SENTENCE IDENTIFICATION IN THE PRESENCE OF SPEECH AND MUSIC MASKERS AMONG MUSICIANS AND NON-MUSICIANS.

SRIVIDYA S

Register Number: 18AUD036

This Dissertation is submitted as part fulfilment

For the Degree of Master of Science in Audiology

University of Mysore, Mysuru

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MANASAGANGOTHRI, MYSURU – 570 006

JULY 2020

CERTIFICATE

This is to certify that this dissertation entitled **'Sentence Identification in the presence of Speech and Music Maskers among Musicians and Non-musicians'** is the bonafide work submitted in part fulfilment for the degree of Master of Science (Audiology) of the student Registration Number: 18AUD036. This has been carried out under the guidance of the faculty of the institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru
July 2020

Prof. M. Pushpavathi
Director
All Indian Institute of speech and Hearing
Manasagangothri, Mysuru-570006

CERTIFICATE

This is to certify that this masters dissertation entitled **'Sentence Identification in the presence of Speech and Music Maskers among Musicians and Non-musicians'** has been prepared under my supervision and guidance. It is also being certified that this dissertation has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru
July 2020

Dr. Ajith Kumar U
Guide
Professor in Audiology
Department of Audiology
All India Institute of Speech and Hearing
Manasagangothri, Mysuru-570006

DECLARATION

This is to certify that this dissertation entitled **'Sentence Identification in the presence of Speech and Music Maskers among Musicians and Non-musicians'** is the result of my own study under the guidance of Dr. Ajith Kumar U, Professor in Audiology, Department of Audiology, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru
July 2020

Srividya S
Register No. 18AUD036

Acknowledgment

Firstly, I thank my ***APPA, AMMA, and all my FAMILY*** *members for their love, support, and showering all their blessings on me.* ***Lakshmi, Balu, Sridi, Anantha, and Raksha -*** *I'm blessed to have a wonderful cousin like you guy's***.**

I'm thankful and feel lucky that I got a chance to work under your guidance, ***AJITH sir****. Your knowledge and calmness in handling situations inspire me sir. You are the best guide sir. Big thanks to* ***Jim sir****, for being their every time I needed, helping in all recordings and clearing all my silly doubts.*

I'm privileged to have got a chance to be your student ***Animesh sir,*** *your lectures made me pursue audiology.* ***Sandeep sir and Nike sir****, I will always be thankful for sowing the interest regarding research in me.*

I thank HOD's ***Sujeeth sir and Prawin sir****, and all the* ***academic and clinical staff*** *of audiology for their constant support and guidance throughout my AIISH journey.*

Thanks for all the singers- ***Aishu, Sahana, Thanushree, and Prabha aunty*** *for being voice of all the songs. I thank all the participants of the study for their time and cooperation.*

This journey would have been incomplete without ***Hannah, Tanuja, Sushma, and Hema****, the best bunch of people. Thanks for making this journey beautiful girls.*

Merina, Rajwinder, and Sushma*, you guys made the clinical posting more fun. I couldn't have got better posting partners than you.*

I thank all my classmates, ***PGions*** *for being throughout this M.Sc journey. I also congratulate all my M.Sc batchmates,* ***BRAINIACS*** *– Corona batch 2020, on successfully completing their dissertation in these tough times. I thank all the seniors and juniors for their timely help during this AIISH journey.*

Abstract

Aim: The main aim of the present study is to compare the perceptual abilities of Musicians and Non-musicians in the presence of different Speech and Music maskers of varying levels of difficulty.

Objective: To measure and compare the SNR required to obtain 50% correct scores (SNR-50) for two types of speech maskers [Lyrical Babble(LB) and Lyric Spectrum Noise(LSN)] and Music maskers [Music Noise(MN) and Music Spectrum Noise(MSN)] in 2 and 4 talker/singer conditions between Musicians(M) and Non-musicians (NM).

Design: Twenty normal hearing individuals aged between 18-35 yrs were divided into two groups as Musicians (M, n=10) and Non-musicians (NM, n=10) based on the musical training they have received (>5years) and their performance in Mini-Profile of Musical Perception Skills (Mini-PROMS) (score of >18 in Musicians group). The normal hearing acuity of participants included in both the group was evaluated by pure tone audiometry, tympanometry, and otoacoustic emissions. To evaluate and compare the effect of different speech and music maskers in 2 and 4 talker/singer conditions, SNR-50 was measured in all eight masker conditions (LB2, LB4, MN2, MN4, LSN2, LSN4, MSN2, and MSN4) on all participants in Musician and Non-musician group.

Results: Results revealed that SNR-50 differs significantly in MN2 and MN4 conditions between Musician and Non-musician group, whereas SNR-50 did not differ significantly between two groups in LB2, LB4, LSN2, LSN4, MSN2, and MSN4 masker conditions. However, reduced SNR-50 scores were noted in the Musician group in the above conditions compared to the Non-musician group.

Conclusion: The presence of music as competing stimuli during the speech perception tasks is difficult in Musicians compared to Non-musicians due to their expertise in music. The music adds more informational content to the masker in the case of Musicians, whereas it serves as just an energetic masker in the case of Non-musicians. Informational masking in Musicians is more due to lyrical and melody content of the signal than its temporal characteristics

Keywords: SNR-50, Speech Perception in Noise, Informational Masking, Music noise, Speech Babble, Musicians, Non-musicians.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Speech and music are everyday acoustic events that often turn out to be a competing background during verbal communication. In these realistic situations, listeners have to focus on speech as the signal of interest while ignoring the competing background and its information content (Shi & Law, 2010). Speech and Music are temporally varying broad band signals falling within similar frequency range and governed by a defined set of rules (Besson & Schön, 2001; Zatorre et al, 2002; Levitin & Menon, 2003). Music is a complex acoustic stimulus that has temporo-spectral characteristics similar to that of speech. The perception of music is a complex process involving appreciation of the timing, sequencing, and anticipation. The perception of the music is mainly driven by the interaction between the pitch and timbre (Krumhansl & Iverson, 1992) and the process of temporal patterning (Rajendran et al., 2018) of the stimulus. The processing of such a stimulus involves many auditory centres from sub cortical level to cortex. There are studies that report the interhemispheric connections involved in perceptual analysis of the music (Skoe, 2017). These similarities in the spectro-temporal characteristics project music as potential masker at par with speech.

It has been proposed that, for simultaneously presented acoustic stimuli, the auditory cortex must strive to successfully separate them into perceptual streams before each can be further processed and this is known as auditory stream segregation (Bregman, 1995). Individuals trained in music have better auditory segregation ability and several other auditory advantages too when compared to non-musicians (Skoe, 2017). Musical training induces both structural and functional changes in the auditory centres. The plasticity of the auditory system changes over time due to musical training. The years of practice also influences the amount of plasticity of the auditory system. Musicians have shown better performances in many auditory skills compared to non-musicians. There are several studies done assessing some of peripheral to cortical auditory processes to prove the same (Bidelman, et al. 2017; Rammsayer & Altenmller, 2006; Zhang, et al. 2015). These auditory processes are assessed by means of many behavioral, physiological and electrophysiological tests. Some of them involve Otoacoustic emissions (OAE), Speech in Noise test (SPIN), Gap detection tests (GDT), Auditory Brainstem Response (ABR), late latency response (LLR), P300 etc.

Results of speech perception in noise studies in musicians are equivocal. Many studies have reported better perception of speech in the presence of noise in musicians (Parbery-Clark, et al. 2009; Swaminathan et al. 2015; Başkent & Gaudrain, 2016) . Parbery-Clark, et al. (2009) reported better performance of the musicians in QuickSIN (speech perception in presence of multi talker babble) and Hearing in noise test(HINT) over non-musicians. The speech perception also depends on the spatial separation between the speech and the masker (Swaminathan et al. 2015). Musician advantage was more when the target and masker were separated spatially than in a collocated situation. These enhanced perceptual skills may be due to improved auditory plasticity (Zendel & Alain, 2012) and auditory cognitive abilities involving attention and working memory(Strait, et al. 2010). The auditory structures for processing music and speech overlap partially and can be one of the reasons for enhanced speech perception in presence of noise in musicians (Besson, et al. 2011). However whether the top-down and bottom – up processes are involved in this is still unclear (Coffey, et al. 2017). Oxenham, et al. (2003) reported the enhanced performance of musicians in the test involving tonal stimuli as both target and competing stimuli. This reflects the superior analytic listening abilities in them which leads to reduced susceptibility to informational masking. Whereas a minimal effect of training was found when the stimuli were having properties of energetic masking.

However, a few studies have reported that the performance of young adults who were trained and not trained in music are comparable in speech in noise test with voiced and whispered speech as target stimuli (Boebinger et al. 2015; Ruggles et al. 2014). Musician's advantage in speech perception in noise is more evident when the difficulty level (complex masker) of the perception task was more (Swaminathan et al. 2015).

The background music encountered in daily life can also act as competing stimulus and affect the effective communication. Russo et al. (2008) reported that the young adults were able to understand speech better in presence of the background music when compared to elderly in word recognition tasks. The differences seen were accounted due to the better focussed attention abilities in young adults. Different musical maskers affect the perception with different difficulty level. Evaluating the performance of speech in presence of music revealed equivalent response as with speech as competing stimuli. However spectrum shaped noise had minimal effect compared to the prior mentioned maskers (Başkent, et al. 2014). While Music was adopted as competing stimuli the expectancy and the hierarchical structure of it affected the perception of the target stimulus. Altering the temporal dynamics of the music masker's affected speech perception (Shi & Law, 2010). However, speech used as masker has more effect on perception of target compared to the music. As the spectral complexity was increased the normal listeners experienced a release from masking and hence scored better with increasing difficulty (Eskander et al. 2011). Perception of speech in presence of competing music stimuli may be modulated by knowledge of music and musicianship.

1.1 Need of the study

Typically, musicians are reported to outperform non-musicians in speech perception in noise tasks. However, effect of competing music on speech perception is relatively unexplored. Since the musicians are trained in the music and has better perception of pitch and rhythm, the effect of competing music on speech perception may be different than non-musicians.

We hypothesize that the presence of music as competing stimuli during the speech perception tasks must be difficult in musicians compared to non musicians due to their expertise in music. The music adds more informational content to the masker in case of musicians, whereas it serves as just an energetic masker in case of non musicians. Therefore, this study is focussed on comparing speech perception in presence of speech babble and music noise between musician and non musicians.

1.2 Aim of the study

To compare the perceptual abilities of Musicians and Non-musicians in the presence of different Speech and Music maskers of varying levels of difficulty.

1.3 Objective of the study

- To measure and compare the SNR required to obtain 50% correct scores (SNR-50) for two types of speech maskers (Lyrical Babble and Lyric Spectrum Noise) in Musicians and non-musicians.
- To measure and compare the SNR required to obtain 50% correct scores (SNR-50) for two types of music maskers (Music Noise and Music Spectrum Noise) in Musicians and non-musicians.

Chapter 2

Review of literature

In everyday communication, there are many background noises, which affect the perception of speech. Background noise such as traffic noise, environmental noises, speech babble, etc. mask the speech signal and make it difficult for an individual to understand the same. Even a soothing sound such as Music in background scores in TV, restaurants, etc. can have an adverse effect on speech understanding at times (Başkent et al., 2014). These competing signals demand extra auditory attention towards the signal of interest, i.e., speech. The impact of different maskers on speech perception varies depending upon the type of noise and its characteristics. The spectrum of noise is one such characteristic that affects perception (Rogers et al., 2006). More similar the spectrum of noise with the signal, more will be the difficulty faced in understanding speech. Even a simple noise such as white noise, broadband noise, or spectrally shaped steady-state noise, they interact with the speech signal's physical properties and affect the comprehension of speech.

Speech babble being very same as the speech of interest concerning the spectrum poses as a potential masker (Başkent & Gaudrain, 2016; Rhebergen et al., 2008). On the other hand, Music, which shares almost a similar spectrum as speech babble, also has an identical adverse effect on speech perception (Besson et al., 2011).

The ability to perceive speech in the presence of noise can vary as additional training in different areas. Many studies done over the years have proved that different forms of training protocols such as computer-based training, musical training (Kraus et al., 2014; Parbery-Clark et al., 2009), abacus training, etc. improve the perception of speech in noise.

2.1 Effect of Musical Training on Speech Perception in Noise

Individuals learning Music has to pay attention to the pitch and other subtle aspects of audition, to appreciate the Music. Studies done over the years have reported increased audition capabilities in such individuals (Skoe, 2017; Zendel & Alain, 2009). Studies have shown that individuals trained in Music have better pitch perception (Marques et al., 2007), temporal resolution (Kumar et al., 2016; Rammsayer & Altenmller, 2006), cognitive abilities such as working memory, attention and reasoning skills (Strait et al., 2012), increased perception of speech in the presence of noise (Parbery-Clark et al., 2009), etc. A study done by Mok & Zuo (2012) revealed better prosody and lexical tone perception of tonal languages such as Cantonese and Mandarin in non-native musicians compared to their counterparts. Moreover, this training effect is not limited to any age group. In their study, Rochette et al. (2014) presented the Better stream segregation abilities in musically trained deaf children. Even Musically trained adults outperformed non-musicians in repeating spectrally degraded signals (Swaminathan & Gopinath, 2013). Furthermore, the sensitivity to rhythm helps a listener understand unfolding speech patterns in degraded listening conditions, and hence the superior rhythm skills of musicians lead to better speech-in-noise perception (Slater & Kraus, 2016). In daily life situations, for the segregation of noise from speech, integration of cognitive, sensory, and linguistic processes is required. Moreover, the above reported studies have provided ample evidence to prove the upheld of the musicians in the same. Therefore, In general, the musicians do not specifically have a better hearing threshold. However, they are excellent listeners due to their increased auditory perceptual abilities, which will help them in a better speech in noise perception. These changes can be mainly contributed to the neuroplastic changes in the auditory system (Schellenberg, 2015).

Two principles mainly explain the Neuroplastic changes in the auditory system. Firstly, Patel (2011) stated the OPERA hypothesis, which explains that improved speech perception in the presence of noise in musicians is a result of enhanced and strengthened neural circuitry involved in speech perception of these individuals. He hypothesized that the speech processing networks undergo adaptive plasticity when the signal meets five conditions, namely, Overlap, Precision, Emotion, Repetition, and Attention. It accounts for the enhanced subcortical coding of speech in the brainstem in musicians.

The second principle, proposed by Strait et al,. (2013) was the Neural resilencing of background noise principle. This explains the speech perception in the presence of noise based on how the auditory brainstem responds to noise stimuli. Their study compared the Frequency Following Response (FFR) between preschool children enrolled for musical training for 2 and 1 year, respectively, using speech stimuli in both quiet and noise conditions. They reported that there was a faster response rate and reduced quite to noise timing difference in children trained for 2year, emphasizing the resilence of neural activity in response to noise. Therefore, they concluded that noise as background stimuli impact auditory brainstem response faster in musicians than non-musicians.

The changes in auditory structures also depend on the time course of the training. A study done by Slater et al., (2015) revealed that the individuals trained for 2 years had enhanced neural processing than those trained for 1year. The aging auditory system also benefits from the musical training due to the strengthening of the underlying neural pathways responsible for the accurate representation of important temporal and spectral features of sound (Parbery-Clark et al., 2012).

2.2 Effect of different Speech maskers on Speech Perception in Musicians

In an everyday situation, the background noises in an individual's environment consist of speech as a background in one or the other from ranging from competing speech from a third talker to cafeteria noises. The amount of effect caused by different background noises depends on its many parameters, such as loudness and spectral characteristics. Davies-Venn et al., (2015), in their study, noticed that the Word Recognition Scores were more mediocre with the noise, which resembled more like speech. Therefore more the spectral similarities between the masker and signal, more difficult are the speech perception task for the individual.

The studies done on evaluating the effect of speech masker on speech perception has given equivocal results (Başkent & Gaudrain, 2016; Boebinger et al., 2015; Parbery-Clark et al., 2009; Ruggles et al., 2014). Parbery-Clark et al., (2009) investigated the musician advantage on speech perception in noise by performing Quick SIN and Hearing in Noise Test (HINT) in 16 Musicians and 15 Non-musicians. They also assessed frequency discrimination ability and working memory. Results revealed that the musicians outperformed non-musicians in both the task, leading to the conclusion that musical training enhanced the ability to perceive speech in noise. They also noted a positive trend with QuickSIN scores, frequency discrimination, and working memory tasks.

On the other hand, the study was done by Boebinger et al., (2015) showed that despite good frequency discrimination ability in 25 musicians who participated in the study, they showed no advantage in perceiving masked speech when compared to the 25 non-musicians. They suggested that the contribution of general cognitive abilities needs to be considered in any investigations involving perceiving speech in noise tests.

2.3 Music as a Background Noise

Music is a pleasurable art form having a time-varying signal which is organized in a parallel hierarchy. Nowadays, the scope of Music as background noise has increased vastly, from serving as a simple background score in a TV program to provide good ambiance in the restaurant. The soothing Music has proven to boost the mood and improve the efficiency in varied everyday tasks (Blood & Zatorre, 2001). Its presence facilitates the performance in different cognitive tasks, including signal detection, visual vigilance, and learning various tasks such as karate (Dalton & Behm, 2007). Its presence increases the motivation, arousal, and also the perceptual ability of the individual.

Even though Music consists of sound with harmony, rhythm and is widely different from annoying and unpleasant noise, it can still pose as a distracter. Furnham & Strbac (2002) reported that the presence of Music affects individuals' performance in many tasks such as attention, vigilance, reading, comprehension, and commonly performed tasks such as driving. The increased stress level and mild aggression induced due to Music deteriorates the performance of the individual.

Similarly, Music as noise can have an adverse effect on speech perception tasks. It shares a very similar spectrum with the speech signal, which makes it a potential masker. However, Music has a broader and dynamic frequency spectrum, including fundamentals and harmonics, and also has greater fluctuation in amplitude and timbre (Gfeller et al., 2012). Unlike a regular broadband noise, Music causes both energetic masking due to its spectral content and informational masking due to its temporal and semantic content. The prior occurs at the peripheral level, whereas the latter at the central level. The instrumental Music had minimal effect compared to Music with lyrics, highlighting the effect of informational masking (Başkent et al., 2014).

Eskridge Elizabeth N. et al., (2012) evaluated the Speech Recognition Threshold (SRT) in 9 normal hearing individuals with four different music maskers, the steady-state noise filtered with a spectral envelope of speech (SSN), steady-state noise filtered with a spectral envelope of Music (MSN), modulated MSN with a temporal envelope of Music (MMSN), and Music with lyrics (MUS). The first two served as a steady-state masker causing more energetic masking, whereas the latter two served as dynamic maskers causing both energetic and informational masking. Results revealed SRTs were much lower with the music-related maskers than with SSN, and release from masking was seen in listeners due to envelope and fine structure cues in the MMSN and MUS maskers than with MSN. They concluded that the temporal envelope and fine structure cues in Music help the listeners to separate it from speech.

Speech Recognition Threshold (SRT) of 49 normal hearing individuals was poor in the presence of orchestra music compare to vocal and piano music as a masker (Gfeller et al., 2012). Therefore, the individual's performance also depends on the hierarchical structure of different types of Music acting as masker due to its varying temporal dynamics (Shi & Law, 2010). The amount of informational making depends on interest and familiarity of the signal (Russo & Pichora-Fuller, 2008). More familiar the signal to the individual in its form, lyrics, or tone, the attention shifts towards the masker and increases future expectancies about the Music.

However, the studies on the effect of Music as masker on musicians are very sparse. According to Oxenham et al., (2003) the amount of information masking is less in musicians than non-musicians when speech was the masker. However, the prior knowledge about the

Music and its aspects such as pitch and rhythm can make it a potential masker for musicians than for non-musicians. Gfeller et al. (2012) even reported that individual ability to separate speech and masker depends on their ability to perceive pitch.

Therefore, this study tries to answer whether music masker increases the amount of information masking in musicians. The study compares across different speech and music masker conditions between musician and non-musician group.

Chapter 3

Methodology

3.1 Participants

A total of 20 normal hearing individuals aged between 18-35 yrs were involved in a study. The participants were divided into two groups as Musicians (M, n=10) and Non musicians (NM, n=10) based on the musical training they have received and their performance in Mini-Profile of Musical Perception Skills (Mini-PROMS) (Zentner & Strauss, 2017). Individuals who score more than 18 in Mini-PROMS were assigned to the musician group. Individuals in the musician group had a minimum of 5 years of experience in Carnatic music (vocal or any instrument) and were practicing music for at least 2-3 hours per day. The individuals having any history of noise exposure, use of ototoxic drugs, or middle ear infections were excluded from the study. Along with this, participants suffering from any neurological or cognitive dysfunction were also excluded from the study. Other inclusion criteria were -

I. No gross otological or neurological history as ascertained through a detailed and structured interview. The individual's musical experience related information was also noted down.

II. Normal hearing sensitivity with air-conduction and bone-conduction pure-tone thresholds being 15 dB HL or lesser at octave frequencies between 250 Hz and 8000 Hz. Threshold estimation was done using the modified Hughson-Westlake procedure (Hughson & Westlake, 1943).

III. ‘A’ type tympanogram with ipsilateral and contralateral acoustic reflex thresholds within 100 dB HL for the frequencies of 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz to ensure normal middle ear functioning

IV. Speech identification scores of 90% or better at 40 dB SL. Phonetically balanced (PB) word list in the native language of the participant was used for the same (Yathiraj & Vandana, 2005).

V. Signal to noise ratio (SNR) of more than 6 dB SPL in transient evoked oto-acoustic emission (TEOAE) to rule out outer hair cell dysfunction

3.2 Test Environment

All the tests, as well as the experiment, were carried out in an air conditioned and sound treated rooms with ambient noise levels within permissible limits (ANSI S3.1-1999, R2013).

3.3 Instrumentation

- A calibrated two channel diagnostic audiometer, GSI-61 (Grason-Stadler Incorporation, USA) with Telephonics TDH 39 supra aural headphones and Radio ear B-71 bone vibrator calibrated as per ANSI (2004) was used for threshold estimation.

- A calibrated GSI-tympstar (Grason-Stadler Incorporation, USA) clinical immittance meter, calibrated as per ANSI 1987, was used for tympanometry and reflexometry.

- ILO 292 DP Echo port system (Otodynamics Inc., UK) was used to assess transient evoked oto-acoustic emissions.

- Adobe Audition 3.0 (Adobe Systems Inc) installed on a Dell Inspiron laptop (Realtek sound card) with MOTU MICROBOOK II external sound card interface was used for recording competing stimuli.

3.4 Stimuli Preparation

3.4.1 Preparation of different maskers

Four different types of maskers were used for the study – Lyrical Babble (LB), Lyric Spectrum Noise (LSN), Music Noise (MN), and Music Spectrum Noise (MSN). Each masker had two variant, according to number of talker / singer (two and four). The preparation of each masker is described below:

Lyrical Babble (LB). Speech babble was recorded in the Kannada language. The lyrics of the song were considered as the individual speech track. The number of talkers in multi-talker babble was two and four to achieve various levels of difficulty. All the talkers' recording was done individually in a sound treated room. The microphone was fixed at a distance of 10 cm from the mouth of the speaker. The recording was done using the Adobe Audition 3.0 software installed in a personal computer, connected to a MOTU MICROBOOK II external sound card interface. Recordings were done at a sampling frequency of 44100 Hz. Post recording, all the individual tracks were first amplitude normalized and then mixed to obtain two talker LB (LB2) and four talker LB (LB4) **.**

Music Noise (MN). The Music Noise (MN) was recorded in a sound treated room. Four classical songs having different raga and rhythm, sung by different, classically trained female singers were recorded individually. The songs were chosen from Carnatic music training protocol, which were more familiar to musician group than the non-musician group. The output from the source was connected to a MOTU MICROBOOK II external soundcard interface connected to a laptop. Stimuli were recorded at a sampling rate of 44100 Hz using Adobe Audition software. Post recording, all the individual tracks were first amplitude normalized and then mixed to obtain MN. The number of singers in Music Noise was two and four to obtain MN2 and MN4 respectively, and to achieve various levels of difficulty.

Lyric Spectrum Noise (LSN). Lyric Spectrum Noise (LSN) with the spectral shape similar to that of LB was generated using a custom Matlab script. To generate LSN2 and LSN4 the spectrum of LB2 and LB4 was used respectively.

Music spectrum noise (MSN). Music Spectrum Noise with the spectral shape similar to that of PM was generated using a custom Matlab script. To generate MSN2 and MSN4 the spectrum of MN2 and MN4 was used respectively.

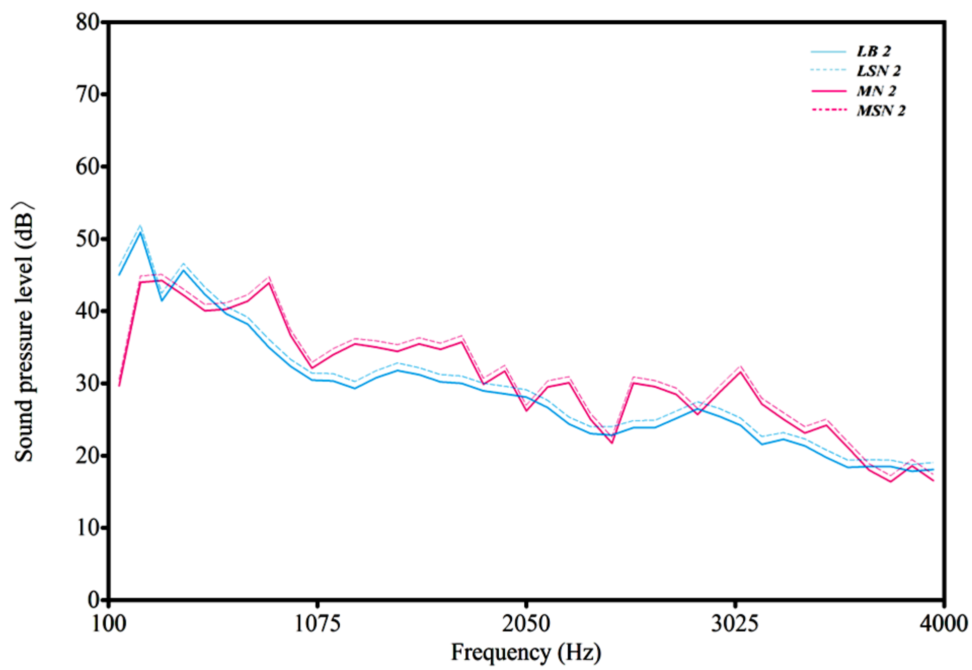
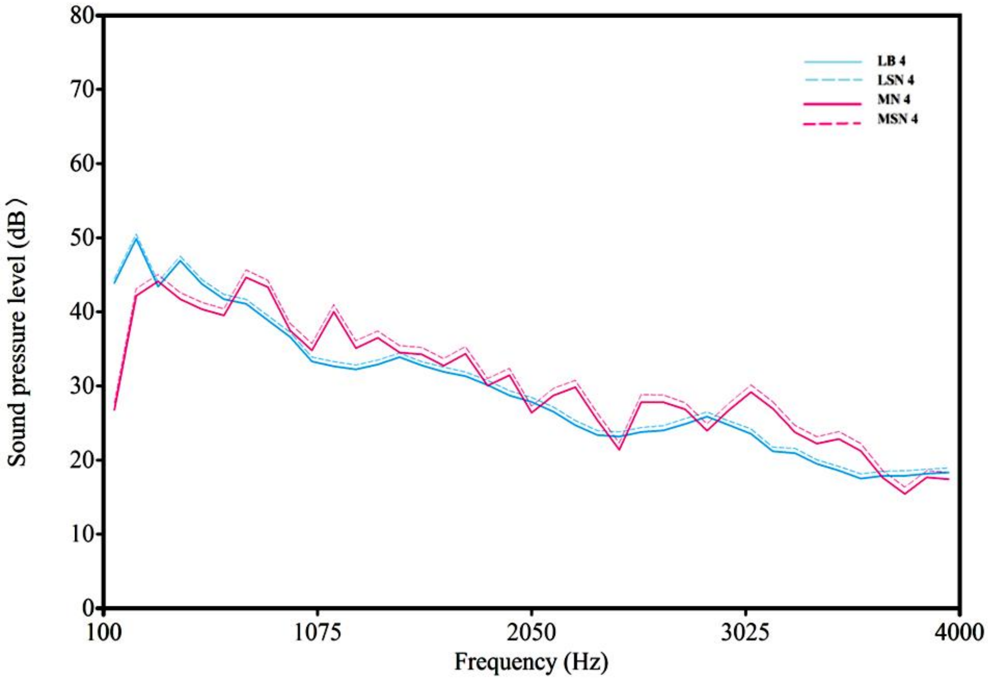


Figure 3.4.1.1.: Long-term average spectra (LTAS) for the two talker / singer maskers (Lyrical Babble 2, Music Noise 2, Lyric Spectrum Noise 2, and Music Spectrum Noise 2)



s

Figure 3.4.1.2.: Long-term average spectra (LTAS) for the four talker/ singer maskers (Lyrical Babble 4, Music Noise 4, Lyric Spectrum Noise 4, and Music Spectrum Noise 4)

The long-term average spectrum (LTAS) of all the 8 makers was analyzed using Praat software. The distribution of sound energy across frequency for 2 talker/ singer and 4 talker/singer maskers are illustrated in Figure 3.4.1.1. and Figure 3.4.1.2. respectively. The spectrum was almost similar in all 8 maskers.

3.4.2 Preparation of stimuli for SNR-50

Sixteen lists from Kannada sentence identification test by Geetha, Kumar, Manjula, and Pavan (2014) was used for estimating the signal to noise ratio required for 50% correct identification (SNR-50). Each list contains ten sentences with equal difficulty level and a steep psychometric function. The maskers (LB2,LB4 MN2, MN4, LSN2, LSN4, MSN2, MSN4) was mixed at different SNRs. The mixing of the maskers and the sentences at different SNRs were done using a custom Matlab function (Gnanateja, 2017). Two lists were used for each of the masker conditions. Within the list, SNR for each of the 10 sentences in the list was progressively reduced from +8 to -8 dB in steps of 2 dB.

3.5 Procedure

3.5.1 Speech perception measure – SNR-50

SNR-50 is the SNR at which a listener gets a 50% correct identification score during a speech reception task. For this, sixteen sentences lists from the Kannada sentence identification test by Geetha, Kumar, Manjula, and Pavan (2014) were used. Each sentence in the sentence list is of similar difficulty level, length, and the number of target keywords in it. Each list consists of 10 sentences with four keywords in each sentence. In each list, the SNR is manipulated, such that across the list, there is a progressive reduction in SNR in from sentence to next at a step size of 2 dB. Thus, within each list, the SNR reduced from +8 to -10. Each sentence has 4 keywords, making to a total of 40 keywords.

Each subject was made to sit comfortably in a quiet, well-lit room. A laptop loaded with the stimuli was used to present stimuli. Stimuli were presented through a calibrated headphone at 70 dB SPL. The participants' task was to repeat the sentences heard verbatim while ignoring the background maskers. Two lists are used for each masker condition. Verbal responses for each list were recorded, and each correctly repeated keyword was awarded a score of 1. The total number of correctly repeated keywords in a list was identified, and the SNR-50 was calculated using the Spearman-Karber equation given by (Finney, 1952);

$$\text{Speech recognition threshold (SNR-50)} = i + 1/2(d) - (d)(\#correct)/(W)$$

where 'i' is the initial presentation level (+8 dB), 'd' is the attenuation/decrement step size (2 dB), 'W' is key words per decrement (4 in this case) and '#correct' is the total number of correct keywords repeated by the participants. The lists used for the different masker conditions was randomized across the participants as well the order of sentences presented within a list was randomized to ensure there are no order effects.

3.6 Statistical Analysis

Shapiro Wilk's test of normality was carried out to check the normality of the entire data. The choice of parametric test will be made in case of normal distribution ($p<0.05$) or else choice of Non parametric test will be made ($p>0.05$) to compare the SNR50 between the two groups. Independent t test or Mann Whitney U test will administered accordingly to check the difference in SNR-50 between Musician and Non-musician groups in all 8 different masker conditions.

Chapter 4

Results

In the present study, the main objective was to compare the effect of different maskers on speech perception scores between the two groups, Musicians, and Non-musicians. A total of 20 individuals participated in the study. As per the inclusion criteria, the Mini PROMPS test was administered on all 20 individuals. The musicians who scored above 18 in the test were considered for the Musician group in the study. Furthermore, for the Non-musician group, there was no such criterion. Each group consisted of 10 participants, and all of them underwent the SNR-50 measure in all eight different masking conditions.

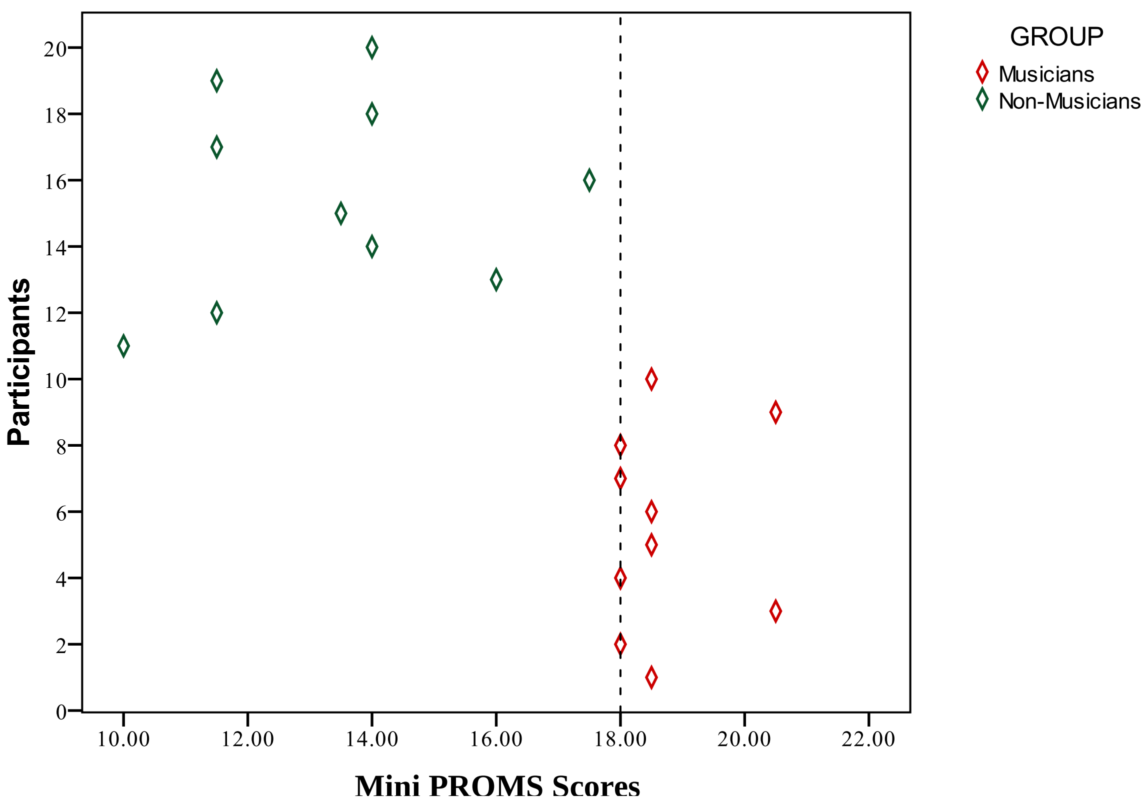


Figure 4.: Distribution of Mini PROMS scores of Participants.

The above Graph (Figure 4.) depicts the distribution of participants' scores in the Mini PROMS test. The scores of the Non-musician group were scattered from 10 to 17.5. Moreover, all the participants in the Musician group had scored beyond 18.

SNR-50 measure was administered on participants in both Musician and Non-musician groups in all eight different masking conditions and was calculated using the Spearman-Karber equation. The descriptive statistics were done on the derived values. To verify the normality of the sample, the Shapiro Wilk test was administered. Even though the result revealed normal distribution due to the limited sample size, a non-parametric test was chosen. To compare the SNR-50 between Musician and Non-musician groups, the Mann Whitney U test was administered.

4.1 Effect of two talker/ singer maskers on SNR-50

All participants underwent the SNR-50 measure in four different two talker/ singer masker conditions (LB2, MN2, LSN2, MSN2). SNR-50 at all the four conditions was compared between Musician and Non-musician groups using Mann Whitney U test.

Results revealed that the SNR-50 was significantly poorer for Musicians than for Non-musicians in MN2 masker condition. Whereas, there was no significant difference between the two groups for LB2, LSN2, and MSN2 masker conditions ($p > 0.05$). Table 4.1. gives the U and p values for various comparisons.

Given below, Figure 4.1. shows the median with Interquartile Range of on SNR-50 between Musician and Non-musician groups along with the individual SNR-50 values for different two talker/ singer masker conditions.

Table 4.1.: *Results of Mann Whitney U test for different two talker/singer masker conditions between musicians and non-musicians.*

Masker Conditions	U value	p value
Lyrical Babble 2 (LB2)	44.50	0.668
Music Noise 2 (MN2)	15.50	0.009*
Lyrical Spectrum Noise 2 (LSN2)	36.00	0.277
Music Spectrum Noise 2 (MSN2)	27.00	0.071

*<0.05 – Significant difference

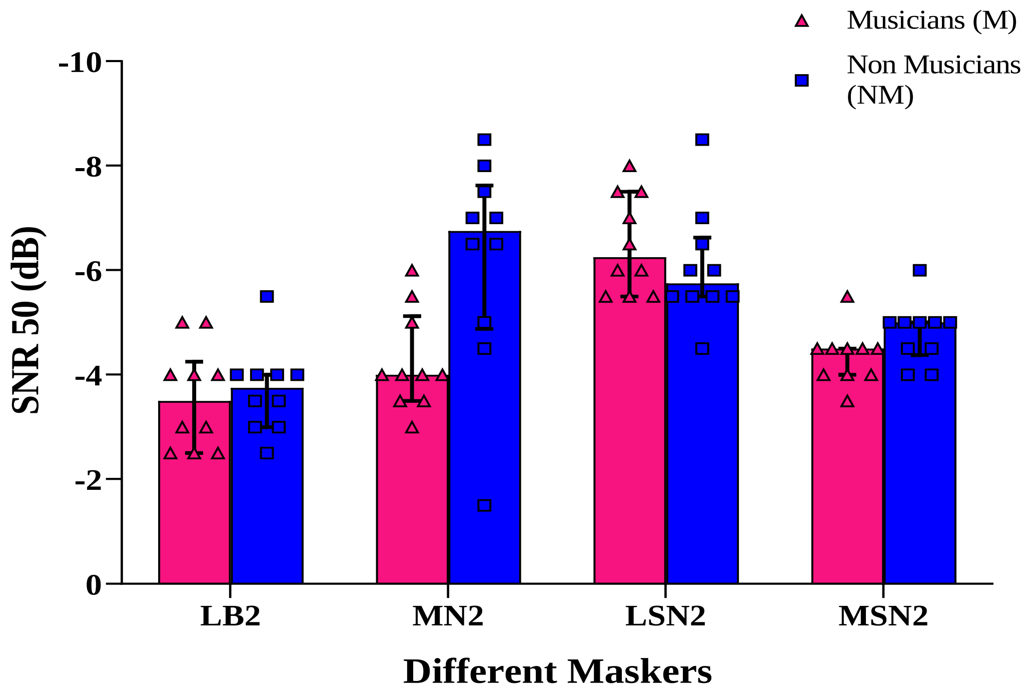


Figure 4.1.: Median with Interquartile Range of SNR-50 for Musician (Pink bar) and Non-musician (Blue bar) group in two talker/singer masker conditions (Lyrical Babble 2, Music Noise 2, Lyric Spectrum Noise 2, and Music Spectrum Noise 2)

4.2 Effect of four talker/ singer maskers on SNR-50

All participants underwent the SNR-50 measure in four different four talker/ singer masker conditions (LB4, MN4, LSN4, MSN4). SNR-50 at all the four conditions was compared between Musician and Non-musician groups using Mann Whitney U test.

Results revealed that the SNR-50 was significantly poorer for Musicians than for Non-musicians) in MN4 masker condition. Whereas, there was no significant difference between the two groups for LB4, LSN4, and MSN4 masker conditions. Table 4.2. Gives the U and p values for various comparisons.

Table 4.2.: *Results of Mann Whitney U test for different four talker/singer masker conditions between musicians and non-musicians.*

Masker Conditions	U value	p value
Lyrical Babble 2 (LB4)	48.00	0.878
Music Noise 2 (MN4)	20.00	0.022*
Lyrical Spectrum Noise 2 (LSN4)	37.50	0.333
Music Spectrum Noise 2 (MSN4)	45.50	0.721

*<0.05 – Significant difference

The Figure 4.2.shows, the median with Interquartile Range of SNR-50 between Musician and Non-musician groups with, individual SNR-50 values for different four talker/ singer masker conditions.

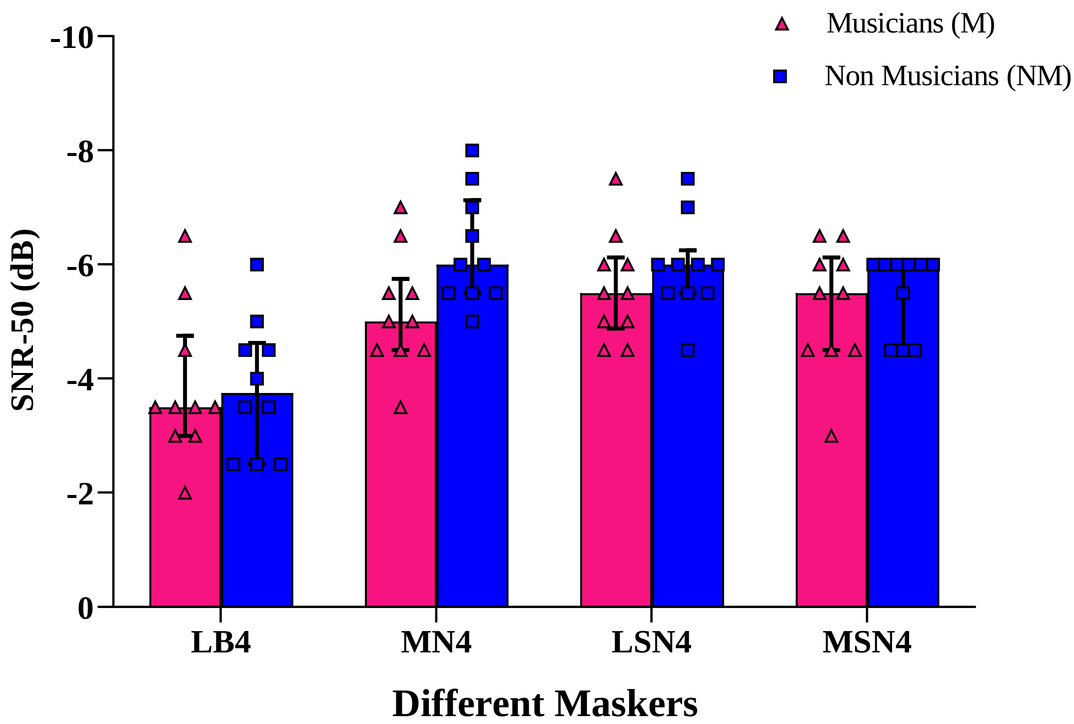


Figure 4.2.: Median with Interquartile Range of SNR-50 for Musician (Pink bar) and Non-musician (Blue bar) group in four talker/singer masker conditions (Lyrical Babble 4, Music Noise 4, Lyric Spectrum Noise 4, and Music Spectrum Noise 4)

Chapter 5

Discussion

The main objective of the study was to measure and compare the SNR required to obtain 50% correct scores (SNR-50) for two types of speech maskers (LB and LSN) and music maskers (MN and MSN) in Musicians and Non-musicians in two and four talker/ singer conditions.

5.1 Effect of Speech Maskers on Speech Perception

Results revealed that there was no significant difference in SNR-50 scores between Musician (M) and Non-musician (NM) group, in both two and four talker/singer conditions of Lyrical Babble (LB) and Lyric Spectrum Noise (LSN). This is in congruence with studies stating that there is no musician advantage for speech in noise perception tasks (Boebinger et al., 2015; Ruggles et al., 2014). Even with enhanced frequency discrimination ability, musicians performed equivalent to non-musicians in speech perception tasks (Fuller et al., 2014).

However, the median scores were less for the Musician group in LB2, LB4, and LSN4 conditions. The reduced scores in LB2 and LB4 conditions could be due to knowledge about the lyrics of the songs in musicians. Even during task at times, some of the participants were repeating the lyric part of LB instead of the target sentence. Furthermore, for LSN4 condition due to the increased number of talkers, the spectrum could have been more flat, providing no or less chance of dip listening.

These results are contrary to results obtained by Başkent & Gaudrain (2016), where the musicians showed overall better intelligibility than non-musicians, confirming a musician's advantage for speech-on-speech perception.

5.2 Effect of Music Maskers on Speech Perception

Results revealed significantly poor SNR-50 scores in the Musician (M) group than the Non-musician (NM) group, in both two and four singer conditions of Music Noise (MN). The presence of MN, impaired the perception of speech more for musically trained individuals than for those who were not trained. This can be explained by the spectro-temporal characteristics of the music signal. When Music acts as the noise, it induces both Energetic Masking (EM) and Informational Masking (IM). The overlapping of the signal and masker at the level of peripheral end organ exciting the same region on basilar membrane causes EM. Whereas, the IM is the result of the actions at several stages of processing, which are beyond the auditory periphery and is intimately connected to perceptual grouping and source segregation, general cognitive processing abilities, attention, and memory (Kidd et al., 2008). IM occurs when maskers are highly similar or confusable with the target, thus producing competition at physiological sites beyond the auditory periphery (J. Swaminathan et al., 2015). The spectral components of the MN interact with the speech leading to EM at the periphery. Whereas, the dynamic temporal characteristics and semantic information of the MN leads to IM. Due to their musical expertise and prior knowledge about the music, musicians are more prone to IM than EM.

However, Oxenham et al., (2003) reported that the amount of IM was less in musicians when different multi-tone masking paradigm and non-musical maskers were used in the study. This clears the point that the semantic information related to the lyric and melody of the MN was the main reason for IM in the Musician group.

The prior knowledge and familiarity of the songs used in the study for creating the MN are also one of the potential factors affecting speech perception of Musician group in MN conditions. The familiarity about the MN drifts the attention towards the masker and increases future expectancies of the same, causing more IM due to cognitive loading and leading to poor speech perception of speech. However, this is contrary to the study by Russo & Pichora-Fuller (2008), where the young individuals used this increased familiarity for better stream segregation and hence scored better in word recognition tasks.

According to Gfeller et al. (2012), individuals' ability to separate speech and Music depends on their ability to perceive the pitch. Even though musicians have better pitch perception, after separating the masker from the speech, the attention shifts towards masker due to the familiarity, therefore, it leads to a more reduced perception of speech in MN's presence in the Musician group.

The Non-musician group had better SNR-50 values in both 2 and 4 singer conditions of MN. This can be accounted as a result of release from masking due to dip listening caused by the dynamic temporal characteristics of the music masker. This is in congruence with the study done by Eskridge Elizabeth N. et al., (2012), where the Speech Recognition Threshold was better in Modulated Music Spectrum Noise and Music Noise conditions having dynamic temporal characteristics compared Music Steady Noise having the only steady spectral characteristic.

Results also revealed that there was no significant difference in SNR-50 scores between Musician and Non-musician group, in both two and four talker/singer conditions of Music Spectrum Noise (MSN). However, the median scores were less for the Musician group in MSN2 and MSN4 conditions. The reduced scores in MSN2 and MSN4 conditions can be due to the masker's temporal modulations, adding IM content in the task. On the other hand, results revealed a statistical difference for MN but not for MSN between Musicians and Non-musicians groups. This highlights the point that IM in Musicians is more due to lyrical and melody content of the signal than temporal characteristics.

Therefore, the presence of music as competing stimuli during the speech perception tasks is difficult in Musicians compared to Non-musicians due to their expertise in music. The music adds more informational content to the masker in the case of Musicians, whereas it serves as just an energetic masker in the case of Non-musicians.

Chapter 6

Summary and conclusion

To main aim of the present study was to compare the perceptual abilities of Musicians and Non-musicians in the presence of Speech and Music maskers of varying levels of difficulty. The main objective of this study was to measure and compare the SNR required to obtain 50% correct scores (SNR-50) for two types of speech maskers [Lyrical Babble (LB) and Lyric Spectrum Noise (LSN)] and Music maskers [Music Noise (MN) and Music Spectrum Noise(MSN)] in two and four talker/singer conditions between Musicians (M) and Non-musicians (NM).

A total of 20 normal hearing individuals aged between 18-35 yrs were involved in a study. The participants were divided into two groups as Musicians (M, n=10) and Non-musicians (NM, n=10) based on the musical training they have received (>5years) and their performance in Mini-Profile of Musical Perception Skills (Mini-PROMS) (score of >18 in Musician group). The normal hearing acuity of participants in both the group was evaluated by pure tone audiometry, tympanometry, and otoacoustic emissions. In order to evaluate and compare the effect of different speech and music maskers in 2 and 4 talker/singer conditions, SNR-50 was measured in all eight masker conditions ((LB2, LB4, MN2, MN4, LSN2, LSN4, MSN2, and MSN4) on all participants in Musician and Non-musician group.

Results revealed a significant difference in SNR-50 between Musician and Non-musician group in MN2 and MN4 conditions, whereas there was no significant difference of SNR-50 between groups in other masker conditions, namely LB2, LB4, LSN2, LSN4, MSN2, and MSN4.

However, reduced SNR-50 scores were noted in the Musician group in above mentioned conditions compared to Non-musician group.

This study demonstrates that the presence of music as competing stimuli during the speech perception tasks is difficult in Musicians compared to Non-musicians due to their expertise in music. The music adds more informational content to the masker in the case of Musicians, whereas it serves as just an energetic masker in the case of Non-musicians.

6.1 Implications of the study

The study throws light on whether the background music affects the speech perception or not, which will further help in the counselling of the Hearing impaired individuals on how to manage in such situations.

Whether the musical training is advantage or disadvantage to encounter a situation with music as background stimulus in everyday life can be answered. This helps to explore usefulness of musical training as a rehabilitative option for individuals with speech perception difficulties.

6.2 Limitations and Future directions

A significant limitation to the present study is the small sample size within each group. Therefore to be more certain of generalization of the results, future studies should include larger sample size within each of the groups.

Even though the comparison of different masker conditions between groups was performed statistically, the same was not performed within group. Including this in future studies would answers whether melody or lyric of music causes more informational masking.

References

American National Standards Institute. (1999). Maximum permissible ambient noise for audiometric test rooms (ANSI S3. 1-1999 [R2013]).

ANSI, A. (2004). S3. 6 Specification for Audiometers. *Society of America, New York, NY, USA*.

Başkent, D., & Gaudrain, E. (2016). Musician advantage for speech-on-speech perception. *The Journal of the Acoustical Society of America*, *139*(3), EL51–EL56.

Başkent, D., van Engelshoven, S., & Galvin III, J. J. (2014). Susceptibility to interference by music and speech maskers in middle-aged adults. *The Journal of the Acoustical Society of America*, *135*(3), EL147–EL153.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*, 94.

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences*, *930*(1), 232–258.

Bidelman, G. M., Schneider, A. D., Heitzmann, V. R., & Bhagat, S. P. (2017). Musicianship enhances ipsilateral and contralateral efferent gain control to the cochlea. *Hearing Research*, *344*, 275–283.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11818–11823. https://doi.org/10.1073/pnas.191355898

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *The Journal of the Acoustical Society of America*, *137*(1), 378–387.

Bregman, A. S., & Ahad, P. A. (1995). *Demonstrations of auditory scene analysis: The perceptual organization of sound*. Auditory Perception Laboratory, Psychology Department, McGill University.

Coffey, E. B. J., Mogilever, N. B., & Zatorre, R. J. (2017). Speech-in-noise perception in musicians: A review. *Hearing Research*, *352*, 49–69. https://doi.org/10.1016/j.heares.2017.02.006

Dalton, B. H., & Behm, D. G. (2007). Effects of noise and music on human and task performance: A systematic review. *Occupational Ergonomics*, *7*(3), 143–152.

Davies-Venn, E., Nelson, P., & Souza, P. (2015). Comparing auditory filter bandwidths, spectral ripple modulation detection, spectral ripple discrimination, and speech recognition: Normal and impaired hearing. *The Journal of the Acoustical Society of America*, *138*(1), 492–503. https://doi.org/10.1121/1.4922700

Eskander, A., Gordon, K. A., Kadhim, L., Papaioannou, V., Cushing, S. L., James, A. L., & Papsin, B. C. (2011). Low Pediatric Cochlear Implant Failure Rate: Contributing Factors in Large-Volume Practice. *Archives of Otolaryngology–Head & Neck Surgery*, *137*(12), 1190–1196. https://doi.org/10.1001/archoto.2011.200

Eskridge Elizabeth N., Galvin John J., Aronoff Justin M., Li Tianhao, & Fu Qian-Jie. (2012). Speech Perception With Music Maskers by Cochlear Implant Users and Normal-Hearing Listeners. *Journal of Speech, Language, and Hearing Research*, *55*(3), 800–810. https://doi.org/10.1044/1092-4388(2011/11-0124)

Finney, D. J. (1952). *Probit Analysis*.

Fuller, C. D., Galvin, J. J., Maat, B., Free, R. H., & BaÅŸkent, D. (2014). The musician effect: Does it persist under degraded pitch conditions of cochlear implant simulations? *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00179

Furnham, A., & Strbac, L. (2002). Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. *Ergonomics*, *45*(3), 203–217. https://doi.org/10.1080/00140130210121932

Geetha, C., Kumar, K. S. S., Manjula, P., & Pavan, M. (2014). Development and standardisation of the sentence identification test in the Kannada language. *J Hear Sci*, *4*(01), 18–26.

Gfeller, K., Turner, C., Oleson, J., Kliethermes, S., & Driscoll, V. (2012). Accuracy of Cochlear Implant Recipients in Speech Reception in the Presence of Background Music. *Annals of Otology, Rhinology & Laryngology*, *121*(12), 782–791. https://doi.org/10.1177/000348941212101203

Hughson, W., & Westlake, H. D. (1943). Manual for program outline for rehabilitation of aural casualties both military and civilian. *Transactions of the American Academy of Opthamology & Otolanyngology*, *Suppl*, 3–15.

Kidd, G., Mason, C. R., Richards, V. M., Gallun, F. J., & Durlach, N. I. (2008). Informational masking. In *Auditory perception of sound sources* (pp. 143–189). Springer.

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Music Enrichment Programs Improve the Neural Encoding of Speech in At-Risk Children. *Journal of Neuroscience, 34*(36), 11913–11918. https://doi.org/10.1523/JNEUROSCI.1881-14.2014

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 739–751. https://doi.org/10.1037/0096-1523.18.3.739

Kumar, P., Sanju, H. K., & Nikhil, J. (2016). Temporal Resolution and Active Auditory Discrimination Skill in Vocal Musicians. *International Archives of Otorhinolaryngology*, *20*(4), 310–314. https://doi.org/10.1055/s-0035-1570312

Levitin, D. J., & Menon, V. (2003). Musical structure is processed in “language” areas of the brain: A possible role for Brodmann Area 47 in temporal coherence. *Neuroimage*, *20*(4), 2142–2152.

Marques, C., Moreno, S., Luís Castro, S., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*(9), 1453–1463.

Mok, P. K. P., & Zuo, D. (2012). The separation between music and speech: Evidence from the perception of Cantonese tones. *The Journal of the Acoustical Society of America*, *132*(4), 2711–2720. https://doi.org/10.1121/1.4747010

O’neil, W., & Guthrie, M. S. (2001). DPOAEs among normal-hearing musicians and non-musicians. *Hearing Review [Serial Online]*.

Oxenham, A. J., Fligor, B. J., Mason, C. R., & Kidd Jr, G. (2003). Informational masking and musical training. *The Journal of the Acoustical Society of America*, *114*(3), 1543–1549.

Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging*, *33*(7), 1483–e1.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, *30*(6), 653–661.

Patel, A. D. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00142

Rajendran, V. G., Teki, S., & Schnupp, J. W. H. (2018). Temporal Processing in Audition: Insights from Music. *Neuroscience*, *389*, 4–18. https://doi.org/10.1016/j.neuroscience.2017.10.041

Rammsayer, T., & Altenmller, E. (2006). Temporal Information Processing in Musicians and Nonmusicians. *Music Perception*, *24*(1), 37–48. https://doi.org/10.1525/mp.2006.24.1.37

Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A. (2008). Prediction of the Intelligibility for Speech in Real-Life Background Noises for Subjects With Normal Hearing: *Ear and Hearing*, *29*(2), 169–175. https://doi.org/10.1097/AUD.0b013e31816476d4

Rochette, F., Moussard, A., & Bigand, E. (2014). Music Lessons Improve Auditory Perceptual and Cognitive Performance in Deaf Children. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00488

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, *27*(3), 465–485. https://doi.org/10.1017/S014271640606036X

Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PloS One*, *9*(1), e86980.

Russo, F., & Pichora-Fuller, M. (2008). Tune In or Tune Out: Age-Related Differences in Listening to Speech in Music. *Ear and Hearing*, *29*(5), 746–760.

Schellenberg, E. G. (2015). Music training and speech perception: A gene–environment interaction. *Annals of the New York Academy of Sciences*, *1337*(1), 170–177.

Shi, L.-F., & Law, Y. (2010). Masking effects of speech and music: Does the masker's hierarchical structure matter? *International Journal of Audiology*, *49*(4), 296–308. https://doi.org/10.3109/14992020903350188

Skoe, E. (2017). The Hearing Power of Music. *The ASHA Leader*. https://leader.pubs.asha.org/doi/abs/10.1044/leader.AEA.22122017.20

Slater, J., & Kraus, N. (2016). The role of rhythm in perceiving speech in noise: A comparison of percussionists, vocalists and non-musicians. *Cognitive Processing*, *17*(1), 79–87. https://doi.org/10.1007/s10339-015-0740-7

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research*, *291*, 244–252.

Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*(1–2), 22–29.

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language*, *123*(3), 191–201. https://doi.org/10.1016/j.bandl.2012.09.001

Strait, D. L., Parbery-Clark, A., O'Connell, S., & Kraus, N. (2013). Biological impact of preschool music classes on processing speech in noise. *Developmental Cognitive Neuroscience*, *6*, 51–60. https://doi.org/10.1016/j.dcn.2013.06.003

Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd Jr, G., & Patel, A. D. (2015). Musical training, individual differences and the cocktail party problem. *Scientific Reports, 5*, 11628.

Swaminathan, S., & Gopinath, J. K. (2013). Music Training and Second-Language English Comprehension and Vocabulary Skills in Indian Children. *Psychological Studies, 58*(2), 164–170. https://doi.org/10.1007/s12646-013-0180-3

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*(1), 37–46.

Zendel, B. R., & Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *Journal of Cognitive Neuroscience*, *21*(8), 1488–1498.

Zendel, B. R., & Alain, C. (2012). The Influence of Lifelong Musicianship on Neurophysiological Measures of Concurrent Sound Segregation. *Journal of Cognitive Neuroscience*, *25*(4), 503–516. https://doi.org/10.1162/jocn_a_00329

Zentner, M., & Strauss, H. (2017). Assessing musical ability quickly and objectively: Development and validation of the Short-PROMS and the Mini-PROMS. *Annals of the New York Academy of Sciences*, *1400*(1), 33–45.

Zhang, L., Peng, W., Chen, J., & Hu, L. (2015). Electrophysiological evidences demonstrating differences in brain functions between nonmusicians and musicians. *Scientific Reports, 5*, 13796.