ARTICLE

The Effect of Background Noise and Music on Speech Recognition Performance of Individuals with Normal Hearing and Hearing Loss

Eun Yeon Kim1, Sang Eun Lee2, Hye Yoon Seol3*

1 Department of Speech Language Pathology, Graduate School of Interdisciplinary Therapy, Myongji University, Seoul, 03674, Korea
2 Department of Music Therapy, Graduate School of Interdisciplinary Therapy, Myongji University, Seoul, 03674, Korea
3 Department of Communication Disorders, Ewha Womans University, Seoul, 03760, Korea

ABSTRACT

This study explores speech recognition characteristics in background noise and music between normal hearing (NH) listeners, hearing aid (HA) users and cochlear implant (CI) users. Sixty individuals participated in the study: 20 with NH, 20 HA users, and 20 CI users. HA and CI users had a Categories of Auditory Performance score of 6 and open set sentence recognition of 85% or higher. They had been using the devices for at least one year. Babble noise (BN), piano solo (PS), piano + violin (PV), and piano + chorus (PC) were presented at +5- and +10 dB signal-to-noise ratio (SNR). The participants were asked to listen and repeat words and sentences from the Korean Standard Sentence List for Adults. At +5- and +10 dB SNRs, CI users performed worse than those with NH on word and sentence recognition in BN, PS, PV, and PC. HA users outperformed CI users in all conditions. Those with NH showed better sentence recognition than HA users across all conditions at +5 dB SNR and better word recognition in PV and BN at +5 dB SNR and in PC at +10 dB SNR. Correlational analysis revealed that the percentage of life with hearing loss before CI was not correlated with sentence and word recognition across all conditions in both SNRs. Statistically significant negative correlations were observed between the duration of deafness and sentence and word recognition in some conditions. Despite individuals with HL performing well on clinical tests, background music can still interfere with communication for those using hearing devices. To accurately evaluate individuals’ communication abilities clinical tools that include background music need to be developed. Studies using different types of music could help develop and standardize such tools for assessing speech and language abilities in individuals with hearing loss.

Keywords: Hearing loss; Audiology; Rehabilitation

*CORRESPONDING AUTHOR:
Hye Yoon Seol, Department of Communication Disorders, Ewha Womans University, Seoul, 03760, Korea; E-mail: seol.helena@ewha.ac.kr

ARTICLE INFO
Received: 11 April 2024 | Revised: 2 May 2024 | Accepted: 15 May 2024 | Published Online: 26 June 2024
DOI: https://doi.org/10.30564/fls.v6i3.6583

CITATION

COPYRIGHT
Copyright © 2024 by the author(s). Published by Bilingual Publishing Group. This is an open access article under the Creative Commons Attribution (CC BY) License (https://creativecommons.org/licenses/by/4.0/).
1. Introduction

Communication can be explained through the speech chain (Denes & Pinson, 2016). When speech is produced, it gets transmitted to the auditory system of both the speaker and the listener in the form of sound waves. The auditory system consists of four parts: the outer ear, middle ear, inner ear, and auditory nerve (Møller, 2012). The outer ear, also known as the pinna and the ear canal, acts like a funnel, gathering sound and transmitting it to the ear drum in the middle ear. In the middle ear, there are also three small bones called the malleus, incus, and stapes, which transmit sound to the inner ear. Additionally, the middle ear contains a structure called the eustachian tube, which helps equalize the air pressure between the ambient pressure and the pressure inside the middle ear. In the inner ear, there is the cochlea, which houses the sensory organ of hearing. The cochlea is a snail-shaped structure, where the base of the cochlea processes high-frequency information, while the apex processes low-frequency information. The sound then gets transmitted to the auditory cortex of the brain through the auditory nerve. At this stage, the listener interprets the meaning of the heard speech using his or her linguistic knowledge. For effective communication, linguistic knowledge and appropriate neuromuscular activity are crucial, but acoustic signals must also be properly transmitted and interpreted by the auditory system; accurate perception of sound is vital for human communication and for individuals to imitate spoken language, auditory perception needs to be connected to motor skills (Prather, 2013).

Hearing loss (HL) is a major public health problem around the world (Diseases & Injuries, 2020; McDaid et al., 2021; Organization, 2017). Numerous studies have revealed that HL impacts various aspects of life negatively, ranging from communication to academic and job performance, and to social isolation (Cunningham & Tucci, 2017; Davis et al., 2016; Kramer, 2008; Wagner-Hartl et al., 2018; Wilson et al., 2017). Kramer et al. 2006 used the Amsterdam Checklist for Hearing and Work to investigate occupational performance between individuals with normal hearing (NH) and HL. A total of 210 participants (60 with NH and 150 with HL) were enrolled in the study and the results showed that HL was correlated with more effort required for listening, meaning that those with HL find listening to be a challenging task. The effort for listening was related to hearing tasks at work, such as sound detection. Additionally, when examining subjective noise levels at work, individuals with HL perceived higher levels of noise than NH listeners, indicating that those with HL are more sensitive to noise at work (Kramer, 2008).

In relation to speech and language development, particularly for children, receiving auditory input and feedback is crucial regardless of the type and degree of HL, as well as whether it is unilateral or bilateral. HL can lead to a lack of auditory stimulation, potentially causing delays in speech and language development (Anne et al., 2017; Ching et al., 2010; Halliday et al., 2017; Lieu et al., 2020; Stelmachowicz et al., 2004). Anne et al. 2017 conducted a systematic review of speech and language characteristics of unilateral HL in children. Among the 13 research studies the authors reviewed, seven studies showed poorer performance on speech and language tests, with these findings being particularly pronounced among children with severe to profound HL (Anne et al., 2017).

With these negative consequences, it is important to manage HL through the use of hearing devices, such as hearing aids (HAs) and cochlear implants (CIs). Both HAs and CIs are devices that aid individuals with HL in perceiving sounds and understanding speech; HAs amplify sounds (acoustic signals), while CIs, as surgically implanted devices, provide a sense of sound through electrical stimulation. Individuals with HL typically start their rehabilitation process through HA use. If HL becomes too severe that it can no longer be compensated by HAs, individuals consider CIs. Unlike HAs, CIs deliver electrical signals, so CI users’ rehabilitation process involves matching these electrical signals to speech, music, and
environmental sounds. Music, commonly encountered in the form of background sounds in various places, plays a role not only in drawing attention but also in strengthening sustained attention. It has been noted in many research studies that music has an impact on depression (Siedliecki & Good, 2006), pain (Siedliecki & Good, 2006), work quality (Lesiuk, 2005), etc. From a linguistic perspective, researchers have reported similarities between language and music (Brandt et al., 2012; McMullen & Saffran, 2004; Temperley, 2022). Music and language are similar in a way that both of them are created based on one's database of sounds that could be chosen to formulate and can be adjusted or modified so that they can sound attractive to others. For example, when adults interact with infants and children, they tend to use infant-directed speech which has characteristics of slower rate, longer pause, and higher pitch (Fernald, 1992). Adults exaggerating sounds could lead to early learning of speech and language as it engages children’s attention (Thiessen et al., 2005). Similar characteristics are observed for music—compared to songs targeting adults, songs for infants and children are usually simple and have repetitive pitch contours (Trainor & Trehub, 1998). Studies have shown that infants can develop their knowledge about the prosodic structures of their native language (Friederici, 2006; Ramus et al., 1999; Wermke et al., 2007). Using this knowledge about prosodic features, infants are known to differentiate between their native language and non-native language (Friederici, 2006). Wermket et al. 2007 examined melodies of infant cries using the Melody Complexity Index and reported that infants who show less than 45% complexity in their cries could experience a language delay (Wermke et al., 2007).

Although many studies have shown the positive effect of music on various aspects of life, music may not always be enjoyable to everyone—HA and CI users could have difficulty with music perception and appreciation. In order for HAs and CIs to enhance speech understanding and quality of life, HA and CI programming based on one’s hearing characteristics and communication needs is required. Appropriately programmed HAs and CIs are known to improve speech understanding and quality of life (Cho et al., 2019; Ketterer et al., 2020; Laske et al., 2009; Mondelli & de Souza, 2012; Seol et al., 2021; Seol & Moon, 2022). Ketter et al. 2020 investigated CI benefit in speech perception, quality of life, tinnitus distress, and psychological comorbidities. Comparing their performance before implantation and 6- and 24 months after implantation, participants were able to understand speech better with CIs and their quality of life significantly improved as well (Ketterer et al., 2020). Mondelli et al. 2012 examined the quality of life before and after three months of HA use for the elderly. Thirty HA users completed the World Health Organization Quality of Life Questionnaire and the results revealed significant improvement in quality of life as all participants reported that their quality of life became good or very good after HA use. The participants were also satisfied with their personal relationships (i.e., friends, relatives, and colleagues) (Mondelli & de Souza, 2012). Although numerous studies have documented HA and CI benefits, optimizing the devices does not always lead to improvement in communication and quality of life; even with HAs and CIs individuals still struggle with speech understanding in noise, phone conversations, and music (Bruns et al., 2016; Jung et al., 2010; Looi et al., 2012; Nasresfahani et al., 2022; Philips et al., 2012; Riley et al., 2018). For example, the perception of music through a CI is influenced by multiple variables and the effects vary widely among individuals (Looi et al., 2012). While CI users may have temporal resolution skills equivalent to those with NH, their frequency-resolution skills are often affected, impacting music perception (Looi et al., 2012). CI sound processors have limited filter bands with fixed center frequencies, potentially hindering the accurate resolution of lower harmonics in complex sounds. Factors, such as electrode insertion depth and processing of complex acoustic signals, can also contribute to inaccuracies in pitch perception. Jung et al. 2010 investigated music perception ability in Korean CI users and reported that, irrespective of positive correlation between speech performance and pitch discrimination ability, CI users showed a
wide range of performance when it came to pitch
discrimination and melody and timbre identification
(Jung et al., 2010).

Although there has been extensive research on
the effects of music on human behavior and the
perception and appreciation of music among CI
users, research on speech recognition performance
in situations where background noise and music are
presented is limited, especially among individuals
with various types of HL wearing hearing devices.
Regardless of HL, when considering conversational
environments in daily life, conversations occur
in both quiet and noisy situations. Therefore, it is
important to measure speech understanding abilities
in various environments using noises encountered in
the real world. This study investigates the effect of
background noise and music on speech recognition
in individuals with NH and HL with a hypothesis
that compared to those with NH, noise and music
may have a negative impact on speech recognition
for HA and CI users.

2. Materials and methods

2.1 Participants

A total of 60 adults were enrolled in the study. Among
the 60 adults, 20 were NH listeners, 20 were
HA users, and 20 were CI users. The HL group
was divided into CI and HA groups depending
on the hearing devices they wore. The CI group
consisted of individuals with sensorineural HL of 70
dB (equivalent to severe HL) or greater who have
been wearing CIs for at least one year (mean = 4.7,
SD = 2.7). The HA group comprised individuals with
sensorineural HL of 26 dB (equivalent to mild HL)
or greater who have been using HAs for at least one
year (mean = 9.1, SD = 6.7). For CI and HA groups,
individuals were eligible to participate in the study if
they achieved a Categories of Auditory Performance
(CAP) score of 6 and demonstrated a performance of
85% or above in Seoul National University Hospital
(SNUH) Everyday Sentence Recognition Test, an open-
set speech test commonly used in the clinical settings
to examine the benefit of hearing devices. The CAP is
an eight-point hierarchical scale that can evaluate an
individual’s auditory performance. The CAP categories
range from 0 to 7, with 0 representing no awareness
of environmental sounds and 7 representing the use
of a telephone with a known speaker (Nikolopoulos
et al., 1999). The CAP score of 6 indicates that after CI
implantation, individuals can understand conversations
without having to read lip movements at least. In
addition, the Percentage of Life (PoL) with HL was
calculated for both CI and HA groups. For CI users,
PoL spans from the onset of HL to the CI surgery, while
for HA users, the PoL covers the time from HL onset to
the use of HAs. The average PoLs were 40.51% for the
CI group and 19.78% for the HA group. Lastly, duration
of deafness (DoD), the period from the point when no
benefit could be obtained from HAs until CI surgery,
was examined for the CI group. The average DoD
was 5.45 years for the CI group. In terms of exclusion
criteria, individuals with a history of neurological
diseases or cleft palate, disorders of the articulatory
organ, sensory disorders, cognitive disorders, behavioral
disorders, and voice disorders were excluded from
the study. All experimental procedures were approved
by the Institutional Review Board (IRB) of Samsung
Medical Center and an informed consent document was
obtained for all participants.

2.2 Speech recognition test

The speech recognition testing was performed
using the Korean Standard Sentence List for Adults
(KS-SL-A). The KS-SL-A consists of a total of 80
sentences and 320 target words (10 sentences and
40 target words * 8 lists). The test sentences were
presented under four conditions: babble noise (BN),
piano solo (PS), piano + violin (PV), and piano +
chorus (PC). The stimuli (test sentences) were
presented at +5- and +10 dB signal-to-noise ratios
(SNRs) through a speaker located 1m in front of the
participant. + 5dB SNR means that the test sentences
were presented 5 dB louder than the noise and music
and +10 dB SNR means that the test sentences were
10 dB louder than the noise and music. In terms of
scoring, percent-correct scores for sentences and
words were obtained based on the correct responses
from the participants.

### 2.3 Statistical analysis

Statistical analysis was performed using the SAS Package (ver. 9.1.3, SAS Institute Inc., Cary, NC, USA). Our results passed the normality test (Kolmogorov-Smirnov test). Repeated Measures ANOVA (RM ANOVA) was performed to examine differences in sentence and word recognition performance between NH, HA, CI groups in BN, PS, PV, and PC at +5- and +10 dB SNRs. Pearson’s correlation was performed to investigate the relationship between PoL, DoD, and speech recognition in CI users.

### 3. Results

#### 3.1 Demographic information

Participant characteristics are described in Table 1. The average ages of the NH, CI, and HA groups were 49.70 (SD = 14.4), 51.97 (SD = 14.3), and 51.08 (SD = 17.1) years, respectively. All participants in the HA and CI groups had post-lingual HL, meaning that HL occurred after these individuals acquired speech and language knowledge and skills. Based on individuals’ hearing test results, the averages of hearing thresholds at four frequencies (500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz) were calculated. Regarding the degree of HL, for adults, the cut-off for NH is 25 dB. If the puretone average falls between 26 and 40 dB, it is considered mild HL; between 41 and 55 dB, moderate HL; between 56 and 70 dB, moderately severe HL, between 71 and 90, severe HL, and if it is 91 dB or greater, it is considered profound HL. The puretone averages were 15 dB in both ears for the NH group, 100 dB in the right ear and 99 dB in the left ear for the CI group, and 67 dB in the right ear and 64 dB in the left ear for the HA group.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of participants</th>
<th>Age (years)</th>
<th>PoL (%)</th>
<th>DoD (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>20 (m = 2, f = 18)</td>
<td>49.70 ± 14.27</td>
<td>40.51± 17.72</td>
<td>5.45± 11.72</td>
</tr>
<tr>
<td>HA</td>
<td>20 (m = 9, f = 11)</td>
<td>51.97 ± 16.35</td>
<td>19.78± NA</td>
<td>NA</td>
</tr>
<tr>
<td>NH</td>
<td>20 (m = 3, f = 17)</td>
<td>51.08 ± 15.49</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note: CI, Cochlear implant; HA, Hearing aid; NH, Normal hearing; PoL, Percentage of the patient’s life with hearing loss before cochlear implant; DoD, Duration of deafness.

#### 3.2 Speech recognition performance

Statistical analysis data regarding speech recognition performance in various listening conditions for both groups are shown in Table 2. At sentence level, for the +5 dB SNR condition, CI users showed significantly lower performance than HA users and NH listeners in BN (24.32 ± 20.78), PS (35.50 ± 31.37), PV (26.63 ± 29.40), and PC (39.88 ± 32.90) conditions (p < 0.05). The HA users’ sentence recognition performance was significantly lower than the NH listeners, but their performance was significantly better when compared to that of the CI users in all test conditions (p < 0.05). The mean performance of the HA users was 54.00 ± 26.64, 76.50 ± 25.81, 60.50 ± 32.20, 64.50 ± 30.00 in BN, PS, PV, and PC conditions respectively. At +10 dB SNR, the CI group showed significantly lower performance than the HA groups except for PC (56.50 ± 25.80). They also showed significantly lower performance (40.00 ± 26.66 for BN, 65.75 ± 26.22 for PS, and 48.00 ± 32.54 for PV) than the NH group in all conditions (p < 0.05). Compared to the NH group, the HA group showed significantly lower performance in BN and PC conditions (63.50 ± 27.77 for BN and 73.50 ± 27.96 for PC). At word level, in the +5 dB SNR condition, CI users showed significantly lower performance than HA users and NH listeners in BN (44.21 ± 25.56), PS (35.50 ± 31.37), PV (39.88 ± 32.90), and PC (51.50 ± 32.13) conditions (p < 0.05). Comparing the HA and CI groups, HA users understood words significantly better than the CI users in BN (75.38 ± 30.15), PS (88.60 ± 17.52), PV (76.50 ± 26.98), and PC
(80.13 ± 29.22) conditions (p < 0.05). Comparing the HA and NH groups, the HA users showed significantly lower performance than the NH listeners in the BN and PV conditions. At +10 dB SNR, statistical differences were observed between the CI (61.18 ± 28.67 for BN, 77.12 ± 28.10 for PS, 65.25 ± 29.56 for PV, and 64.12 ± 26.13 for PC) and HA groups (82.25 ± 21.91 for BN, 93.37 ± 15.40 for PS, 90.13 ± 18.54 for PV, and 82.13 ± 25.58 for PC) and between the CI and NH groups (97.62 ± 2.86 for BN, 98.50 ± 3.07 for PS, 99.00 ± 1.70 for PV, and 99.50 ± 1.02 for PC) in all test conditions (p < 0.05). Statistical difference between the HA and NH groups was observed only in the PC condition (p < 0.05).

### Table 2. Speech recognition performance for NH, HA, and CI groups.

<table>
<thead>
<tr>
<th>Condition</th>
<th>+5 dB SNR</th>
<th>+10 dB SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CI</td>
<td>HA</td>
</tr>
<tr>
<td>PS Sentence</td>
<td>35.50 (± 31.37)</td>
<td>76.50 (± 25.81)</td>
</tr>
<tr>
<td></td>
<td>57.88 (± 28.48)</td>
<td>88.60 (± 17.52)</td>
</tr>
<tr>
<td>PV Sentence</td>
<td>26.63 (± 29.40)</td>
<td>60.50 (± 32.20)</td>
</tr>
<tr>
<td></td>
<td>39.88 (± 32.90)</td>
<td>76.50 (± 26.98)</td>
</tr>
<tr>
<td>PC Sentence</td>
<td>33.50 (± 27.58)</td>
<td>64.50 (± 30.00)</td>
</tr>
<tr>
<td></td>
<td>51.50 (± 32.13)</td>
<td>80.13 (± 29.22)</td>
</tr>
<tr>
<td>BN Sentence</td>
<td>24.32 (± 20.78)</td>
<td>54.00 (± 26.64)</td>
</tr>
<tr>
<td></td>
<td>44.21 (± 25.56)</td>
<td>75.38 (± 30.15)</td>
</tr>
</tbody>
</table>

Note: PS, piano solo; PV, piano and violin; PC, piano and chorus; BN, babble noise.

a, CI & HA: significant difference between speech recognition (p < 0.05); b, CI and NH group: significant difference between speech recognition (p < 0.05); c, HA and NH group: significant difference between speech recognition (p < 0.05).

### 3.3 Speech recognition performance according to the period of HL

For CI users, the relationships between PoL and speech recognition performance, between DoD and speech recognition performance, as well as duration of CI use and speech recognition performance were examined for all testing conditions at +5- and +10 dB SNR. Correlational analysis results are shown in Table 3. Correlational analysis revealed that for both words and sentences, no significant correlations were observed between PoL and speech recognition performance for BN, PS, PV, and PC conditions at +5- and +10 dB SNR. However, at +5 dB SNR, there was a statistically significant strong negative correlation between DoD and speech recognition performance in the BN (Rho = −0.5665, p < 0.01) conditions for sentences. Statistically significant strong negative correlation was also found for words in the BN (Rho = −0.4754, p < 0.05) and PC (Rho = −0.5391, p < 0.01) conditions. At +10 dB SNR, statistically significant strong negative correlations at the sentence level were observed between DoD and speech recognition performance in both PS (Rho = −0.4875, p < 0.05) and PV (Rho = −0.5673, p < 0.01) conditions. At the word level, statistically significant strong negative correlations were also observed between DoD and speech recognition performance in both PS (Rho = −0.5056, p < 0.05) and PV (Rho = −0.5127, p < 0.01) conditions. Regarding the correlation between the duration of CI use and sentence and word recognition performance, no significant results were observed irrespective of SNR and testing conditions.
4. Discussion and conclusion

The current study explores how sentence and word recognition performance is impacted by several types of sounds including BN and music in people with NH, HA users, and CI users. The findings of the study showed that in terms of speech recognition performance, as expected, CI and HA users showed lower word and sentence recognition performance than those with NH regardless of the SNRs except in the PC condition at +10 dB SNR. Correlational analysis also revealed that longer DoD correlates with a decline in speech recognition performance in various listening conditions which was consistent with findings in previous studies to some extent (Bernhard et al., 2021; Moon et al., 2012). Bernhard et al. 2021 conducted a systematic review of the effect of DoD on hearing performance. In this systematic review, the DoD ranged from 0.1 to 77 years and the age at implantation ranged from three months to 14 years. The results showed negative correlations between DoD and monosyllabic word and sentence perception performance. In addition, the authors performed further analysis with subgroups: time of post-implantation testing and mean DoD. CI users who came to post-implantation follow-up sessions less than 12 months had a strong negative correlation between monosyllabic and sentence perception. When analyzing mean DoD data with a cut-off of 12 years, CI users with a DoD of 12 years exhibited poorer outcomes in monosyllabic word and sentence tests, meaning that the longer the DoD, the fewer opportunities for auditory input, leading to a decline in speech understanding ability (Bernhard et al., 2021). Comparing CI and HA users, also similar to pre-existing studies, the HA group showed better performance in all conditions for words and sentences except in one condition (Agelfors, 1996; Flynn et al., 1996; Kaandorp et al., 2015; Meilijson & Spitzer, 2015). In 1995, Agelfors examined speech perception performance in individuals with profound HL who were using HAs and CIs. The speech testing included a consonant test (/aCa/ syllables), a suprasegmental test, and a connected discourse tracking test. For the suprasegmental testing, prosodic contrasts were applied (number of syllables, vowel length, juncture, tone, and word emphasis) and the participants were required to choose between two options. The connected discourse testing involved a speaker reciting passages from a book (sentence by sentence) and the participants were asked to repeat the verbatim. The results showed that HA users showed better performance than CI users for suprasegmental and connected discourse tests. Kaandorp et al. 2015 performed speech testing in quiet and noise on 12 young adults with NH, 24 CI users, and 24 HA users. The speech testing involved consonant-vowel-consonant (CVC) monosyllables and sentences. The results revealed that there was a significant variation in the ability to recognize both CVCs and sentences and generally, CI users exhibited lower sentence recognition than CVCs, unlike HA users who did not show such a difference in performance. NH listeners performed significantly better than HA and CI users (Kaandorp et al., 2015). In our study, for sentence recognition, the HA group showed significantly better performance than the CI users in all conditions except for the PC condition at +10 dB SNR. For word recognition, statistical difference between the HA and NH groups was observed for all conditions at +5 dB SNR, but only in the PC condition at +10 dB SNR. Statistical analysis also revealed that CI users showed poorer performance than HA users and those with NH. Considering that only individuals who showed good performance on clinical tests (CAP score of 6 and 85% or higher on an open-set speech test) were able to participate in the study, the findings of the study illustrate that even background sounds may still hurt speech recognition ability in people with HL who use hearing devices and demonstrate good performance on clinical testing. The reduced sentence and word recognition performance observed in the CI group under the two SNRs, as presented in Table 2, illustrates that background sounds have a more negative impact on the CI group compared to the HA and NH groups. In contrast, the HA group, in the +5 dB SNR condition, demonstrated significant
performance differences in sentence recognition scores compared to the NH group, while no performance differences were observed at the word level in some conditions, such as PS and PC. Even in the +10 dB SNR condition, significant differences with the NH group were only observed in some conditions, such as PC and babble noise. This infers that even if HA users do not hear all words in a sentence, they could exhibit performance comparable to individuals with NH if they hear keywords within a sentence. In addition, for all testing conditions, the HA group outperformed the CI group, even in the +10 dB SNR condition for PS and PV, where sentence recognition performance did not significantly differ from the NH group. However, in the PC condition, where speech sounds were embedded in the background music, the HA group’s sentence and word recognition performance was significantly lower compared to the NH group. For HA users, the relative impact of music on speech perception may not be significantly negative, but for CI users, background music may act as a noise interfering with speech perception. In other words, for HA users, if they are communicating with others in an environment with only instrumental music (no chorus) at +10 dB SNR, they may not experience significant difficulty with speech recognition. Yet, for CI users to engage in communication, they may need to be in an environment that is free from background music or noise. These findings highlight the importance of creating an environment without background noise, especially in places where multitalker conversations are present, to achieve successful communication with those with HL. This is part of communication strategies that can be used for individuals regardless of the presence of HL to achieve effective communication (Helvik et al., 2007; Stephens et al., 1999; Wilson et al., 1998). According to Tye-Murray et al. 1992, communication strategies include repair strategies, corrective strategies, and anticipatory strategies (Tye-Murray, 1992). Repair strategies are tactics that people can use when they misunderstand words. For example, they could ask the speaker to repeat, simplify, clarify, or rephrase what he or she said. They could also ask the speaker to provide key or specific information. Corrective strategies can be used to modify the speaker’s behavior or the communication settings. The listener can ask the speaker to adjust his or her pace and volume of speech. People can also modify the environment, such as having good lighting, reducing any background noise, and talking in the same room. In good light, people can benefit from visual information, such as lip movements, facial expressions, and gestures. Reducing background noise and talking in the same room allow individuals to hear clean signals for communication rather than distorted signals. Lastly, anticipatory strategies involve reviewing potential vocabulary so that individuals can be ready for upcoming conversations. People can also benefit from taking turns in the conversations, paying attention to each other, grabbing the conversational partner’s attention before speaking and so on. There are many communication strategies and they have been known to be beneficial for communication, especially for those with HL. In 1999, Stephens and colleagues investigated the most commonly used communication strategy that people with HL utilized among the five communication strategies through questionnaires (avoidance of conversation, interruption, pretending to have heard or understood, asking for repetition, and positioning self in order to hear) (Stephens et al., 1999). The results showed that people with HL used avoidance and request for repetition strategies most frequently, followed by pretending to have heard or understood, positioning self in order to hear, and interruption, in that order. The importance of communication strategy usage has become even more pronounced during the Covid-19 pandemic (Eby et al., 2020; Mansutti et al., 2023; Moon et al., 2022; Seol et al., 2023). During the pandemic, most countries enforced social distancing and mask measures to prevent infection. In this process, individuals with HL experienced greater challenges with communication as visual cues, such as facial expressions and lip movements, were lost and face masks filtered high frequency sounds that are important for speech understanding. Mansutti et al. 2023 examined communication issues that people...
with HL were experiencing during the pandemic and strategies that people could use to improve communication. Face coverings (i.e., face mask, and face shield etc.), physical and social distancing, and accessibility to information led to communication challenges. The authors mentioned that various strategies could improve communication challenges. For example, for those with HL, utilizing hearing devices (HAs and CIs) could be helpful. Along with ensuring face-to-face communication and avoiding shouting, individuals could also benefit from the use of, such as portable amplifiers (Mansutti et al., 2023). In sum, HL can lead to communication breakdown and delay in speech and language development, so it is vital to engage in the rehabilitation process through HAs and CIs. However, wearing HAs and CIs does not always lead to improved speech understanding. HA and CI users still struggle with understanding speech in noisy environments, and particularly CI users also report difficulties in enjoying and listening to music. This study is meaningful as it investigates speech recognition performance using noises encountered in the real-world, such as music, chorus, and babble. Additionally, to reflect real-world communication settings, sentence and word recognition performance was measured in environments with various listening conditions (+5- and +10 dB SNRs). In most cases, CI users showed lower word and sentence performance compared to HA users and those with NH. Despite showing high performance in speech tests commonly used in clinical settings, with scores of 85%, CI users still struggled to understand sentences and words in the presence of noise and music. Taken together, findings of the study implicate that the longer the duration of HL, the more speech recognition ability tends to decline. Since HL can lead to speech and language development delays in children, it is important to intervene at an appropriate time, but considering the fact that intervention does not always lead to improvement in speech recognition performance and music perception, it is necessary to assess speech recognition ability in various aspects (i.e., background noise and music) and encourage individuals to participate in rehabilitation.

In terms of limitations and future directions, subsequent studies utilizing a higher sample size and incorporating various participant and noise and music characteristics would be beneficial in assessing communication abilities across diverse environments and characteristics. Within music, considering a range of genres and types could be explored and investigating individuals’ performance not only at the level of words and sentences but also at the level of phonemes may reveal differences. Currently, there are no tools available to evaluate speech performance using music as background noise. Findings from studies using various types of music could provide data that could aid in the development and standardization of tools for comprehensively investigating the speech and language abilities of individuals with HL in the future. Besides, the use of objective assessments (i.e., electroencephalography) along with subjective measures would provide information about biomarkers that can contribute to individuals’ speech recognition performance.

Author Contributions

E.Y.K. and H.Y.S. conceptualized the study, E.Y.K. and S.E.L. conceived the experiments, E.Y.K., and H.Y.S. reviewed the concept. E.Y.K. and S.E.L. conducted the experiments, and E.Y.K. analyzed the results. E.Y.K. wrote the main paper and all authors reviewed the manuscript.

Conflict of Interest

Declaration of conflict of interest.

Data Availability Statement

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

Funding

This research received no external funding.
References


Bruns, L., Mürbe, D., Hahne, A., 2016. Understanding music with cochlear implants. Scientific Reports. 6(1), 32026.


DOI: https://doi.org/10.1044/2016_JSLHR-L-16-0297


DOI: https://doi.org/10.1080/15017410600687073


DOI: https://doi.org/10.3109/00016480903380521


DOI: https://doi.org/10.3109/14992027.2014.945623


DOI: https://doi.org/10.1097/MAO.0000000000002590


DOI: https://doi.org/10.1080/14992020802310887


DOI: https://doi.org/10.1097/MAO.0b013e31819bd7e6


DOI: https://doi.org/10.1177/03057356050650


DOI: https://doi.org/10.1001/jama.2020.17647


DOI: https://doi.org/10.1055/s-0032-1329222


DOI: https://doi.org/10.1111/jocn.16572


DOI: https://doi.org/10.1080/14992027.2021.1883197


DOI: https://doi.org/10.1525/mp.2004.21.3.289


DOI: https://doi.org/10.1001/archotol.130.5.556


DOI: https://doi.org/10.3109/030053640000097


DOI: https://doi.org/10.1146/annurev-linguistics-031220-121126


DOI: https://doi.org/10.1207/s15327078inf0701_5


DOI: https://doi.org/10.1080/21642850.2014.900450


DOI: https://doi.org/10.3390/ijerph15010098


DOI: https://doi.org/10.1080/02699200701659243


DOI: https://doi.org/10.1016/S0140-6736(17)31073-5


DOI: https://doi.org/10.1179/136132898805577250