Voice Telephony for Individuals with Hearing Loss: The Effects of Audio Bandwidth, Bit Rate and Packet Loss

Linda Kozma-Spytek

Gallaudet University
Washington DC, USA

{linda.kozma-spytek, paula.tucker, christian.vogler}@gallaudet.edu

Paula Tucker

Gallaudet University
Washington DC, USA

Christian Vogler

Gallaudet University
Washington DC, USA

ABSTRACT

This paper describes three studies conducted with a total of 114 individuals with hearing loss and 12 hearing controls, with the goal of investigating the impact of audio quality parameters on the accessibility of voice telecommunications. Three categories of parameters are covered: (1) narrowband (NB) versus wideband (WB) audio; (2) encoding audio at varying bit rates, ranging from typical rates used in today's telecom networks to the highest quality supported by these audio codecs; and (3) absence of packet loss to worst-case packet loss in VoIP telephony. With WB audio, individuals with hearing loss exhibit better speech recognition, expend less perceived mental effort, and rate speech quality higher than with NB audio. Bit rate affects speech recognition for NB audio, and speech quality ratings for both NB and WB audio. Packet loss affects all of speech recognition, mental effort, and speech quality ratings. WB versus NB audio also affects hearing individuals, especially under packet loss.

Author Keywords

Voice telephony; telecommunications; hearing loss; hard of hearing; cochlear implant; hearing aid; wideband audio; narrowband audio; audio codec; bit rate; packet loss.

CSS Concepts

Human-centered computing~Empirical studies in accessibility • Human-centered computing~Accessibility design and evaluation methods.

INTRODUCTION

Telephone communication for individuals with hearing loss has long been problematic for a variety of reasons. Problems include difficulty hearing the telephone ring, willingness to answer the telephone and difficulty listening to conversation on the phone [37]. Comments we have received from individuals with hearing loss indicate that telephone communication is challenging at best, particularly for those who want to both listen and speak for themselves. They report that when they cannot make out individual words and hence fail to piece together conversations through listening

alone that it is frustrating, tiring, stressful and embarrassing. In spite of this, voice telephone use remains very important. Many report the need for better sound quality and better clarity particularly on cell phones, emphasizing that simply making speech louder is not enough [28]. Fundamentally, speech understanding by people with hearing loss is significantly less robust [7] compared to that of hearing people, and telecommunications has been for the most part voice-only.

To improve access, the Federal Communications Commission administers the popular captioned telephone relay service, which provides audio accompanied by captions generated by human operators. This service has been beset by increasing cost [10], and does not offer the same degree of privacy and effective communication as direct voice calls that do not involve a third-party operator. Improving the accessibility of voice calls for people with hearing loss could be both high-impact and cost-effective.

Figure 1. Narrowband versus wideband audio power spectrum from the IEEE Harvard sentence set: “Cats and dogs each hate the other.” There is approximately 20% more speech information available in the frequencies > 3400 Hz.

The characteristics of the audio transmissions on the telecommunications networks may disproportionately impact the accessibility of the telephone system. In voice calls the audio quality of the transmitted speech is typically reduced compared to face-to-face communication. This reduced quality is due in part to the audio bandwidth transmitted over various telephone networks. Many phone networks do not transmit sounds outside approximately 300–3,400 Hz. Some acoustic cues important for speech intelligibility are at these higher frequencies; access to them

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
ASSETS '19, October 28–30, 2019, Pittsburgh, PA, USA
© 2019 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-6676-2/19/10 \$15.00
<https://doi.org/10.1145/3308561.3353796>

increases the amount of speech that may be available by approximately 20% [20] (Figure 1).

The limitations of narrowband audio (also called standard definition voice) originate from the analog public switched telephone network (PSTN), which has long formed the backbone of telecommunications. Even as PSTN gets replaced with digital mobile and VoIP systems that are, in principle, capable of supporting wider frequency ranges, it still forms the lowest common denominator for cross-carrier calls. Modern VoIP and wireless mobile telephone environments pose additional challenges for people with hearing loss. Both are susceptible to packet loss during network transmission, which degrades sound quality. Additionally, mobile telephony, in the face of network bandwidth constraints, uses a narrowband speech encoding that results in compromised sound quality compared to that delivered by the PSTN and fixed VoIP handsets. The most common codec used for mobile networks (AMR-NB) has significantly reduced the bit rate from 64 kbps on VoIP networks (with quality comparable to PSTN) to rates of 5.90-12.20 kbps via lossy compression techniques based on perceptual coding.

In mainstream telephony, ratings of mobile network sound quality have consistently demonstrated poorer quality for speech than for that experienced on the PSTN. Intelligibility of speech is not typically assessed, because for hearing individuals, speech understanding can remain robust even under severe signal degradation and so does not serve as a meaningful differentiator of system performance. For people with hearing loss, however, intelligibility may be negatively impacted, and consequently, access to the telephone system also may be impacted.

One simple way to improve sound quality, and also potentially improve access for people with hearing loss, is to use wideband encoding (also called high definition voice). Wideband audio codecs typically extend the audio bandwidth from the narrowband 300 Hz-3.4 kHz to 50 Hz-7 kHz, although the effective bandwidth will be circumscribed by other factors, such as the handset characteristics on both the send and receive ends. The 3rd Generation Partnership Project (3GPP) standards group adopted the AMR-WB codec for implementing wireless mobile wideband audio.

AMR-WB supports bit rates from 12.65 to 23.85 kbps and has been deployed by carriers in the US in their 4G networks and handsets operating Voice over Long-Term Evolution (VoLTE). The improvement in data speeds accompanying the deployment of LTE provides the opportunity to increase the bandwidth for speech at very little additional cost in terms of data usage. Unfortunately, cross-carrier interoperability has not been a priority, and calls between different carriers typically fall back to narrowband audio. Additionally, wideband encoding techniques have been employed in point-to-point VoIP calls (e.g., Skype), business telephone systems and videoconferencing systems.

Many people with hearing loss have access to the wideband audio frequencies above the narrowband cutoff of 3400 Hz. Among hearing devices used by individuals with hearing loss, all three cochlear implant (CI) systems available in the US can process audio and deliver information from approximately 200 Hz up to 8.5 kHz [22] depending on the system. Hearing aids (HA) have, likewise, extended their effective audio bandwidth and are being fit for individuals who demonstrate sound audibility and benefit from access to this extended frequency region [2][4][35].

Therefore, even though the deployment has been driven by sound quality considerations, wideband audio may also confer telecommunications accessibility benefits for individuals with hearing loss who either prefer or need to use voice communication where they both listen and speak for themselves. These benefits may go beyond improved speech quality and extend to better speech recognition and reduced expenditures of mental effort.

However, to date there has been little understanding of the purported benefits – in particular, the impact on objective measures of speech recognition or subjective judgments of either sound quality or mental effort by individuals with hearing loss. Likewise, the effect of possible degradations due to common network impairments in the VoIP and mobile worlds, such as packet loss, on these same objective and subjective measures is unknown.

Purpose: This paper aims to provide this understanding through a series of three experiments conducted over three years. Overall, we address the following research questions for people with hearing loss in each respective experiment: (1) Does wideband (WB) audio increase speech recognition and decrease mental effort compared to narrowband (NB) audio? (2) Does the data rate (i.e. bit rate) of the NB and WB audio codecs impact these two measures, and how does it compare to the quality of NB audio on PSTN and VoIP? (3) What is the impact of packet loss, characteristic for mobile and IP networks, on speech recognition, subjective mental effort and speech quality ratings; and how do people with hearing loss compare to hearing people in these categories?

The three experiments, described below, followed an iterative design approach. Each successive experiment built on questions raised by the findings of the previous one, and also replicated key findings from prior experiments. Experiment 1 found that wideband audio benefits people with hearing loss under optimal listening conditions. Experiment 2 replicated this finding, and tested the effect of varying bit rates, which shows that low bit rate narrowband mobile and IP telecommunications are especially challenging compared to old landline phones. Experiment 3 again confirmed the benefit of wideband audio and showed that those benefits still hold under less-than optimal listening conditions due to packet loss, typically found in mobile and IP telecommunications networks.

RELATED WORK

Increasing the telephone audio bandwidth has been shown to increase speech quality and, in some cases, intelligibility in noise for hearing individuals [11]. Artificial bandwidth extension, primarily for cochlear implant users, has received limited attention, with some promise for improved speech recognition and quality [14][25][30]. Higher bit rates for hearing individuals, as measured either directly or through a predictive algorithm, are associated with higher perceived audio quality [1][13]. Likewise, network impairments, such as bursty packet loss, that occur to varying degrees in mobile and VoIP telephony environments are known to degrade perceived audio quality for hearing individuals [1][13]. No systematic investigation has been conducted to date on audio bandwidth, bit rate and network packet loss for individuals with hearing loss, in spite of the prevalence of hearing loss. In the United States, the prevalence of hearing loss is about 12.7% or 30 million people 12 years or older [23], increasing with age to nearly 63.1% of adults 70 years or older [24].

There is some past work on technical requirements related to accessible telephone systems for people with hearing loss. Such work includes the impact of audio-video sync for video telecommunications on listening and lip-reading [21][26], sign language video communications [6][8][36], signal to noise ratios for coupling digital wireless telephones with hearing devices [17][18], and connectivity options between phones and hearing devices more generally [27][32][39]. Other research has focused on building audio enhancements directly into telecommunications systems [19][33], in order to match the hearing loss profiles of individuals with their use of telephones. However, to our knowledge, this is the first study to explore the impact of technical parameters on telecommunications for people with hearing loss.

AUDIO QUALITY PARAMETERS COMPARISON

We conducted a series of three related experiments on voice telecommunications accessibility for individuals with hearing loss over the course of three years. The goal was to better understand the impact of a variety of technical parameters related to audio quality in mobile and VoIP telephony environments on audio-only telecommunications by individuals with hearing loss. The audio quality parameters investigated included codec audio bandwidth (i.e., narrowband vs wideband audio), codec bit rate and the impact of packet loss. Each experiment utilized the same basic experimental design and protocol, which allowed us to both replicate the codec audio bandwidth condition and strengthen our protocol in the second and third experiments. In the third and final experiment, the protocol was extended to include a group of individuals without hearing loss. This provided a means of direct comparison of results between individuals with hearing loss and hearing individuals, whose quality of experience is most often considered by industry.

Each experiment used a within-subjects repeated measures design and consisted of one-hour test sessions, during which paid participants listened to stimuli (available at

<https://bit.ly/2JLF9Pj>) and provided recognition and ratings data. Participant recruitment took place through Hearing Loss Association of America and other institutions serving people with hearing loss. All participants were required to be fluent English-speaking adults, 18 years of age or older. Participants with hearing loss were also required to be daily hearing device users, as well as regular users of the voice telephone. For all participants, we ensured that they were able to access the higher frequencies provided by wideband audio. In the first experiment, only daily CI users were recruited, who are known to perceive sound up to at least 7 kHz [22]. In the subsequent two experiments, both daily HA and CI users were recruited. All participants for these subsequent experiments were required to pass a hearing screening for audibility of the higher frequencies (4 kHz and 5 kHz) available through wideband audio codecs. For participants with hearing loss, the screening was completed while using their hearing devices. Participants with hearing loss also completed an intake survey at the beginning of their test session, which included questions about their degree of hearing loss, hearing device use, and communication during face-to-face and telephone conversations.

Experiment 1: Codec Audio Bandwidth

The research question addressed in Experiment 1 was: Does the use of wideband (WB) encoding compared to narrowband (NB) encoding increase speech recognition and decrease expenditures of mental effort for individuals with hearing loss who have access, through their hearing devices, to frequency information above 3,400 Hz? Testing of NB versus WB telephone speech was completed with a group of 42 cochlear implantees, who as noted in the introduction, can access the frequencies above the NB cutoff of 3,400 Hz.

Participants

Forty-two individuals with hearing loss participated in the study. Of the 42 individuals with hearing loss, 29 were women and 13 were men, with an average age of 57.5 years (ranging from 22-86 years). All participants had at least two to three years of self-reported hearing device use. All individuals used their cochlear implants during testing, with 24 bilateral CI users, 8 unilateral CI users and 10 individuals who used a CI in one ear and a HA in the other ear. Self-reported hearing loss ranged from severe to profound across both ears. Most participants (37) reported profound hearing loss in the ear(s) with a CI.

Materials

Stimuli for the experiment were drawn from the Computer Assisted Speech Perception Evaluation and Training tool, or CASPER [3]. CASPER is a system for evaluation of speech in audio, visual (lipreading) and combined modes that consist of 72 sentence sets. These sets consist of related topics and are designed to be representative of conversational speech. Eight sentence sets were used to prepare the stimuli, with 2 different sets used for each test condition.

Stimulus preparation: For each study participant, speech recognition was tested using CASPER sentence sets in two

audio-only conditions, narrowband telephone speech (AMR NB at a bit rate of 12.20 kBit/s) and wideband telephone speech (AMR WB at a bit rate of 23.85 kBit/s), which are the respective highest-quality rates for both codecs.

The original CASPER files are encoded PCM audio sampled at 22,050 Hz. The audio stream was demultiplexed and downsampled to PCM audio at a sampling rate of 16 kHz, thereby limiting the audio bandwidth to 8 kHz. Then each stimulus was transcoded to AMR-NB and AMR-WB via ffmpeg with the OpenCore-AMR libraries [31] at the respective bit rates. Because the resulting samples are not widely supported by operating system playback and audio editing software, each AMR-encoded stimulus was losslessly converted back to PCM audio, again using ffmpeg with the OpenCore-AMR libraries. Finally, each file was equalized in a post-processing step in Adobe Audition, in order to ensure consistent volume levels across all stimuli.

All sentence sets were processed for each of the two test conditions, and the sentence sets used for each condition was counterbalanced across subjects. This was done to guard against the effects of possible differences in intelligibility, either inherent or as a result of the processing.

Method

Participants repeated the sentences that they heard through the microphone of their CI; using a test set up that simulated cell phone listening via speakers with an average presentation level set so that the frequent peaks of speech measured between 64-70 dB SPL on a sound level meter at the location of the listener's head. This range was selected because it represents both the average level of conversational speech during typical face-to-face communication at one meter and the nominal levels for conversational gain used in telephony for one-ear and two-ear listening [5]. The participants' responses were scored as the number of words correctly repeated per set (out of 102 total words) for each condition. Each administration of one sentence set per condition took approximately 5 minutes. Presentation of conditions was counterbalanced across subjects.

Following the completion of testing for one sentence set/one condition, the Subjective Mental Effort Questionnaire (SMEQ) [34] was administered. The SMEQ provides a post-task rating of the mental effort an individual expends in completing a task. It consists of a single numeric scale from 0 to 150, with nine labels from "Not at all hard to do" to "Tremendously hard to do." Participants moved a slider on a computer to the point in the scale that represented their judgment of task difficulty. The software calculated the scale value selected by the participant. Higher values indicate greater perceived task difficulty.

Results

Results showed access to WB telephone audio improved speech recognition (in quiet) over speech recognition using typical NB telephone audio. In addition, WB audio resulted in lowering the perceived mental effort expended during

completion of the speech recognition task compared to that expended for NB audio.

Speech Recognition: A paired t-test was performed between the conditions of NB and WB audio. WB audio resulted in better speech recognition compared to NB audio. The difference in percent words correct was statistically significant [$t(41)=4.33$, $p=9.364 \times 10^{-5}$]. The observed average of differences between WB audio and NB audio was 10.4%. Table 1 below shows mean results.

Codec	%Words Correct	Std Dev	Std Error
AMR-WB 23.85	75.7	23.4	3.6
AMR-NB 12.20	65.2	26.4	4.1

Table 1. Mean percent words correct statistics for wideband WB x NB audio (n=42).

Codec	SMEQ	Std Dev	Std Error
AMR-WB 23.85	38.2	29.0	4.5
AMR-NB 12.20	54.5	32.4	5.0

Table 2. Mean subjective mental effort statistics for WB x NB audio (n=42). Higher scores imply more effort.

Mental Effort: A paired t-test was performed between the conditions of NB and WB audio. WB audio resulted in lower perceived mental effort compared to NB audio. The difference in perceived mental effort (SMEQ) was statistically significant [$t(41)=3.63$, $p=7.863 \times 10^{-4}$]. The observed average of differences between WB audio and NB audio was 16.3 points. Table 2 above shows mean results.

Experiment 2: Codec Audio Bandwidth and Bit Rate

Experiment 1 suggests that at optimum quality (i.e., the highest available data rates) under carefully controlled conditions (i.e., simulated phone audio over speakers at known levels), WB audio results in significantly better speech understanding, as well as significantly lower expended mental effort. The follow-up research question for Experiment 2 was if this effect could be replicated not just at the highest data rates, but also at lower data rates more representative of mobile networks, and how the quality of mobile audio codecs compares to that found on the PSTN.

Additionally, the careful control of the sound levels was relaxed by switching from a simulated phone environment to a real wireless phone, and letting participants self-select their most comfortable listening level (MCL) on the phone. We also expanded the participant pool from CI users to a mix of CI and HA users. As mentioned previously, all participants were required to pass a listening test with tones 4 and 5 kHz to ensure that they had access to the WB audio frequencies.

Testing of narrowband versus wideband telephone speech at various bit rates was completed with a group of 36 cochlear implant and hearing aids users.

Participants

Thirty-six individuals with hearing loss participated in the study. Of the 36 individuals with hearing loss, 23 were women and 13 were men, with an average age of 48 years (ranging from 18-73 years). All participants had at least one year of self-reported hearing device use. Fourteen individuals used their CIs during testing, while the other 22 used their HAs. Self-reported hearing loss ranged from mild to profound across both ears. In the test ear, most hearing aid users reported moderately-severe or severe hearing loss, while three reported moderate hearing loss, one reported mild hearing loss and one reported not knowing their degree of hearing loss. All CI users reported profound hearing loss.

Materials

As in Experiment 1, stimuli for Experiment 2 were drawn from CASPER. However, six different sentence sets, each spoken by two speakers (1 male and 1 female), were used to prepare the stimuli, with 2 different sets used for each test condition. Each participant received one of the two speakers across all conditions. A separate set of CASPER sentences were used to train participants on the procedure and to establish an individual's MCL for telephone listening.

Stimulus preparation: The conditions with associated codecs and bit rates used for this experiment are shown in Table 3 below. G.711 u-law was the baseline NB audio condition providing toll quality speech (i.e., quality comparable to that of a long-distance call placed over the PSTN). The next four conditions used AMR-NB and AMR-WB audio at the maximum bit rates for the highest possible quality (as in Experiment 1), and bit rates more typical of wireless mobile networks at the time of testing. The sixth condition was a low-pass (LP) filtered version of AMR-WB audio at the highest quality (i.e., 23.65 kbps) with a cut off frequency of 3500 Hz. This condition examined whether any change in performance with the AMR-WB codec at 23.65 kbps, compared to the others, was due to the higher bit rate or extended audio bandwidth used for encoding. The 3500 Hz cutoff was realized through an 8th-order Butterworth filter, in order to produce a high rate of roll off and a narrow transition band while having a maximally flat filter.

All sentence sets were processed in a similar way as in Experiment 1 for each of the six test conditions (cf. Experiment 1: Codec Audio Bandwidth), except for an additional preprocessing step. A send mask was applied to each sample prior to transcoding, to mimic the frequency response of wireless phone microphones. The sentence sets for each condition were counterbalanced across subjects.

Method

Participants' preferred ear and self-selected speech MCL for telephone listening were used in all test conditions. An iPhone 4S was used for presentation of stimuli. A custom app developed by the researchers was used to control presentation of all stimuli in the correct order, and to control the phone settings. To minimize the risk of RF interference with the hearing device, no mobile or WiFi network

connections were active on the phone during testing. The phone was placed in a normal use position at the microphone of a participant's hearing device. An adjustable stand was used to position and hold the handset. The stand was used in order to reduce fatigue, maximize coupling and maintain consistent handset positioning, thereby providing a constant input level to a hearing device's microphone during the course of testing. A Velcro headband was loosely placed around the participant's head and the phone to assist the listener in maintaining the relative positioning of their hearing device's microphone and the phone's speaker for best-case acoustic coupling, as shown in Figure 2.

Condition	Codec	Bandwidth	Bit Rate (kBit/s)
1	G.711 u-law	NB	64.00
2	AMR-NB	NB	5.90
3	AMR-NB	NB	12.20
4	AMR-WB	WB	12.65
5	AMR-WB	WB	23.85
6	AMR-WB	WB (lowpass filtered)	23.85

Table 3. Experimental conditions of audio codec and bit rate.

Figure 2. Experimental setup with iPhone 4S. The crossbar is holding the phone in the position that the participant selected for listening during the MCL setting. The Velcro headband holds the head immobile relative to the crossbar.

A Bluetooth keyboard paired with the phone was used by the testers to interact with the phone; no on-phone buttons were used. Prior to the start of testing, all participants received training on the entire procedure, with instructions provided both verbally and in writing.

The speech MCL was established at the beginning of testing and locked on the phone for the remainder of testing. While participants held the phone to their ear or hearing device's microphone, the volume control (VC) setting of the phone was set at its mid-point. Each participant then listened to the telephone speech indicating whether they wanted the tester to increase or decrease the VC so that the speech level was

comfortably loud. The VC setting was adjusted up and down several times to converge on a consistent MCL judgment. The phone was then placed in the stand. The VC setting for the MCL judgment was confirmed and locked.

As in Experiment 1, participants listened to and then repeated each sentence that they heard, and researchers scored their responses for the number of words correctly repeated in each sentence. Following presentation of all 12 sentences for a given condition, the Mean Opinion Score (MOS) was administered. The MOS is an absolute category rating for speech quality on a 5-point scale from excellent/5 to bad/1. Participants selected the category which best represented the overall quality of speech they experienced when listening to the sentences for a given condition. Lastly, audibility of third octave band noises centered at 150, 250, 4k, 5k Hz was tested to reconfirm the results of the audibility screening individuals were required to pass in order to participate in the study. As in Experiment 1, presentation of conditions was counterbalanced across subjects.

Results

Overall, results showed that WB audio again provided benefits over NB audio for both speech recognition and speech quality. Additionally, lower bit rates for a given audio bandwidth produced lower ratings of speech quality, and for NB audio, poorer speech recognition also resulted.

Figure 3. Mean percent words correct and standard error for narrowband and wideband audio, grouped by the bit rates for telephony (low bit rate), and the best-quality bit rates allowed by the respective codecs (high bit rate). WB-LP is the lowpass-filtered wideband audio encoding to mimic narrowband audio at the higher bit rates supported by AMR-WB.

It should be noted that overall performance levels were high for the speech recognition task. Between 5 to 15 participants for each audio codec condition had maximum or near maximum speech recognition scores of 98 – 100 % correct. Given the high proportion of participants having maximal scores, it is possible that a ceiling effect [9] occurred for speech recognition, especially for WB audio codec conditions. Although the independent variable of codec did have an effect on the dependent variable of speech recognition as evidenced by the significant post hoc tests

described below, the differences between the means in most cases were small.

Figure 4. Mean speech quality (MOS) and standard error for narrowband and wideband audio, grouped by the bit rates for telephony (low bit rate), and the best-quality bit rates allowed by the respective codecs (high bit rate). WB-LP is the lowpass-filtered wideband audio encoding to mimic narrowband audio at the higher bit rates supported by AMR-WB.

Figure 5. Mean percent words correct and standard error for varying bit rates, grouped by NB and WB audio.

Figure 6. Mean speech quality (MOS) and standard error for varying bit rates, grouped by NB and WB audio.

Pair	Bit Rate	% Words Correct df(35)	MOS df(35)
NB 5.9 x WB 12.65	Low	$t=5.00$, $p<0.0001$	$t=4.79$, $p<0.0001$
NB 12.2 x WB 23.85	High	$t=2.94$, $p<0.0057$	$t=6.81$, $p<0.0001$
NB 12.2 x WB-LP	High	$t=1.81$, $p<0.0792$	$t=0.96$, $p<0.3430$
WB 23.85 x WB-LP	High	$t=1.38$, $p<0.1776$	$t=5.05$, $p<0.0001$

Table 4. Pairwise comparisons (n=36) with post hoc significance testing for speech recognition and MOS of AMR-WB x AMR-NB audio, cf. the graphs in Figure 3-Figure 4.

Pair	Band width	% Words Correct df(35)	MOS df(35)
NB 5.9 x NB 12.2	NB	$t=3.03$, $p<0.0046$	$t=2.25$, $p<0.0309$
NB 5.9 x G.711 64.0	NB	$t=2.16$, $p<0.0378$	$t=2.58$, $p<0.0142$
NB 12.2 x G.711 64.0	NB	$t=0.57$, $p<0.5721$	$t=0.27$, $p<0.7859$
WB 12.65 x WB 23.85	WB	$t=0.77$, $p<0.4474$	$t=2.94$, $p<0.0057$

Table 5. Pairwise comparisons (n=36) with post hoc significance testing for speech recognition and MOS of varying bit rates, corresponding to the graphs in Figure 5-Figure 6.

Speech Recognition and Speech Quality Ratings: A one-way ANOVA with repeated measures was performed separately for the speech recognition ($F(5,175)=8.93$, $p<10^{-6}$) and speech quality (MOS) ($F(5,175)=17.3$, $p<10^{-6}$) data, indicating in both cases that the means for the six audio codec conditions were not all equal. Further post hoc significance testing was done on select pairwise comparisons to examine the effects of audio bandwidth and bit rate.

Audio Bandwidth: For the same coding strategy (AMR), WB audio @ 12.65 and 23.85 kbps produced better speech quality and speech recognition than NB audio @ 5.90 and 12.20 kbps, respectively. Filtering the higher bit rate WB audio (23.85 kbps), reduced speech quality, but not speech recognition compared to unfiltered WB audio of the same bit rate, while no difference was observed between the filtered WB audio and AMR-NB @ 12.20 kbps (Figure 3-Figure 4 and Table 4).

Bit rates: Higher bit rates produced better speech quality for both NB and WB audio than their lower bit rate counterparts. For AMR-NB audio, a higher bit rate also produced better speech recognition. The much higher bit rate for G.711 NB audio (64 kbps) produced neither better speech quality nor speech recognition than AMR-NB @ 12.20 kbps, although G.711 showed better speech recognition and quality than AMR-NB @ 5.90 kbps (Figure 5-Figure 6 and Table 5).

Experiment 3: Audio Codec Bandwidth and Network Packet Loss

Experiment 2 showed that the WB audio advantage over NB still holds with bit rates typical of mobile telephony, with a real phone as opposed to simulated, and when participants are allowed to position the phone and determine MCL themselves. The follow-up research question for Experiment 3 was to determine the impact of network impairments, on speech recognition, mental effort, and sound quality ratings. An additional question was how these impacts differ between people with hearing loss and hearing people.

In the previous experiment, a ceiling effect was observed for the wideband audio stimuli for a number of participants. This suggests that the CASPER sentence sets were not challenging enough for these participants. Therefore, for this experiment, we employed the IEEE Harvard sentence set [15], which is commonly used in voice telephony audio testing. The sentences have low predictability and pose much greater challenges for speech recognition. We partnered with AT&T Labs to encode the Harvard stimuli, so as to ensure that the audio encoding exactly mirrored the behavior of real-world mobile handsets and networks, and that the packet loss mirrored conditions found in the real world.

Testing of NB vs WB telephone speech at various levels of packet loss was completed with a group of 36 cochlear implant and hearing aids users and 12 hearing controls.

Participants

Thirty-six individuals with hearing loss participated in the study, along with twelve hearing individuals all younger than 50 years old. Of the 36 individuals with hearing loss, 25 were women and 11 were men, with an average age of 51 years (ranging from 22-79 years). All participants had at least two years of self-reported hearing device use. Twelve individuals used their CIs during testing, while the other 24 used their HAs. Self-reported hearing loss ranged from mild to profound across both ears. In the test ear, most HA users reported moderately-severe or severe hearing loss, while all CI users reported profound hearing loss.

Materials

Stimuli for this experiment were drawn from the IEEE Harvard sentence lists [15], which as mentioned above pose a much greater listening challenge than CASPER. These sentences are a collection of 72 lists of 10 sentences that are phonetically balanced, using specific phonemes at the same frequency they appear in English. Because these sentences date back to the 1940s and language use patterns have shifted since then, lists containing words that might have been offensive or unfamiliar to participants today were screened out. In the end, twelve sentence lists were used to prepare the stimuli. Pairs of lists were combined to create six sets of 20 sentences each with, on average, 157 words per set.

Recordings of all Harvard sentences were provided by Sense Synergy and included two male and two female speakers per sentence. Within a set of 20 sentences in a single test condition, five sentences were spoken by each of the four

speakers, and each participant received a different mix of speakers across all conditions. Furthermore, the presentation order of sentences in each set was randomized across test conditions and participants. Additionally, other sentences were used to train participants on the procedure and to establish an individual's MCL for telephone listening.

Stimulus preparation: The conditions associated with codecs, bit rates and packet loss are shown in Table 6 below. Two baseline conditions and four conditions of reduced audio quality due to packet loss were prepared, for a full factorial design with audio bandwidth and packet loss as the factors. The baseline conditions included NB and WB audio with no packet loss at typical data rates used in mobile networks, determined in consultation with AT&T. The other four conditions involved two levels of bursty packet loss, 3% and 20%, for each baseline condition. The worst-case quality levels in a managed mobile network and over unmanaged Wi-Fi Internet calling, respectively, determined the percentages of packet loss selected. These conditions set upper/lower bounds on the effects of packet loss on speech quality likely to be experienced in mobile calling.

Condition	Codec	Band width	Bit Rate (kBit/s)	Packet Loss
1	AMR-NB	NB	5.90	0%
2	AMR-NB	NB	5.90	3%
3	AMR-NB	NB	5.90	20%
4	AMR-WB	WB	12.65	0%
5	AMR-WB	WB	12.65	3%
6	AMR-WB	WB	12.65	20%

Table 6. Experimental conditions of audio codec, bit rate, and packet loss, with a full factorial design across bandwidth and packet loss.

AT&T Labs carried out all signal processing for speech coding and injection of packet loss. To prepare the stimuli for processing, the silences at the beginning and end of each sentence were deleted, and the sentences within a given set were concatenated. Each concatenated sentence set was processed using the NB and WB encoding strategies. Silence suppression (DTX) was hardcoded to be off in both the narrowband and wideband codec implementations used to process all stimuli, which means comfort noise generation and voice activity detection were also disabled. Since the silence preceding and following each sentence was removed before the sentences were joined together, it is unlikely there is any impact of having DTX off. Packet loss concealment (PLC) was provisioned on, and the same technique, a form of waveform substitution, was used for both codecs. Conceptually, with this technique, the last received packet is repeated in place of each lost packet until another packet is received. When more than one packet is lost in a row, as is the case with bursty packet loss, the signal level of each substituted packet is reduced. This reduction continues

progressively for each lost packet until a level approximating that of comfort noise is reached or a new packet is received.

Packet loss was introduced using the Gilbert-Elliott model [12] with Gamma set at 0.8. This model utilizes a two-state Markov model approach and is widely used to generate impairments that simulate transmission failures in real-time services over telecom networks. Within this model, lower levels of Gamma produce more random packet loss distributions, while higher levels of Gamma produce more bursty distributions of packet loss. Packet loss in both mobile and VoIP networks has been characterized as bursty [16], rather than random. Therefore, a higher level of Gamma was selected in order to simulate the bursty nature of packet loss in these telephony environments. Following processing, the sentences in each set were separated and losslessly converted to PCM audio to support playback on an iPhone. 100 ms of silence was added to the beginning and end of each sentence to avoid any clipping induced by playback delays, and levels were equalized as in the previous experiments.

As in the previous experiments, sentence sets were processed for each of the six test conditions, and the sentence set used for each condition was counterbalanced across subjects. This was done to guard against the effects of possible differences in intelligibility, either inherent or as a result of the temporal distribution of lost packets, among sentence sets.

Method

As in Experiment 2, participants used their preferred ear, and self-selected speech MCL and phone positioning (in the case of hearing participants, to their ear). The setup of the hardware and software was identical to that of Experiment 2, as well. The same custom app on an iPhone 4S was used and controlled through a Bluetooth keyboard. As before, the phone was held in position with a stand and a Velcro headband (cf. Experiment 2: Codec Audio Bandwidth and Bit Rate - Method; and Figure 2).

The procedure was analogous to that of the previous experiment (cf. Experiment 2: Codec Audio Bandwidth and Bit Rate). Prior to testing, all participants received training on the entire procedure. The speech MCL was established and locked for the remainder of the testing. For each study participant, speech recognition was tested using one set of IEEE sentences for each of the six audio quality conditions via listening to and repeating sentences, providing ratings; followed by reconfirming audibility of band noises. Participants provided both SMEQ and MOS ratings.

Each administration of one condition took seven minutes. Presentation of conditions was counterbalanced across subjects. To guard against bias, a double-blind procedure was used in which neither researchers administering the experiment, nor the participants, were aware of which conditions were being evaluated for any given sentence set.

Results

Overall, results showed for people with hearing loss that WB audio continues to confer benefits over NB audio in the areas

of speech recognition, perceived mental effort, and speech audio quality ratings. Hearing participants saw benefits, too, especially under packet loss conditions. Additionally, participants with hearing loss had poorer speech recognition, higher perceived mental effort and lower ratings of speech audio quality, compared to their hearing counterparts, regardless of audio codec bandwidth and degree of packet loss. There was no ceiling effect present for participants with hearing loss, in contrast to Experiment 2. The detailed results are shown in Table 7.

Speech Recognition: For participants with hearing loss, a repeated measures, two-way ANOVA for words correct showed significant main effects of the factors audio bandwidth ($F(1,35) = 16.8, p < 0.000$) and packet loss ($F(2,70) = 278, p < 0.000$), but no significant interaction between the two factors ($F(2,70) = 1.44, p < 0.243$).

For the hearing participants, a repeated measures, two-way ANOVA for words correct showed significant main effects of the factors audio bandwidth ($F(1,11) = 16.0, p < 0.002$) and packet loss ($F(2,22) = 31.9, p < 0.000$) and a significant interaction between the two factors ($F(2,22) = 5.82, p < 0.009$); see Figure 7.

Figure 7. Main and interaction effects plot for % words correct with respect to the packet loss and audio bandwidth factors, for both hearing (H) participants and participants with hearing loss (HL). There were significant main effects for packet loss and audio bandwidth for both H and HL, but only a significant interaction for H.

Mental Effort: For participants with hearing loss, a repeated measures, two-way ANOVA for SMEQ ratings showed significant main effects of the factors audio bandwidth ($F(1,35) = 4.44, p < 0.042$) and packet loss ($F(2,70) = 54.1, p < 0.000$), but no significant interaction between the two factors ($F(2,70) = 1.75, p < 0.181$).

For the hearing participants, a repeated measures, two-way ANOVA for SMEQ ratings showed significant main effects of the factors audio bandwidth ($F(1,11) = 14.0, p < 0.003$) and packet loss ($F(2,22) = 47.3, p < 0.000$), but no significant interaction between the two factors ($F(2,22) = 0.329, p < 0.723$); see Figure 8.

Mean Opinion Score: For participants with hearing loss, a repeated measures, two-way ANOVA for MOS ratings showed significant main effects of the factors audio bandwidth ($F(1,35) = 7.75, p < 0.009$) and packet loss ($F(2,70) = 58.2, p < 0.000$) and a significant interaction between the two factors ($F(2,70) = 3.22, p < 0.046$).

Figure 8. Main and interaction effects plot for SMEQ with respect to packet loss and audio bandwidth, for both hearing (H) participants and participants with hearing loss (HL). There were significant main effects for packet loss and audio bandwidth for both H and HL, but no significant interaction.

For the hearing participants, a repeated measures, two-way ANOVA for MOS ratings showed significant main effects of the factors audio bandwidth ($F(1,11) = 22.6, p < 0.001$) and packet loss ($F(2,22) = 77.0, p < 0.000$), but no significant interaction between the two factors ($F(2,22) = 1.06, p < 0.363$); see Figure 9.

Figure 9. Main and interaction effects plot for MOS with respect to the packet loss and audio bandwidth factors, for both hearing (H) participants and participants with hearing loss (HL). There were significant main effects for packet loss and audio bandwidth for both H and HL, but only a significant interaction for HL.

LIMITATIONS

Study limitations are related to the stimuli and data collection methods. With regard to stimuli, we used a limited set of audio codecs; other codecs should be explored. The packet loss conditions used static percentages of packet loss, while real-world packet loss rates vary over time. The stimuli were

receptive listening tasks, and may not reflect performance during phone conversations between two people. In terms of methods, we used microphone coupling (i.e., the audio goes from phone speaker to hearing device microphone) and a fixed handset position. We did not explore other coupling methods, such as telecoil or Bluetooth coupling, or the variations that can occur in the acoustic signal received at the hearing device's microphone when the handset is held at the ear without the fixed positioning we used in the experiments. Addressing each of these limitations is important to understanding how our findings translate to real-world use.

CONCLUSIONS AND FUTURE WORK

We have shown for 68 cochlear implant and 46 hearing aid users, for a total of 114 participants with hearing loss, that wideband audio confers significant benefits in the areas of speech recognition, reduced perceived mental effort, and perceived audio quality. These findings have been replicated across three studies. The codec and bit rate typical of wireless mobile telephony – AMR-NB 5.90 – results in both poorer speech recognition and poorer perceived audio quality than higher-quality narrowband encodings prevalent on VoIP and PSTN. For wideband audio, the typical AMR-WB 12.65 supports recognition similar to higher quality bit rates, but perceived audio quality is still better with higher bit rates. Packet loss results in significantly degraded performance; however, the advantage of wideband over narrowband audio still holds, especially on managed networks that limit packet loss to 3%. Finally, we have also confirmed that among hearing people, wideband audio confers benefits, especially under degraded network conditions.

Overall, these studies strongly suggest that ubiquitous wideband audio in telecommunications is a win-win for hearing people and people with hearing loss alike – speech quality and accessibility are both significantly improved over narrowband. Implementers of telecommunications systems should favor wideband audio over narrowband audio, and technical/policy standards should make wideband support required (e.g., as already done in Section 508 in the US, and ETSI EN301 549 in Europe), rather than recommended (e.g., the next-generation emergency calling specification in North America [29]). Telecommunication providers also should make sustained efforts to ensure interoperability of wideband

audio calls across environments and carriers, which has languished in the US since a 2014 petition in front of the FCC [38]. If the narrowband audio codec AMR-NB must be supported in legacy telecom, bit rates higher than 5.90 kBit/s should be considered. Additionally, managed networks with quality of service guarantees, so as to minimize packet loss, can potentially offer better accessibility for voice calls than services that operate on the open internet.

Now that we better understand the impact of speech coding and network impairments, appropriate follow-up studies should explore these technical parameters in more ecologically valid tasks and environments. Future work also should consider what bit rates are appropriate for people with hearing loss for the Opus codec, which is widely implemented in web-based audio and video calling, and adopted as a requirement in the 2017 Section 508 Refresh. Additionally, the interaction between wideband audio and synchronized video should be considered, as well as the performance of wideband audio in noisy environments that are typical of mobile phone use. Lastly, although this paper provides clear evidence toward improving receptive telephone listening, the relationship between listening and the ability of people with hearing loss to hold phone conversations is poorly understood, and merits further investigation using conversational-type tasks.

ACKNOWLEDGMENTS

This paper was developed under a grant from the AT&T Foundation, and under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90RE5003). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this paper do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government. Laurie Garrison of AT&T Labs carried out all signal processing for speech coding and injection of packet loss. Norman Williams developed the software for presenting the stimuli and recording participant responses. Rosemary Johnson assisted in preparing the stimuli for encoding, data collection and participant interaction. Mary Garvert assisted in the data collection and participant interaction.

Cond.	HL % Words Correct	% Std Err	HL SMEQ	SMEQ Std Err	HL MOS	MOS Std Err	NH % Words Correct	% Std Err	NH SMEQ	SMEQ Std Err	NH MOS	MOS Std Err
NB 0%	77.6	3.5	50.3	6.4	3.4	0.2	98.1	0.8	14.3	3.1	3.8	0.2
NB 3%	73.6	3.3	55.8	5.4	3.1	0.2	95.3	1.1	26.2	3.9	3.3	0.2
NB 20%	42.9	3.1	83.8	5.8	2.0	0.2	83.0	2.6	60.6	7.1	2.2	0.3
WB 0%	81.2	3.4	37.1	5.2	3.9	0.2	99.2	0.9	2.3	0.9	4.7	0.1
WB 3%	77.3	3.5	47.6	5.7	3.4	0.2	98.5	0.9	14.8	3.0	3.9	0.2
WB 20%	51.2	3.8	84.1	5.3	2.0	0.2	91.5	1.7	45.2	5.3	2.6	0.3

Table 7. Results for testing audio bandwidth and packet loss on participants with hearing loss (HL) and hearing participants (NH).

REFERENCES

- [1] 3GPP TR 26.935 V8.0.0 Technical Report. 2008. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Packet-switched conversational multimedia applications; Performance characterisation of default codecs (Release 8). Retrieved May 1, 2019 from <https://www.3gpp.org/ftp/Specs/3GPP/TS/26/935/TS26935-800.pdf>
- [2] Doug L. Beck and Jes Olsen. 2008. Extended bandwidths in hearing aids. *Hearing Review* 15(11), 22-26.
- [3] Arthur Boothroyd. 1987. CASPER: Computer-Assisted speech perception evaluation and training. In 10th annual conference on Rehabilitation Technology, Washington, DC.
- [4] Marc A. Brennan, Ryan McCreery, Judy Kopun, Brenda Hoover, Joshua Alexander, Dawna Lewis, and Patricia G. Stelmachowicz. 2014. Paired comparisons of nonlinear frequency compression, extended bandwidth, and restricted bandwidth hearing aid processing for children and adults with hearing loss. *Journal of the American Academy of Audiology* 25(10), 983-998. <https://doi.org/10.3766/jaaa.25.10.7>
- [5] James R. Bress, Steven Whitesell, and Tony Jasionowski. 2013. Telecommunications Industry Association (TIA): [Developing Standards for Accessibility](http://standards.tiaonline.org/standards/committees/documents/DevelopingStandardsForAccessibility.pdf). Presented to the TR-41.3.14 Accessibility Working Group. Retrieved May 1, 2019 from <http://standards.tiaonline.org/standards/committees/documents/DevelopingStandardsForAccessibility.pdf>
- [6] N. Cherniavsky, A. Cavender, R. Ladner, and E. A. Riskin. 2007. Variable frame rate for low power mobile sign language communication. *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*, pp. 163-170. ACM. <https://doi.org/10.1145/1296843.1296872>
- [7] Teresa YC Ching and Harvey Dillon. 2013. A brief overview of factors affecting speech intelligibility of people with hearing loss: implications for amplification. *American Journal of Audiology* 22, 306-309. [https://doi.org/10.1044/1059-0889\(2013\)12-0075](https://doi.org/10.1044/1059-0889(2013)12-0075)
- [8] J. Chon, S. Whittle, E. A. Riskin and R. Ladner. 2011. Improving compressed video sign language conversations in the presence of data loss. *2011 Data Compression Conference*, pp. 383-392. IEEE. <https://doi.org/10.1109/DCC.2011.45>
- [9] A. Ebrahimi-Madiseh, R. H. Eikelboom, D.M., Jayakody, and M.D. Atlas. 2016. Speech perception scores in cochlear implant recipients: An analysis of ceiling effects in the CUNY sentence test (Quiet) in post-lingually deafened cochlear implant recipients. *Cochlear Implants International*, 17(2), 75-80. <https://doi.org/10.1080/14670100.2015.1114220>
- [10] Federal Communications Commission. 2018. 2018-2019 TRS Rate Order. *FCC #DA-18-680*. Retrieved May 1, 2019 from <https://www.fcc.gov/document/2018-2019-trs-rate-order>
- [11] Hans W. Gierlich. 2005. Wideband speech communication-The quality parameters as perceived by the user. In *Proceedings of Forum Acusticum*, Vol. 2005.
- [12] Gerhard Haßlinger and Oliver Hohlfeld. 2008. The Gilbert-Elliott model for packet loss in real time services on the Internet. In *14th GI/ITG Conference-Measurement, Modelling and Evaluation of Computer and Communication Systems* 1-15 VDE.
- [13] Christian Hoene, Berthold Rathke, and Adam Wolisz. 2003. On the importance of a VoIP packet. In *ISCA Tutorial and Research Workshop on Auditory Quality of Systems*, 55-62.
- [14] Yi Hu, Qudsia Tahmina, Christina Runge, and David R. Friedland. 2013. The perception of telephone-processed speech by combined electric and acoustic stimulation. *Trends in amplification*, 17(3), 189-196. <https://doi.org/10.1177/1084713813512901>
- [15] Institute of Electrical and Electronics Engineers. 1969. IEEE recommended practice for speech quality measurements. *IEEE transactions on audio and electroacoustics*, 17(3), 225-246.
- [16] Wenyu Jiang and Henning Schulzrinne. 2000. Modeling of packet loss and delay and their effect on real-time multimedia service quality. In *Proc. NOSSDAV*.
- [17] Stephen Julstrom, Linda Kozma-Spytek, and Scott Isabelle. 2011. Telecoil-mode hearing aid compatibility performance requirements for wireless and cordless handsets: magnetic signal-to-noise." *Journal of the American Academy of Audiology* 22(8), 528-541. <https://doi.org/10.3766/jaaa.22.8.5>
- [18] Stephen Julstrom and Linda Kozma-Spytek. 2014. Subjective assessment of cochlear implant users' signal-to-noise ratio requirements for different levels of wireless device usability. *Journal of the American Academy of Audiology* 25(10), 952-968. <https://doi.org/10.3766/jaaa.25.10.4>
- [19] Anna Chi Shan Kam, John Ka Keung Sung, Tan Lee, Terence Ka Cheong Wong, and Andrew van Hasselt. 2017. Improving mobile phone speech recognition by personalized amplification: application in people with normal hearing and mild-to-moderate hearing loss. *Ear and Hearing* 38(2), e85-e92. <http://dx.doi.org/10.1097/AUD.0000000000000371>

- [20] Mead Killion and H. Gustav Mueller. 2010. Twenty Years Later: A New Count-the-Dots-Method. *The Hearing Journal* 63:1, 10-17.
- [21] Linda Kozma-Spytek, Paula Tucker, and Christian Vogler. 2013. Audio-visual speech understanding in simulated telephony applications by individuals with hearing loss. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, No. 6. ACM. <https://doi.org/10.1145/2513383.2517032>
- [22] Charles J. Limb, and Alexis T. Roy. 2014. Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hearing Research* 308, 13-26. <https://doi.org/10.1016/j.heares.2013.04.009>
- [23] Frank R. Lin, John K. Niparko, and Luigi Ferrucci. 2011. Hearing Loss Prevalence in the United States. *Arch Intern Medicine*, 171(20), 1851–1853. <https://doi.org/10.1001/archinternmed.2011.506>
- [24] Frank R. Lin, Roland Thorpe, Sandra Gordon-Salant, and Luigi Ferrucci. 2011 Hearing Loss Prevalence and Risk Factors Among Older Adults in the United States. *J Gerontol A Biol Sci Med Sci*. 66A(5), 582–590. <https://doi.org/10.1093/gerona/glr002>
- [25] Chuping Liu, Quien-Je Fu, and Shrikanth S. Narayanan. 2009. Effect of bandwidth extension to telephone speech recognition in cochlear implant users. *The Journal of the Acoustical Society of America*, 125(2), EL77–EL83. <https://doi.org/10.1121/1.3062145>
- [26] Georgios Mantokoudis, Claudia Dähler, Patrick Dubach, Martin Kompis, Marco D. Caversaccio, and Pascal Senn. 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>
- [27] Steven C. Marcum, Erin M. Picou, and Thomas Steffens. 2017. Avoiding disconnection: An evaluation of telephone options for cochlear implant users. *International Journal of Audiology* 56(3), 186-193. <https://doi.org/10.1080/14992027.2016.1247502>
- [28] National Academies of Sciences, Engineering, and Medicine (NASEM). 2016. Hearing health care for adults: priorities for improving access and affordability. Retrieved from <http://nationalacademies.org/hmd/reports/2016/Hearing-Health-Care-for-Adults.aspx>
- [29] National Emergency Number Association. 2016. Detailed Functional and Interface Standards for the NENA i3 Solution. *NENA-STA-010.2-2016*. Retrieved from https://www.nena.org/page/i3_Stage3
- [30] W. Nogueira, J. Abel, and T. Fingscheidt. 2019. Artificial speech bandwidth extension improves telephone speech intelligibility and quality in cochlear implant users. *The Journal of the Acoustical Society of America*, 145(3), 1640-1649. <https://doi.org/10.1121/1.5094347>
- [31] OpenCore-AMR. 2019. Library of OpenCORE Framework implementation of Adaptive Multi Rate Narrowband and Wideband (AMR-NB and AMR-WB) speech codec. Retrieved April 27, 2019 from <https://sourceforge.net/projects/opencore-amr/>
- [32] Erin M. Picou and Todd A. Ricketts. 2013. Efficacy of hearing-aid based telephone strategies for listeners with moderate-to-severe hearing loss. *Journal of the American Academy of Audiology* 24(1), 59-70. <https://doi.org/10.3766/jaaa.24.1.7>
- [33] Christina M. Roup, Gayla L. Poling, Kimberly Harhager, Ashok Krishnamurthy, and Lawrence L. Feth. 2011. Evaluation of a telephone speech-enhancement algorithm among older adults with hearing loss. *Journal of Speech, Language, and Hearing Research* 54(5), 1477-1483. [https://doi.org/10.1044/1092-4388\(2011/10-0181\)](https://doi.org/10.1044/1092-4388(2011/10-0181))
- [34] Jeff Sauro and Joseph S. Dumas. 2009. Comparison of three one-question, post-task usability questionnaires. *Proceedings of the SIGCHI conference on human factors in computing systems*. 1599-1608. ACM. <https://doi.org/10.1145/1518701.1518946>
- [35] Angeline Seeto and Grant D. Searchfield. 2018. Investigation of Extended Bandwidth Hearing Aid Amplification on Speech Intelligibility and Sound Quality in Adults with Mild-to-Moderate Hearing Loss. *Journal of the American Academy of Audiology* 29(3), 243-254. <https://doi.org/10.3766/jaaa.16180>
- [36] J.J. Tran, B. Flowers, E. A. Risken, R. Ladner and J. O. Wobbrock. 2014. Analyzing the intelligibility of real-time mobile sign language video transmitted below recommended standards. *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, pp. 177-184. <https://doi.org/10.1145/2661334.2661358>
- [37] Vanessa Vas, Michael A. Akeroyd, and Deborah A. Hall. 2017. A Data-Driven Synthesis of Research Evidence for Domains of Hearing Loss, as Reported by Adults with Hearing Loss and Their Communication Partners. *Trends in Hearing*, 21, 1-25. <https://doi.org/10.1177/2331216517734088>
- [38] Voice Communication Exchange Committee (VCXC). 2014. Petition for Rulemaking. *Filed with the FCC*, GN Docket 13-5. February 25, 2014. Retrieved from <https://www.fcc.gov/ecfs/filing/6017604148>
- [39] Jace Wolfe, Mila Morais Duke, Erin Schafer, George Cire, Christine Menapace, and Lori O'Neill. 2016. Evaluation of a wireless audio streaming accessory to

improve mobile telephone performance of cochlear implant users. *International Journal of Audiology* 55 (2), 75-82.

<https://doi.org/10.3109/14992027.2015.1095359>

GLOSSARY

AMR-NB – Adaptive Multirate Narrowband, used to encode narrowband audio on cellular networks.

AMR-WB – Adaptive Multirate Wideband, used to encode wideband audio on cellular networks.

CI – shorthand for cochlear implant.

DTX – shorthand for discontinuous transmission, a method of saving bandwidth during moments of silence.

G.711 u-law (Toll-Quality Speech) – uncompressed narrowband audio transmitted via internet-based telephony. Comparable to best-case analog phone networks (PSTN).

HA – shorthand for hearing aid.

MCL (Most Comfortable Loudness Level) – The volume setting deemed most comfortable by a person.

Packet Loss Concealment (PLC) – a method to smooth over packet loss; for example, by repeating and attenuating the audio information from the previous packet.

PCM – Pulse Code Modulation, lossless encoding of audio.

Public Switched Telephone Network (PSTN) – the analog wired phone and first- to third-generation cellular networks.

RF Interference (Radio Frequency Interference) – interference caused in the hearing device by the electromagnetic emissions of digital wireless phones.

VoIP (Voice over Internet Protocol) – internet-based telephony, as well as fourth-generation cellular telephony

VoLTE (Voice over Long-Term Evolution) – internet protocol-based transmission of voice over fourth-generation cellular networks.