprehension scores for all four scenarios with 50, 60, 70, and 80 dB SPL WBN. At 50 dB SPL WBN, A-GSM resulted in an average speech comprehension of 54.4% (16.0–99.1%, 95% CI 39.4 to 69.5%) compared to 68.7% (29.2–100.0%, 95% CI 54.1 to 83.3%) for B-GSM, 75.6% (37.7–99.1%, 95% CI 62.6–88.6%) for A-VoIP and 71% (17.0–97.2%, 95% CI 56.0 to 85.9%) for B-VoIP. A-VoIP demonstrated statistically significant increased speech comprehension compared to A-GSM ($p < 0.008$, effect size $r = 1.3$). B-GSM and B-VoIP demonstrated superior speech comprehension scores compared to A-GSM, however, without statistically significant differences.

At 60 dB SPL WBN, A-GSM resulted in an average speech comprehension of 44.6% (0–93.4%, 95% CI 27.0– 62.3%) compared to 61.9% (17.0–100%, 95% CI 46.7–77.1%) for B-GSM, 71.3% (20.8–99.1%, 95% CI 56.5–86.1%) for A-VoIP, and 70.0% (24.5–100%, 95% CI 53.4–86.5%) for B-VoIP. A-VoIP ($p < 0.005$, effect size $r = 1.3$) and B-VoIP ($p < 0.001$, effect size $r = 1.4$) demonstrated statistically significant increased speech comprehension compared to A-GSM. B-GSM demonstrated superior

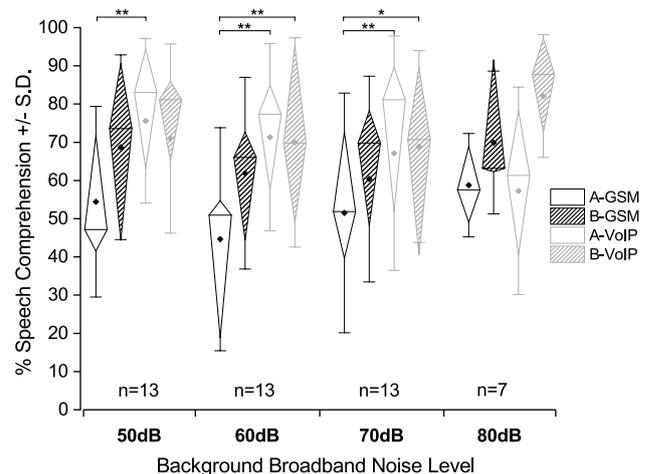


Fig. 4 Speech comprehension ± standard deviation in percentage of correctly verbalized 20 Hochmair-Schulz-Moser (HSM)-test sentences (106 words) in cochlear implant users (box plot 25%, 50%, 75%, Diamond = mean). The speech signal was presented at a constant 65 dB SPL with broadband background noise levels of 50 ($n = 13$), 60 ($n = 13$), 70 ($n = 13$), and 80 dB ($n = 7$) for mobile telephony (GSM) with acoustic signal transmission (A-GSM), GSM with Bluetooth transmission via Wireless Phone Clip (WPC) (B-GSM), Voice over Internet Protocol (VoIP) telephony with acoustic transmission (A-VoIP), and VoIP with Bluetooth transmission via WPC (B-VoIP). Level of significance $*(p < 0.05)$, $** (p < 0.01)$

speech comprehension scores without statistical significance compared to A-GSM.

At 70 dB SPL WBN, A-GSM resulted in an average speech comprehension of 51.5% (0–96.2%, 95% CI 32.6–70.5%) compared to 60.4% (9.4–100%, 95% CI 44.1–76.6%) for B-GSM, 67.1% (6.6–99.1%, 95% CI 48.6–85.6%) for A-VoIP, and 68.9% (27.4–99.1%, 95% CI 53.7–84.0%) for B-VoIP. A-VoIP ($p < 0.003$, effect size $r = 1.4$) and B-VoIP ($p < 0.02$, effect size $r = 1.2$) demonstrated statistically significant increased speech comprehension compared to A-GSM. B-GSM demonstrated superior speech comprehension scores without statistical significance compared to A-GSM.

At 80 dB SPL WBN 6 of the 13 included participants demonstrated 0% speech comprehension in all four scenarios. To eliminate confounding, these 6 imperceptive participants were not included in the data analysis at 80 dB SPL WBN. The remaining 7 participants demonstrated an average speech comprehension of 58.8% (39.6–79.2%, 95% CI 46.3–71.3%) for A-GSM compared to 69.9% (40.6–94.3%, 95% CI 52.7–87.2%) for B-GSM, 57.3% (6.6–88.7%, 95% CI 32.2–82.4%) for A-VoIP, and 82.1% (52.8–98.1%, 95% CI 67.2–96.9%) for B-VoIP. No statistically significant differences were observed.

A-VoIP demonstrated superior speech comprehension at 50, 60, and 70 dB SPL WBN. B-VoIP demonstrated superior speech comprehension at 60, 70, and 80 dB SPL WBN. At WBN levels of 50, 60, and 70 B SPL, B-GSM and B-VoIP demonstrated no difference compared to A-VoIP. At 80 dB SPL WBN, however, B-GSM and B-VoIP demonstrated increased speech comprehension compared to A-GSM und A-VoIP. Speech comprehension did not correlate with the PESQ score in our population.

Mean opinion score

The Mean Opinion Score (MOS) represents the subjectively perceived audio signal quality of the participants with 1 as the lowest score and 5 as the highest. The participants rated A-VoIP (MOS 3.2) as the modality with the subjectively best signal quality, followed by B-VoIP (MOS 3.1). B-GSM resulted in a MOS of 2.8, whereas A-GSM resulted in a MOS of 2.3 (Fig. 3).

Discussion

This study investigates cochlear implant users' speech comprehension for VoIP telephony with (B-VoIP) and without (A-VoIP) WPC compared to conventional mobile telephony (A-GSM and B-GSM). Overall, VoIP provided superior voice quality and improved speech comprehension

for CIUs; however, it appears that available accessories do not fully support the extended frequency band provided by VoIP.

The frequency response measurements demonstrated a broader frequency range for A-VoIP alone. All other scenarios (A-GSM, B-BSM, B-VoIP) showed a narrow band transmission not supporting frequencies > 4 kHz. Application of the WPC resulted in a reduced signal loss in low frequencies (70–200 Hz), which are considered irrelevant for speech comprehension and may actually be interpreted by the user as additional background noise. Furthermore, the application of the WPC demonstrated lower signal intensities over a narrower frequency range resulting in increased signal loss. Restrictions of frequency range and signal intensity are often due to the device design and chosen hardware components. This has also been illustrated in previous studies, where different telephones or mobile devices were assessed in terms of voice quality and intelligibility [9]. In addition, telecommunication providers offering GSM apply speech codecs, which compress voice signals and restrict high-frequency transmission using 13kBit/s transmission rates (ITU recommendations, G-series). The compressed speech signals have a crucial impact on speech discrimination in the presence of noise since the high-frequency component of speech is absent [6, 7].

PESQ measurements were performed to quantify speech signal quality; however, PESQ scores were neither consistent with speech comprehension results nor with MOS scores. This discrepancy was also found in a previous study analyzing Digital Enhanced Cordless Technology (DECT) phones and their sound quality [9]. This raises the question: given its original purpose of measuring sound signal quality in conventional acoustic telephony for a normal hearing population, is the PESQ-score an adequate test of sound signal quality for VoIP in CIUs? MOS scores (subjective perceived voice quality), contrary to PESQ-scores, were consistent with the speech comprehension results. However, no statistically significant difference was observed. The results suggest a loss of speech signal quality by application of the WPC for conventional as well as for VoIP telephony.

In contrast to the PESQ scores, VoIP telephony demonstrated statistically significant improved speech comprehension at 50 (A-VoIP), 60 (A- and B-VoIP), and 70 dB SPL (A- and B-VoIP) compared to A-GSM. Although additional video transmission results in a further 8.5% improvement in speech comprehension, this study excluded visual cues to specifically examine the effect of the WPC [13]. Overall, the application of the WPC resulted in improved speech comprehension at all tested background noise levels compared to A-GSM but only outperformed A-VoIP at a background noise level of 80 dB

SPL WBN. This finding was expected because accessories such as the WPC offer improved coupling to the implant while reducing background noise. In addition, the speech processor's built-in microphone was still active to enable CIUs to continue to be aware of their surroundings while calling. CIUs have the option to switch off the built-in microphones to block or reduce background noise. The WPC, however, did not transmit the full frequency range and full intensity of the VoIP software.

Based on previous studies and on the current data, we can conclude that signal quality is reduced in GSM (due to signal restrictions from the providers) [1, 9], improves with accessories like the WPC in environmental background noise and is superior with VoIP, for which no additional benefit from the WPC can be observed.

Strengths and limitations

This is the first study investigating the benefit of VoIP telephony using a phone clip for direct coupling to the cochlear implant; however, our study had also limitations:

We calculated the required sample size ($n = 8$) based on data from a previous study done by our group [11]. A post-hoc power-analysis demonstrated statistical power of more than 80% for the statistically significant results of our study; however, a higher number of study participants would have been preferable to increase the overall power of this study.

All participants used the WPC for the first time and we cannot exclude a learning effect through repeated use. We also applied the phone clip only monaurally, and we would expect an increased benefit with binaural transmission to the implants [25]. Furthermore, only one VoIP application and one phone accessory were tested. Other accessories, such as induction neck loops, have not been tested. Built-in microphones remained switched on as this was the standard user setting. Other mapping strategies and microphone settings adapted to the WPC may have resulted in better voice and speech comprehension. Additionally, not all CIUs have access to high-speed mobile internet, which is a pre-requisite for utilizing VoIP. As previously discussed, there was also a discrepancy between measured and subjectively perceived voice quality. Measurement tools adapted to cochlear implant recipients are lacking.

Clinical implications

Professionals dealing with CIUs should advise the use of VoIP for distant communication. Smartphone applications in conjunction with fast and reliable mobile internet connections improve the overall communication experience of CIUs. Adding a phone accessory such as a WPC may be recommended

for mobile calls in the presence of background noise such as street traffic, railway stations, etc. Future accessory developments for CIUs should support a broad frequency phone transmission range.

Conclusion

Speech comprehension was best using a mobile phone application (A-VoIP) that took advantage of improved voice signal quality offered by newer technology. Accessories such as a WPC provide improvement with environmental background noise. Mobile VoIP calls both with and without WPC accessories result in superior speech comprehension compared to conventional mobile phone calls.

Acknowledgements The authors express their gratitude to the Foundation Besser-Hören-Schweiz, Allmendstrasse 11, CH-6312 Steinhausen, Switzerland and Cochlear AG, Peter Merian-Weg 4, CH-4052 Basel, Switzerland for funding of this study in equal parts. We additionally thank Cochlear AG additionally for providing the Wireless Phone Clip. Furthermore, we thank Jérémie Guignard (Cochlear AG) for his help in setting up the equipment measurements and Catherine S. Reid, MD for critical review and language correction of the manuscript.

Author contributions All authors contributed to the study and meet all authorship qualifying standards.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest This research was partially funded by Cochlear AG, Peter Merian-Weg 4, CH-4052 Basel, Switzerland, who is the producer of the Wireless Phone Clip. The authors have no further conflict of interest to declare. The authors have no financial interest to declare.

Ethical approval Ethics approval was obtained from the local institutional review board (ethical commission of the canton of Bern) (KEK-BE #2016-0135). The study was conducted in full compliance with the ethical standards as stated in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Written informed consent including consent for publication was obtained from all participants.

References

1. Tan BY, Gluth MB, Statham EL, Eikelboom RH, Atlas MD (2012) Mobile and landline telephone performance outcomes among telephone-using cochlear implant recipients. *Otolaryngol Head Neck Surg* 146(2):283–288. <https://doi.org/10.1177/0194599811424120>
2. Adams JS, Hasenstab MS, Pippin GW, Sismanis A (2004) Telephone use and understanding in patients with cochlear implants. *Ear Nose Throat J* 83(2):96
3. Galindo J, Lassaletta L, Mora RP, Castro A, Bastarrica M, Gavilan J (2013) Fine structure processing improves telephone

- speech perception in cochlear implant users. *Eur Arch Oto-Rhino-Laryngol* 270(4):1223–1229. <https://doi.org/10.1007/s00405-012-2101-9>
4. Kepler LJ, Terry M, Sweetman RH (1992) Telephone usage in the hearing-impaired population. *Ear Hear* 13(5):311–319
 5. Wu CM, Liu TC, Wang NM, Chao WC (2013) Speech perception and communication ability over the telephone by Mandarin-speaking children with cochlear implants. *Int J Pediatr Otorhinolaryngol* 77(8):1295–1302. <https://doi.org/10.1016/j.ijporl.2013.05.015>
 6. Liu C, Fu QJ, Narayanan SS (2009) Effect of bandwidth extension to telephone speech recognition in cochlear implant users. *J Acoust Soc Am* 125(2):77–83. <https://doi.org/10.1121/1.3062145>
 7. Milchard AJ, Cullington HE (2004) An investigation into the effect of limiting the frequency bandwidth of speech on speech recognition in adult cochlear implant users. *Int J Audiol* 43(6):356–362
 8. Baer T, Moore BC, Kluk K (2002) Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *J Acoust Soc Am* 112(3 Pt 1):1133–1144. <https://doi.org/10.1121/1.1498853>
 9. Mantokoudis G, Koller R, Guignard J, Caversaccio M, Kompis M, Senn P (2017) Influence of telecommunication modality, internet transmission quality, and accessories on speech perception in cochlear implant users. *J Med Internet Res* 19(4):e135. <https://doi.org/10.2196/jmir.6954>
 10. Mantokoudis G, Dubach P, Piffner F, Kompis M, Caversaccio M, Senn P (2012) Speech perception benefits of internet versus conventional telephony for hearing-impaired individuals. *J Med Internet Res* 14(4):e102. <https://doi.org/10.2196/jmir.1818>
 11. Mantokoudis G, Kompis M, Dubach P, Caversaccio M, Senn P (2010) How internet telephony could improve communication for hearing-impaired individuals. *Otol Neurotol* 31(7):1014–1021. <https://doi.org/10.1097/MAO.0b013e3181ec1d46>
 12. Guignard J, Senn P, Koller R, Caversaccio M, Kompis M, Mantokoudis G (2019) Mobile internet telephony improves speech intelligibility and quality for cochlear implant recipients. *Otol Neurotol* 40(3):e206–e214. <https://doi.org/10.1097/MAO.0000000000002132>
 13. Mantokoudis G, Dahler C, Dubach P, Kompis M, Caversaccio MD, Senn P (2013) Internet video telephony allows speech reading by deaf individuals and improves speech perception by cochlear implant users. *PLoS One* 8(1):e54770. <https://doi.org/10.1371/journal.pone.0054770>
 14. Cray JW, Allen RL, Stuart A, Hudson S, Layman E, Givens GD (2004) An investigation of telephone use among cochlear implant recipients. *Am J Audiol* 13(2):200–212
 15. Ito J, Nakatake M, Fujita S (1999) Hearing ability by telephone of patients with cochlear implants. *Otolaryngol Head Neck Surg* 121(6):802–804. <https://doi.org/10.1053/hn.1999.v121.a938864>
 16. Duke M, Wolfe J (2015) Evaluation of speech recognition over the telephone with and without the Cochlear Wireless Phone Clip. Clinical evidence with Cochlear™ Wireless Accessories. Cochlear Ltd.,
 17. Marcrum SC (2015) Wireless streaming with the Cochlear Wireless Phone Clip improves speech understanding and reduces listening effort during telephone use in noise. Clinical evidence with Cochlear™ Wireless Accessories. Cochlear Ltd.,
 18. Robier M, Bakhos D, Pawelczyk T, Lescanne E (2015) Evaluation of benefit provided by the Cochlear Wireless Phone Clip. Clinical evidence with Cochlear™ Wireless Accessories. Cochlear Ltd.,
 19. Hochmair-Desoyer I, Schulz E, Moser L, Schmidt M (1997) The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. *Am J Otol* 18(6 Suppl):S83
 20. Taal CH, Hendriks RC, Heusdens R, Jensen J (2011) An evaluation of objective measures for intelligibility prediction of time-frequency weighted noisy speech. *J Acoust Soc Am* 130(5):3013–3027. <https://doi.org/10.1121/1.3641373>
 21. ITU (2001) P.862: Perceptual evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs., vol P.862. International Telecommunication Union (ITU), Geneva
 22. ITU (2005) Revised annex A—reference implementations and conformance testing for ITU-T Recs P.862, P.862.1 and P.862.2 P.862 Amendmend 2, vol P.862.2. International Telecommunication Union, Geneva
 23. Chu KC, Choi CT (2012) Noise reduction using wavelet thresholding of multitaper estimators and geometric approach to spectral subtraction for speech coding strategy. *Clin Exp Otorhinolaryngol*. <https://doi.org/10.3342/ceo.2012.5.S1.S65>
 24. Sen D, Lu W (2012) Objective evaluation of speech signal quality by the prediction of multiple foreground diagnostic acceptability measure attributes. *J Acoust Soc Am* 131(5):4087–4103. <https://doi.org/10.1121/1.3699262>
 25. Litovsky RY, Goupell MJ, Godar S, Grieco-Calub T, Jones GL, Garadat SN, Agrawal S, Kan A, Todd A, Hess C, Misurelli S (2012) Studies on bilateral cochlear implants at the University of Wisconsin's binaural hearing and speech laboratory. *J Am Acad Audiol* 23(6):476–494. <https://doi.org/10.3766/jaaa.23.6.9>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.