Effect of Prolonged Non-Traumatic Noise Exposure on Unvoiced Speech Recognition

by

Mengchao Zhang

Bachelor of Education, East China Normal University, 2010

Master of Education, East China Normal University, 2013

Submitted to the Graduate Faculty of the

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Mengchao Zhang

It was defended on

June 4, 2020

and approved by

Catherine Palmer, Associate Professor, Department of Communication Science and Disorders

Bharath Chandrasekaran, Professor, Department of Communication Science and Disorders

Richard Stern, Professor, Department of Electrical and Computer Engineering, Carnegie Mellon University

Dissertation Director: Christopher Brown, Associate Professor, Department of Communication Science and Disorders

Effect of Prolonged Non-Traumatic Noise Exposure on Unvoiced Speech Recognition

Mengchao Zhang, MEd

University of Pittsburgh, 2020

Animal models in the past decade have shown that noise exposure may affect temporal envelope processing at supra-threshold levels while the absolute hearing threshold remains in the normal range. However, human studies have failed to consistently find such issue due to poor control of the participants' noise exposure history and the measure sensitivity. The current study operationally defined non-traumatic noise exposure (NTNE) to be noise exposure at dental schools because of its distinctive high-pass spectral feature, non-traumatic nature, and systematic exposure schedule across dental students of different years. Temporal envelope processing was examined through unvoiced speech recognition interrupted by noise or by silence. The results showed that people who had systematic exposure to dental noise performed more poorly on tasks of temporal envelope processing than the exposed people. The effect of high-frequency NTNE on temporal envelope processing was more robust inside than outside the spectral band of dental noise and was more obvious in conditions that required finer temporal resolution (e.g faster noise modulation rate) than in those requiring less fine temporal resolution (e.g. slower noise modulation rate). Furthermore, there was a significant performance difference between the exposed and the unexposed groups on tasks of spectral envelope processing at low frequency. Meanwhile, the two groups performed similarly in tasks near threshold. Additional analyses showed that factors such as age, years of musical training, non-dental noise exposure history and peripheral auditory function were not able to explain the variance of the performance in tasks of temporal or spectral envelope processing. The findings from the current study support the general assumptions from

animal models of NTNE that temporal and spectral envelope processing issues related to NTNE likely occur in retro-cochlear sites, at supra-threshold levels, and could be easily overlooked by clinically routine audiologic screening.

Table of Contents

Preface ... **xix**

1.0 Background ... **1**

1.1 Statement of the problem ... **1**

1.2 Non-traumatic noise exposure in animal studies ... **3**

1.2.1 Auditory nerve fiber, its classes and functions in temporal envelope coding **3**

1.2.2 Effect of non-traumatic noise exposure on auditory nerve fibers ... **5**

1.2.3 Effect of non-traumatic noise exposure on central auditory system ... **6**

1.2.4 Summary for animal studies ... **7**

1.3 Non-traumatic noise exposure in human studies ... **8**

1.3.1 Conventional behavioral and electrophysiological measures for testing temporal envelope processing ... **9**

1.3.2 Review of human studies about the effect of non-traumatic noise exposure on temporal envelope processing ... **10**

1.3.3 Summary for human studies ... **12**

1.4 Rationales for the population and the measures to assess temporal envelope processing in relation to non-traumatic noise exposure ... **14**

1.4.1 Introducing the population of interest ... **14**

1.4.2 Introducing the measure ... **16**

1.4.2.1 What is unvoiced speech? ... **16**

1.4.2.2 Advantages of unvoiced speech recognition ... **17**

1.4.2.3 Potential challenges of using unvoiced speech recognition ... **18**

1.4.3 Summary of the proposed methodology 19

2.0 Experiment 1: Non-Traumatic Noise Exposure on Unvoiced Speech Recognition in Noise 20

2.1 Introduction 20

2.2 Methods 22

2.2.1 Participants..22

2.2.1.1 Overview of the screening procedures and inclusion criteria 22

2.2.1.2 Demographic screening 23

2.2.1.3 General noise exposure questionnaire 24

2.2.1.4 Dental noise exposure questionnaire... 27

2.2.1.5 Audiologic screening.. 28

2.2.1.6 Eligible participants.. 30

2.2.2 Stimuli 32

2.2.2.1 Target speech 32

2.2.2.2 Noise maskers... 33

2.2.3 Procedures 37

2.3 Results... 39

2.3.1 Unvoiced speech recognition in quiet 39

2.3.2 Unvoiced speech recognition in temporally modulated noise: SRT 41

2.3.3 Unvoiced speech recognition in temporally modulated noise: masking release 44

2.3.4 Unvoiced speech recognition in spectrally modulated noise: SRT 46

2.3.5 Unvoiced speech recognition in spectrally modulated noise: masking release 48
2.4 Interim discussion.......... 49
2.4.1 The relationship between non-traumatic noise exposure and temporal envelope processing..........50
2.4.2 The relationship between non-traumatic noise exposure and spectral envelope processing..........53
2.4.3 Summary of this experiment..........56
3.0 Experiment 2: Sensitivity of low-pass and high-pass filtered unvoiced speech recognition to broadened temporal integration window.......... 58
3.1 Introduction 58
3.2 Methods 59
3.2.1 Participants..........59
3.2.2 Stimuli 59
3.2.3 Procedures 61
3.3 Results.......... 62
3.3.1 Unvoiced speech recognition in quiet as a function of the temporal window 62
3.3.2 Unvoiced speech recognition in 32-Hz TMN as a function of the temporal window 63
3.4 Interim discussion.......... 64
4.0 Experiment 3: Non-Traumatic Noise Exposure on Silence-Interrupted Unvoiced Speech Recognition 67

4.1 Introduction 67

4.2 Methods 69

4.2.1 Participants...69

4.2.2 Stimuli 69

4.2.3 Procedures 71

4.3 Results.. 72

4.4 Interim discussion... 73

5.0 Additional analyses: Contributions of Demographic and Audiologic Factors to Temporal and Spectral Envelope Processing 77

5.1 Effect of NTNE on peripheral auditory functions and extended high frequency listening... 78

5.1.1 Raw data cleaning 79

5.1.1.1 Raw data 79

5.1.1.2 Dimension reduction.. 79

5.1.2 Effect of NTNE on peripheral auditory functions and extended high frequency listening 85

5.1.2.1 Effect of NTNE on middle ear muscle reflex....................................... 85

5.1.2.2 Effect of NTNE on DPOAE 85

5.1.2.3 Effect of NTNE on extended high-frequency pure-tone average 86

5.1.3 Summary of the effect of NTNE on peripheral auditory functions and extended high-frequency listening...87

5.2 Contributions of demographic factors and peripheral audiologic screening outcomes to temporal and spectral envelope processing outside the exposed noise band............ 88

5.2.1 Correlation and regression factors/predictors 88

5.2.2 Contributions to temporal envelope processing outside noise band 89

5.2.3 Contributions to spectral envelope processing outside noise band 92

5.3 Interim Discussion 94

6.0 Discussion.. 97

6.1 Summary of the current findings... 97

6.2 Effect of non-traumatic noise exposure on temporal envelope processing 99

6.2.1 Current finding vs. past findings...99

6.2.2 Speculations on the form of degraded temporal envelope processing related to NTNE 101

6.2.3 Poor temporal envelope processing manifests at higher modulation rates in noise- and silence-interrupted speech recognition 102

6.2.4 On the frequency-specific effect of non-traumatic noise exposure on temporal envelope processing..105

6.3 Probing the mechanisms of poorer spectral envelope processing.......................... 105

6.4 Limitations and future directions 106

7.0 Conclusion 110

Appendix A Subject demographic form 112

Appendix B Noise exposure questionnaires.. 114

Appendix C Diagnostic criteria and statistics of the hierarchical multiple linear regression model.. 116

Appendix C.1 Assumption check for outcome variables of LPF unvoiced speech tests 118

Appendix D Summary of findings and potentially affected auditory processing from animal models of NTNE 120

Bibliography 123

List of Tables

Table 1. Inclusion criteria for demographic screening 24

Table 2. Inclusion criteria for audiologic screening 30

Table 3. Descriptive and inferential statistics of the demographic information for the control and the experimental groups 31

Table 4. Passband frequencies of the SMNs 36

Table 5. Test conditions of unvoiced speech in noise and their abbreviations 39

Table 6. Simple effect comparisons between the two groups when filtering is controlled ... 40

Table 7. Simple effect comparisons between the two filtering conditions when group is controlled 41

Table 8. Cut-off frequencies of the lower and the upper sides of the band-pass filters for LPF and HPF unvoiced speech 60

Table 9. Parameters of the silence-interrupted unvoiced speech 70

Table 10. Test conditions of interrupted unvoiced speech their abbreviations 71

Table 11. Audiologic data used as outcome variables 79

Table 12. Exploratory Factor Analysis of the Audiologic Exams 82

Table 13. Audiologic data used as outcome variables 84

Table 14. Outcome variables for correlation/regression analysis and their orders entering the HMLR model 89

Table 15. Correlation between outcome variables of temporal envelope processing outside the frequency band of the noise and demographic or audiologic predictors 90

Table 16. HMLR model for LP48 91

Table 17. Correlation between outcome variable of spectral envelope processing outside the frequency band of the noise and demographic or audiologic predictors 92

Table 18. HMLR model for LPSMN 93

Table 19. Summary of the study conclusions 98

Appendix Table 1. Assumptions of multiple linear regression 116

Appendix Table 2. Summary of findings and potentially affected basic auditory processing and temporal processing from animal models of NTNE 120

Appendix Table 3. Summary of findings and potentially affected spectral processing from animal models of NTNE (see Table 20 for source study) 122

List of Figures

Figure 1. Schematic rate-level functions of ANFs with high (blue) and low (red) spontaneous rates 4

Figure 2. Spectra of the environmental noise at dental clinics and preclinics (replotted from Fernandes et al., 2006) and of the commonly used devices at dental clinics (replotted from Choosong et al., 2001) 15

Figure 3. Sample page of the subject background form (see Appendix A for higher resolution) 23

Figure 4. Sample noise exposure history questionnaire (see Appendix B for higher resolution) 25

Figure 5. Sample dental noise exposure questionnaire (see Appendix B for higher resolution) 27

Figure 6. Speech intelligibility as a function of cut-off frequencies of LPF and HPF unvoiced speech (n = 5) 33

Figure 7. Spectra of the full-band UN (solid line), LPF UN (dashed line) and HPF UN (dotted line) 34

Figure 8. Time-domain waveforms of the full-band UN (black waves in both panels), of LPF 16-Hz TMN (grey wave in the upper panel), and of HPF 16-Hz TMN (grey wave in the lower panel) 35

Figure 9. Spectra of the full-band UN, LPF SMN, and HPF SMN 36

Figure 10. Performance of LPF and HPF unvoiced speech in quiet: dark grey, CTL group; light grey, EXP group (error bars: 95% CI) 40

Figure 11. SRTs of unvoiced speech in UN (0Hz), 16-Hz and 32-Hz TMNs under each filtering condition: left panel, LPF; right panel, HPF; dark grey, CTL group; light grey, EXP group. Error bars: 95% CI. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. 42

Figure 12. SRTs of unvoiced speech in UN (0Hz) and TMN within each group: left panel, CTL group; right panel, EXP group; drak grey, LPF; light grey, HPF. Error bars: 95% CI. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.. 43

Figure 13. Masking release of unvoiced speech in TMN under each filtering condition. Error bars: 95% CI. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.. 45

Figure 14. Masking release of unvoiced speech in TMN within group. Error bars: 95% CI. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. 45

Figure 15. SRTs of unvoiced speech in SMN under each filtering condition. Error bars: 95% CI. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.. 47

Figure 16. SRTs of unvoiced speech in SMN within each group. Error bars: 95% CI. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. 48

Figure 17. Masking release of unvoiced speech in SMN for clustered filtering conditions (left panel) and clustered group (right panel) (error bars: 95% CI) 49

Figure 18. Speech intelligibility of temporally smoothed LPF (dark grey) and HPF (light grey) unvoiced speech in quiet as a function of temporal integration window. Error bars: 95% CI. Uns: unsmoothed. 62

Figure 19. SRT of temporally smoothed LPF (dark grey) and HPF (light grey) unvoiced speech in 32-Hz SAM noise as a function of temporal window. The horizontal bracket indicates a significant difference ($p < 0.05$) between the 128-Hz condition and another given smoothed condition. Error bars: 95% CI. Uns: unsmoothed.................................. 63

Figure 20. Intelligibility of temporally interrupted speech in quiet as a function of interruption rate from past studies. Dotted line: natural speech. Dashed line: filtered natural speech. Solid line: noise-vocoded speech. 67

Figure 21. Pilot study on the interrupted unvoiced speech intelligibility as a function of rate. Left panel: results from 3 listeners at stage 1. Right panel: results from an additional 6 listeners at stage 2. 70

Figure 22. Time domain waveforms of the original sentence and the silence-interrupted sentence (interruption rate 12 Hz)... 71

Figure 23. Recognition accuracy (%) of silence-interruped unvoiced speech under each filtering condition. Error bars: 95% CI. *, p < 0.05; **, p < 0.01; ***, p < 0.001. 72

Figure 24. Recognition accuracy (%) of silence-interruped unvoiced speech under within each group. Error bars: 95% CI. *, p < 0.05; **, p < 0.01; ***, p < 0.001. 73

Figure 25. Headcount of the participants at various scores. Top panels: LPF. Bottom panels: HPF. From left to right: 12, 24, and 48 Hz. Within each panel, blue: CTL group; red: EXP group. 74

Figure 26. Matrix of correlation coefficients among the outcomes of the audiologic screenings. Axis labels are in the form of 'test + test frequency (kHz)'. AR: acoustic reflex. A: audiogram. Labels without a test letter belong to DPOAE. Negative correlation coefficients were set to 0 only for the purpose of visual clarity. Boxes with white solid lines encircled the items for different tests. Boxes with dotted lines encircled the items within the tests that were grouped together for analysis. 80

Figure 27. Matrix of correlation coefficients among the new outcome variables of the audiologic screenings 84

Figure 28. Acoustic reflex amplitude between groups at each test frequency and on average. 85

Figure 29. DPOAE amplitude between groups at each test frequency and on average at low frequency (LF), high frequency (HF) and extendend high frequency (EHF) regions...... 86

Figure 30. Audiogram at extended high frequencies between groups 87

Figure 31. Modulation spectra in normalized modulation index (upper panels) and the PDFs of modulation energy (lower panels) of unvoiced speech (colored) and noise (grey). Left panels are for LPF speech and right panels HPF speech. Speech modulation spectra were processed in four channels with one color representing the modulation spectra from the same channels, and three lines in the same color represent the modulation spectra from three different sets of speech stimuli. The modulation spectra for each TMN was processed in one channel with the most prominent peak at the rate of modulation. The darkest grey stems represented the modulation spectra of 4-Hz TMN and the lightest grey 32-Hz TMN. 103

Figure 32. Unpublished pilot data of SRT of LPF (dark grey) and HPF (light grey) unvoiced speech in TMN of different modulation rates (N = 25). Error bars: 95% CI. 104

Appendix Figure 1. Subject background form (Page 1)... 112

Appendix Figure 2. Subject background form (Page 2)... 113

Appendix Figure 3. General noise exposure questionnaire 114

Appendix Figure 4. Dental noise exposure questionnaire.. 115

Appendix Figure 5. SPSS output for test of normality for LPSMN and LP48 118

Appendix Figure 6. SPSS plot of standardized residuals against standardized predicted value for homoscedasticity. Left: LP48. Right: LPSMN... 118

Appendix Figure 7. SPSS output for multicollinearity. Left: LP48. Right: LPSMN. LifetimeDNE: dental noise exposure. LifeNDNE: non-dental noise exposure. LEipsiAR: left ear ipsilateral acoustic reflex. LEDP0.5_2.8K: low frequency DPOAE. LEDP3.5_8.8K: high frquency DPOAE... 119

Preface

It is the year 2020, A.D. The world has not come to an end. The lovely Pitt is still this low-key university in a low-key city. It is extraordinarily beautiful walking from the intersection of Bellefield and Fifth, near Heinz Chapel, and to the Cathedral of Learning at 8:30 on a regular sunny morning in September. The sun is just about right to light up the Cathedral, quiet and mesmerizing like an elvish kingdom in Middle Earth.

The year 2020 marks the seventh year of my sojourn in the United States. Along with my pursuit of the Ph.D. degree, I have been blessed with many great mentors, professors, and friends. I would like to first express my sincerest gratitude to my dissertation committee chair and advisor, Dr. Christopher Brown. As my dissertation committee chair, he has been supportive in every way I can think of. I thank him for his constructive advice, enthusiasm in sharing his expertise and knowledge, and for tirelessly encouraging me to try, to challenge myself, and to see things differently.

I cannot make this far without the distinguished members of my dissertation committee. I would like to thank Dr. Richard Stern for his generous support, selflessly spending time to provide instructions and honest critiques. And I especially thank him for being there and patiently giving me advice during the toughest time of my Ph.D. study. I would like to thank Dr. Catherine Palmer for her kind encouragement, many timely reminders, and for generously allowing me to use her lab space and equipment. I would not have finished my data collection so smoothly if it were not for her support. I would like to thank Dr. Bharath Chandrasekaran for his helpful insights and many meaningful perspectives on doing research. I thank him for always emphasizing the importance of storytelling in research.

I would also like to thank my first Ph.D. advisor, Dr. Deborah Moncrieff. I would not have come to Pitt if it were not for her kind acceptance. She has been supportive and protective during her time mentoring my Ph.D. study and after her relocation. Without her, there was no way I could have made it this far, and I would never have had so much interest in auditory physiology and neuroscience. I will always be grateful for her being there on my Ph.D. journey.

I would like to thank many other great professors at Pitt and at CMU. I want to thank Dr. Connie Tompkins for teaching me the principles of scientific research, especially the importance of theory. I want to thank Dr. Laurie Heller for always caring about my progress in my Ph.D. study. I want to thank Dr. Michael Dickey for selflessly giving me advice on various aspects of my Ph.D. study. I would like to thank Dr. Deborrah Johnston, Ms. Michelle Parfitt, Dr. Ruth Auld, and colleagues at the DePaul School of Hearing and Speech for their constant and kind support to me during my Ph.D. journey. I would also like to thank my participants and my students, for whom I maintain my passions in researching and teaching.

I would like to thank my parents, who have been supporting me with unwavering confidence. I would like to thank my close relatives, my dear friends and colleagues I met in Shanghai, and my dear friends and colleagues I met at Pitt and in greater Pittsburgh. I cannot make this far without them constantly thinking about me, praying for me, and checking in on me.

Lastly, I would like to thank my maternal grandparents. Thank you for loving me unconditionally as Christ does. You two are always my soft spot.

1.0 Background

1.1 Statement of the problem

Sound is often represented in the form of a wave that records the pressure change in the medium over time. Sound is also represented by the mechanic vibrations of ear drum attached to the ossicle chain in the middle ear, through the tonotopic response of the cochlea and then by the electrical pulses of the auditory nerves that are stimulated at a fixed phase of the pressure variation; the representation from the auditory nerve is refined in the central auditory system (CAS) and eventually interpreted by the perceiver. If any parts along the auditory pathway fail to represent the sound as designed, it will likely decrease the accuracy of the final interpretation.

For many decades, noise exposure has been considered 'safe' or 'non-traumatic' as long as the absolute hearing threshold maintains within the normal range (Saunders et al., 1985; Eggermont, 2017). However, animal studies in the past decade have shown that non-traumatic noise exposure, even at daily-activity levels, could negatively impact the integrity of the auditory pathway (Kujawa & Liberman, 2009; Lin et al., 2011; Furman et al., 2013; Valero et al., 2017; Munguia et al., 2013; Sheppard et al., 2017; Pienkowski & Eggermont, 2009; Pienkowski & Eggermont, 2010a, 2010b; Pienkowski et al., 2011; Zhou & Merzenich, 2012; Fernandez et al., 2020). Naturally, many researchers proposed that non-traumatic noise exposure (NTNE) would compromise the representation of sound, causing poor auditory perception that is overlooked by routine audiologic exams. One of the frequently proposed auditory processing skills affected by NTNE was temporal envelope processing.

The reason that temporal envelope processing was highly likely affected by NTNE stemmed from the mechanism of NTNE. Briefly, NTNE was found to selectively damage the synapses between the inner hair cells (IHC) and the group of auditory nerve fibers (ANF) that were responsible for encoding intensity at supra-threshold levels, but the synapses of other ANFs that encoded near-threshold sound were less affected (Kujawa & Liberman, 2009; Lin et al., 2011; Furman et al., 2013). One of the consequences of such selective loss of high-intensity ANFs is that the auditory system loses its ability to faithfully represent the amplitude changes of a sound at supra-threshold levels, limiting the detection, the discrimination and the use of amplitude changes of the sound. These findings prompted the current study to ask:

1) Do people who had a specific type of NTNE perform differently on tasks of temporal envelope processing from people who were unexposed?

Furthermore, cortical-level studies have shown that NTNE even at moderate sound level produced maladaptive long-term changes. These changes were observed in both temporal processing or spectral processing and were found to be frequency-specific and supra-threshold (Pienkowski & Eggermont, 2010a; Munguia et al., 2013; Pienkowski et al., 2013; Sheppard et al., 2017; Pienkowski, 2018). Therefore, the current study continued to ask:

2) Do people who had a specific type of NTNE also perform differently on tasks of spectral envelope processing from people who were unexposed?

3) If there is a difference between the exposed and the unexposed, will the difference be frequency specific?

4) Will the difference occur at only supra-threshold levels?

Unfortunately, findings from past human studies examined the effect of NTNE only on temporal envelope processing and the conclusions have been inconsistent. In short, the past human studies used tasks based on detection or discrimination of amplitude modulation of sound to examine temporal envelope processing, which may indicate issues like off-frequency listening or central auditory compensation which may decrease the effect size. The other critical reason of inconsistent finding in human studies was the lack of control for listeners' noise exposure history, specifically, their highly variable exposure history such as noise exposure schedule, exposure intensity and noise spectral composition may further reduce the effect size.

The primary purpose of this study, therefore, was to examine the effect of NTNE on envelope processing, especially temporal envelope processing, while exerting more control over measures and exposure. The study also examined whether the effect of NTNE on envelope processing was frequency specific and occurred only at supra-threshold level.

Section 1.2 of this document provides a review on the mechanism of NTNE and its relation to poorer temporal envelope processing in animal studies. Then, section 1.3 provides a review on the human studies that examined the relationship between noise exposure and temporal envelope processing and a summary on the issues of conventional measures of temporal envelope processing. Lastly, section 1.4 provides the rationales for the population of interest and the measures that were used in the current study.

1.2 Non-traumatic noise exposure in animal studies

1.2.1 Auditory nerve fiber, its classes and functions in temporal envelope coding

Auditory nerve fibers (ANF) are found at the bases of the hair cells inside the cochlea. The afferent fibers carry the information into the CAS from inner hair cells while the efferent fibers bring information from the CAS back to the hair cells. About 90% of the afferent ANFs are heavily myelinated type I fibers. The type I ANFs are often classified based on the spontaneous rates (Sachs and Abbas, 1974; Liberman, 1978). The ANFs with high spontaneous rates (HSR, > 18 spike/sec) are largest in quantity (60%), have low thresholds and low saturation levels. The ANFs with low spontaneous rates (LSR, < 18 spike/sec) are fewer in quantity, have high thresholds (up to 40 dB SPL) and high saturation levels (Figure 1).

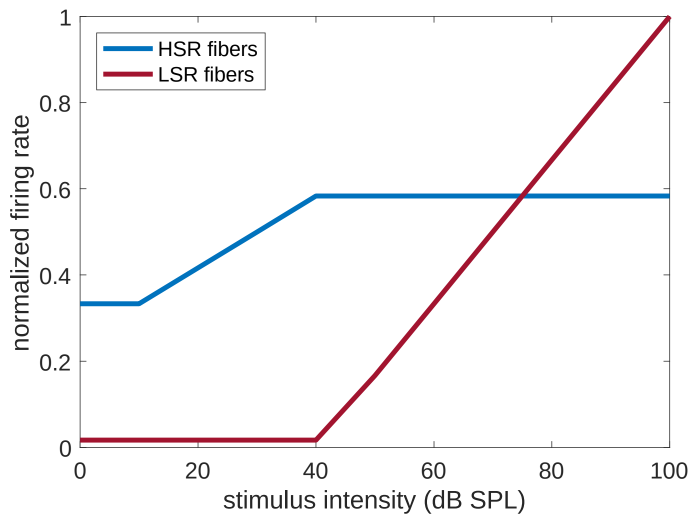


Figure 1. Schematic rate-level functions of ANFs with high (blue) and low (red) spontaneous rates

The ANFs with different dynamic ranges allow sound amplitude and amplitude change over time (i.e. temporal envelope) to be encoded properly at all intensities. To near-threshold, low-intensity sound, fibers with HSR actively respond and continue to increase firing rates with increasing stimulus level until firing rates saturate at around 30 dB above threshold (Sachs and Abbas, 1975), above which increasing stimulus level does not change the firing rates. Meanwhile, the saturation level of the fibers with HRS just hits the thresholds of fibers with LSR and these fibers start to increase their firing rates with further increase of stimulus level, not showing saturation even at fairly high intensities (e.g. 100 dB SPL). The collaboration between the two classes of ANFs ensures the proper encoding of the temporal envelope of a sound at a wide range of intensities.

1.2.2 Effect of non-traumatic noise exposure on auditory nerve fibers

Animal researches over the past decade have found that noise exposure could undo the collaboration among the ANFs with different spontaneous rates through selective damage to the ANFs with LSR while sparing the hair cells and the other ANFs with HSR (Kujawa & Liberman, 2009; Lin et al., 2011; Furman et al., 2013). Specifically, animals were exposed to brief and intense noise which would normally produce an acute threshold shift in distortion product otoacoustic emission (DPOAE) and wave I of auditory brainstem response (ABR), suggesting a temporary dysfunction in the outer hair cells (OHC) and the auditory nerves. But the thresholds of both measures returned to baseline if the animals were given enough time in quiet, leaving a misconception that the functions of the OHCs and the ANFs were recovered. However, ABR wave I amplitude at supra-threshold levels have shown irreversible and drastic decrease at higher intensities, indicating that the NTNE had potentially affected the ability of the ANFs to encode amplitude at high intensities (Kujawa & Liberman, 2009; Lin et al., 2011; Furman et al., 2013; Hickox & Liberman, 2014). Furthermore, Shaheen et al. (2015) observed significantly lower envelope following response (EFR) from the ANFs with characteristic frequencies (CF) at the affected spectral regions, indicating that the affected ANFs were unable to properly phase-lock to the temporal envelope of the sound.

Furthermore, the effect of NTNE seemed to be frequency specific. Greater loss of synapses consistently occurred at and above the frequency band of the noise than below the frequency band of the noise in studies using octave-band noise (OBN) (Kujawa & Liberman, 2009; Lin et al., 2011; Furman et al., 2013; Chen et al., 2019), or at the higher frequency side than at the lower frequency side in studies using broad-band noise (BBN) (Liu et al. 2012; Shi et al., 2013; Shi et al., 2016).

1.2.3 Effect of non-traumatic noise exposure on central auditory system

The effect of NTNE on temporal envelope processing is not confined to short-term noise exposure at intense levels (Fernandez et al., 2020) and is certainly not confined to the ANFs. Many studies of cat and rodent models found a detrimental effect of the noise exposure that was presented at daily environmental levels as long as the exposure last long enough. Michael Merzenich, Jos Eggermont and many others' works discovered long-lasting maladaptive changes at the subcortical and cortical auditory nuclei after the animals were exposed to prolonged moderate-level noise (e.g. 65 to 80 dB SPL for 4 to 12 weeks). These changes occurred in both temporal and spectral processing, including increase spontaneous activity (Pienkowski & Eggermont, 2009; Munguia et al., 2013; Pienkowski et al., 2013), reduced inhibitory neurons (Zhou & Merzenich, 2012; Munguia et al., 2013; Kamal et al., 2013; Lau et al., 2015), broadened frequency tuning curves (Zhou et al., 2011; Zhou & Merzenich, 2012; Kamal et al., 2013), disrupted excitation duration and peak latency (Pienkowski & Eggermont, 2009), disrupted cortical tonotopic representation for neuronal CFs in and outside the frequency band of the noise (Pienkowski & Eggermont, 2009; Pienkowski & Eggermont, 2010a, 2010b; Pienkowski et al., 2011, 2013) and so on. As Fernandez et al. (2020) estimated that continuous noise exposure at 75 to 85 dB SPL over 8 hours a day could already result in 10 to 40% loss of auditory nerve synapses, it is possible that the pathological effect of the prolonged moderate level of noise exposure would begin from the ANFs and extend into the auditory cortex, causing issues of temporal envelope processing throughout the central auditory pathway.

Like the studies concerning the ANFs, the effect of NTNE on the CAS also appeared to be frequency-specific, but the findings were more complicated. For example, prolonged moderate-level noise exposure usually increased the overall spontaneous activity of the auditory cortex regardless of the neuronal CFs and the bandwidth of the noise (Pienkowski & Eggermont, 2009; Munguia et al., 2013; Pienkowski et al., 2013). Meanwhile, the driven response was often found decreasing within the frequency band of the noise and increasing outside the frequency band of the noise (Pienkowski & Eggermont, 2009; Pienkowski & Eggermont, 2010b; Munguia et al., 2013; Pienkowski et al., 2013). The extent of decreased driven response was also affected by the steepness of the noise spectral slope. In summary, the effect of NTNE on the CAS depended on the spectral features of the noise, such as noise bandwidth, noise band center frequency, steepness of the noise spectral slope, etc., and also showed frequency specificity.

1.2.4 Summary for animal studies

Due to the different experimental conditions across the animal models of the ANFs and of the CAS, it is difficult to form a comprehensive model to describe the configuration and the extent of pathological changes that can be predicted by the level, the schedule, and the spectral features of the exposed noise. This lack of a unified theory on NTNE does pose challenges to generate a clear hypothesis on the auditory perceptual consequences in relation to noise exposure. However, it *is* evident based on the animal studies that noise exposure could spare the capacity of the auditory system to detect and encode soft-level sound but not spare temporal envelope processing or even other auditory processing of the ANFs and of the CAS at supra-threshold levels. And these supra-threshold processing issues could be long-lasting or even irreversible, and they could occur post noise exposure with short duration at intense levels or with long duration at moderate levels. The effect of noise exposure on the ANFs was more apparent at and above the frequency band of the noise (or at the higher end of the frequency band of the noise) while the effect on the auditory cortex was detrimental at the frequency band of the noise (e.g. decreased cortical representation). In other words, the effect of noise exposure on the auditory units with CFs within the frequency band of the noise was not entirely identical to the units with CFs outside the frequency band of the noise. Therefore, human studies should be careful to search for participants who had similar noise exposure profiles that had involved a noise with a distinctive spectral band, to test them with measures specifically designed for the noise spectrum and to avoid any confounding contributions of the unaffected auditory units from off-frequency regions.

1.3 Non-traumatic noise exposure in human studies

In this section, conventional measures of temporal envelope processing are briefly explained first. Then, past human studies that have used these conventional measures to examine the relationship between noise exposure and temporal envelope processing are reviewed. In short, human studies have not consistently shown supra-threshold envelope processing issues among people with higher noise exposure. The potential pitfalls of these studies are discussed at the end of the section.

1.3.1 Conventional behavioral and electrophysiological measures for testing temporal envelope processing

Temporal envelope processing reflects the ability to encode amplitude change over time. Conventional behavioral measures assess temporal envelope processing by examining the individual's ability to detect or discriminate amplitude modulation (Moore, 2013). In an amplitude modulation detection task, the listener is presented with a set of at least two tones (or noise bursts), one of which is processed to contain structured amplitude fluctuation that the listener needs to pick out. The amount of amplitude fluctuation is controlled through the depth of modulation. The smallest detectable fluctuation is taken as the amplitude modulation detection threshold. In an amplitude modulation discrimination task, the listener is instructed to discriminate whether pairs of amplitude modulated (AM) tones (or noise bursts) are identical. The smallest detectable amplitude fluctuation to differentiate the two sound tokens is taken as the discrimination threshold.

Electrophysiological measures, such as the EFR, assess temporal envelope processing by quantifying the ability of the brain to phase lock the amplitude modulation. To record the EFR in an animal, electrodes are placed near the auditory units of interest, and the voltage change of the electrodes in response to a stimulus is measured. The stimulus can be an AM tone of which the carrier frequency coincides with the CF of the interested ANF and the envelope is modulated by another lower frequency sinusoid. The strength of phase-locking to the envelope and the frequency of the phase-locking response can be used to reflect the temporal envelope precision of the ANFs with CF identical to the carrier tone (Bharadwaj et al., 2015; Shinn-Cunningham et al., 2017). In human listeners, the EFR is recorded non-invasively at far-field sites, usually with the recording electrodes placed on the scalp, and responses from a wide range of auditory neurons and nuclei are recorded.

The above measures have been utilized in human studies that examined the relationship between noise exposure and temporal envelope processing. The next section reviews the findings from these studies.

1.3.2 Review of human studies about the effect of non-traumatic noise exposure on temporal envelope processing

Kumar et al. (2012) compared temporal processing skills and speech in noise between train drivers with normal audiograms and age-matched individuals without excessive noise exposure. The results showed that the noise-exposed groups, especially between age 30 to 50 years old, performed significantly worse than the age-matched controls on modulation detection at higher modulation frequencies (60 and 200 Hz), on duration pattern test and on speech in noise test, supporting deficiency of temporal envelope processing among individuals with noise exposure but normal hearing thresholds.

Stone et al. (2008) as well as Stone and Moore (2014) examined the amplitude modulation discrimination for people with excessive noise exposure but normal hearing thresholds and found the discrimination only at near-threshold level but not at supra-threshold levels was poorer in the highly exposed group. This finding did support poorer temporal envelope processing in noise-exposed human listeners, but the occurrence of the issue near threshold contradicted the prediction that the deficits should occur well above threshold.

Paul et al. (2017, 2018) examined AM detection and EFR between young adults of 18 to 19 years old with high or low noise exposure history. They presented the AM stimuli in narrow-band noise and the noise was presented at different levels as a way to engage the ANFs with different SRs. Their primary outcome did not consistently support the relationship between noise exposure and the performance of AM detection or EFR strength (Paul et al., 2018).

Füllgrabe et al. (2020) used frequency discrimination limen and AM regularity test to examine the relationship between noise exposure and temporal processing. The carrier frequency of the target stimuli was chosen at 4 kHz as they suspected this frequency to be most susceptible to noise exposure at least based on literature from noise-induced hearing loss. Their finding did not support the relationship between noise exposure and temporal processing among people with normal hearing thresholds from 0.25 to 4 kHz.

A few studies (Stamper & Johnson, 2015a, 2015b; Bramhall et al., 2017) examined the amplitude of ABR wave I among human listeners and found reduced wave I amplitude in individuals with noise exposure but normal audiograms. Their results indirectly supported the selective loss of high-intensity ANFs and the potential existence of temporal envelope processing issues.

Stamper and Johnson (2015a, 2015b) examined the relationship between personal noise exposure history and the pure-tone audiogram, the amplitude of DPOAE, and the amplitude of ABR wave I and wave V, among young adults (age 19 to 28). The results showed a negative relationship between wave I amplitude and noise exposure experience, but the relationship was only held for female subjects, not for male subjects.

Bramhall and her colleagues (2017) measured DPOAEs and ABRs among young veterans with or without significant noise exposure and young civilians with or without a history of firearm use. The four groups were found similar in pure-tone audiograms and DPOAE input/output functions, but the two groups without excessive noise exposure showed greater amplitudes of ABR wave I than the two groups with excessive noise exposure.

Liberman and his colleagues (2016) examined ABR wave I amplitude as well as compressed or reverberated speech in noise between people with and without excessive noise exposure. They found that the group with more noise exposure performed significantly more poorly than the less exposed group on compressed and reverberated word recognition in noise, which may indicate some issues in general temporal processing. However, their study did not find reduced ABR wave I amplitude in the group with more exposure, which seemed to contradict the notion that the loss of high-intensity ANFs was responsible for poorer temporal processing skills.

Prendergast et al. (2017a, 2017b) and Yeend et al. (2017) were the studies that have tested a large number of participants (N > 100) but did not find the relationship between noise exposure and temporal envelope processing. In specific, amplitude modulation detection, EFR and ABR wave I amplitude at soft and high levels were tested to evaluate temporal envelope processing and the loss of high-intensity ANFs. They also tested other types of auditory processing skills, including frequency discrimination, sound localization, speech recognition in noise, and so on. The lack of a strong correlation between noise exposure and various auditory perceptual skills raised some concerns and debates on the existence of such selective ANF loss in human listeners. The inconsistency among the human studies also prompted researchers to reconsider the sensitivity of the conventional measures and the optimal population for studying this issue.

1.3.3 Summary for human studies

The potential causes regarding the lack of consistency in studies of noise exposure and temporal envelope processing among human listeners are discussed below.

The issues of using amplitude modulation detection or discrimination to assess temporal envelope processing are possibly two. First, off-frequency responses brought by the spread of excitation for high-intensity stimuli will recruit the responses of normal units which then overwhelm the responses of abnormal units, causing the testing outcomes to appear normal. (Ruggero et al., 2000). Second, the CAS is exceptional at coding AM sound starting even from the cochlear nucleus, which is the central nuclei that the ANFs connect to (Frisina, Smith & Chamberlain, 1990; Rhode & Greenberg, 1994). For example, octopus neurons in the cochlear nucleus are sensitive to AM sound and receive input from ANFs across a wide range of CFs (Rhode et al., 1983), so the detection of amplitude modulation can still be signaled through auditory units at off-frequency regions. The same argument can be used in the case of EFR. When EFR to AM tones were measured at near-field (i.e. ANF) and far-field (i.e. scalp) sites in animals with NTNE, the response collected through near-field recording showed much poorer performance for the exposed animals than for the controls, but there was no significant group difference through far-field recording (Chen et al., 2019), supporting the potential confounding contribution from off-frequency regions and the CAS.

Furthermore, since the effect of NTNE is frequency-specific (e.g. Kujawa & Liberman, 2009; Lin et al., 2011; Furman et al., 2013; Chen et al., 2019), showing more damage to the auditory units with CFs within or above the frequency band of the noise, it is important to develop a task that can manifest the encoding capacity of only the affected auditory units while keeping the unaffected units away from 'participating' the encoding of the stimulus. Recruiting participants from a wide range of noise exposure profiles, unfortunately, will not help to determine the frequency regions that are most likely affected by NTNE. Therefore, the current study concluded that when studying the effect of NTNE in human listeners, it is important to search for participants who had similar noise exposure profiles; the spectral features of the exposed noise should be as distinctive from environmental noise as possible; the measures should be carefully designed to reflect as much as possible the effect within the noise spectrum while avoiding the confounding contributions of the unaffected auditory units. Having said that, the next section introduces the population of interests and the measure that may meet these criteria.

1.4 Rationales for the population and the measures to assess temporal envelope processing in relation to non-traumatic noise exposure

1.4.1 Introducing the population of interest

The population of interest in the current study is the dental school students. Dental school students are chosen for several reasons. First, the noise exposure at a dental office is spectrally distinctive from the noise exposure at daily environments. The spectra of daily environmental noise usually appear low-pass filtered (LPF) with sound energy concentrating below 2 kHz (Can et al., 2010; Bořil et al., 2012; Rämö et al., 2013; Albert & Decato, 2017). The spectra of the dental noise, on the other hand, usually appear high-pass filtered (HPF) with sound energy concentrating above 4 kHz (Fernandes et al., 2006; Choosong et al., 2011) (Figure 2). This distinction between the environmental and the dental noise spectra allows the frequency-specific measures of envelope processing to be designed.

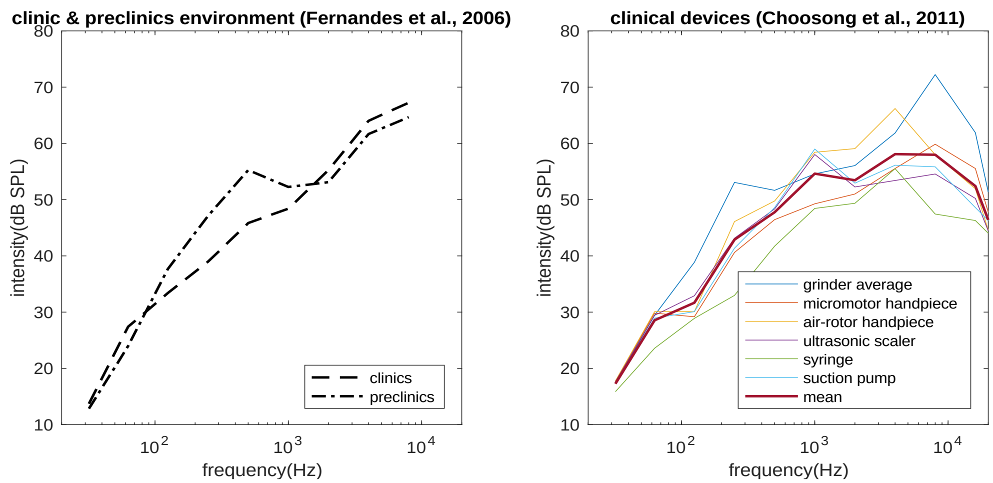


Figure 2. Spectra of the environmental noise at dental clinics and preclinics (replotted from Fernandes et al., 2006) and of the commonly used devices at dental clinics (replotted from Choosong et al., 2001)

Moreover, modern-day dental devices produce sound at non-traumatic levels as the overall levels at dental pre-clinics and clinics normally occur below 85 dB SPL (Chen et al., 2013; Al-Dujaili et al., 2014; Myers et al., 2016; Ai et al., 2017; da Cunha et al., 2017; Ahmed & Ali, 2017) and the levels of most of the handheld devices occur below 80 dB SPL (Bahannan et al., 1993; Nassiri et al., 1993; Setcos et al., 1998; Fernandes et al., 2006; Mojarad et al., 2009; Elmehdi, 2010; Kadanakuppe, et al., 2011; Qsaibati & Ibrahim, 2014; Yousuf et al., 2014; Myers et al., 2016; Goswami et al., 2017). Noise exposure dose measurement also has shown that the 8-hour time-weight average (TWA) level of most dental students ranges from 70 to 80 dB SPL (Choosong et al., 2011; Burk & Neitzel, 2016), which is below legislative standards (85 dBA for 8 hours, NIOSH, 1998).

Therefore, if the amount of environmental noise exposure, as well as other demographic and audiologic factors, can be controlled between young dental professionals and young non-dental listeners, neither having hearing loss, then any observable group difference in an auditory perceptual task may be related to the additional non-traumatic exposure of dental noise.

1.4.2 Introducing the measure

As was mentioned earlier, temporal envelope processing was conventionally assessed through behavioral tasks, like amplitude modulation detection and discrimination, or electrophysiological measures, like EFR, but these measures cannot avoid the confounding contributions from off-frequency auditory units and the CAS. Here, unvoiced speech recognition is introduced as the measure of the current study to assess temporal envelope processing in people with or without NTNE.

1.4.2.1 What is unvoiced speech?

Speech can be decomposed into temporal envelopes and temporal fine structures (TFS) (Rosen, 1992; Moore, 2019). When the TFS is replaced by noise or sinusoids through vocoding techniques, speech remains highly intelligible despite that recognition only depends on the spectro-temporal envelopes of the signal (e.g. Shannon et al., 1995; Dorman et al., 1997; Dorman & Loizou, 1998; Loizou et al., 1999). Unvoiced speech is a type of such envelope-based speech that is entirely noise-excited and preserves spectro-temporal envelopes at high fidelity. An example of unvoiced speech in real life is whispered speech where periodic glottal pulses are replaced by aperiodic noise as the excitation source, but the vocal track and the articulators work as filters that shape the source excitation spectrally and temporally to form various speech tokens (Fant, 1970). Unvoiced speech can also be synthesized through extracting spectro-temporal envelopes from natural utterances and exciting the envelopes with random noise (Kawahara & Irino, 2005; Kawahara et al., 2009).

1.4.2.2 Advantages of unvoiced speech recognition

Speech recognition is an object-based listening task where each frequency region has its unique contribution to intelligibility (ANSI, S3.5-1997). If noise exposure selectively affects envelope processing in a limited frequency region, auditory neurons in other frequency regions will be occupied in processing information within their own CFs and so off-frequency contribution can be greatly reduced, allowing the compromised envelope processing in a specific frequency to emerge. Moreover, the failure to encode information at lower-level auditory centers cannot be compensated by the processing at higher-level auditory centers, for there is no available information to begin with. Consequently, degraded envelope processing in a specific frequency region will lead to poorer recognition of the speech information in that frequency region.

Since the current study was interested in envelope processing in both temporal and spectral domains, the other advantage of unvoiced speech was that the target ability can be easily emphasized toward one of the two types of processing through presenting the unvoiced speech in either temporally or spectrally challenging environments. Specifically, the study presented the speech in maskers that were either temporally or spectrally notched to force the listeners to use the information in the temporal or spectral domains, which is also called 'glimpsing' or 'dip-listening' (Cooke, 2006). If poorer performance is observed in one of the challenged processing domains, it indicates less efficient use of the information in that specific domain.

Lastly, given that unvoiced speech is acoustically redundant and that the noise in a dental office is high frequency centered, the study proposed to high-pass filter and low-pass filter the unvoiced speech so the performance may primarily reflect the processing in and outside the spectral region of the exposed noise, respectively, potentially unraveling the frequency-specific effect of NTNE on envelope processing.

1.4.2.3 Potential challenges of using unvoiced speech recognition

One potential issue for using unvoiced speech to reflect temporal envelope processing is that unvoiced speech in quiet could be insensitive to a broader temporal window. Many studies have shown that modulation rate above 30 Hz contributed little to the intelligibility of natural speech in quiet (Drullman et al., 1994; Hou & Pavlovic, 1994; Arai et al., 1999; Chi et al., 2005; Elliott & Theunissen, 2009; Chait et al., 2015). A study that used noise-vocoded speech with rich spectral channels also has shown that vocoded speech intelligibility decreased drastically only when the modulation below 16 Hz was removed (Van der Horst et al., 1999).

Therefore, the current study should use unvoiced speech recognition in acoustic challenging conditions. As was described in the previous section, one of the tasks was to test unvoiced speech recognition in temporally modulated noise, examining how well the listeners could use the speech information in the temporal gaps where signal-to-noise (SNR) was favorable for speech understanding. Poorer temporal envelope processing, in theory, will lead to an incomplete representation of the signal and, more importantly, unprecise representation of the temporal gaps of the noise, preventing the information in the gaps from being efficiently accessed by the listeners and hence leading to poorer recognition performance in this type of noise. The other task for temporal envelope processing was to present unvoiced speech interrupted by silent gaps at various rates. Based on Jin & Nelson's studies (2006, 2010), the ability to recognize speech that was interrupted by noise or by silence was supported by the same temporal processing mechanism, as the performances of these two tasks were highly correlated among listeners with hearing impairments and listeners with normal hearing. If temporal envelope processing is affected by NTNE, the faithful encoding of the interrupted unvoiced speech envelopes should be negatively impacted, like the unvoiced speech in temporally modulated noise.

1.4.3 Summary of the proposed methodology

The current project proposed to study dental professionals as the population of interest because of the distinctive spectral features and the non-traumatic nature of the dental noise. The study also proposed to assess temporal envelope processing by using unvoiced speech recognition under various acoustic challenging conditions. The stimuli were proposed to be high-pass or low-pass filtered to examine any frequency-specific effect of NTNE on temporal envelope processing. The methodology of the current study should help avoid off-frequency listening and consider central auditory contributions, improving the sensitivity of the task to detect temporal envelope processing issues related to NTNE.

2.0 Experiment 1: Non-Traumatic Noise Exposure on Unvoiced Speech Recognition in Noise

2.1 Introduction

Experiment 1 examines the listeners' ability to recognize unvoiced speech in quiet and in different types of noise. Noise maskers included steady-state unmodulated noise (UN), temporally modulated noises (TMN) with modulation rate at 16 Hz or 32 Hz, and spectrally modulated noise (SMN). All stimuli were HPF and LPF so that the performance could ideally reflect as much as possible the effect of non-traumatic dental noise exposure in and outside the dental noise spectrum, respectively.

Unvoiced speech intelligibility depends on both the spectral and temporal envelope cues (Kawahara & Irino, 2005; Kawahara et al., 2009; Irino et al., 2012). When unvoiced speech is presented in TMN, a listener is forced to weigh the temporal envelope cues more heavily for speech understanding (Feston & Plomp, 1990; George et al., 2006; Jin & Nelson., 2006). For people who have been exposed to non-traumatic noise, the internal representations of the speech and the TMN can be inaccurately encoded at supra-threshold levels, causing the listeners to have lower access to the speech information in the noise temporal gaps and impairing their speech recognition performance. Furthermore, because the temporal gaps at higher modulation rates are briefer and require finer temporal resolution to be accurately encoded, poorer temporal envelope processing should be more susceptible to noise with higher modulation rates than noise with lower modulation rates (Dubno et al., 2003). Therefore, it is additionally hypothesized that the difference between the dental-noise-exposed and the unexposed groups will be more obvious in faster-modulated TMN than in slower-modulated TMN.

The current experiment also used unvoiced speech recognition in SMN as a task to reflect spectral envelope processing after NTNE. Animal studies have argued that the effect of NTNE on the auditory system was not constrained to the temporal domains, such as broadened frequency tuning curve (Zhou et al., 2011; Zhou & Merzenich, 2012; Kamal et al., 2013) and reorganization of cortical tonotopic maps (Pienkowski & Eggermont, 2009; Pienkowski & Eggermont, 2010a, 2010b; Pienkowski et al., 2011, 2013). Additionally, the ANFs with LSR are the primary task force to encode spectral envelopes at high intensities as the ANFs with HSR saturate in firing rates (Sachs & Young, 1979; May et al., 1996; Reiss et al., 2011). With these high-intensity fibers damaged after NTNE, the rate-profile representation of spectral envelope by the auditory fibers will likely appear smoothed out at high intensities. Given that no human studies specifically explored the effect of NTNE on spectral envelope processing, the current experiment explored this issue by comparing unvoiced speech in SMN between the dental-noise-exposed and unexposed groups.

Lastly, the unvoiced speech was presented in quiet and in UN as two baseline conditions. In quiet, unvoiced speech recognition is hypothesized to show no difference between the exposed and the unexposed groups, because speech intelligibility in quiet is usually unharmed without the modulated components above 30 Hz, indicating that unvoiced speech recognition in quiet could still be robust in the presence of poorer temporal resolution. Likewise, it is hypothesized that there will be no difference in unvoiced speech recognition in UN between the two groups because there are no spectral or temporal gaps to assess the speech information under favorable SNRs for speech understanding.

2.2 Methods

2.2.1 Participants

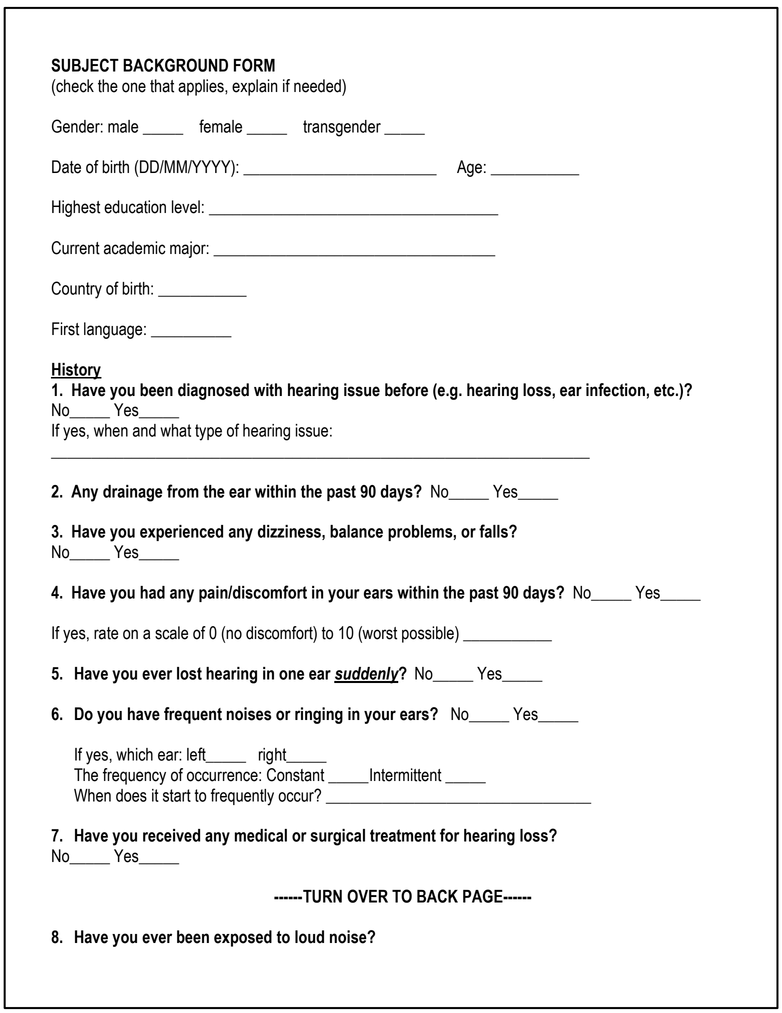
2.2.1.1 Overview of the screening procedures and inclusion criteria

Participants were recruited through being approached in person, in flyers, via emails, and Pitt+Me service (a research participant recruitment program offered through Clinical and Translational Science Institute at the University of Pittsburgh). All participants first received demographic screening by filling out a background form and a questionnaire that surveyed their general noise exposure history. Participants in the experimental group filled out an additional questionnaire that surveyed their dental noise exposure history. Participants who passed the demographic screening then received audiologic screening, which included tympanometry of a 226-Hz tone in each ear, the ipsilateral acoustic reflex from 0.5 to 4 kHz at the left ear, DPOAE from 1 to 8 kHz in each ear and pure-tone audiometry from 0.25 to 8 kHz in each ear. Participants must have met all of the demographic and audiologic criteria to be eligible for the formal experiments.

Section 2.2.1.2 introduces the demographic information form and the inclusion criteria of demographic screening. Sections 2.2.1.3 and 2.2.1.4 introduce the details of the general noise exposure questionnaire and the dental noise exposure questionnaire, respectively. Section 2.2.1.5 introduces the details of the audiologic screening. Section 2.2.1.6 describes the inclusion criteria and the demographics of the participants who passed the screenings and completed the entire experiments.

2.2.1.2 Demographic screening

Demographic information was collected with a subject background form (sample page in Figure 3 and enlarged form in Appendix A). The questions in the form covered issues that would determine the eligibility of the participants demographically, such as the participants' hearing-related health history, past medical history, and so on. Participants who met the criteria 1 to 6 in Table 1 then filled out the general noise exposure questionnaire for determining eligibility for criteria 7 and 8. Those who passed all the demographic inclusion criteria were allowed to proceed to the audiologic screening.



SUBJECT BACKGROUND FORM
(check the one that applies, explain if needed)

Gender: male _____ female _____ transgender _____

Date of birth (DD/MM/YYYY): ________________________ Age: ___________

Highest education level: ____________________________________

Current academic major: ___________________________________

Country of birth: ___________

First language: __________

History
1. Have you been diagnosed with hearing issue before (e.g. hearing loss, ear infection, etc.)?
No_____ Yes_____
If yes, when and what type of hearing issue:
____

2. Any drainage from the ear within the past 90 days? No_____ Yes_____

3. Have you experienced any dizziness, balance problems, or falls?
No_____ Yes_____

4. Have you had any pain/discomfort in your ears within the past 90 days? No_____ Yes_____

If yes, rate on a scale of 0 (no discomfort) to 10 (worst possible) ___________

5. Have you ever lost hearing in one ear *suddenly*? No_____ Yes_____

6. Do you have frequent noises or ringing in your ears? No_____ Yes_____

If yes, which ear: left_____ right_____
The frequency of occurrence: Constant _____Intermittent _____
When does it start to frequently occur? _________________________________

7. Have you received any medical or surgical treatment for hearing loss?
No_____ Yes_____

------TURN OVER TO BACK PAGE------

8. Have you ever been exposed to loud noise?

Figure 3. Sample page of the subject background form (see Appendix A for higher resolution)

Table 1. Inclusion criteria for demographic screening

Category	**Inclusion criteria**
Demographic screening	1. Age: 22 to 30 years old 2. Male and female 3. English as the first language 4. No hearing issues in the past or at the time of screening, including otological disorders, ear infection, otitis media, or hearing loss. 5. No neurological, neurophysiological or neuropsychological disorders, and no brain trauma. 6. Perception of tinnitus allowed for the experimental group only if it occurred since starting dental school 7. No history of frequent impulse noise exposure 8. Occupational noise exposure cannot exceed NIOSH standard (85 dBA for 8 hours, NIOSH, 1998)

2.2.1.3 General noise exposure questionnaire

The questionnaire used for collecting general noise exposure history was revised from the Exposed Noise and Hearing Disorders of Conscripts (ENHDC) by Jokitulppo et al. (2006). The original ENHDC collects information on a wide range of noisy activities that can be used to calculate lifetime noise exposure in equivalent sound exposure level (L_{eq}). The estimate of L_{eq} was used to determine whether the daily noise exposure was non-traumatic.

The revised ENHDC (Figure 4) preserved the overall structure and the noisy activities from the original questionnaire but also added noisy activities mentioned in other more recent noise exposure questionnaires (Henry et al., 2015; Johnson et al., 2017). In specific, 1) ‘attending concerts and festivals or other musical events’ was re-worded into two activities, 'attending concerts and festivals with acoustic system' and 'attending concerts and festivals with amplified system'; 2) ‘listening to a portable stereo’ was deleted; 3) ‘attending or participating outdoor sports events (e.g. football)', 'attending car or truck race', 'riding in private airplanes', 'attending any other events with amplified public music or announcement system' were added to the category of leisure time noise activities (Henry et al., 2015; Johnson et al., 2017); 4) ‘any work involving power tools, such as chainsaws, etc.’ and ‘'any work involving driving heavy equipment or machinery, such as trucks, lawn mowers, leaf blowers, tractors, etc.’ were added to occupational noise exposure (Henry et al., 2015; Johnson et al., 2017). The revised ENHDC was able to quantify noise exposure in the recent one year, three years, and over the participant’s lifetime.

Exposed Noise and Hearing Disorders of Conscripts (revised)

Subject ID: ____________ Date: ____________

The purpose of this questionnaire is to determine your personal noise exposure except for the major-related noise. For example, if you are a dentist, your major-related noise will be the noise from equipment used in your dental practice. If you are a violinist, your major-related noise will be the sound practicing with your violin alone and in a group. Since there is a separate questionnaire for major-related noise exposure, we would like you to exclude the major-related noise exposure when completing this questionnaire.

Please estimate the sound level of your leisure-time activities and non-major occupational noise exposure using a rating of 1 to 5. Please also mark the exposure time per time, how many times a week, how many weeks over the period of last 1 year, last 3 to 2 years, and before 3 years.

To estimate the sound level of each item, you write:

1 to refer to the sound level of normal conversation,
2 to refer to the sound level of loud conversation,
3 to refer to the sound level at which you must shout over a distance of one meter (over table) to be heard,
4 to refer to the normal sound level of pub music that makes you shout to be heard over a near distance,
5 to refer to loud pub music that makes communication almost impossible.

Activity		estimation of loudness (fill in 1 to 5)	hours/week	#weeks in the past 52 weeks (1 year)	#weeks from 3 years ago to 1 year ago (total 104 weeks)	weeks/year x # years from birth till 3 years ago	% time using hearing protection device
Leisure time noise part 1	Listening to a home stereo/radio via loudspeakers						
	Listening to a home stereo/radio via headphones						
	Watching television						
	Playing computer games						
	Watching movies or the theatre						
	Going to nightclubs or pubs						
	attending classical music concerts and festivals						
	attending rock/pop music concerts and festivals						
	Attending motor sports						
	Using tools indoors (name machine)						
	Using tools outdoors (name machine)						
	Attending or participating indoor sports events (e.g. ice-hockeys)						
	Attending or participating outdoor sports events (e.g. football)						
	Exercising to music						
	Attend car/truck races						
	Attend any other events with amplified public announcement (PA)/music systems						
	Ride in or pilot small aircraft/private airplanes						
	Other, name:						
	Other, name:						
Leisure time noise part 2 (music major student, do not answer)	Playing in a band or orchestra						
	Practicing a musical instrument (name instrument: ____________)						
Occupational noise	Any work involving power tools, chainsaws, or other shop tools						
	Any work using drive heavy equipment or loud machinery (such as tractors, trucks, or farming or lawn equipment like mowers/leaf blowers)						
Activity		#shots/session	sessions/week	#weeks in the past 52 weeks (1 year)	#weeks from 3 to 1 year ago (total 104 weeks)	# years before 3 years ago	% time using hearing protection device
Impulse noise	Fireworks						
	Use of firearms, such as rifles, pistols, shotguns, etc. for hunting, sports shooting						

Figure 4. Sample noise exposure history questionnaire (see Appendix B for higher resolution)

To compute the lifetime noise exposure for a participant in the control group, first, the exposure dose for each noisy activity was computed using EQ 1 (NIOSH, 1998; Neitzel et al., 2004):

$$D_i = \frac{C_i}{(8760*N_i)/2^{(L_i-79)/3}} \times 100 \qquad \text{(EQ 1)}$$

Here, *N* is the number of years which the noise exposure needs to be computed out of. *C* is the number of hours participating in that activity during the time specified by number of years (i.e. *N*). *L* is the average sound pressure level of that activity. And *i* represents the ordinal number of a particular noisy activity. To compute noise exposure over the past one year, then *N* was assigned to be 1 and the information from the general noise exposure questionnaire under columns 'estimate of loudness', 'number of hours per week' and 'number of weeks over past 52 weeks (1 year)' was needed (Figure 4). To compute lifetime noise exposure, *N* was assigned 1, 2 and the number that equals to participant's current age subtracted by 3 to represent the total time duration 'in the past one year', 'in the past 3 to 1 years (2 years)' and 'from birth to 3 years ago', respectively. On the other hand, *C* was the actual time spent on each noisy activity, which should be estimated for each of the three stages of life. Then, the dose, D_i, for each activity at each stage of life was computed using EQ 1.

Second, the number of hours participating in other regular activities was calculated by subtracting the total hours of noisy activity from the total hours of lifetime. These activities were pre-determined to occur at 64 dB SPL on average (Johnson et al., 2017). In this way, the dose for regular activity was also computed using EQ 1.

Lastly, the doses of all activities, noisy and regular, were summed and converted into L_{eq} with EQ 2 to quantify lifetime noise exposure in sound pressure level:

$$L_{eq} = \left[10 \times log_{10}\left(\frac{\sum D_i}{100}\right)\right] + 79 \qquad \text{(EQ 2)}$$

2.2.1.4 Dental noise exposure questionnaire

The dental noise exposure questionnaire surveyed the frequency and duration that a participant used the dental devices during their study at the graduate program of dentistry at the University of Pittsburgh (Figure 5). The listed devices covered those from pre-clinics, clinics, and laboratories. The intensity and spectral analyses of these devices have also been conducted to provide a more accurate estimate of the L_{eq} for a dental student.

Dental equipment use questionnaire

Subject ID: ____________ Date: ____________

Category	Equipment	1st year			2nd year			3rd year			4th year		
		# hours/week using this device	# weeks	% time HPD	# hours/week using this device	# weeks	% time HPD	# hours/week using this device	# weeks	% time HPD	# hours/week using this device	# weeks	% time HPD
Student handpieces	Turbine (high speed)												
	Contra-angle hand-piece												
	Straight hand-piece												
Clinic handpieces (not used until 3rd year)	Ultrasonic scaler												
	Turbine												
	Contra-angle hand-piece												
	Straight hand-piece												
Lab-and-clinic equipment	Polishing equipment												
	Vibrating equipment												
	Lathe equipment 3000												
	Stone trimmer												
	Suction pump (low volume)												
	Suction pump (high volume)												
	Air-water syringe												
	Sandblaster												

Figure 5. Sample dental noise exposure questionnaire (see Appendix B for higher resolution)

The way to estimate the L_{eq} for participants of the experimental group followed the general procedures described for estimating general noise exposure L_{eq}. The dose of exposure to dental noise was calculated as a separate source of noisy activity. The total hours and the exposure dose of regular non-noisy activity were re-calculated by subtracting the total hours of dental and non-dental noisy activities from total hours of the individual’s lifetime.

2.2.1.5 Audiologic screening

The audiologic screening included otoscopic exam, tympanometry, acoustic reflex, DPOAE, and pure-tone audiogram. The otoscopic exam was conducted at both ears using a handheld Welch Allyn otoscopy. An eligible participant should show no occlusion of the ear canal and intact eardrums (Table 2).

Tympanogram was tested at both ears with a 226-Hz tone presented through a testing probe from GSI Tympstar Middle Ear Analyzer. Participants were asked to restrain movement, swallowing or conversation during the test. Tester compared the measurements of compliance, middle ear pressure and ear canal volume for each ear with inclusion criteria for tympanogram (Table 2)

Acoustic reflex was also conducted using GSI Tympstar Middle Ear Analyzer. Ipsilateral and contralateral acoustic reflexes were measured at each ear with a probe tone of 0.5, 1, 2, and 4 kHz presented at 95 dB SPL. The responses (in ml) of ipsilateral stimulation at the left ear were used to determine eligibility (Table 2) because the stimuli in the unvoiced speech tests were only presented at the left ear.

DPOAE was measured through Intelligent Hearing System (IHS). The frequency range of the DPOAE spanned from 1 to 8 kHz with 3 frequency points per octave, so F2 frequencies were 1104, 1392, 1753, 2207, 2783, 3506, 4419, 5566, 7012, 8838, 11133, 14028, 17671 Hz. The F2/F1 ratio was 1.22. The presentation levels for L1 and L2 were 65 and 55 dB SPL, respectively. Eligibility was determined based on the SNR from 1 to 8 kHz (Table 2). Responses from 8 to 16 kHz were also recorded but was only used to analyze the effect of dental noise exposure on high-frequency hearing (Chapter 4).

The audiometric screening was also conducted. Pure tones from 0.25 to 8 kHz were generated through the Otosuite software installed on a desktop outside the booth. The desktop was connected to a Madsen Astera 2 Otometrics Audiometry that presented the tones through a pair of ER-3 insert earphones to the participants. The participants sat inside a soundproof booth and pressed a handheld bottom to indicate response to the tone. At each frequency, a tone was presented at 25 or 30 dB HL, decreased by 10 dB if it was heard, or increase by 5 dB if no response was given. Absolute thresholds were not searched in this case, but participants who were able to hear at or below 20 dB HL at each frequency were considered eligible (Table 2). Additionally, absolute thresholds were searched and recorded for tones at 12.5, 14, and 16 kHz at both ears, which were used to analyze the effect of dental noise exposure on high-frequency hearing (Chapter 4). In this case, the lowest intensity where participants gave 2 out of 3 correct responses was recorded as the absolute threshold at that frequency.

Table 2. Inclusion criteria for audiologic screening

Category	Inclusion criteria
Audiologic screening	• Otoscopic exam (both ears): no occlusion, intact ear drum with clear cone of light (Roeser et al., 2007) • Tympanometry to 226-Hz tone (both ears): compliance between 0.3 ml to 1.8 ml, middle ear pressure between -150 daPa to +150 daPa, ear canal volume between 0.6 cc to 2.0 cc (Roeser et al., 2007) • Ipsilateral acoustic reflex to pulsed tones at 0.5, 1, 2 and 4 kHz (left ear): response ≥ 0.02 ml (Roeser et al., 2007) • DPOAE with f2 from 1 to 8 kHz, 3 points/octave (both ears), L1 = 65 dB SPL and L2 = 55 dB SPL: SNR ≥ 6 dB for 80% of the test points from 1 to 8 kHz (Roeser et al., 2007; Hall, 2000) • Pure-tone audiogram at 0.25, 0.5, 1, 2, 3, 4, 6 and 8 kHz (both ears): threshold at each frequency ≤ 20 dB HL (ANSI, 2004)

2.2.1.6 Eligible participants

A total of 86 participants from ages 22 to 30 were screened. Thirty-four out of 86 were excluded: 1) thirty-two failed the demographic screening, 2) two passed the demographic screening but failed the audiologic screening (ear canal completely occluded by ear wax). Eventually, 52 participants passed both the demographic and the audiologic screening and completed the unvoiced speech experiments.

Table 3. Descriptive and inferential statistics of the demographic information for the control and the experimental groups

Variable	Group	Mean ± SD	Range	F-statistics (p-value)
Age	CTL	24.6 ± 2.1	22 - 30	1.847 (0.180)
	EXP	25.3 ± 1.7	23 - 29	
Lifetime L_{eq} (dental and non-dental noise combined)	CTL	75.1 ± 5.6	66.5 - 87.7	8.240 (0.006)
	EXP	78.7 ± 2.7	72.9 - 82.0	
Lifetime L_{eq} (non-dental noise only)	CTL	75.1 ± 5.6	66.5 - 87.7	0.028 (0.867)
	EXP	75.3 ± 4.3	65.2 - 81.7	
Lifetime L_{eq} (dental noise only)	CTL	0	0	13164.561 (0.000)
	EXP	74.6 ± 3.4	67.5 - 80.7	
Years of musical training	CTL	3.4 ± 4.7	0 - 18	0.292 (0.591)
	EXP	2.7 ± 3.8	0 - 11	

Participants in the experimental (EXP) group were twenty-five second to fourth-year graduate students from the School of Dental Medicine at the University of Pittsburgh (female, n = 15) with average age 25.3 years (standard deviation [SD] = 1.7). Five participants were the second-year dental students (female, n = 4), twelve the third-year students (female, n =7), and eight the fourth-year students (female, n = 4). Participants in the control (CTL) group were twenty-seven graduate students and professionals with at least bachelor's degrees (female, n = 25) of average age 24.6 years (SD = 2.1). Table 3 shows the descriptive data of age, lifetime L_{eq} to non-dental noise, to dental noise, to both types of noise combined, and the number of years of musical training. A one-way analysis of variance (ANOVA) was conducted for each outcome variable. The lifetime L_{eq} with dental and non-dental noise combined is significantly higher for the EXP group than for the CTL group by about 3.6 dB, $F(1, 50) = 8.240$, $p = 0.006$. When considering the lifetime L_{eq} with non-dental noise exposure only, there is no significant difference between the two groups, p

> 0.05. There is also no significant difference between the two groups in age or in the years of musical training, $p > 0.05$.

For the EXP group, the L_{eq} of lifetime dental noise exposure systematically and significantly increased with the number of years at dental school, where the L_{eq} of the 2nd, the 3rd, and the 4th year students were 71.2 dB SPL (SD = 1.4), 74.4 dB SPL (SD = 3.2), 77.1 dB SPL (SD = 2.6), respectively. Only 3 out of 25 participants in the EXP reported experience wearing earplugs when they were using the student handpieces. Three different participants in EXP reported experience with ringing in the ear, but the tinnitus occurred intermittently, randomly or at night, and did not seem to relate to the experience with using drills at dental school.

2.2.2 Stimuli

2.2.2.1 Target speech

The source speech stimuli were IEEE sentences (Rothauser, 1969) spoken by an adult female in standard English sampled at 44.1 kHz (spectrum spanning from 0.08 to 12 kHz). The IEEE sentences are advantageous because there are a large number of sentences in the corpus to meet the needs of the current study conditions, and all sentence lists are phonetically balanced, low in contextual cues, and ecologically more valid than words.

The speech stimuli were first processed into unvoiced utterance using TANDEM-STRAIGHT. TANDEM-STRAIGHT is a source-filter vocoder that was written in MATLAB (MathWorks, Natick, MA). In essence, the vocoder generated an unvoiced version of a naturally uttered speech token through extracting the envelopes of the natural utterance and exciting the envelopes with random noise (Kawahara & Irino, 2005; Kawahara et al., 2009). The resulting unvoiced token preserved high spectral richness and sounded like whisper speech. Unvoiced tokens were generated following the instructions from Kawahara et al. (2009) under the default settings of the vocoder.

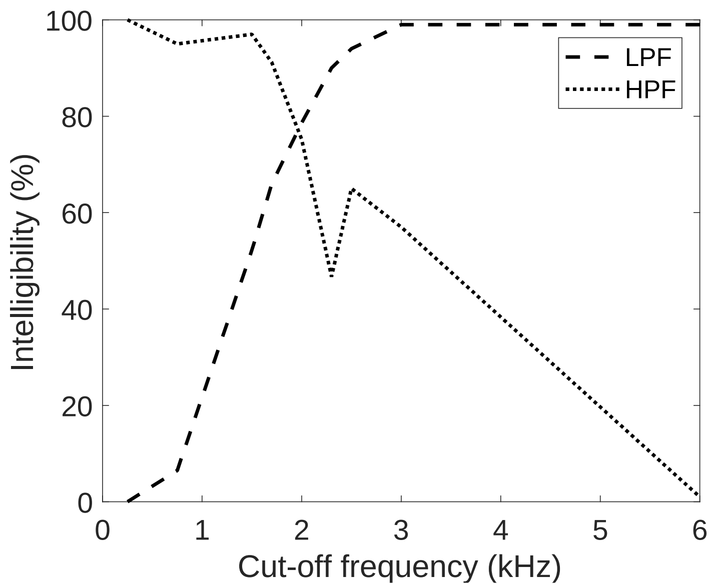


Figure 6. Speech intelligibility as a function of cut-off frequencies of LPF and HPF unvoiced speech (n = 5)

Unvoiced speech tokens were then processed into the LPF and the HPF versions of the speech using a 40th-order Butterworth infinite impulse response (IIR) band-pass filter. Through a pilot study, the cut-off frequencies of the filters were determined to achieve the narrowest passbands where young normal-hearing listeners scored 90% accuracy within twenty IEEE sentences. The passbands were determined to be 0.08 to 2.3 kHz for the LPF speech and 1.7 to 12 kHz for the HPF speech (Figure 6).

2.2.2.2 Noise maskers

There were four types of maskers, UN, 16-Hz TMN, 32-Hz TMN, and SMN. A full-band masker with spectral shape matching the long-term average spectrum of the IEEE sentences was first generated. Next, the masker was LPF or HPF, using the same filter specifications for the filtered unvoiced speech, to produce the filtered unmodulated masker (i.e. the filtered UN, Figure 7).

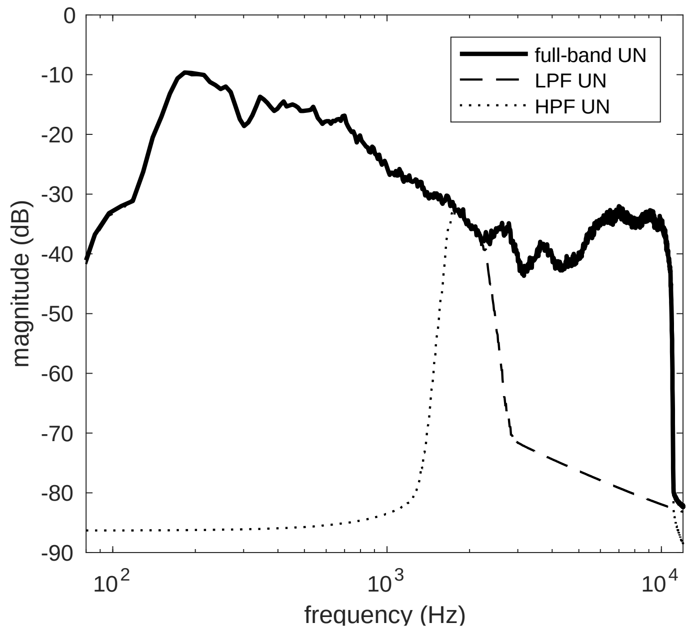


Figure 7. Spectra of the full-band UN (solid line), LPF UN (dashed line) and HPF UN (dotted line)

To generate the filtered TMN, the unfiltered unmodulated speech-shaped noise was multiplied in the time domain by a sine wave with amplitude from 0 to 1 and then was LPF or HPF. The current study used 16-Hz and 32-Hz sine waves to modulate the unmodulated noise, creating the slower-modulated and the faster-modulated TMNs (Figure 8). The modulation rates were determined through a pilot study in which 16-Hz and 32-Hz TMN showed a greater amount of masking release than the TMNs modulated at 1, 2, 4, 8, or 64 Hz. These two modulation rates were also used in Grose et al. (2009) that used natural speech in TMN to compare temporal envelope processing between young and older adults.

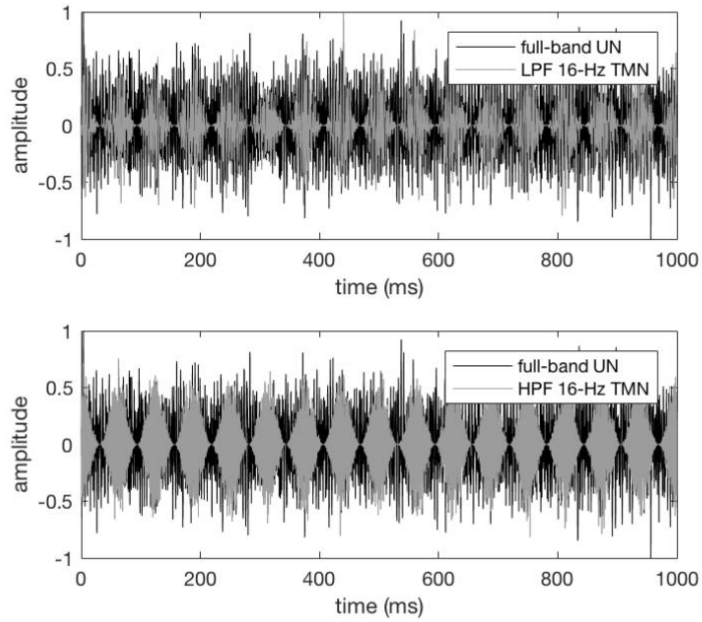


Figure 8. Time-domain waveforms of the full-band UN (black waves in both panels), of LPF 16-Hz TMN (grey wave in the upper panel), and of HPF 16-Hz TMN (grey wave in the lower panel)

Spectrally modulated noise is a type of noise with spectral gaps in the frequency domain. The SMN used in the current study contained spectral gaps that were 3 equivalent rectangular bands (ERB) wide and interleaved with passbands which were also 3 ERBs wide (Figure 9). The spectral gap width was determined based on previous studies (Peters et al., 1998; Hall et al., 2012) and also through a pilot study where 3-ERB gapped SMN produced robust amount of masking release with the lowest variation out of 1-, 2-, 3-, 4-, 6- and 8-ERB gapped SMNs. The SMNs were obtained by passing the UN through a bank of 40_{th}-order Butterworth IIR band-pass filters. The passband cut-off frequencies of the LPF and the HPF SMNs are shown in Table 4.

Table 4. Passband frequencies of the SMNs

Filtering condition	Passband No.	Lower band cut-off frequency (Hz)	Upper band cut-off frequency (Hz)
LPF	Passband 1	80	198
	Passband 2	360	585
	Passband 3	894	1322
	Passband 4	1913	2729
HPF	Passband 1	1913	2729
	Passband 2	3856	5413
	Passband 3	7562	10530

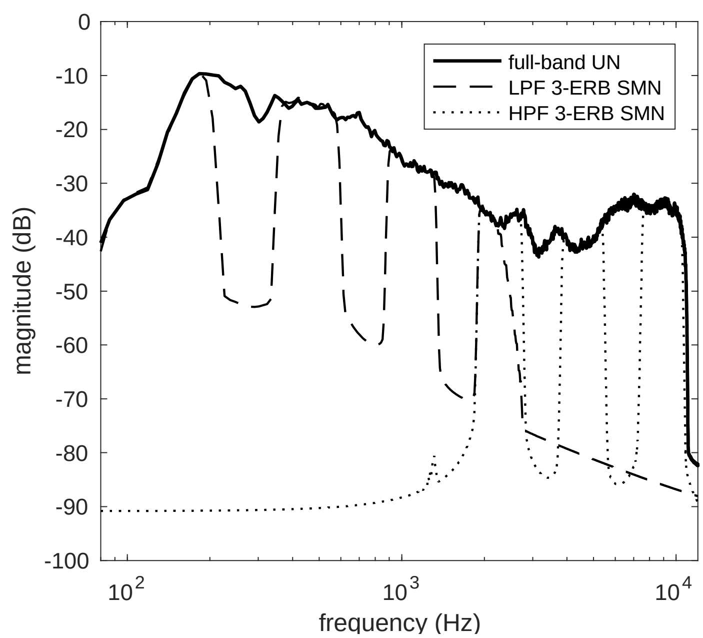


Figure 9. Spectra of the full-band UN, LPF SMN, and HPF SMN

Indeed, ERB, octave, Hz, Q3 (width of the tuning curve at 3 dB above threshold), and Q10 (width of the tuning curve at 10 dB above threshold) are all units to describe frequency selectivity and are interchangeable with each other (Patterson, 1976; Glasberg & Moore, 1990). However, since ERB was designed to express the output of the auditory filter, that was originally described with two back-to-back exponential functions with a rounded top, now in form of a rectangle-shaped filter, the width of the auditory filter can be quantified in a more simplified and straightforward manner than in other units. Although using ERB deliberately ignores any level-dependent change of auditory filter width, the current study only presented the task either in quiet or in noise at a fixed level, so the level-dependent change of auditory filter width is not a concern for the current study.

2.2.3 Procedures

All tasks were conducted in a soundproof booth. Stimuli were controlled through MATLAB scripts on a MacBook Pro, which connected to a pair of AKG K240 MKII supra-aural headphones. All the stimuli of the unvoiced speech perception tasks were presented monaurally to the left ear. Before the testing, 40 filtered unvoiced sentences (20 for each filtering condition) were presented to the participants to repeat to familiarize themselves with the task. Naturally uttered sentences were presented as feedbacks at the end of each repetition. Participants who scored lower than 90% for the first 20 sentences at a given filtering condition were presented with an additional 10 sentences. Participants who did not achieve 80% accuracy even after given the additional sentences were excluded from the study, which the current study has not encountered during data collection.

The familiarity session was followed by unvoiced speech recognition in quiet. The performance was measured in absolute speech recognition threshold (ASRT), the softest sound pressure level where the participant achieved 50% accuracy. The ASRT was measured through a one-down-one-up adaptive procedure (Levitt, 1971). The level of the sentence was elevated if the participant repeated incorrectly and reduced if the participant repeated correctly. The first sentence started at 0 dB SPL where the participant could not perceive the sentence and was presented repeatedly until the participant gave a correct response. The rest of the sentences were presented only once. The step size started with 4 dB and turned 2 dB after two reversals. Correctly repeating 3 or more key words was scored as a correct response and correctly repeating 2 or fewer keywords was scored as an incorrect response. The omission of the ending 's' was counted correct but the omission of 'ed' or replacing phonemes was considered incorrect. Three IEEE lists were used in one block. A total of two blocks were used, one for each filtering condition (LPF, HPF). The sequence of the filtering conditions was randomized. A block stopped: (1) at the end of 20 sentences if the participant achieved at least 10 reversals, (2) at the sentence where 10 reversals were achieved if the participant did not achieve 10 reversals within 20 sentences and did not complete 30 sentences, or (3) at the 30th sentence if the participant did not achieve 10 reversals. If it was the third case, one more block of the same condition was given, and the average of the ASRTs from both blocks under the same condition was taken as the final performance. The ASRT of a given block was taken as the average of the sound pressure levels at all but the first 2 reversals. No feedback was given during the testing. After completing unvoiced speech recognition in quiet, participants first complete either unvoiced speech recognition in noise or silence-interrupted unvoiced speech test (Chapter 3).

Unvoiced speech recognition in noise was measured by speech recognition threshold (SRT), which is the SNR where the participant achieved 50% accuracy. The SRT in noise was determined through a one-down-one-up adaptive procedure. The noise level was fixed at 65 dB SPL and the sentence level was adaptively varied. The SNR of the first sentence was -4 dB. The first sentence was presented repeatedly until the participant gave a correct response. The rule of changing step size, the scoring criteria, and the condition stopping rule of the unvoiced speech in noise tasks were identical to those of the unvoiced speech in quiet. The SRT was taken as the average of the SNRs of all but the first 2 reversals. There were 8 conditions (i.e. 8 blocks) for the unvoiced speech in noise (Table 5) and each condition used 3 IEEE lists. If a participant was not able to achieve 10 reversals by the end of the third list, one more block of the same condition was tested, and the average of the two SRTs from the same condition was taken as the final performance. The sequence of the conditions was randomly determined by a random number generator function in MATLAB. No feedback was given during the testing.

All the data were analyzed in IBM SPSS® Statistics 26.0.

Table 5. Test conditions of unvoiced speech in noise and their abbreviations

Noise	**Filtering conditions**	
	LPF	**HPF**
Unmodulated	LPUN	HPUN
Spectrally modulated	LPSMN	HPSMN
Temporally modulated at 16 Hz	LPTMN16	HPTMN16
Temporally modulated at 32 Hz	LPTMN32	HPTMN32

2.3 Results

2.3.1 Unvoiced speech recognition in quiet

The ASRT measured from unvoiced speech recognition in quiet was analyzed through 2 (group) × 2 (filtering) mixed-model ANOVA. Based on the Shapiro-Wilk test of normality, the performance of LPF speech in quiet of the CTL group violates the normality assumption, $W(27) = 0.842$, $p = 0.001$. Nonetheless, the assumption of normality is loosened in this case for the sample size is large enough (Wickens & Keppel, 2004). The group means and SDs of unvoiced speech in quiet are plotted in Figure 10. There is no significant main effect of filtering, $F(1, 50) = 3.637$, p

= 0.062, η_{p2} = 0.068, of group, $F(1, 50) = 2.757$, $p = 0.103$, η_{p2} = 0.052, or interaction between filtering and group, $F(1) = 0.360$, $p = 0.551$, η_{p2} = 0.007.

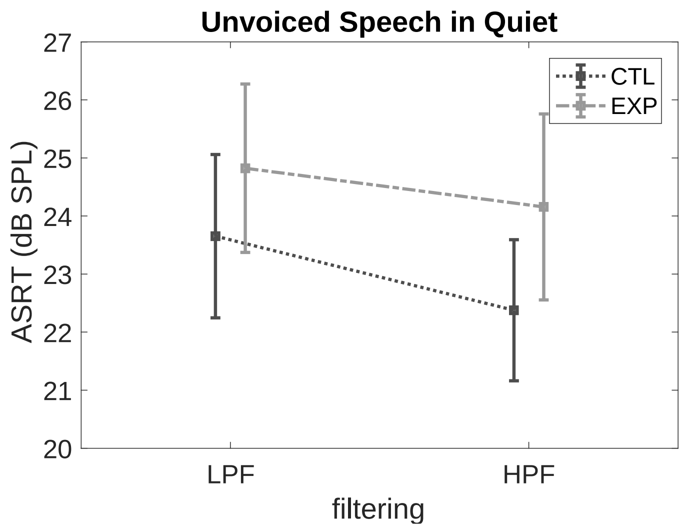


Figure 10. Performance of LPF and HPF unvoiced speech in quiet: dark grey, CTL group; light grey, EXP group (error bars: 95% CI)

Table 6. Simple effect comparisons between the two groups when filtering is controlled

Filtering condition	**Group**	**Mean ± SD (in dB SPL)**	**Mean difference (CTL – EXP)**	**F-statistics (p-value)**
LPF	CTL	23.7 ± 3.7	-1.2	1.286 (0.262)
	EXP	24.8 ± 3.7		
HPF	CTL	22.4 ± 3.2	-1.8	3.069 (0.086)
	EXP	23.2 ± 3.7		

A simple effects analysis reveals no significant difference in ASRT between the two groups when filtering is controlled, $p > 0.05$ (Table 6). There is also no significant difference in ASRT between the two filtering conditions within any given group, $p > 0.05$ (Table 7).

Table 7. Simple effect comparisons between the two filtering conditions when group is controlled

Group	Mean difference (LPF - HPF)	F-statistics (p-value)
CTL	1.3	3.270 (0.077)
EXP	0.7	0.822(0.369)

2.3.2 Unvoiced speech recognition in temporally modulated noise: SRT

The SRT measured from unvoiced speech recognition in noise was analyzed through 2 (group) × 2 (filtering) × 4 (masker) mixed-model ANOVA. Here, speech recognition in SMN is included in the analysis to avoid the inflation of the overall type I error, but the performances under SMN is presented in the next sections (2.3.4 and 2.3.5). Based on the Shapiro-Wilk test of normality, data of unvoiced speech in unmodulated and in temporally modulated noise of both groups satisfy the assumption of normality, $p > 0.05$.

The Mauchly's test of sphericity shows that the assumption of homogeneity of variance is violated for the interaction between the two within-subject factors (filtering × masker), $\chi^2(5) = 13.231$, $p = 0.021$. The ANOVA shows that there is significant main effects of group, $F(1, 50) = 6.584$, $p = 0.013$, $\eta_p^2 = 0.116$, of filtering, $F(1, 50) = 20.292$, $p = 0.000$, $\eta_p^2 = 0.289$, and of masker, $F(1, 50) = 50.889$, $p = 0.000$, $\eta_p^2 = 0.504$. There are also significant interactions between group and masker, $F(3, 150) = 2.741$, $p = 0.045$, $\eta_p^2 = 0.052$, and between filtering and masker after Greenhouse-Geisser correction, $F(2.545, 127.231) = 5.851$, $p = 0.002$, $\eta_p^2 = 0.105$. There is no significant interaction between group and filtering or across the three factors.

To examine whether non-traumatic dental noise exposure affects temporal envelope processing, the first simple effects analysis compares the performance between the two groups when filtering and masker are controlled. As is shown in Figure 11, there is no significant group difference in SRTs of LPF unvoiced speech in UN or in TMNs, $p > 0.05$, though the mean SRTs of the CTL group appear lower than those of the EXP group for 16-Hz (mean difference [MD] = 0.8 dB, standard error [SE] = 0.8 dB) and 32-Hz TMNs (MD = 1.0 dB, SE = 0.6 dB). There is also no significant group difference in SRTs of HPF unvoiced speech in UN or in 16-Hz TMN, $F(1, 50) = 3.481$, $p = 0.068$, $\eta_{p2} = 0.065$, though the mean SRT of the CTL group appears lower than that of the EXP group (MD = 1 dB, SE = 0.6 dB). However, the SRT of the CTL group is significantly lower than that of the EXP group for 32-Hz TMN (MD = 3.1 dB, SE = 0.8 dB), $F(1, 50) = 14.112$, $p = 0.000$, $\eta_{p2} = 0.220$.

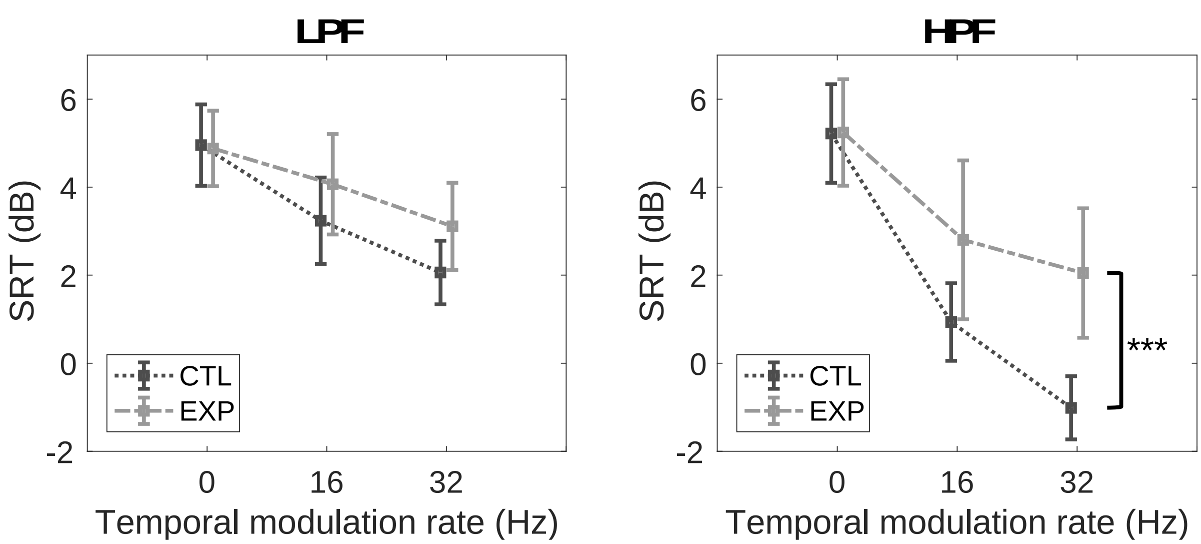


Figure 11. SRTs of unvoiced speech in UN (0Hz), 16-Hz and 32-Hz TMNs under each filtering condition: left panel, LPF; right panel, HPF; dark grey, CTL group; light grey, EXP group. Error bars: 95% CI. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

The second simple effects analysis compares the performance between the two filtering conditions within each group. The pattern of performance when listening to LPF and HPF unvoiced speech in TMN may provide information from another perspective on whether the two groups perceived cues at low- and high-frequency-centered speech differently. As is shown in Figure 12, the CTL group show increasing mean SRT difference between the two filtering conditions with increasing modulation rate of the noise: at 16-Hz, the SRT for the HPF speech is significantly lower than that for the LPF speech by 2.3 dB on average (SE = 0.7 dB), $F(1, 50) = 10.290$, $p = 0.002$, $\eta_{p2} = 0.171$; at 32-Hz, the SRT for the HPF speech is even lower than that for the LPF speech by 3.1 dB on average (SE = 0.7 dB), $F(1, 50) = 19.351$, $p = 0.000$, $\eta_{p2} = 0.279$. The EXP group show a similar trend that the SRTs for the HPF speech appear lower than those for the LPF speech in TMNs, but these mean differences are smaller in magnitude compared to those observed in the CTL group and are not statistically significant for any modulation rate, $p > 0.05$.

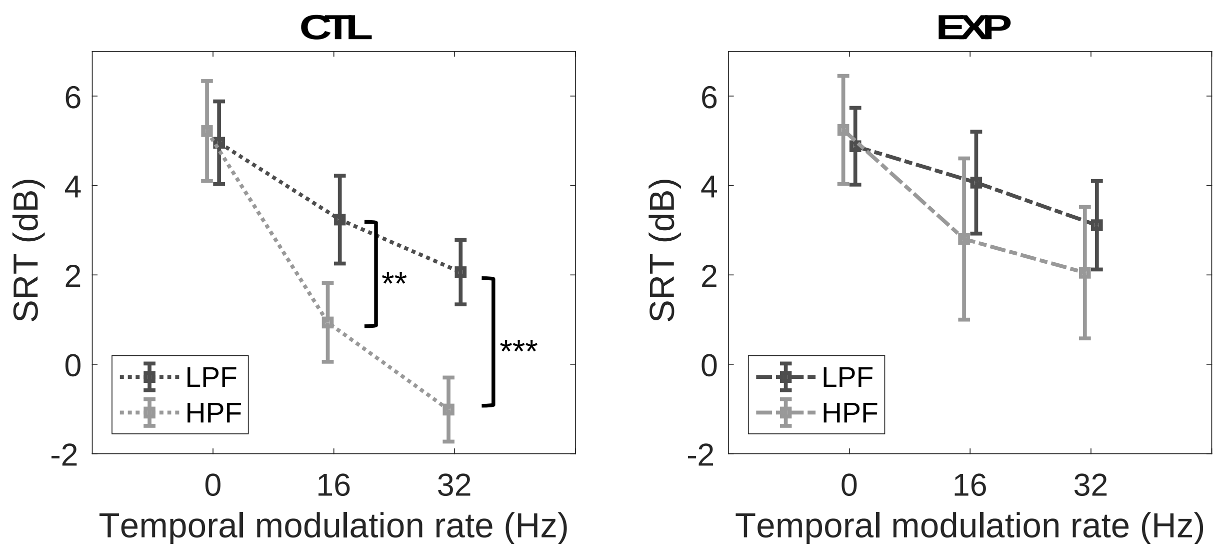


Figure 12. SRTs of unvoiced speech in UN (0Hz) and TMN within each group: left panel, CTL group; right panel, EXP group; drak grey, LPF; light grey, HPF. Error bars: 95% CI. *, *p* < 0.05; **, *p* < 0.01; *, *p* < 0.001.**

2.3.3 Unvoiced speech recognition in temporally modulated noise: masking release

The amount of masking release (MR) was also analyzed as a dependent variable under the effect of noise exposure (group), speech filtering, and temporal modulation rate (masker). Conventionally speaking, MR quantifies the benefit that a listener may take to assist speech understanding by listening for the information in the temporal gaps where SNR is favorable. The amount of MR was computed by subtracting the SRT in the UN by the SRT in the TMNs for each participant.

A 2 (group) × 2 (filtering) × 3 (masker) mixed-model ANOVA was conducted to analyze the MR. Mauchly's test of sphericity shows that the assumption of homogeneity of variance is violated for the interaction between the two within-subject factors (filtering × masker), $\chi_2(2)$ = 12.037, p = 0.002. There are significant main effects of group, $F(1, 50) = 10.988$, $p = 0.002$, η_{p2} = 0.180, of filtering, $F(1, 50) = 19.022$, $p = 0.000$, η_{p2} = 0.276, and of masker, $F(1, 50) = 11.916$, p = 0.000, η_{p2} = 0.192. There are no significant 2-way interactions between any pairs of the three factors nor 3-way interactions among the three factors.

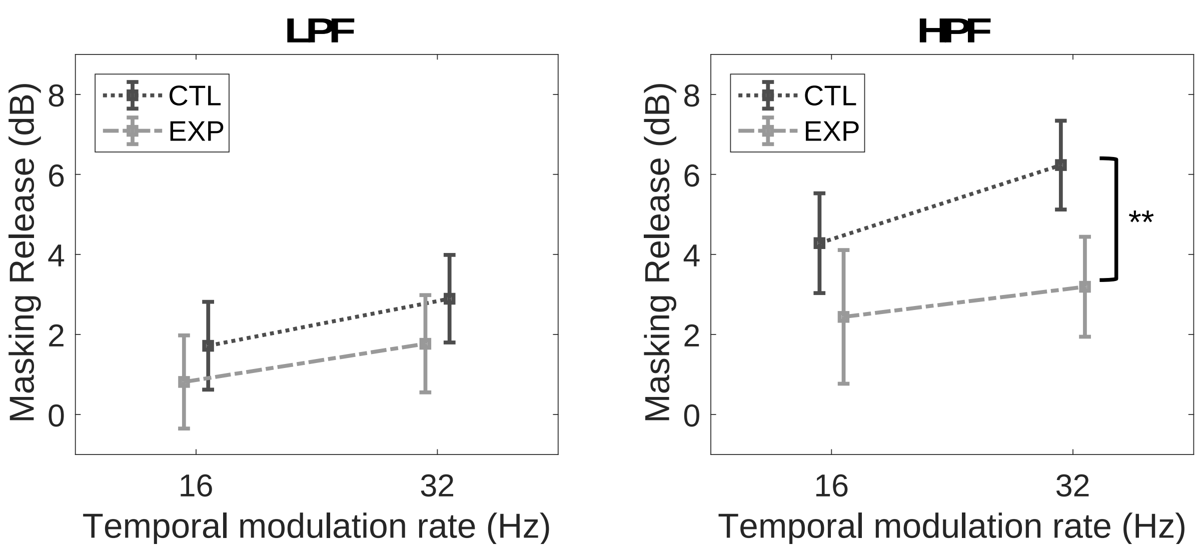


Figure 13. Masking release of unvoiced speech in TMN under each filtering condition. Error bars: 95% CI. *, *p* < 0.05; **, *p* < 0.01; *, *p* < 0.001.**

The simple effects analysis was conducted to examine whether the outcome patterns observed for SRT would occur for MR. First, the performance between the two groups was compared when filtering and masker were controlled. As is shown in Figure 13, there is no significant group difference in MR for LPF speech in 16-Hz or 32-Hz TMNs, $p > 0.05$, though the mean MR of the CTL group appears larger than that of the EXP group for 16-Hz (mean difference [MD] = 0.9 dB, standard error [SE] = 0.8 dB) and 32-Hz TMNs (MD = 1.1 dB, SE = 0.8 dB). There is also no significant group difference in MR for HPF speech in 16-Hz TMN, $F(1, 50) = 3.060$, $p = 0.086$, $\eta_{p2} = 0.058$, though the mean MR of the CTL group appears larger than that of the EXP group (MD = 1.8 dB, SE = 1.1 dB). However, like the pattern observed for SRT, the MR of the CTL group is significantly larger than that of the EXP group in 32-Hz TMN (MD = 3.0 dB, SE = 0.8 dB), $F(1, 50) = 12.790$, $p = 0.001$, $\eta_{p2} = 0.204$.

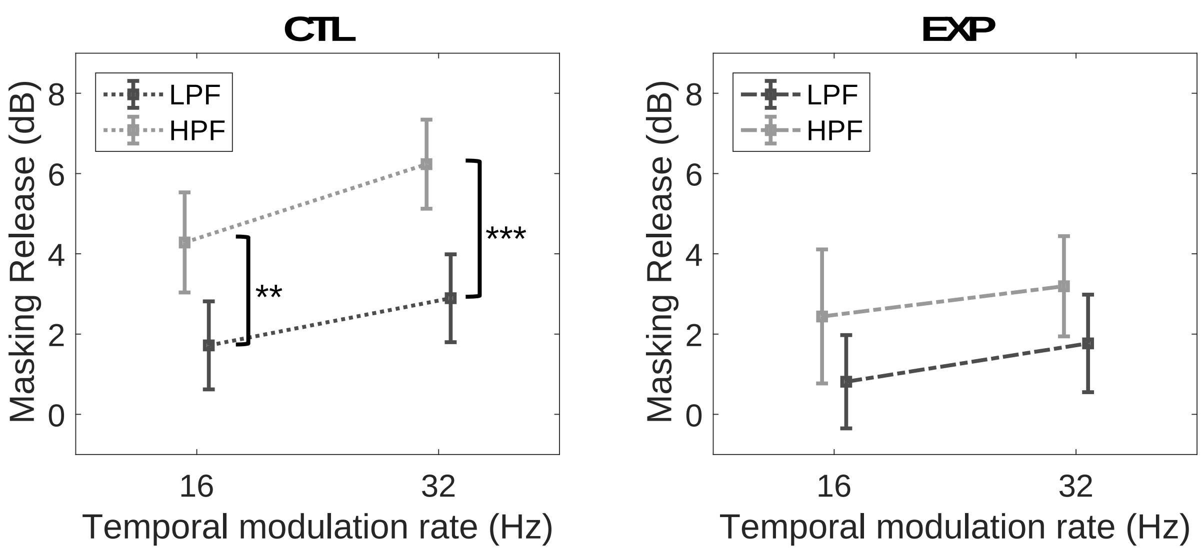


Figure 14. Masking release of unvoiced speech in TMN within group. Error bars: 95% CI. *, *p* < 0.05; **, *p* < 0.01; *, *p* < 0.001.**

Second, the performance between the two filtering conditions within each group was also analyzed for each modulation rate (Figure 14). The CTL group show increasing mean MR difference between the two filtering conditions with increasing modulation rate of the noise: at 16-Hz, the MR for the HPF speech is significantly larger than that for the LPF speech by 2.6 dB on average (SE = 0.7 dB), $F(1, 50) = 7.967$, $p = 0.007$, $\eta_{p2} = 0.137$; at 32-Hz, the MR for the HPF speech is even larger than that for the LPF speech by 3.3 dB on average (SE = 0.7 dB), $F(1, 50) = 16.424$, $p = 0.000$, $\eta_{p2} = 0.247$. The EXP group once more show a similar trend that the MR for the HPF speech appears larger than that for the LPF speech, but the mean MR differences are smaller in magnitude compared to those observed in the CTL group (16-Hz by mean of 1.6 dB; 32-Hz by mean of 1.4 dB) and are not statistically significant for any modulation rate, $p > 0.05$.

2.3.4 Unvoiced speech recognition in spectrally modulated noise: SRT

Performance of unvoiced speech in SMN was analyzed along with the other types of maskers through the 2 (group) × 2 (filtering) × 4 (masker) mixed-model ANOVA in section 2.3.2. The simple effects analysis first shows that the SRT of LPF unvoiced speech in SMN is significantly lower for the CTL group than for the EXP group (MD = 1.5 dB, SE = 0.6 dB), $F(1, 50) = 6.853$, $p = 0.012$, $\eta_{p2} = 0.121$, but there is no significant group difference for the SRT of the HPF speech in SMN (MD = 0.9 dB, SE = 1.1 dB), p > 0.05, despite that the average SRTs of both filtering conditions appear lower for the CTL group than for the EXP group (Figure 15).

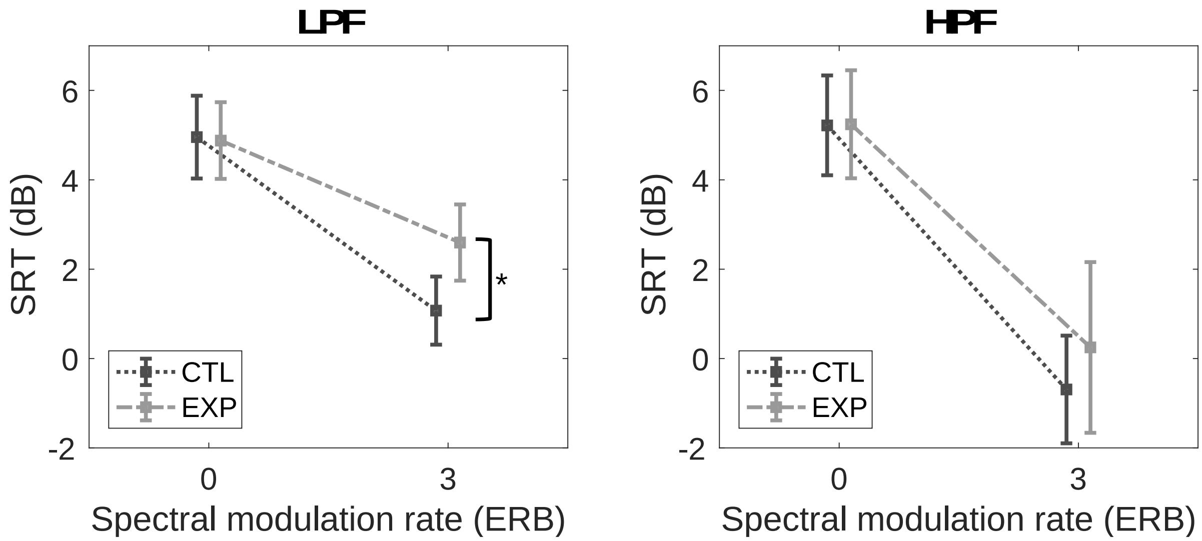


Figure 15. SRTs of unvoiced speech in SMN under each filtering condition. Error bars: 95% CI. *, p < 0.05; **, p < 0.01; ***, p < 0.001.

The simple effects analysis also compared the SRTs for the two filtering conditions within each group. Both groups score significantly lower SRT for the HPF speech than for the LPF speech and the difference between the filtering conditions appear larger for the EXP group than for the CTL group (Figure 16). For the CTL group, the SRT for the LPF condition is higher than that for the HPF conditions by 1.8 dB on average (SE = 0.7 dB), $F(1, 50) = 5.741$, $p = 0.020$, $\eta_{p2} = 0.103$. For the EXP group, the SRT for the LPF condition is higher than that for the HPF conditions by 2.3 dB on average (SE = 0.8 dB), $F(1, 50) = 9.413$, $p = 0.003$, $\eta_{p2} = 0.158$.

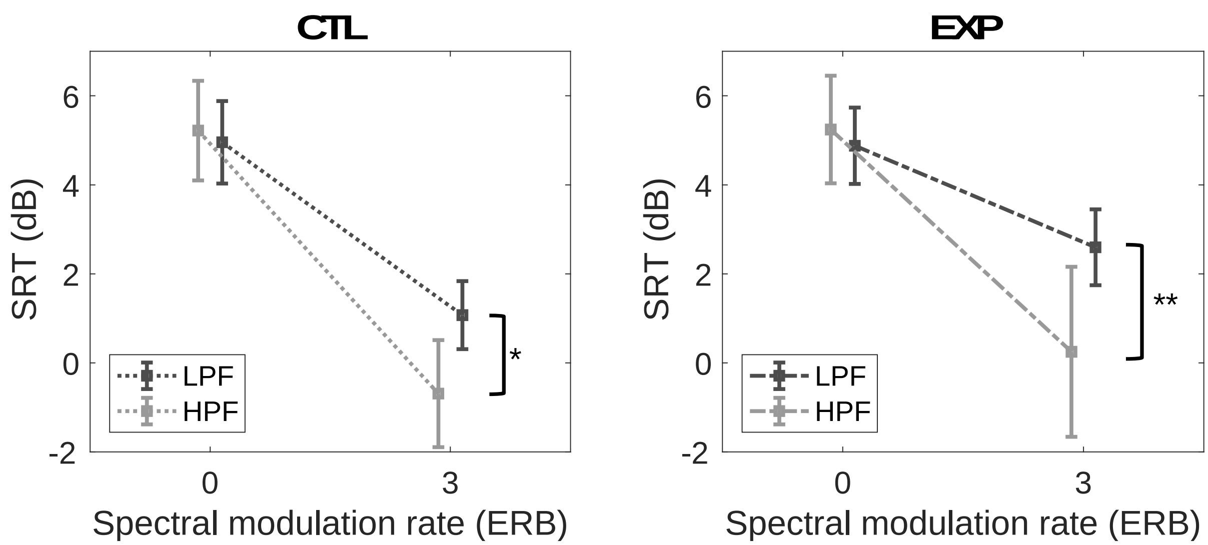


Figure 16. SRTs of unvoiced speech in SMN within each group. Error bars: 95% CI. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

2.3.5 Unvoiced speech recognition in spectrally modulated noise: masking release

The amount of MR was also computed for the unvoiced speech recognition in SMN and analyzed through the 2 (group) × 2 (filtering) × 3 (masker) mixed-model ANOVA in section 2.2.3. The simple effects analysis shows that despite that the amount of MR is on average smaller for the EXP group than for the CTL group under both filtering conditions (LPF, MD = 1.6 dB, SE = 0.8 dB; HPF, MD = 0.9 dB, SE = 1.2 dB), the group difference in MR does not reach statistical significance for the LPF condition, $F(1, 50) = 3.853$, $p = 0.055$, $\eta_{p2} = 0.072$, or for the HPF condition, $F(1, 50) = 0.606$, $p = 0.440$, $\eta_{p2} = 0.012$ (Figure 17, left panel). The mean difference between the two filtering conditions reach statistical significance within both groups and the result patterns are similar. For the CTL group, the MR for the HPF condition is significantly larger than for the LPF condition by mean of 2.0 dB (SE = 1.0 dB), $F(1, 50) = 4.185$, $p = 0.046$, $\eta_{p2} = 0.077$.

For the EXP group, the MR for the HPF condition is significantly larger than for the LPF condition by mean of 2.7 dB (SE = 1.0 dB), $F(1, 50) = 6.935$, $p = 0.007$, $\eta_{p2} = 0.122$.

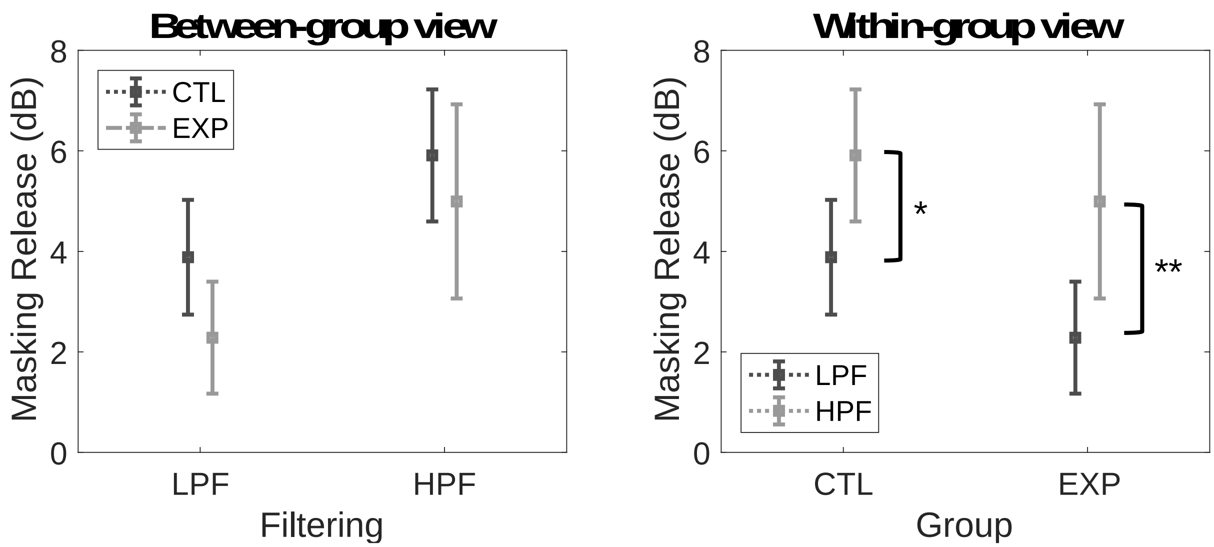


Figure 17. Masking release of unvoiced speech in SMN for clustered filtering conditions (left panel) and clustered group (right panel) (error bars: 95% CI)

2.4 Interim discussion

This chapter examines the relation of NTNE to temporal envelope processing by comparing unvoiced speech recognition performance between a group of dental school students (i.e. the EXP group) and a group of young adults with similar demographic and audiologic profiles but without the dental noise exposure (i.e. the CTL group). The speech stimuli were low-pass or high-pass filtered in an attempt to reveal any frequency-specific effect of high-frequency dental noise exposure on temporal and spectral envelope processing. The results showed that:

(1) NTNE was related to poorer temporal envelope processing, affecting the EXP group in their ability to recognize unvoiced speech when the stimuli were constrained in the dental noise

band and presented in TMN. The negative impact of NTNE on temporal envelope processing appeared more obvious when the modulation rate of the noise was at 32 Hz than at 16 Hz.

(2) NTNE was also related to poorer spectral envelope processing, affecting the EXP group in their ability to recognize unvoiced speech when the stimuli were constrained outside the frequency band of the noise and presented in SMN. This is a novel finding as no previous studies have explored the relationship between NTNE and spectral resolution.

(3) The performances of the EXP and the CTL groups were not significantly different from each other for unvoiced speech recognition in quiet, indicating that the effect of NTNE on envelope processing occurs only at supra-threshold levels, but not at near-threshold levels, which supports the findings of animal studies of NTNE.

(4) There was no group difference in the performance of recognizing LPF unvoiced speech in TMN or HPF unvoiced speech in SMN. This suggests that NTNE spared temporal envelope processing outside the frequency band of the noise and spectral envelope processing inside the frequency band of the noise, but the insensitivity of the measures could not be excluded.

2.4.1 The relationship between non-traumatic noise exposure and temporal envelope processing

Experiment 1 has shown that the EXP group performed significantly more poorly than the CTL group in TMN for HPF conditions. This indicates that people exposed to non-traumatic noise may be less efficient in using the temporal envelope information within the frequency band of the noise exposure than people who were not exposed. Although the SRT and the MR for 16-Hz HPF TMN did not show a statistically significant difference between the groups, the outcome patterns across the two filtering conditions within each group appeared inconsistent between the EXP and the CTL groups. Specifically, the CTL group performed significantly better in the HPF conditions than in the LPF conditions for both 16-Hz and 32-Hz TMNs, but the EXP performed similarly between the two filtering conditions for both TMNs, indicating that the EXP group essentially perform differently from the CTL group in 16-Hz TMN for the HPF condition. The inefficiency of using temporal envelope cues in temporally fluctuated noise was more noticeable when the modulation rate of the noise was higher.

However, this relationship between NTNE and poorer temporal envelope processing is less clear at frequency regions outside the frequency band of the noise. When the unvoiced speech was LPF, the recognition performance of the CTL group was on average better than the EXP group at both noise modulation rates, but the group difference at neither rate reached statistical significance. This non-significant result can be explained by a lack of negative impact of NTNE on temporal envelope processing outside the frequency band of the noise. However, an alternative explanation is that perhaps unvoiced speech at lower frequency does not carry temporal envelope cues that are as important for speech intelligibility. In other words, if there is a group difference in temporal envelope processing outside the frequency band of the noise, LPF unvoiced speech in TMN was not sensitive enough to reveal it.

Let's consider the possibility of the alternative explanation for now. The low-pass filtering process in the current study retained the speech information between 80 to 2300 Hz, which means that a listener could loosely identify the following phonetic features in the LPF unvoiced speech: (1) most of the vowels except the second formant (F2) of /i/, (2) most of the sonorant consonants except /j/, (3) all the stop consonants (/b/, /p/, /d/, /t/, /g/, /k/) followed by vowels other than /i/, (4) the affricatives /tʃ/ (/dʒ/), and (5) most of the fricatives which would be perceived but perhaps not discriminated. All monophthongs and diphthongs except those containing /i/ might be identified because most of the vowel F2 are below 2.5 kHz (Peterson & Barney, 1952; Hillenbrand et al., 1995). Similarly, most of the sonorant consonants might be also identified because they will sound like unvoiced vowels after the unvoicing process and so the possible indiscriminative sonorant might just be /j/ due to its resemblance to /i/. Stop consonants can be identified because their place of articulation is determined by the second formant transitions, which mostly span below 2500 Hz if the consonant is not followed by /i/ (Cooper et al., 1952). The voicing feature of the stop consonants could no longer be signaled by first formant transition due to the unvoicing process but could still be signaled by voice onset time (VOT), which is a temporal envelope cue (Liberman et al., 1958; Rosen, 1992; Benkí, 2001). Affricative /tʃ/ (/dʒ/) may also be identified based on its manner of articulation, which is a temporal envelope cue (Rosen, 1992), and further discrimination between voiceless /tʃ/ and voiced /dʒ/ is based on the duration of frication, which is also a temporal envelope cue after unvoicing (Cole & Cooper, 1975). Fricatives like /ʃ/ (/ʒ/), /f/ (/v/), /θ/ (/ð/) might be *perceived* because their spectral energy could span from 1.5 or 2 kHz up to 8 kHz, but they might not be precisely *identified* because the spectral cues for identification usually reside above 2.3 kHz. The coarse distinction among the obstruent consonants with three manners of articulation (i.e. stops, fricatives and affricatives) might be preserved because the manner of articulation is encoded as temporal envelope cues (Rosen, 1992). Lastly, /s/ (/z/) might be greatly attenuated and unperceived because its spectral energy concentrates above 4 kHz. Based on these acoustic analyses of the phonetics, it does not appear that the intelligibility of LPF unvoiced speech was single-handedly dominated by spectral envelope cues. Instead, both spectral envelope cues and temporal envelope cues contributed significantly to recognizing LPF unvoiced speech. Therefore, it is likely that poorer temporal envelope processing may not necessarily affect LPF unvoiced speech intelligibility in TMN due to the existence of spectral cues.

Based on these concerns, Experiment 2 was conducted in Chapter 3 to explore the theoretical relationship between temporal envelope processing and LPF vocoded speech recognition. Experiment 2 aims to address the extent to which poor temporal envelope processing will degrade the intelligibility of LPF vocoded speech when there are little spectral envelope cues. In short, the results from Experiment 2 show that poor temporal envelope processing negatively impacts both LPF and HPF vocoded speech in TMN but more drastic degradation in performance was observed for the HPF than the LPF conditions. Some past studies have shown that temporal resolution of the auditory system depends on frequency region and that temporal resolution is poorer at lower frequencies than at higher frequencies (Viemeister, 1979; Formby & Muir, 1988; Plack & Moore, 1990; Yost & Sheft, 1997), which might explain the non-significant group difference for the LPF condition. In summary, Experiment 1 concludes that the impact of NTNE on temporal envelope processing outside the frequency band of the noise is minor or negligible compared to that inside the frequency band of the noise.

2.4.2 The relationship between non-traumatic noise exposure and spectral envelope processing

This chapter also explores the relation of NTNE to spectral envelope processing. Spectral envelope processing could, in theory, be at risk if high-intensity ANFs are selectively damaged by NTNE, but no past studies have researched this topic in human subjects. The result showed that the EXP group performed significantly more poorly than the CTL group for unvoiced speech recognition in SMN only in the LPF condition, but not in the HPF condition. This result could suggest that the impact of NTNE on spectral envelope processing might occur only outside the frequency of the exposed noise. But just like the argument concerning temporal envelope processing, an alternative explanation for the result is that the HPF unvoiced speech does not possess the spectral details which require exceptional spectral resolution from the auditory system. In other words, the task of HPF unvoiced speech recognition in SMN is not sensitive enough to pick up spectral envelope processing issues at high frequencies.

Let's examine the alternative explanation from the acoustics of the phonetics once again. The high-pass filtering process retained the speech information between 1.7 to 12 kHz, which means that a listener could roughly identify the following phonetics in the HPF unvoiced speech: (1) almost all the fricative and affricative consonants such as /s/ (/z/), /ʃ/ (/ʒ/), /f/ (/v/), /θ/ (/ð/), /tʃ/ (/dʒ/), (2) the F2 and F3 of vowels like /i/, /I/, /e/ for they have high-frequency F2 (Peterson & Barney, 1952; Hillenbrand et al., 1995), (3) the stops consonants /b/, /p/, /d/, /t/, /g/, /k/ only when they are followed by vowels with high-frequency F2, such as /i/, /I/, /e/ (Delattre et al., 1955). Meanwhile, most of the vowels with F2 at lower frequencies (e.g. /u/, /o/, /a/, /æ/) and the sonorant consonants (e.g. /m/, /l/, /n/, /ŋ/, /j/, /w/) were greatly attenuated by high-pass filtering, leaving the intelligibility of the HPF unvoiced speech mostly dependent on the obstruent consonants and a small number of vowels. Discrimination of the obstruent consonants across different manners (i.e. stops, fricatives and affricatives) should be largely unaffected in the HPF condition because discriminating the manner of articulation depends primarily on temporal envelope cues (Rosen, 1992). Discriminating voiced and voiceless pairs within a given manner could no longer depend on the vibration of the vocal fold, so voicing cue can only depend on VOT, which is also a temporal envelope cue (Rosen, 1992). The place of articulation within stop consonants can be discriminated based on the F2 trajectories but only when the consonants are followed by /i/, /I/, /e/. The place of articulation within fricatives or within affricatives is based on the spectral envelopes and spectral bandwidth (e.g. /s/ spanned from 4 to 8 kHz, /ʃ/ from 2 to 8 kHz, /f/ and /θ/ from 1.5 to 8 kHz). In summary, for HPF unvoiced speech recognition, the intelligibility depends heavily on the accurate encoding of the temporal envelope for the manner of articulation and voicing cues and limitedly on the spectral envelope cues when identifying the place of articulation for stop consonants followed by /i/, /I/, /e/, or for fricatives. Furthermore, the spectral resolution to identify the place of articulation for fricatives is not so demanding, because most of the fricatives and affricatives possess a broadband frequency response. It is possible that the only occasions that need fine spectral envelope coding for the HPF condition are to identify stop consonants when they were followed by /i/, /I/, /e/. Based on these analyses, it is suggested that the intelligibility of HPF unvoiced speech is single-handedly dominated by temporal envelope cues, which is different from the situation of LPF unvoiced speech. This discrepancy between the LPF and HPF speech where the relative contributions of temporal and spectral envelope cues to intelligibility are different may also explain why the SRT for the HPF speech generally appeared much better than for the LPF speech in TMN. For now, it is concluded that even if there is an issue in spectral envelope processing at high frequencies after dental noise exposure, the task of HPF unvoiced speech recognition in SMN will be inconclusive to reveal the issue.

With that being concluded, it is worth thinking whether the integrity of spectral resolution at high frequencies is necessary to human listeners. The speech cues at higher frequency regions do not require high spectral resolution and spectral resolution of the auditory system is usually poor at high frequency in the first place, which may have naturally led humans to produce sound with little requirement of fine spectral resolution at high frequencies. Nonetheless, if future study wishes to examine spectral resolution at high frequency, conventional methods that assess auditory filter width or frequency tuning curve will be more appropriate than HPF unvoiced speech recognition in SMN.

2.4.3 Summary of this experiment

The current experiment shows that the EXP group who have been exposed over time to dental noise performed more poorly than those without the dental noise exposure on speech recognition that requires the use of supra-threshold temporal or spectral envelope cues. The result suggests that there is a relationship between NTNE and poorer supra-threshold temporal envelope processing, supporting the general assumption of the animal models of NTNE. The noise-induced temporal envelope processing issue occurred more noticeably within the frequency band of the noise and was more obvious when the task had a higher reliance on temporal resolution. The lack of group differences for unvoiced speech recognition in the LPF condition may suggest that the negative impact of NTNE on temporal envelope processing is frequency-specific but may also result from the insensitivity of the LPF unvoiced speech recognition in TMN to poorer temporal envelope processing. Therefore, Experiment 2 was conducted in Chapter 3 to examine whether the LPF and HPF unvoiced speech recognition in TMN was equally sensitive to reflect poorer temporal resolution. Furthermore, in Experiment 3, silence-interrupted unvoiced speech test, which was suggested to share the same mechanism as speech interrupted by noise (Jin & Nelson, 2010), was used to examine the effect of dental noise exposure on temporal envelope processing. The experiment provided some additional pieces of evidence on the effect of NTNE on temporal envelope processing and on the frequency specificity of NTNE.

The result of the current experiment also suggests that there is a relationship between NTNE and poorer supra-threshold spectral envelope processing, which was based on an extension from the general hypothesis of the animal model and was a novel finding in human studies. The noise-induced spectral envelope processing issue occurs in frequency region outside the frequency band of the noise, but not within the frequency band of the noise. Based on linguistic accounts, the absent effect of NTNE within the frequency band of the noise may likely result from the insensitivity of the measure at high frequency.

3.0 Experiment 2: Sensitivity of low-pass and high-pass filtered unvoiced speech recognition to broadened temporal integration window

3.1 Introduction

Experiment 1 showed that poorer temporal envelope processing was related to NTNE as young listeners with dental noise exposure performed significantly worse on unvoiced speech recognition in TMN than those without the exposure when the test stimuli were constrained within the dental noise band. However, since there was no group difference in speech recognition in TMN outside the dental noise band, it was unclear whether the non-significant difference resulted from lack of noticeable impact of NTNE on temporal envelope processing outside the frequency band of the noise, or from the insensitivity of LPF unvoiced speech task to poorer temporal envelope processing. To attempt to clarify this issue, the current experiment whether speech perception with primarily temporal envelope cues would be sensitive to degraded temporal envelope processing and whether the sensitivity was similar at LPF and HPF conditions.

Temporal envelope processing was manipulated by paradigmatically smoothing the temporal envelopes of the speech using temporal integration windows of different lengths after the unvoiced speech was passed through a 4-channel vocoder. The vocoding process greatly reduced the spectral resolution of the signal, so intelligibility was predominantly reliant upon temporal envelope cues. The stimuli were LPF or HPF and presented in quiet or in TMN. 32-Hz TMN was used in this simulation study because briefer temporal gaps that result from the higher modulation rates are more readily filled in by the smoothing process.

3.2 Methods

3.2.1 Participants

Participants were ten newly recruited female young adults whose mean age was 21.9 years (range 21 to 23 years) and were tested to have audiometric thresholds ≤ 20 dB HL from 0.25 to 8 kHz. All participants were undergraduate students from the University of Pittsburgh. Because there were not enough sentence stimuli for each participant to complete all conditions in Experiment 2, five of them completed HPF conditions, and the rest completed LPF conditions.

3.2.2 Stimuli

The speech stimuli in Experiment 2 were LPF or HPF unvoiced sentences with smoothed temporal envelopes. Unvoiced IEEE sentences from Experiment 1 were first passed through a four-channel vocoder (40th-order Butterworth IIR filters) with cut-off frequencies listed in Table 8. The temporal envelope of the signal in each channel was then extracted by half-wave rectification and low-pass filtering (40th-order Butterworth IIR filter with cut-off frequency at 256 Hz). MATLAB function `filtfilt` was used to avoid any phase delay caused by low-pass filtering. The TFS of the signal in each channel was also extracted through Hilbert transform for later speech synthesis. The extracted envelopes were convolved with a Hamming window to create smoother envelopes. The size of the Hamming window was 128, 64, 48, 32, 16, 8 or 4 Hz. Note that the shape of the temporal window is not perfectly symmetric in time domain (Moore et al., 1988). Hamming window was chosen for simplicity reason and it was chosen over rectangular window because the stop-band frequency response of Hamming window is relatively lower

(Oppenheim & Schafer, 2010), meaning that the envelope would be less contaminated by unwanted modulations from the stopband. The smoothed envelopes were used to modulate the TFSs extracted earlier from the same band. The modulated signals were filtered again with the corresponding band-pass filters in the first step and were summed to produce the filtered temporally smoothed unvoiced speech. The maskers were LPF or HPF 32-Hz SAM noises that were temporally smoothed in the same way as the speech stimuli.

In total, there were 8 smoothing conditions under each filtering: one unsmoothed condition and seven smoothed conditions (128-Hz, 64-Hz, 48-Hz, 32-Hz, 16-Hz, 8-Hz and 4-Hz windows). The unsmoothed version was the filtered unvoiced speech used in Experiment 1 without the channel-vocoding process, which had the richest spectral and temporal envelope details among all the versions. The smoothed version by a 128-Hz window should retain the rich temporal envelope details but lose the spectral resolution.

Table 8. Cut-off frequencies of the lower and the upper sides of the band-pass filters for LPF and HPF unvoiced speech

Filtering condition	Channel No.	Lower cut-off	Upper cut-off
LPF	1	80	305
	2	305	669
	3	669	1270
	4	1270	2300
HPF	1	1700	2652
	2	2652	4178
	3	4178	6812
	4	6812	12000

3.2.3 Procedures

Experiment 2 contained two tasks. Before the first task, participants were randomly assigned to one of the two filtering conditions (LPF or HPF) and completed a familiarization session with 20 filtered unvoiced IEEE sentences. Then, they completed the first task, which was filtered unvoiced speech recognition in quiet. Participants repeated to 80 filtered unvoiced IEEE sentences where every 10 sentences (50 keywords) were presented for a smoothing condition (8 conditions in total). For the unsmoothed condition, target stimuli were filtered unvoiced speech without the vocoding process. All conditions in quiet were randomized. Sentences were presented at 65 dB SPL. The scoring criteria followed those described in Experiment 1. The percent correct was computed as the outcome variable for unvoiced speech recognition in quiet.

The second task examined the effect of temporal envelope smoothing on unvoiced speech recognition in 32-Hz SAM noise. Again, there were eight temporal smoothing conditions (unsmoothed, 128-, 64-, 48-, 32-, 16-, 8- and 4-Hz windows). All conditions were randomized. The testing procedures followed those described in Experiment 1, except that the highest presentation level of the sentence was set at 91 dB SPL (i.e. 16 dB SNR) to prevent any potential damage or discomfort in hearing. Like Experiment 1, SRT was used as the outcome measure.

3.3 Results

3.3.1 Unvoiced speech recognition in quiet as a function of the temporal window

Despite small sample size, the Shapiro-Wilk test shows that the data from Experiment 2 follow a normal distribution, $p > 0.05$. A 2 (filtering) by 8 (window) mixed-model ANOVA was performed on the intelligibility of temporally smoothed unvoiced speech in quiet (Figure 18). The results show that there are no main effects of window or filtering, $p > 0.05$, and simple effects multiple comparisons show no significant difference in performance between any of the two window sizes at any given filtering condition, $p > 0.05$. There is also no significant difference in performance between the two filtering conditions at any given window size, $p > 0.05$.

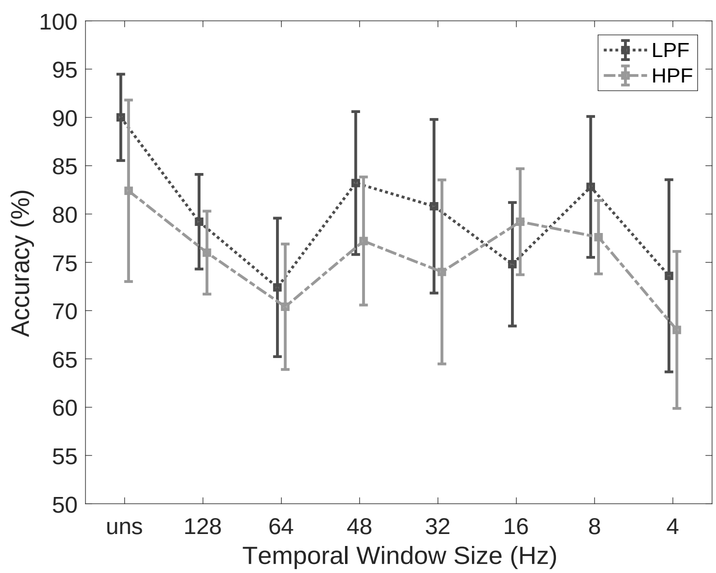


Figure 18. Speech intelligibility of temporally smoothed LPF (dark grey) and HPF (light grey) unvoiced speech in quiet as a function of temporal integration window. Error bars: 95% CI. Uns: unsmoothed.

3.3.2 Unvoiced speech recognition in 32-Hz TMN as a function of the temporal window

A 2 (filtering) by 8 (window) mixed-model ANOVA was conducted for the SRT of temporally smoothed unvoiced speech in 32-Hz SAM noise. The results show that there are significant main effects of window, $F(7, 56) = 71.300$, $p = 0.000$, $\eta_{p2} = 0.899$, of filtering, $F(1, 8) = 10.209$, $p = 0.013$, $\eta_{p2} = 0.561$, and significant interaction between window and filtering, $F(7) = 2.919$, $p = 0.011$, $\eta_{p2} = 0.267$ (Figure 19).

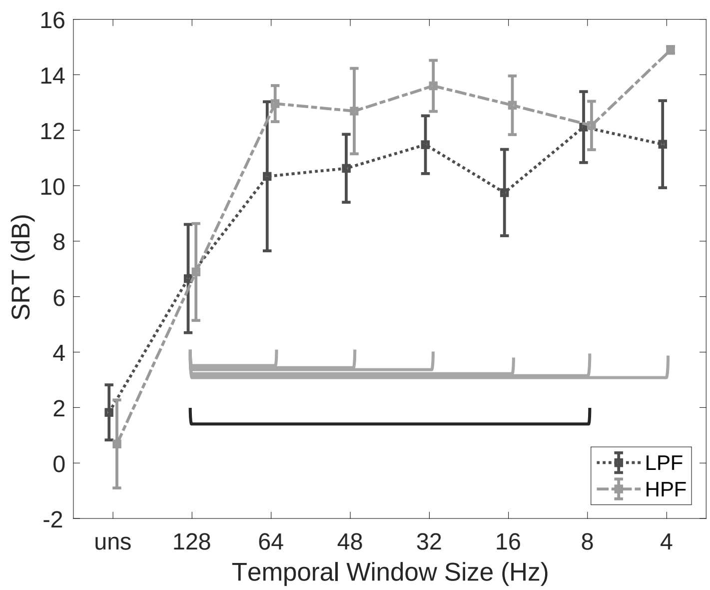


Figure 19. SRT of temporally smoothed LPF (dark grey) and HPF (light grey) unvoiced speech in 32-Hz SAM noise as a function of temporal window. The horizontal bracket indicates a significant difference (*p* < 0.05) between the 128-Hz condition and another given smoothed condition. Error bars: 95% CI. Uns: unsmoothed.

The simple effects multiple comparisons show that there is an initial increase of SRTs from unsmoothed condition to 128-Hz window for both filtering conditions (LPF: MD = 4.8 dB, $p = 0.025$; HPF: MD = 6.2 dB, $p = 0.005$). The mean SRT continues to increase for both filtering conditions from 128-Hz to 64-Hz window. The difference of SRTs between 128-Hz and 64-Hz windows is significant for HPF speech (MD = 6.1 dB, $p = 0.042$), but not for the LPF speech recognition probably due to large standard deviation and small sample size (MD = 3.7 dB, p = 0.583). For the HPF condition, the SRT for the 128-Hz window is significantly lower than the SRTs for all the other broader windows. For the LPF condition, the SRT for the 128-Hz window is only significantly lower than for the 8-Hz window (MD = 5.5 dB, p = 0.037), though the mean SRT for 128-Hz appears lower compared to all the other broader windows. When comparing the SRTs between the filtering conditions at a given window size, the SRTs of LPF speech on average appear lower than those of HPF speech and significant difference is found for 32 Hz (MD = 2.1 dB, p = 0.017), 16 Hz (MD = 3.2 dB, p = 0.011) and 4Hz (MD = 3.4 dB, p = 0.003) windows.

3.4 Interim discussion

Based on the discussion of Experiment 1, it was concluded that the intelligibility of HPF unvoiced speech depended predominantly on temporal envelope cues while the intelligibility of LPF unvoiced speech was determined by both spectral and temporal envelope cues. The reduced proportion of temporal contribution to LPF speech recognition (in contrast to the HPF condition) was interpreted as LPF unvoiced speech recognition being less sensitive to poorer temporal envelope processing, so the lack of group difference for the LPF speech in TMN cannot fully exclude that NTNE could still affect temporal processing outside the frequency band of the noise. This experiment, therefore, was conducted to examine whether LPF unvoiced speech was as sensitive to poorer temporal envelope processing as HPF speech. To do this, spectral details were greatly reduced, and intelligibility of both filtered speeches relied only on temporal envelopes.

It was first observed that, for unvoiced speech recognition in quiet, increasing temporal window did not lead to a monotonic decrease of speech intelligibility for either filtering conditions, suggesting that even when intelligibility of vocoded speech is contributed by only temporal envelope cues, speech in quiet is not sensitive to poorer temporal envelope processing, which explains why people with complaints of hearing in noise could show no signs of hearing loss or degraded intelligibility in an acoustically amiable environment. This is consistent with many studies which have examined the effect of temporal window size on natural or vocoded speech recognition and found that modulation rate above 20 to 30 Hz contributes negligibly to speech intelligibility in quiet (Drullman et al., 1994; Hou & Pavlovic, 1994; Arai et al., 1999; Chi et al., 2005; Elliott & Theunissen, 2009; Chait et al., 2015). For example, Hou and Pavlovic (1994) studied temporal window size from 130 to 33 Hz on natural speech intelligibility and found that speech intelligibility decreased by around 10% when temporal window increased from 130-Hz to 60-Hz but did not continue to decrease with increasing temporal window size, which resembled the finding from the current study. Souza and Rosen (2009) examined the effect of temporal resolution on vocoded speech recognition in quiet and no significant difference was found in the performance of recognizing speech with temporal envelope filtered below 30 or 300 Hz. In short, the current results suggest that neither LPF nor HPF vocoded unvoiced speech recognition is capable of reflecting poorer temporal envelope processing if it is presented in quiet.

For vocoded unvoiced speech recognition in 32-Hz TMN, it was observed that at first when spectral resolution was degraded (i.e. at 128-Hz condition), the SNR needed to achieve 50% of intelligibility was similar for the two filtering conditions (Figure 18). But once the window size doubled (i.e. 64 Hz), which is poorer than temporal resolution related to the tone in gap test at low frequency (i.e. 13 ms or 77 Hz based on the tone in gap test; Plack & Moore, 1990), listeners showed poorer performance on average for both filtering conditions, but the increase of SRT was greater for the HPF condition than the LPF condition. There was a trend that the SRT for the LPF condition increased with increasing size of the window, but the mean SRTs of the LPF condition always appeared better than those of the HPF. In other words, when speech intelligibility of both filtering conditions depends only on temporal envelopes, the magnitude of performance degradation in response to a broadened temporal window is larger for the HPF than for the LPF speech. This suggests that LPF unvoiced speech recognition in 32-Hz TMN is not as sensitive as the HPF condition to demonstrate poorer temporal envelope processing. However, it also does not exclude that LPF unvoiced speech recognition in 32-Hz TMN cannot reflect poor temporal processing at all, because when the window size increased from 128 Hz to 64 Hz which is beyond the normal low-frequency temporal resolution, performance of the SRT of LPF unvoiced speech recognition did get worse. Therefore, the results from Experiment 1 and 2 thus far suggest that NTNE affects temporal envelope processing within the frequency band of the noise more severely than outside the frequency band of the noise, but it is still an open question how severe the effect is outside the frequency band of the noise: is there a minor effect or simply no effect? In Experiment 3, temporal envelope processing after non-traumatic dental noise exposure was examined with an experiment using the perception of unvoiced speech interrupted by silent gaps.

4.0 Experiment 3: Non-Traumatic Noise Exposure on Silence-Interrupted Unvoiced Speech Recognition

4.1 Introduction

As it is still unclear through unvoiced speech recognition in noise whether NTNE affects temporal envelope processing outside the exposed noise band, the current experiment used silence-interrupted unvoiced speech recognition to assess temporal envelope processing.

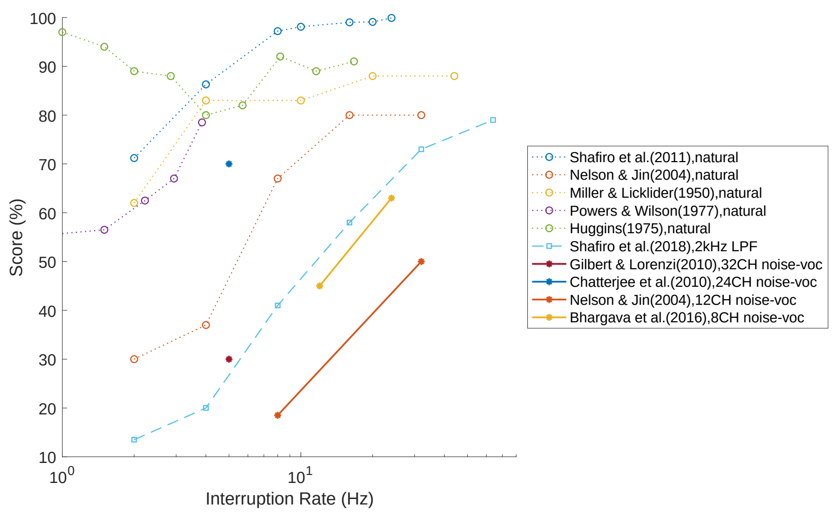


Figure 20. Intelligibility of temporally interrupted speech in quiet as a function of interruption rate from past studies. Dotted line: natural speech. Dashed line: filtered natural speech. Solid line: noise-vocoded speech.

Studies on the intelligibility of interrupted natural speech began several decades ago (Miller & Licklider, 1950) and have been expanding since then (Huggins, 1975; Powers & Wilcox, 1977; Bashford et al., 1988; Bashford et al., 1992; Bashford et al., 1996; Nelson & Jin, 2004; Jin & Nelson, 2010; Wang & Humes, 2010; Shafiro et al., 2011; Bhargava & Baskent, 2012; Shafiro et al., 2018). It has been consistently observed that speech intelligibility increases with increasing interruption rate (Figure 20, solid lines) and that when the gaps are filled with noise, the speech intelligibility improved (e.g. Miller & Licklider, 1950; Bashford et al., 1996). Previously, it was argued that recognizing speech in TMN depends on the listener to make use of the information in the temporal gaps where the SNR was favorable for speech understanding, also called ‘glimpsing’ (‘dip-listening) (Cooke, 2006). Here, recognizing interrupted speech is similar to recognizing speech in TMN, as the listener glimpses pieces of speech segments to make sense of the speech. Therefore, it was proposed that recognizing silence-interrupted speech and speech in TMN was supported by the same temporal processing mechanism. This hypothesis was supported by a paper examining silence and noise interrupted natural speech (Jin and Nelson, 2010). Jin and Nelson (2010) examined the performances of these two tasks of people with normal hearing and with hearing impairments and found a strong correlation ($r = 0.8$ to 0.9) between the two tasks when they examined the group with hearing impairments alone and when the two groups were combined. Despite the fact that there are no published reports directly comparing silence and noise interrupted unvoiced speech recognition, the current study considered the silence interrupted unvoiced speech test a viable alternative of the noise interrupted unvoiced speech test to reflect temporal envelope processing. It was expected that NTNE would affect temporal envelope processing within the frequency band of the noise, resulting in a larger group difference for the HPF condition than for the LPF condition, just as Experiment 1. Meanwhile, whether there would be a group difference for the LPF condition is an exploratory question here.

4.2 Methods

4.2.1 Participants

Participants from Experiment 1 also participated in this experiment (see section 2.2.1). In short, the EXP group were 25 second-to fourth-year dental students and the CTL groups were 27 young adults without dental noise exposure. Participants in neither group showed signs of hearing impairments or other auditory-related health issues.

4.2.2 Stimuli

Filtered unvoiced IEEE sentences from Experiment 1 were temporally interrupted by silent gaps at interruption rates 12, 24, or 48 Hz. The shape of the gating was a square wave. Two pilot studies were conducted to examine intelligibility of speech interrupted at rate 2, 4, 8, 12, 16, 24, 32, 48, and 64 Hz. The pilot results showed that when interruption rates were lower than 10 Hz, the intelligibility was at floor level, potentially because entire syllables or phonemes were masked by silent gaps (Houtgast & Steeneken, 1985; Rosen, 1992). At rates above 16 Hz, the interruptions occur within phonemic boundaries and the listener has access to multiple 'looks' of the phoneme (Houtgast & Steeneken, 1985; Poeppel, 2003) and hence the intelligibility improved monotonically (Figure 21). At 64 Hz, the intelligibility dropped again as listeners started to perceive buzz-like sound brought by the 64-Hz interruption which greatly interfered with the completion of the task.

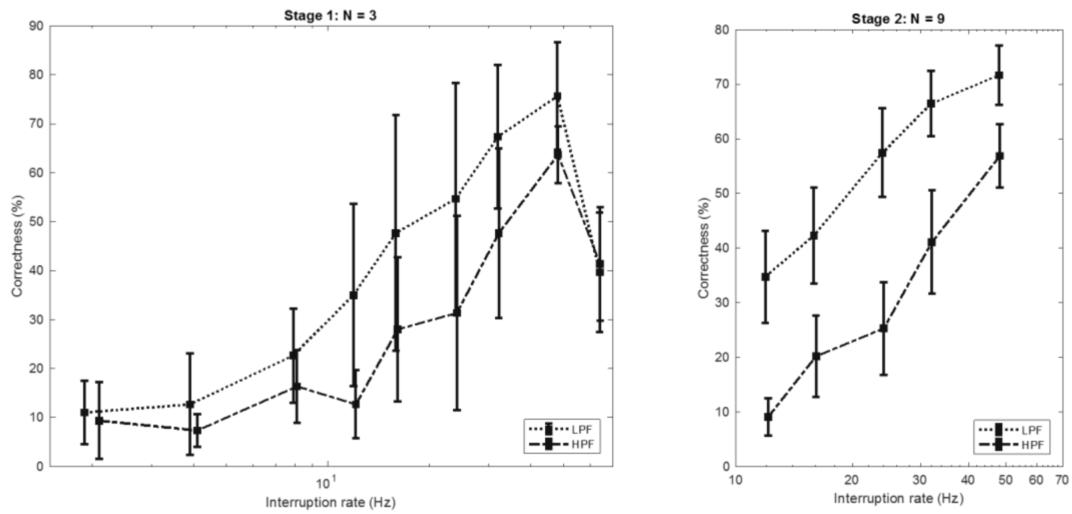


Figure 21. Pilot study on the interrupted unvoiced speech intelligibility as a function of rate. Left panel: results from 3 listeners at stage 1. Right panel: results from an additional 6 listeners at stage 2.

Therefore, 12, 24, and 48 Hz were chosen as the interruption rates. The processing parameters used in the current study are given in Table 9. In short, periodic square wave at rate 12, 24 or 48 Hz was used to interrupt the filtered unvoiced sentence with a 50% duty cycle. Each square wave contained 5-ms raised-cosine ramp to reduce the onset transient sound. Figure 22 shows an example of an interrupted sentence at an interruption rate of 12 Hz.

Table 9. Parameters of the silence-interrupted unvoiced speech

Parameters	**Values**
Duty cycle	50%
Interruption rate (Hz)	12, 24, 48
Window shape	Square wave
Ramp of each window	5-ms raised cosine
Gating depth	100%

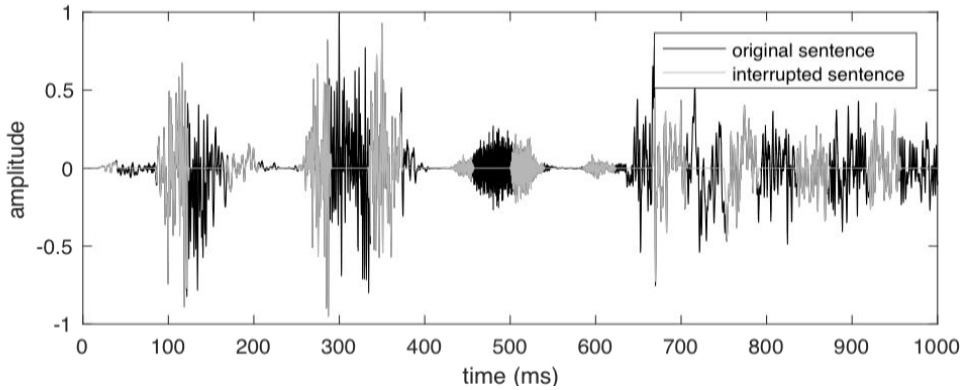


Figure 22. Time domain waveforms of the original sentence and the silence-interrupted sentence (interruption rate 12 Hz).

4.2.3 Procedures

There were six conditions for the interrupted unvoiced speech task (Table 10). The order of the conditions was randomized for every participant. Twenty-five sentences were presented in each condition with no feedback. The first five sentences of each condition were presented to get the participant ready for the condition and the intelligibility was computed based on the last 20 sentences of each condition. Sentence presentation was 62 dB SPL to compensate for the loss of half the speech information.

Table 10. Test conditions of interrupted unvoiced speech their abbreviations

Interruption rate	**Filtering conditions**	
	LPF	**HPF**
12	LP12	HP12
24	LP24	HP24
48	LP48	HP48

4.3 Results

Performance of silence-interrupted unvoiced speech recognition in noise was analyzed through 2 (group) × 2 (filtering) × 3 (rate) mixed-model ANOVA. Mauchly's test of sphericity shows that the assumption of homogeneity of variance is satisfied for rate and interaction between speech and rate. The ANOVA shows that there are significant main effects of filtering, $F(1, 50) = 497.995$, $p = 0.000$, $\eta_{p2} = 0.909$, and of rate, $F(1, 50) = 400.271$, $p = 0.000$, $\eta_{p2} = 0.889$. There are also significant interactions between group and rate, $F(2, 100) = 3.188$, $p = 0.045$, $\eta_{p2} = 0.060$, and between filtering and rate, $F(2,100) = 9.954$, $p = 0.000$, $\eta_{p2} = 0.116$.

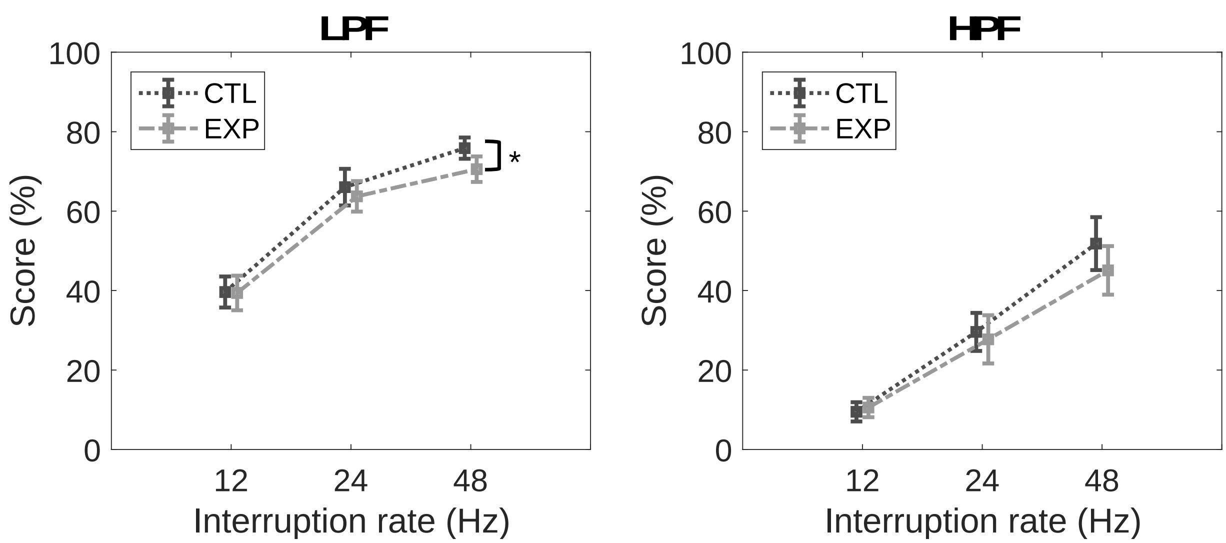


Figure 23. Recognition accuracy (%) of silence-interruped unvoiced speech under each filtering condition. Error bars: 95% CI. *, *p* < 0.05; **, *p* < 0.01; *, *p* < 0.001.**

The simple effects analysis first examines the performance between the two groups when filtering and rate are controlled. As is shown in Figure 23, there is no significant group difference in recognition accuracy for either filtering condition when the interruption rates are lower (i.e. 12 Hz and 24 Hz). At 48-Hz, the scores of the CTL group are on average higher than the EXP group at both filtering conditions (LPF: MD = 5.3%, SE = 2.1%; HPF: MD = 6.7%, SE = 4.6%), but statistical significance is only found for the LPF condition, $F(1, 50) = 6.243$, $p = 0.016$, $\eta_{p2} = 0.111$. Meanwhile, it is also shown that the performance for the LPF conditions is consistently better than for the HPF conditions within both groups (Figure 24).

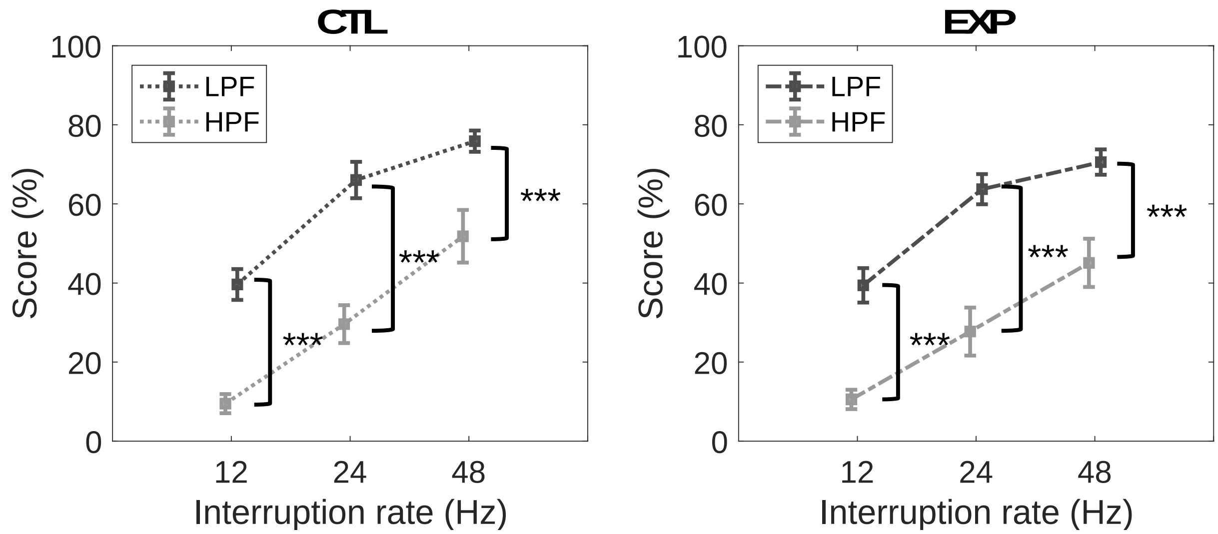


Figure 24. Recognition accuracy (%) of silence-interruped unvoiced speech under within each group. Error bars: 95% CI. *, *p* < 0.05; **, *p* < 0.01; *, *p* < 0.001.**

4.4 Interim discussion

The current experiment shows that when recognizing unvoiced speech that was temporally interrupted by silent gaps, the EXP group who were exposed to dental noise scored slightly lower on average than the CTL group who were unexposed when the interruption rate is high. This pattern was found within the frequency band of the noise without statistical significance but also outside the frequency band of the noise with statistical significance. The magnitude of score difference, which is 5 to 6% at the highest rate, does not seem large or clinically relevant. For this reason, the number of participants was counted for each accuracy level (Figure 24).

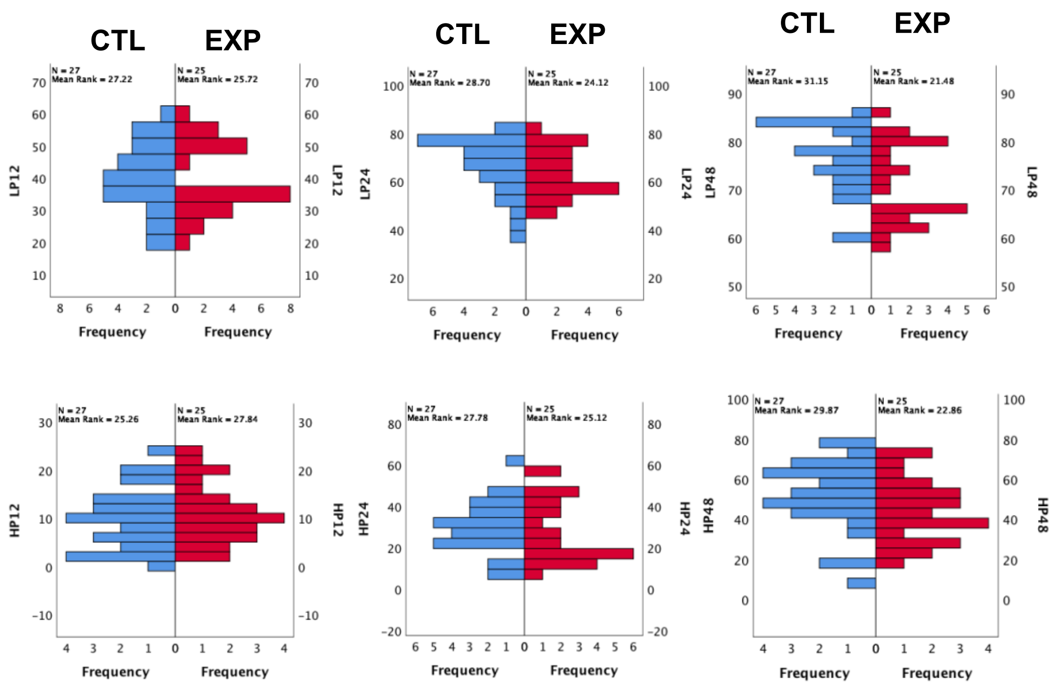


Figure 25. Headcount of the participants at various scores. Top panels: LPF. Bottom panels: HPF. From left to right: 12, 24, and 48 Hz. Within each panel, blue: CTL group; red: EXP group.

Through visual inspection, it can be observed in Figure 25 (right panels) that for the 48-Hz condition, there are more participants in the CTL group who score high in accuracy than in the EXP group, which may have contributed to the better mean scores in the CTL group. However, for the HPF 48-Hz condition, there are a few participants from both groups who score a lot worse than the high scorers within their own group. One participant from the CTL group score only 8% for the HPF 48-Hz condition. In other words, for some reasons, a few participants performed near-floor for the HPF condition even when the interruption gaps were relatively small, and these floor performances significantly increased the individual variability within the group and rendered the between-group difference statistically non-significant.

So, does the non-significant group difference in the HPF 48-Hz condition undo the group difference observed for the HPF speech in 32-Hz TMN (HPTMN32)? Not so much. First, the experiment does show that despite large individual variability, the mean difference between the two groups grew with increasing interruption rate. This suggests that under the conditions which require finer temporal resolution, the EXP group are less capable of keeping up with the CTL group, just like the findings from Experiment 1 that the group difference grows more obvious when the modulation rate of the TMN increases from 16-Hz to 32-Hz. Second, Pearson correlation shows that the performance of HPTMN32 is significantly correlated with the performance of interrupted speech at 48-Hz for both the LPF (LP48), $r(50) = -0.355$, $p = 0.010$, and the HPF conditions (HP48), $r(50) = -0.382$, $p = 0.005$, meaning that faster silence-interrupted speech shares the temporal processing mechanism as the faster noise-interrupted speech to some extent. But the correlation coefficients are not as high as those observed in Jin and Nelson (2010), suggesting that some other factors may account for the large variations for HP48.

The large variations for HP48 might result from the high difficulty level of the HPF interrupted speech compared to the LPF condition. As was discussed in earlier chapters, HPF unvoiced speech recognition predominantly depends on temporal envelope cues while LPF unvoiced speech recognition could depend on both temporal and spectral envelope cues. Interrupting the HPF and the LPF unvoiced speech with the same rate and duty cycle might disrupt the HPF speech recognition more severely, which is inferred from the results that listeners performed more poorly for the HPF than for the LPF conditions at all interruption rates. It may be more ideal in future studies to only use 48-Hz as the interruption rate but with various duty cycles.

In addition to the task being too difficult, the similar top-down abilities between the two groups might have contributed to the non-different HP48 performances. Many previous studies have found that interrupted speech recognition not only depended on the listeners' ability to hear the remained phonemes, but also on the top-down linguistic knowledge and context cues to make sense from the remained phonemes (Bashford et al. 1992; Sivonen et al. 2006; Grossberg and

Kazerounian 2011). Meanwhile, this study has purposefully avoided top-down influences by using the target sentences with low context cues and by recruiting the participants with comparable educational backgrounds. If there were top-down influences, the two groups would not have shown similar performance in some tasks (e.g. HPUN) and different in others (e.g. HPTMN32). Therefore, for a task that requires both bottom-up and top-down processes, the lack of a group difference in top-down abilities may have ameliorated the difference in the bottom-up abilities.

In summary, Experiment 3 in this chapter shows a subtle difference in temporal envelope processing between the EXP and the CTL groups in the LPF condition. Through the silence-interrupted unvoiced speech test, the group difference is more obvious at faster interruption rates than slower rates, echoing the findings from unvoiced speech recognition in TMN. Combining the findings from this and the previous two experiments, there is more confidence to suggest that 1) poorer temporal envelope processing is related to NTNE, 2) the effect of NTNE is relatively small outside the exposed noise band compared to inside the frequency band of the noise, and 3) the negative impact of NTNE on temporal envelope processing is more obvious when the task required finer temporal resolution. With these conclusions established, the next chapter aims to address whether NTNE is first related to degradation in peripheral auditory function or audibility, which might act as some intermediate factors to contribute to poorer temporal envelope processing.

5.0 Additional analyses: Contributions of Demographic and Audiologic Factors to Temporal and Spectral Envelope Processing

Since it has been established in the previous chapters that there was an effect of NTNE on temporal and spectral envelope processing under certain conditions, this chapter aimed to address some follow-up research questions by examining the effect of NTNE on peripheral auditory function, on hearing sensitivity as well as examining the relative contributions of demographic and audiologic factors to temporal and spectral envelope processing.

The first follow-up analysis aims to address whether NTNE could negatively impact peripheral auditory function at standard and extended high frequencies (EHF) (section 5.1), and if yes, whether these defected peripheral auditory functions could meaningfully explain the poorer temporal envelope processing discovered in the previous experiments (i.e. HPTMN16, HPTMN32). The reason to explore EHF listening was that there were some animal studies using NTNE discovering that OHC loss did occur post NTNE but only at EHFs, almost 2 octaves above the exposed noise band (Liberman et al., 2015; Fernandez et al., 2020). There was also a study in human listeners by Prendergast et al., (2017) that showed significantly poorer EHF hearing threshold among listeners with high noise exposure than those with low exposure. As auditory functions at EHFs are usually not tested in audiologic clinics, the potential damage from NTNE to EHF regions may also be considered as an indicator of noise-related hearing issues that have been hidden or largely neglected in routine audiologic exams. It should be noted that animal models have only suggested the NTNE-induced hidden hearing issues to originate from retro-cochlear sites and to occur at supra-threshold levels. Therefore, it is hypothesized that there would be no clinically meaningful difference between the EXP and the CTL groups on the outcomes of peripheral auditory screening at standard audiometric frequencies (e.g. acoustic reflex amplitude from 0.5 to 4 kHz, DPOAE amplitude from 1 to 8 kHz).

Second, the previous chapters showed that high-frequency NTNE could negatively impact spectral envelope processing outside the frequency band of the noise, but the magnitude of impact (MD = 1.8 dB) was not as large as for temporal envelope processing within the frequency band of the noise (MD = 3.1 dB). Additionally, the high-frequency NTNE was found to have a negligibly small effect on temporal envelope processing outside the frequency band of the noise. Such reduced effects of NTNE outside the frequency band of the noise could be related to the acoustic nature of the LPF unvoiced speech. However, since LPF unvoiced speech was constrained below 2.3 kHz, which is around the cut-off frequency of the environmental noise, it was also questioned whether the lack of significant group difference in low frequencies was due to the large individual differences within a group and/or the similar between-group environmental noise exposure profile. For this reason, **the second follow-up analysis (section 5.2) examines whether demographic factors, such as non-dental noise exposure history, and peripheral audiologic function may explain the variation of performances in spectral and temporal envelope processing outside the frequency band of the noise.**

5.1 Effect of NTNE on peripheral auditory functions and extended high frequency listening

In this section, data from audiologic screenings are first reduced in dimensions so that the outcome variables are relatively independent from each other (section 5.1.1). Then, the effect of NTNE (i.e. EXP vs CTL groups) on these audiologic variables is analyzed (section 5.1.2).

5.1.1 Raw data cleaning

5.1.1.1 Raw data

In Experiment 1, participants' audiologic exam outcomes were recorded during the audiologic screening to determine their eligibility for the speech experiments. Here, the outcomes from these audiologic screening were used as the raw data.

Table 11. Audiologic data used as outcome variables

Test category	**Outcome variables and test frequencies**
Middle ear muscle reflex	Amplitude of left ipsilateral acoustic reflex (ml) to pulsed tones at 0.5, 1, 2 and 4 kHz
Inner ear OHC functions	SNR (dB) for left-ear DPOAE at f2 frequencies 0.55, 0.7, 0.88, 1.1, 1.4, 1.7, 2.2, 2.8, 3.5, 4.4, 5.5, 7, 8.8, 11.1, 14, 17 kHz. L1 = 65 dB SPL, L2 = 55 dB SPL.
High frequency sensitivity	Left-ear pure-tone absolute threshold (dB HL) at 12.5, 14 and 16 kHz

5.1.1.2 Dimension reduction

There were 16 frequency points for the raw data of DPOAE. If the data were directly used for ANOVA to compare the group difference, it could increase the number of false negatives. Therefore, before ANOVA was conducted, the dimension of the DPOAE data was reduced to a clinically explainable level.

Figure 26 shows the correlation coefficient matrix of the different levels of acoustic reflex (axis labels starting with 'AR'), DPOAE (axis labels with only digits), and EHF pure-tone average (axis labels starting with 'A'). The outcomes at different frequencies of DPOAE are relatively strongly correlated to outcomes at adjacent frequencies. Additionally, the acoustic reflex amplitudes at 0.5 to 1 kHz are moderately correlated with DPOAE amplitude at 1.7 and 2.2 kHz. If these factors are entered into a multiple regression model without dimension reduction, it will inevitably cause severe collinearity issues. Therefore, exploratory factor analysis was conducted to provide some rationales on which test and what test frequencies could be grouped together.

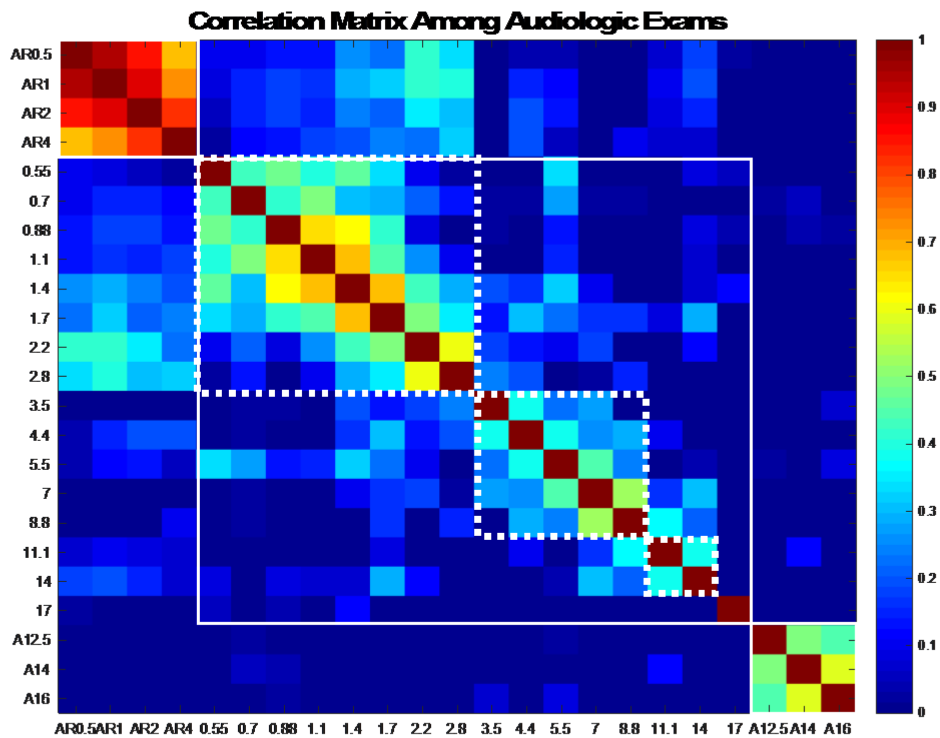


Figure 26. Matrix of correlation coefficients among the outcomes of the audiologic screenings. Axis labels are in the form of 'test + test frequency (kHz)'. AR: acoustic reflex. A: audiogram. Labels without a test letter belong to DPOAE. Negative correlation coefficients were set to 0 only for the purpose of visual clarity. Boxes with white solid lines encircled the items for different tests. Boxes with dotted lines encircled the items within the tests that were grouped together for analysis.

First, Kaiser-Meyer-Olkin (KMO) measure was used to verify the sampling adequacy for the analysis, KMO = 0.583, Bartlett's test of sphericity $\chi_2(253) = 699.985$, $p = 0.000$, indicating that the sample size is mediocre for factor analysis. Four frequencies from acoustic reflex, 16 frequencies from DPOAE and 3 frequencies from EHF pure-tone threshold were factor analyzed using principal component analysis with Varimax (orthogonal) rotation and criterion eigenvalue greater than 1. The analysis yields 7 factors that explain 73.9% of the variance for these variables:

Factor 1, the acoustic reflex amplitude, accounted for 15.9% of the variance. Factor 2, the low-frequency DPOAE, accounted for 14.8% of the variance. Factor 3, the EHF pure-tone threshold, accounted for 9.5% of the variance. Factor 4, the mid-frequency DPOAE, accounted for 8.9% of the variance. Factor 5, the low-to-mid-frequency DPOAE, accounted for 8.5% of the variance. Factor 6, the high-frequency DPOAE, accounted for 8.3% of the variance. Lastly, Factor 7, the EHF DPOAE, accounted for 7.9% of the variance. The only unused variable is DPOAE amplitude at 17 kHz, which did not meaningfully contribute to any factor.

Table 12. Exploratory Factor Analysis of the Audiologic Exams

Test	Frequency (kHz)	Factor						
		1 AR	2 low frequency DPOAE	3 HF PTA	4 mid frequency DPOAE	5 low to mid frequency DPOAE	6 high frequency DPOAE	7 EHF DPOAE
Acoustic reflex	**0.5**	.899	.071	-.091	-.018	.181	-.136	.144
	1	.923	.109	-.041	.082	.211	-.073	.139
	2	.947	.095	-.080	.079	.081	.017	.058
	4	.855	.086	-.215	-.081	.037	.156	-.070
DPOAE	**0.55**	.016	.746	-.043	.022	-.153	-.040	-.012
	0.7	.103	.662	.103	-.018	.061	.273	-.134
	0.88	.082	.807	.123	.014	.026	-.247	.089
	1.1	.061	.812	-.107	-.141	.152	.035	.005
	1.4	.109	.743	-.244	.251	.320	-.191	.025
	1.7	.119	.557	-.117	.214	.488	.072	.271
	2.2	.236	.195	-.281	.080	.741	.038	.018
	2.8	.281	.026	-.071	.079	.793	.136	-.244
	3.5	-.136	-.085	.107	.728	.375	-.192	.001
	4.4	.159	-.065	-.075	.672	.133	.284	-.032
	5.5	.095	.357	-.079	.714	-.273	.240	-.099
	7	-.209	.009	-.336	.515	.013	.451	.277
	8.8	-.038	-.103	-.068	.161	.040	.805	.220
	11.1	.105	-.073	.100	-.015	-.166	.286	.710
	14	.072	.088	-.179	-.042	.007	.109	.826
	17	-.018	-.017	-.140	-.072	-.069	-.708	-.056

Table 12 Continued

EHF PTA	**12.5**	-.177	.014	.585	-.070	-.068	.181	-.513
	14	-.132	-.019	.834	-.264	-.140	.062	.055
	16	-.166	-.070	.829	.162	-.130	-.049	-.101

The factor analysis correctly assigned the measurements from acoustic reflex and EHF pure-tone threshold to their original tests and created five levels of DPOAE variables, each covering a narrow frequency region. However, there were also some frequencies contributing to both newly loaded DPOAE factors. Specifically, response at 1.7 kHz contributed to both factor 2 (low-frequency DPOAE) and factor 5 (low-to-mid-frequency DPOAE) while response at 7 kHz contributed to both factor 4 (mid-frequency DPOAE) and factor 6 (high-frequency DPOAE).

For simplicity, factor 2 and factor 5 were combined into low-frequency DPOAE while factor 4 and factor 6 were combined into high-frequency DPOAE. The resulting factors for the peripheral audiologic functions and hearing sensitivity are listed in Table 13. In Figure 27, the correlation coefficient matrix of the new outcome variables of the audiologic screening shows that the correlations between different pairs of new variables are fairly low, allowing them to be analyzed as predictors of independent source in later multiple regression analyses.

Table 13. Audiologic data used as outcome variables

Test category	Outcome variables and test frequencies	New outcome variable label
Middle ear muscle reflex	Average amplitude of left ipsilateral acoustic reflex (ml) to pulsed tones at 0.5, 1, 2 and 4 kHz	Low-to-mid-frequency acoustic reflex (LMF AR)
Inner ear OHC functions	Average DPOAE amplitude (dB SNR) at f2 of 0.55, 0.7, 0.88, 1.1, 1.4, 1.7, 2.2, 2.8 kHz	Low frequency DPOAE (LF DPOAE)
	Average DPOAE amplitude (dB SNR) at f2 of 3.5, 4.4, 5.5, 7, 8.8 kHz	High frequency DPOAE (HF DPOAE)
	Average DPOAE amplitude (dB SNR) at f2 of 11.1 and 14 kHz	EHF DPOAE
High frequency sensitivity	Average pure-tone absolute threshold (dB HL) at 12.5, 14 and 16 kHz	EHF pure-tone average (PTA)

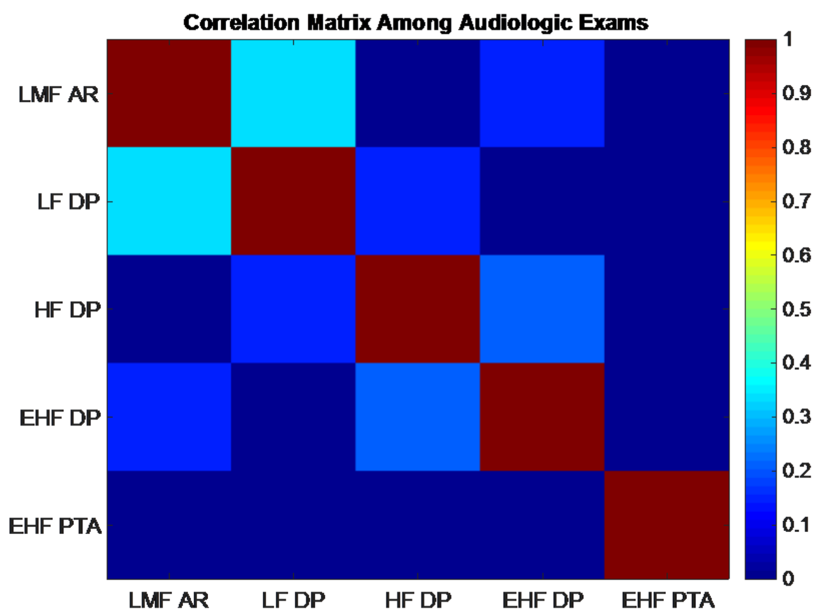


Figure 27. Matrix of correlation coefficients among the new outcome variables of the audiologic screenings

5.1.2 Effect of NTNE on peripheral auditory functions and extended high frequency listening

5.1.2.1 Effect of NTNE on middle ear muscle reflex

The effect of NTNE on middle ear muscle reflex was analyzed through one-way ANOVA. The independent variable was group (EXP vs CTL) and the dependent variable was low-to-mid-frequency (LMF) acoustic reflex (AR) amplitude, which was the mean amplitude of acoustic reflex from 0.5 to 4 kHz.

The result shows that while the mean reflex amplitude appears larger for the CTL group (mean = 0.12 ml, SD = 0.08 ml) than for the EXP group (mean = 0.09 ml, SD = 0.05), the difference was not statistically significant, $F(1,50) = 2.266$, $p = 0.139$ (Figure 28).

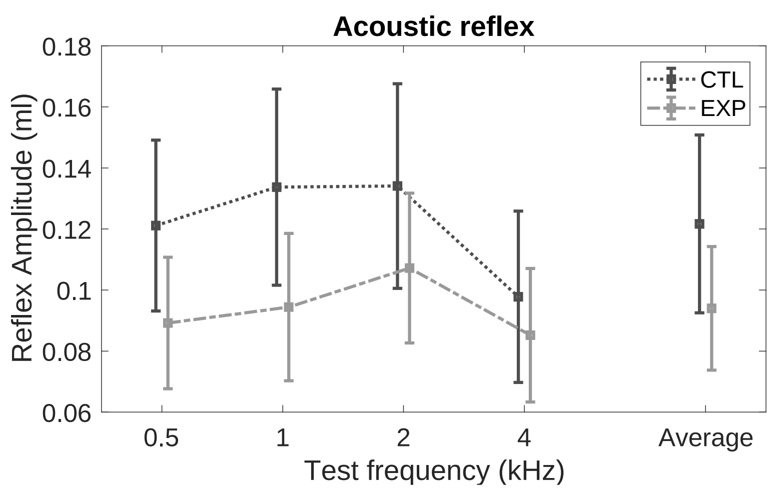


Figure 28. Acoustic reflex amplitude between groups at each test frequency and on average.

5.1.2.2 Effect of NTNE on DPOAE

The effect of NTNE on DPOAE amplitude was analyzed through a 2 (group) × 3 (frequency) mixed-model ANOVA. Mauchly’s test of sphericity shows that the assumption of homogeneity of variance is violated for frequency, $\chi^2(2) = 7.643$, $p = 0.022$. The result shows that there is significant main effects of frequency after Greenhouse-Geisser correction, $F(1.748, 87.380) = 105.025$, $p = 0.000$, $\eta_{p2} = 0.677$. There is no significant main effect of group or interaction between group and frequency, $p > 0.05$. The simple effects multiple comparisons also do not show a statistically significant group difference at any given frequency condition, $p > 0.05$ (Figure 29).

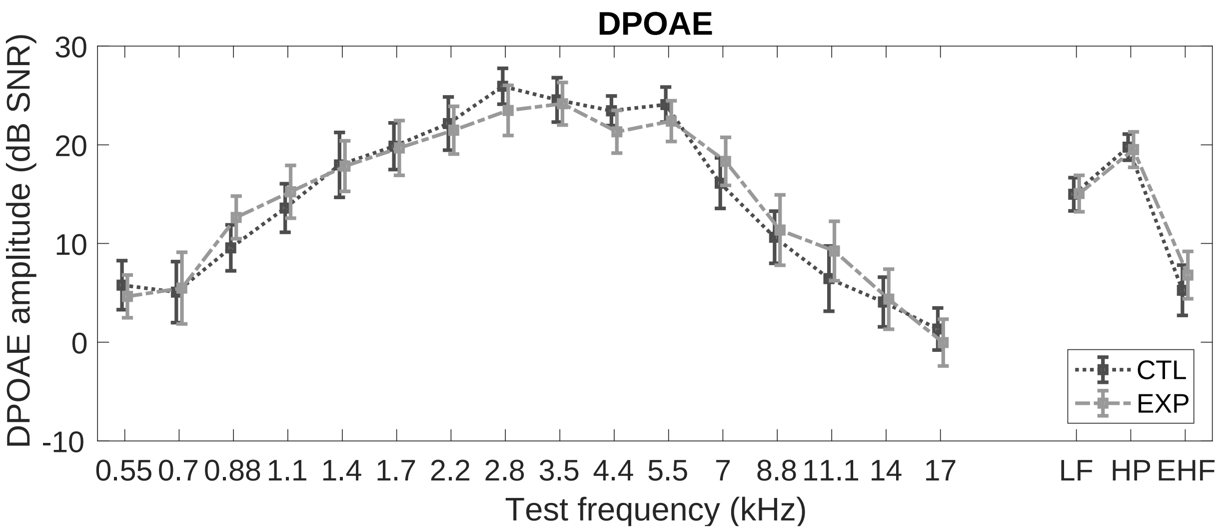


Figure 29. DPOAE amplitude between groups at each test frequency and on average at low frequency (LF), high frequency (HF) and extendend high frequency (EHF) regions

5.1.2.3 Effect of NTNE on extended high-frequency pure-tone average

The effect of NTNE on the EHF pure-tone average (PTA) thresholds was analyzed through one-way ANOVA. The independent variable was group and the dependent variable was EHF PTA, which was the mean of the absolute thresholds from 12.5 to 16 kHz.

The result shows that while the EHF PTA appears lower for the CTL group (mean = 6.6 dB HL, SD = 10.9 dB) than for the EXP group (mean = 9.2 dB HL, SD = 10.1 dB), the difference was not statistically significant, $F(1,50) = 0.792$, $p = 0.378$ (Figure 30).

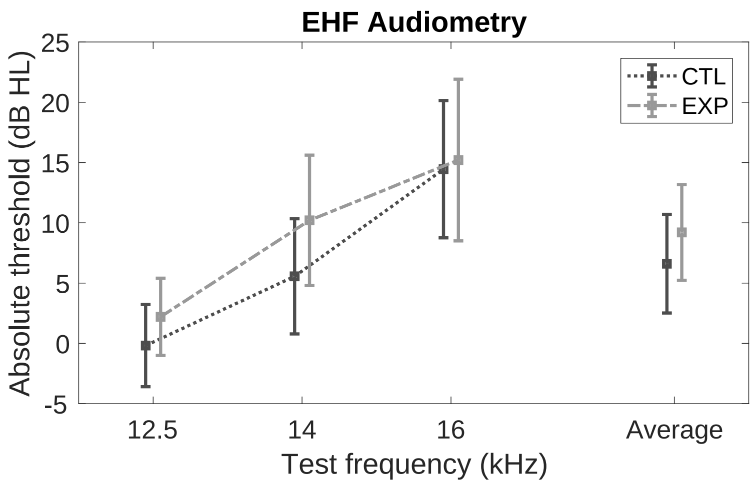


Figure 30. Audiogram at extended high frequencies between groups

5.1.3 Summary of the effect of NTNE on peripheral auditory functions and extended high-frequency listening

As is shown in the results, the middle ear acoustic reflex response, outer hair cell functions or high-frequency hearing sensitivity appeared unrelated to whether a listener has been exposed to non-traumatic high-frequency dental noise. Therefore, no further multiple regression analysis was conducted to examine whether the peripheral audiologic status after NTNE meaningfully contributed to the group difference in temporal envelope processing inside the frequency band of the noise.

5.2 Contributions of demographic factors and peripheral audiologic screening outcomes to temporal and spectral envelope processing outside the exposed noise band

One of the questions remaining from previous chapters is why the effect of high-frequency NTNE is relatively small on spectral and temporal envelope processing outside the frequency band of the noise. To understand how much the exposure to dental noise may account for the variance of spectral and temporal envelope processing compared to other factors, simple linear correlation and hierarchical multiple linear regression (HMLR) were used to analyze the contributions of various demographic and peripheral audiologic factors to temporal and spectral envelope processing outside the frequency band of the noise. The outcome variables of interest were the SRTs of LPF unvoiced speech recognition in 16-Hz TMN, in 32-Hz TMN, in SMN as well as the performance of silence-interrupted LPF unvoiced speech recognition at 48 Hz.

5.2.1 Correlation and regression factors/predictors

Table 14 lists the factors or predictors for the correlation and linear regression analyses. To build the HMLR model, the sequence of factors being added to each step of the model was demographic factors first and audiologic factors next. For demographic factors, the basic information such as participants' age and years of music training was added in step 1, and lifetime noise exposure histories, including L_{eq} of non-dental and dental noise exposure, were added in step 2. Then, audiologic factors including the outcomes from middle-ear to inner-ear screening measures were added, the sequence of which follows the anatomical structure for sound transmission in the peripheral auditory system. The EHF DPOAE and PTA were excluded from the model because the information in these frequencies does not contribute to speech intelligibility and should not provide a theoretically meaningful interpretation of the speech-related outcome variables. This HMLR model was used for all the outcome variables from unvoiced speech tests.

Table 14. Outcome variables for correlation/regression analysis and their orders entering the HMLR model

Step No.	Category	Factors (unit)
1	Demographic: basic	Age (years), years of musical training (years)
2	Demographic: noise exposure	Lifetime dental noise exposure L_{eq} (dB SPL) Lifetime non-dental noise exposure L_{eq} (dB SPL)
3	Middle ear function	LMF AR (ml)
4	Inner ear OHC function	LF DP (dB SNR) HF DP (dB SNR)

5.2.2 Contributions to temporal envelope processing outside noise band

Previous chapters of the study showed that temporal envelope processing outside the frequency band of the noise was not significantly different between the two groups, supported by the findings that there was no group difference in LPF speech in 16-Hz (LPTMN16) or 32-Hz TMNs (LPTMN32) and that there was a significant but small group difference in 48-Hz interrupted LPF speech (LP48). The analysis here examines whether other demographic or audiologic factors than dental noise exposure could account for the variance of LPTMN16, LPTMN32, and LP48.

Based on the Pearson correlation, there are no significant correlations between any two of the predictors so multicollinearity should not occur (see Appendix C.1 for diagnostic statistics). However, there are also no significant correlations between the outcome variables of the LPF unvoiced speech recognition in TMNs (i.e. LPTMN16, LPTMN32) and any of the predictors. In other words, there are no linear relationships between any of the demographic or audiologic predictors and LPF unvoiced speech recognition performance in 16-Hz TMN or in 32-Hz TMN.

The lack of a linear relationship between the predictors and the outcome variable rendered further multiple linear regression futile. The results suggest that demographic information like age, years of musical training and lifetime noise exposure level, and peripheral audiologic functions like middle ear muscle reflex and OHC integrity do not account for the variance of LPF unvoiced speech recognition performance in TMN.

On the other hand, there is a significant correlation between dental noise exposure Leq and LP48, $r(50) = -0.332$, $p = 0.016$, and a marginally significant correlation between acoustic reflex amplitude and LP48, $r(50) = 0.275$, $p = 0.049$ (Table 15).

Table 15. Correlation between outcome variables of temporal envelope processing outside the frequency band of the noise and demographic or audiologic predictors

		Demographic predictors				Audiologic predictors		
		Age	Years of Music	Non-dental L_{eq}	Dental L_{eq}	LMF AR	LP DP	HP DP
LPTMN	*r*	-.165	-.232	.051	.145	-.157	-.117	-.256
	p	.243	.097	.720	.305	.266	.410	.067
LPTMN32	*r*	.130	.205	.080	.234	-.135	-.138	-.146
	p	.360	.145	.571	.096	.341	.329	.303
LP48	*r*	.068	.239	-.088	-.332*	.275*	-.004	-.062
	p	.633	.089	.537	.016	.049	.976	.664

In this case, the HMLR model was built for LP48. Tests of normality, homoscedasticity, and multicollinearity are not violated (see Appendix C.1 for details). The analysis summary is shown in Table 16. When age and years of musical training were entered (Model 1), the model does not significantly predict the performance of LP48, $F(2, 49) = 0.045$, $p = 0.956$. When lifetime non-dental noise exposure L_{eq} and dental noise exposure L_{eq} are entered, they improve the prediction but are not significant, R^2 change = 0.110, $F(2, 47) = 3.115$, $p = 0.054$. However, the t-test for individual regression coefficients shows that lifetime dental noise exposure L_{eq} is a significant predictor for LP48, but the lack of significant contribution from non-dental noise exposure L_{eq} may have watered down the effect. Additional predictors from middle ear acoustic reflex and DPOAE amplitude for Model 3 and Model 4 do not significantly increase the predictability of the model, $p > 0.05$. The final model (Model 4) explains 21.9% of the variance in LP48, but lifetime dental noise exposure is the only significant predictor.

Table 16. HMLR model for LP48

	Unstandardized Coefficients		Standardized Coefficients	Individual regression coefficients	R-squared	R-squared change
	B	SE	ß	t-statistics		
Model 1					0.059	0.059
Age	.175	.575	.042	.304		
Music	.440	.262	.234	1.678		
Model 2					0.169	0.110
Age	.454	.577	.110	.787		
Music	.372	.255	.198	1.455		
Dental L_{eq}	-.072	.029	-.339	-2.478*		
Non-dental L_{eq}	-.033	.221	-.020	-.147		
Model 3					0.203	0.035
Age	.508	.572	.123	.887		
Music	.260	.265	.138	.980		
Dental L_{eq}	-.064	.029	-.302	-2.195*		
Non-dental L_{eq}	-.108	.225	-.068	-.481		
LMF AR	24.317	17.219	.203	1.412		

Table 16 continued

Model 4					0.219	0.015
Age	.375	.616	.091	.609		
Music	.343	.283	.182	1.210		
Dental L_{eq}	-.062	.030	-.292	-2.062*		
Non-dental L_{eq}	-.111	.231	-.070	-.483		
LMF AR	25.891	18.467	.217	1.402		
LF DP	-.135	.271	-.076	-.496		
HF DP	-.202	.279	-.101	-.722		

*p<0.05; ** p<0.01

5.2.3 Contributions to spectral envelope processing outside noise band

Following the similar procedures for analyzing the outcome variables for temporal envelope processing, the Pearson correlation was first conducted to examine whether there was a linear relationship between the demographic or audiologic factors and the performance of LPF unvoiced speech in SMN (LPSMN). The result shows that only lifetime dental noise exposure L_{eq} is significantly correlated with LPSMN.

Table 17. Correlation between outcome variable of spectral envelope processing outside the frequency band of the noise and demographic or audiologic predictors

		Demographic predictors				Audiologic predictors		
		Age	Years of Music	Non-dental L_{eq}	Dental L_{eq}	LMF AR	LP DP	HP DP
LPSMN	*r*	.060	.121	-.223	.337*	-.076	.055	-.090
	p	.675	.393	.111	.015	.592	.698	.525

Tests of normality, homoscedasticity, and multicollinearity are not violated for LPSMN (see Appendix C.1 for details), so the HMLR model is built for LPSMN. The analysis summary is shown in Table 18. The HMLR model shows that lifetime dental noise exposure L_{eq} is the only significant predictor for LPSMN, while age, years of musical training, non-dental noise exposure, middle ear reflex, or inner ear OHC functions fail to account for the variance of LPSMN. The final model (Model 4) explains 19.9% of the variance in LPSMN.

Table 18. HMLR model for LPSMN

	Unstandardized Coefficients		Standardized Coefficients	Individual regression coefficients	R-squared	R-squared change
	B	SE	ß	t-statistics		
Model 1					0.017	0.017
Age	.054	.162	.047	.330		
Music	.060	.074	.116	.813		
Model 2					0.185	0.168
Age	-.088	.158	-.078	-.559		
Music	.061	.070	.117	.871		
Dental L_{eq}	.022	.008	.368	2.717**		
Non-dental L_{eq}	-.102	.060	-.230	-1.682		
Model 3					0.185	0.000
Age	-.088	.160	-.077	-.550		
Music	.060	.074	.116	.811		
Dental L_{eq}	.022	.008	.369	2.647*		
Non-dental L_{eq}	-.102	.063	-.231	-1.626		
LMF AR	.190	4.812	.006	.039		

Table 18 continued

Model 4					0.199	0.014
Age	-.092	.172	-.081	-.536		
Music	.079	.079	.152	.995		
Dental L_{eq}	.021	.008	.366	2.554*		
Non-dental L_{eq}	-.097	.065	-.221	-1.511		
LMF AR	-.374	5.166	-.011	-.072		
LF DP	.007	.076	.014	.092		
HF DP	-.068	.078	-.124	-.876		

*p<0.05; ** p<0.01

5.3 Interim Discussion

This chapter provides additional analyses on whether NTNE negatively impacts peripheral auditory function at standard audiometric and EHF regions (section 5.1) and whether demographic and peripheral audiologic factors could account for the variance of performance in spectral and temporal envelope processing outside the frequency band of the noise (section 5.2).

The results first show that there was no significant difference between the EXP and the CTL groups on acoustic reflex amplitude, on DPOAE amplitude at regularly tested frequencies, on DPOAE amplitude at EHFs, or on EHF pure-tone averages. The hypothesis that acoustic reflex may be negatively impacted by NTNE originated from the argument that the damage of high-intensity ANFs (i.e. fibers with LSR) would reduce the capacity to encode high-intensity sound, which would consequently reduce the amount of stimulation in the CAS to the stapedius muscles (Valero et al., 2017; Valero et al., 2018). However, it should be noted that since the acoustic reflex is usually stimulated by high-intensity sound, the range of excitation on the basilar membrane will be broad and many ANFs from less noise-impacted regions may respond to the stimulus, causing confounding contributions from off-frequency regions. This explanation is supported by the current study where significantly high correlations are among the acoustic reflex amplitudes at different frequencies. The current finding suggests that acoustic reflex is theoretically and empirically insensitive as a screening tool to reflect the loss of high-intensity ANFs at a restrained frequency region.

A small number of animal studies have suggested that OHCs at frequencies almost 2 octaves above the exposed noise band show irreversible damage after NTNE (Liberman et al., 2015; Fernandez et al., 2020). Therefore, studies have suggested that EHF hearing thresholds may serve as an early indicator for noise-related hidden hearing issues. Furthermore, many studies on age-related hearing deficits have argued that EHF thresholds (above 8 kHz) start to elevate well before there are signs of elevated thresholds at routinely tested speech frequencies (i.e. 0.25 to 8 kHz) (see Rodríguez Valiente, 2014 for review). As some animal studies have further shown that noise exposure at an early age could accelerate the onset of age-related hearing loss (Sergeyenko et al., 2013; Fernandez et al., 2015), the idea that EHF listening may be related to NTNE was reinstated. However, the current analyses did not find a significant group difference in EHF hearing thresholds. One potential explanation is that the lifetime environmental noise exposure L_{eq} for both groups are similar (Table 3) and that the dental noise exposure L_{eq} is not traumatic for the EXP group (only 74 dB SPL on average). Eggermont and his colleagues who exposed cats with prolonged moderate-level noise exposure have not found abnormal hearing sensitivity post exposure (e.g. Pienkowski & Eggermont, 2009). Furthermore, the current participants were all below 30 years old so age-related hearing issues like high-frequency hearing loss should be eliminated. If time allows, a longitudinal study will be more suitable to address whether there is an early onset of EHF issues for people with NTNE, but it is beyond the scope of this study.

This chapter also shows that apart from dental-noise exposure history, neither the demographic factors nor the outcomes from peripheral audiologic screening could explain the variance of the performance on spectral and temporal envelope processing outside the frequency band of the noise. Taken collectively, the findings from the current study consolidate the conclusions that poor temporal and spectral envelope processing related to NTNE likely occurs at super-threshold levels, at retro-cochlear sites and can be overlooked by clinically routine audiologic screenings.

6.0 Discussion

6.1 Summary of the current findings

The current study aimed to address the following research questions:

1) Do people who had a specific type of NTNE perform differently on tasks of temporal envelope processing from people who were unexposed?

2) Do people who had a specific type of NTNE perform differently on tasks of spectral envelope processing from people who were unexposed?

3) If there is a difference between the exposed and the unexposed, will the difference be frequency-specific?

4) Will the difference occur at only supra-threshold levels?

The current study operationally defined NTNE as exposure to noise in the dental school clinics, because of its distinctive high-pass spectral feature, non-traumatic nature, and systematic exposure schedule across dental students of different years. Temporal envelope processing was examined using unvoiced speech recognition in TMN and silence-interrupted unvoiced speech recognition. Spectral envelope processing was examined using unvoiced speech recognition in SMN. And all the stimuli were either low-pass or high-pass filtered to examine the frequency-specific effects of NTNE. Unvoiced speech recognition in quiet was also tested to examine spectro-temporal processing near-threshold.

As is summarized in Table 19, people who were exposed to a specific type of NTNE performed poorly on tasks of temporal envelope processing compared to people who were unexposed. The effect of NTNE on temporal envelope processing is more robust within the spectral band of the exposed noise than outside the band. The results also suggest that the negative effect of NTNE on temporal envelope processing is more obvious in conditions that required finer temporal resolution (e.g. faster modulation rates for TMN or silence interruption) than in those requiring less fine temporal resolution (e.g. slower modulation rates for TMN or silence interruption).

Table 19. Summary of the study conclusions

Processing ability / Frequency specificity	Near-threshold listening	Supra-threshold temporal envelope processing	Supra-threshold spectral envelope processing
Effect of NTNE within noise band	Not affected	Affected; more obvious when requiring finer temporal resolution	Unclear, lack of measure sensitivity
Effect of NTNE outside noise band	Not affected	Mildly affected, unexplained by demographic or audiologic factors	Affected, unexplained by demographic or audiologic factors

Furthermore, NTNE is also associated with reduced spectral envelope processing. The effect was found outside the exposed noise band, supporting findings from animal studies that the spectral receptive field and neural responses are enhanced outside the noise band (e.g. Pienkowski & Eggermont, 2009). Although the current measure is insensitive to spectral envelope processing inside the frequency band of the noise, the negative effect of NTNE on spectral envelope processing is a novel finding which could be further explored with other measures of spectral resolution.

The current study also examined the effect of NTNE on perceptual processing near threshold, but found no significant group difference, which supports the hypothesis from animal studies that the effect of NTNE is largely supra-threshold.

Lastly, additional analyses showed that factors such as age, years of musical training, non-dental environmental noise exposure history, and peripheral auditory screening outcomes did not significantly contribute to the variance of the performance in tasks of temporal or spectral envelope processing outside the frequency band of the noise. And there was no group difference in the outcome variables for EHF listening. These findings suggest that the supra-threshold temporal and spectral envelope processing relates to NTNE occurred retro-cochlear and could be neglected by clinically routine audiologic screenings.

6.2 Effect of non-traumatic noise exposure on temporal envelope processing

6.2.1 Current finding vs. past findings

One of the significant findings of the current study is that people who have been exposed to high-frequency non-traumatic dental noise perform significantly more poorly than people without such exposure to recognize unvoiced speech in amplitude-modulated noise, especially at high frequency which is inside the frequency band of the exposed noise. This is the first time that unvoiced speech recognition is used to examine and support the effect of NTNE on temporal envelope processing.

Early studies that attempted to examine noise-related temporal envelope processing have been inconclusive due to issues in their measures and/or choice of participants. To measure temporal envelope processing, most studies have used amplitude modulation detection (psychophysical) or EFR (electrophysiological). Findings from Prendergast and colleagues did not support the existence of noise-related temporal envelope processing issues (Prendergast et al., 2017a, 2017b), but their measures were not controlled to eliminate off-frequency listening or central contribution to AM encoding. Paul et al. (2017, 2018) examined AM detection and EFR between young adults of 18 to 19 years old with high or low noise exposure history. Their measures controlled for off-frequency listening by presenting the stimulus in narrow-band noise and the noise was presented at different levels as a way to engage the ANFs with different SRs. Their primary outcome showed no relationship between noise exposure and the performance of AM detection or EFR strength. At one point, their analysis seemed to support that the higher exposed group showed lower EFR strength when the stimulus was presented at 40 dB spectrum level, but the result was soon contradicted by the non-significant group difference for stimulus at 45 dB spectrum level (Paul et al., 2018). Yeend et al. (2017) and Füllgrabe et al. (2020) controlled the issue of off-frequency listening by presenting the AM targets in threshold-equalizing noise and have found no relation between NTNE and temporal processing. The measures used in both studies were designed to reflect temporal processing only at high frequency, as noise-induced hearing loss often occurs with a notch on the audiogram at 4 kHz, but they did not further examine the temporal processing within the frequency band of the exposed noise, which should be primarily environmental noise with energy below 2 kHz in their cases. As the current findings have successfully shown that NTNE is related to poor temporal envelope processing and the effect of exposure is more obvious to temporal envelope processing inside the frequency band of the noise, future studies may consider examining temporal processing using the unvoiced speech recognition in fast modulated noise alongside with the refined AM-based measures (e.g. AM detection, EFR in noise).

6.2.2 Speculations on the form of degraded temporal envelope processing related to NTNE

Poor temporal processing takes different forms of physiological changes in animal studies (see Appendix Table 2 for a summary). Studies at the auditory nerve level have suggested that the loss of high-intensity fibers most likely smoothens the temporal envelope (e.g. Shaheen et al., 2015) but has not specified that the smoothing effect is a result of broader temporal integration window. Thus, there is no direct evidence from animal studies at ANFs to suggest that temporal gaps are inaccurately encoded by the remaining low-intensity fibers with HSR. Meanwhile, animal studies at cortical levels have shown that neural phase-locking to pulse onsets diminished with increasing pulse rate more drastically for animals with NTNE (Zhou et al., 2011; Liu et al., 2012; Shi et al., 2013), suggesting that temporal gaps are not precisely encoded after NTNE and that the impact of NTNE in central auditory physiology could evolve in its form from the ANFs to the auditory cortex. Other research groups have also observed that cortical regions with CFs outside the frequency band of the noise showed greater ringing, while regions with CFs inside the frequency band of the noise no longer responded to their original CFs (Pienkowski & Eggermont, 2009; Pienkowski & Eggermont, 2010a, 2010b; Pienkowski et al., 2011, 2013). These results suggest a potential mixture of issues like broader temporal integration window and information misrepresentation.

The current study did not aim to provide a precise locus of lesions or to describe the exact form of degraded temporal processing. However, based on the glimpsing model (Cooke, 2006), interrupted speech recognition requires a listener to use the pieces of speech information that are present in the noise or silent gaps. If the auditory system has a poor temporal resolution in the form of greater forward masking or broader temporal integration window, there will be greater masking of the speech segments in the temporal gaps of the noise as a result, causing poor speech recognition performance in TMN (Jin & Nelson, 2006). Similarly, for silence interrupted speech, a broader temporal window may increase the amount of forward masking from the preceding to the following speech fragments and so degrade speech intelligibility (Jin and Nelson, 2010). Furthermore, under the broadened temporal window hypothesis, conditions of higher modulation rates will be more susceptible to smoothing effects, causing degraded temporal resolution. These assumptions of degraded speech perception due to broadening of the temporal window are supported by the current findings that the group with NTNE performed more poorly than the group without the NTNE in noise- or silence-interrupted speech recognition, and that the effect of NTNE becomes more obvious when the task has higher demands on temporal resolution. If the physiological effect of NTNE is simply information misrepresentation, the magnitude of group differences should not depend on the modulation rate.

6.2.3 Poor temporal envelope processing manifests at higher modulation rates in noise- and silence-interrupted speech recognition

The current study shows that different interruption rates for interrupted speech recognition are not equally sensitive to poorer temporal processing. Specifically, it was shown that poor temporal envelope processing was more obvious when the interruption rates were around 32 to 48 Hz instead of 12 to 24 Hz or even lower rates.

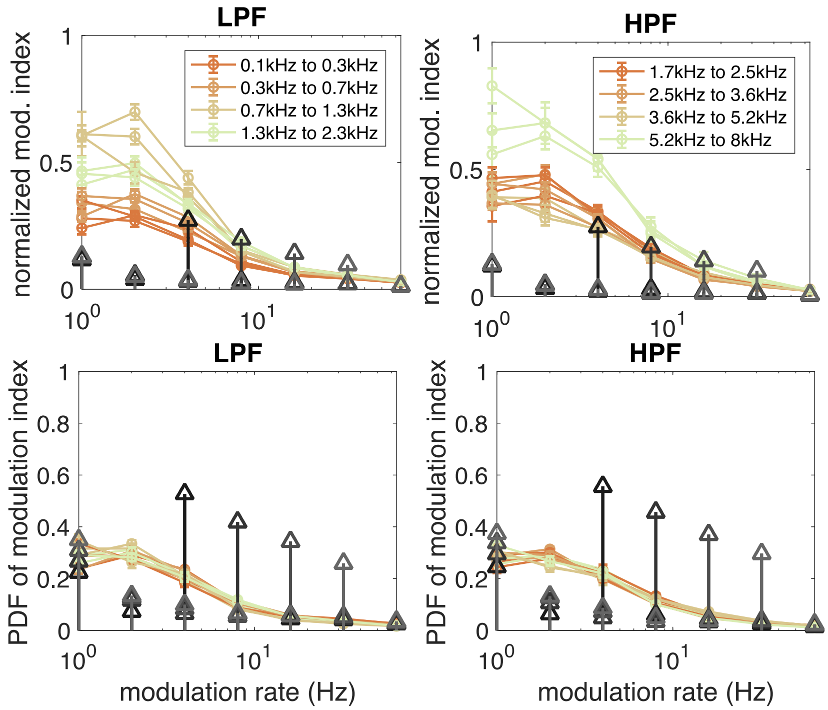


Figure 31. Modulation spectra in normalized modulation index (upper panels) and the PDFs of modulation energy (lower panels) of unvoiced speech (colored) and noise (grey). Left panels are for LPF speech and right panels HPF speech. Speech modulation spectra were processed in four channels with one color representing the modulation spectra from the same channels, and three lines in the same color represent the modulation spectra from three different sets of speech stimuli. The modulation spectra for each TMN was processed in one channel with the most prominent peak at the rate of modulation. The darkest grey stems represented the modulation spectra of 4-Hz TMN and the lightest grey 32-Hz TMN.

The reason that temporal processing is less affected at low rates is likely related to the time scale of syllable and phoneme production in English. At low rates, an entire syllable can be masked or silenced by an interruption, making temporal processing a less effective remediation (Figure 31). For example, if we imagine speech intelligibility at extremely low rates, interruptions may mask entire words, while other whole words will be intact, and speech intelligibility is then a function of the amount of speech information that is present rather than as a function of temporal resolution (Figure 32).

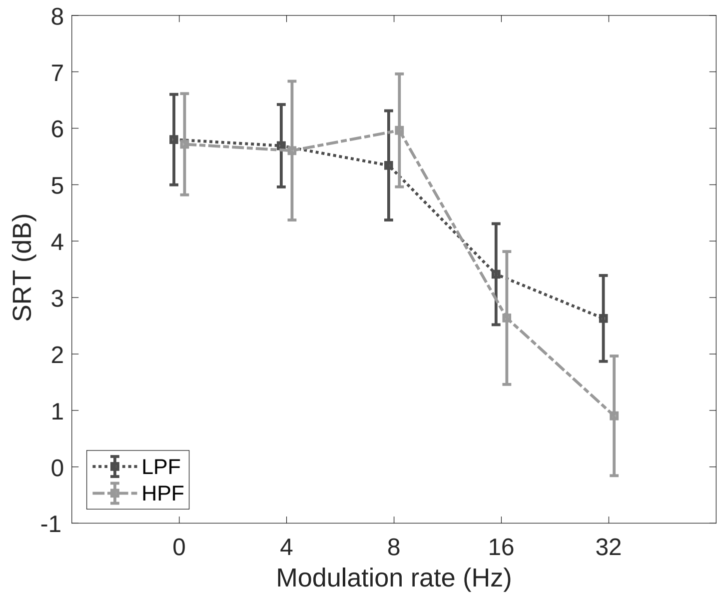


Figure 32. Unpublished pilot data of SRT of LPF (dark grey) and HPF (light grey) unvoiced speech in TMN of different modulation rates (N = 25). Error bars: 95% CI.

On the other hand, when the interruption rate is at 16 to 24 Hz, the listener has multiple looks at the syllable. If temporal resolution is poor, the listener will have less access to the segments within the syllables, resulting in poor recognition performance. Then, as the interruption rate further increases to 32 to 48 Hz, the multiple looks occur within the phonemic boundaries and speech recognition requires even finer temporal resolution. If future studies examine temporal processing using interrupted speech recognition, it should be considered to use faster modulation rates that are within phonemic boundaries and requires finer temporal resolution, such as 32 to 48 Hz.

6.2.4 On the frequency-specific effect of non-traumatic noise exposure on temporal envelope processing

The current study found that the effect of high-frequency dental noise exposure on temporal envelope processing was less severe at low frequency than at high frequency, suggesting that the effect of NTNE on temporal processing may be more obvious inside than outside the frequency band of the noise.

The small effect of noise exposure outside the frequency band of the noise may be a result of the similar L_{eq} of non-dental noise exposure between the EXP and the CTL groups. In other words, because the noise exposure of interest here is high-frequency, the small group difference at low frequency region is likely the result of relatively similar exposure dose to low-frequency noise between the groups. Experiment 2 suggests that the lack of substantial group difference of temporal processing at low frequencies could also result from the lower demand of temporal resolution at low frequencies (Viemeister, 1979; Shailer & Moore, 1983; Formby & Muir, 1988; Plack & Moore, 1990). Therefore, the current study does not clarify this issue and more researches are needed to address the frequency-specific effect of NTNE with different center frequency.

6.3 Probing the mechanisms of poorer spectral envelope processing

The current study shows that the effect of NTNE is related to poor spectral envelope processing. Specifically, the effect was found outside the frequency band of the noise. The mechanism of poorer spectral envelope processing may be similar to that of poorer temporal envelope processing. That is, broadened auditory filter smoothed the spectral gaps of the noise with unwanted energy, decreasing the SNR in the spectral gaps and deteriorating speech intelligibility (Baer & Moore, 1993; Boothroyd et al., 1996; Gnansia et al., 2009).

Animal models of NTNE have also suggested similar issues both at ANFs and at cortical levels (see Appendix Table 3 for summary). At the auditory nerve level, a lack of contributions from the ANFs with LSR would lead to a smoothed spectrum, causing lower spectral contrast (Sachs & Young, 1979; May et al., 1996; Reiss et al., 2011). At cortical levels, neuronal frequency tuning curves were found broadened after NTNE, which was an evident physiological counterpart to broadened auditory filters (Zhou et al., 2011; Zhou & Merzenich, 2012; Kamal et al., 2013).

With this novel finding concerning reduced spectral processing post NTNE, future studies may consider using a conventional notched noise paradigm (NNP) to examine spectral resolution in people with NTNE. The conventional NNP will be more advantageous than unvoiced speech recognition in SMN, because the NNP is designed to avoid off-frequency listening and allows measurement of auditory filter at any center frequencies. The NNP is more time-consuming than unvoiced speech recognition in SMN, but it can provide spectral resolution at higher frequencies, which the speech-based method cannot.

6.4 Limitations and future directions

Due to the limitation of measure sensitivity, there are still some unaddressed questions from the current study that are worth exploring in the future.

The first issue is the frequency-specific nature of NTNE. Indeed, the current study found that the effect of NTNE was more obvious inside than outside the frequency band of the noise, but this conclusion was established based on high-frequency noise exposure. In other words, the conclusion concerning the frequency-specific effect of NTNE from this study is limited to high-frequency noise exposure. Future studies should consider examining the effect of low-frequency noise exposure on temporal processing at both low and high frequencies.

The second issue is the selection of a sensitive measure to reflect changes in temporal envelope processing at low frequencies. Future studies may use the vocoded speech recognition used in Experiment 2, which ensures that speech intelligibility relies on temporal envelope processing. The potential downside of using vocoded speech with low spectral details is that such speech does not show the benefit of masking release in TMN (Nelson et al., 2003; Qin & Oxenham, 2003; Stickney et al., 2004; Ihlefeld et al., 2010; Fogerty et al., 2016), which causes difficulty in quantifying the ability to use the information in temporal gaps. Therefore, degraded temporal resolution can only be demonstrated by comparison with a control group. The second potential measure for temporal envelope processing that may control for off-frequency contribution is psychoacoustic tasks based on detecting or using across-frequency modulation, such as co-modulation masking release (CMR). In CMR tasks, amplitude modulation with the same modulation rate is perceptually grouped across frequency regions to form a sound object, improving the detection of sound with uncorrelated temporal envelope (Hall & Grose, 1990). If the ability to encode amplitude modulation at a given frequency region is reduced, the grouping across frequency may become less efficient, and hence the ability to detect the uncorrelated sound is affected (Ernst et al., 2010). The third optional measure is to use the AM-based measures, such as AM detection and EFR presented in narrow-band noise, so that off-frequency listening can be eliminated. Although many past studies using these measures have not found the relationship between noise exposure and temporal envelope processing (Prendergast et al., 2017a, 2017b; Paul et al., 2017, 2018; Yeend et al., 2017; Füllgrabe et al., 2020), these non-significant results may have suffered from poor choice of participants in terms of controlling their noise exposure background or the lack of correspondence between the exposure frequency and the test frequency. The current study shows that dental school students may be a viable population to study the perceptual consequences of NTNE, as dental noise exposure is highly distinctive from the daily environmental noise exposure in spectral domain and is rarely experienced by people outside dental professions.

The third issue is whether NTNE would only impact spectral envelope processing outside the frequency band of the noise. The degradation of spectral resolution outside the frequency band of the noise is not limited to studies of NTNE. Normally speaking, large inhibitory response areas have been observed in the neuronal spectral receptive fields (SRF) from the cochlear nucleus in the CAS, even by single-tone stimulation (Young & Brownell, 1976; Rhode, 1999), which is different from the response of ANFs where single tone stimulation only produces all-excitatory SRF (Liberman & Dodds, 1984; Narayan et al., 1998). Many studies of traumatic noise exposure have shown that inhibitory areas in the SRF of the central auditory neurons turned into large excitatory areas post traumatic noise exposure (Cai et al., 2009; Ma & Young, 2006; Li et al., 2015; Ropp et al., 2014). Such changes in excitatory-inhibitory response area are considered to be a compensatory mechanism of the central auditory system to tune up the gain and to maintain neural internal homeostasis in the absence of sound input (Eggermont, 2017), which was proposed to account for noise-related tinnitus (Kaltenbach et al., 2004; Munguia et al., 2013; Auerbach et al., 2014; Noreña, 2015). Interestingly, Eggermont and his colleagues have also observed similar issues in the cat model of NTNE. They discovered that cortical response was enhanced to stimuli right outside the spectral edges of the noise but was reduced to stimuli within the frequency band of the noise. This increased response to edge frequency is considered a form of excitatory-inhibitory imbalance and echoes the finding in the current study that spectral envelope processing of people with NTNE is worse outside the frequency band of the noise. However, it is still unclear from the current study whether poor spectral processing is limited to frequencies outside the noise spectrum. To study this question in the future, people with noise exposure at low frequency should be tested with measures like LPF unvoiced speech recognition in SMN.

This study provides a new and flexible way to assess temporal envelope processing at high frequency, which is the task of HPF unvoiced speech recognition in 32-Hz TMN. The duration of administering this condition is about 4 minutes. The HPF unvoiced speech is perceived to resemble natural whisper speech, so participants have enough experience with this type of speech in real life. As the audiologic research and clinical community are exploring the potential use of the tests that are already in audiologic clinics to identify noise-related hidden hearing deficits, such as middle-ear muscle reflex, OAE, ABR wave I amplitude, medial olivocochlear reflex (MOCR) (Bhatt, 2017; Guest et al., 2018; Smith et al., 2019; Bhatt & Wang, 2019; Shehorn et al., 2020; Bramhall et al., 2020), the current study provides a new measure potentially as a part of a test battery of hidden hearing deficits and the measure can be administered rather quickly in clinical settings.

7.0 Conclusion

Animal models from the past decade have suggested that noise exposure can affect temporal envelope processing at supra-threshold levels even though the absolute hearing thresholds and the peripheral auditory function appear normal. This issue has been considered a form of hidden hearing deficits that may account for the phenomenon that people without hearing impairments show difficulty listening in acoustically challenging environments. However, human studies have failed to consistently find the relationship between noise exposure and temporal envelope processing or other auditory processing tasks among listeners without hearing impairments, an outcome that has very likely resulted from poor control of the participants' noise exposure history and the measure sensitivity.

The current study examined the effect of NTNE by comparing temporal envelope processing of young adults who studied at dental schools to that of young adults without dental noise exposure. Temporal envelope processing was examined through unvoiced speech recognition interrupted by noise or by silence. The results show that the group with dental noise exposure performed more poorly than the group without the exposure. The effect of high-frequency NTNE on temporal envelope processing was found to be more robust within the spectral band of the dental noise than outside the spectral band. The effect was also more obvious in conditions that required finer temporal resolution (e.g faster noise modulation rate) than in those requiring less fine temporal resolution (e.g. slower noise modulation rate). Furthermore, the study examined spectral envelope processing between the two groups and found that the exposed group did not perform as well as the unexposed group when spectral processing was examined at low frequencies. On the other hand, factors such as age, years of musical training, non-dental noise exposure history, and the outcomes from peripheral auditory function screening were not able to explain the variance of the performance in the tasks of temporal and spectral envelope processing. And the two groups performed similarly when the task of spectro-temporal envelope processing was presented near threshold. In conclusion, the findings from this study contribute a new piece of evidence to the field of auditory research that NTNE is related to temporal and spectral envelope processing issues, and that as the animal models have suggested, such noise-related listening issues occur in retro-cochlear sites, at supra-threshold levels, and could be easily overlooked by clinically routine audiologic screening.

Appendix A Subject demographic form

SUBJECT BACKGROUND FORM
(check the one that applies, explain if needed)

Gender: male _____ female _____ transgender _____

Date of birth (DD/MM/YYYY): ________________________ Age: ___________

Highest education level: ____________________________________

Current academic major: ___________________________________

Country of birth: ___________

First language: __________

History
1. Have you been diagnosed with hearing issue before (e.g. hearing loss, ear infection, etc.)?
No_____ Yes_____
If yes, when and what type of hearing issue:
____

2. Any drainage from the ear within the past 90 days? No_____ Yes_____

3. Have you experienced any dizziness, balance problems, or falls?
No_____ Yes_____

4. Have you had any pain/discomfort in your ears within the past 90 days? No_____ Yes_____

If yes, rate on a scale of 0 (no discomfort) to 10 (worst possible) ___________

5. Have you ever lost hearing in one ear *<u>suddenly</u>*? No_____ Yes_____

6. Do you have frequent noises or ringing in your ears? No_____ Yes_____

If yes, which ear: left_____ right_____
The frequency of occurrence: Constant _____Intermittent _____
When does it start to frequently occur? _________________________________

7. Have you received any medical or surgical treatment for hearing loss?
No_____ Yes_____

------TURN OVER TO BACK PAGE------

8. Have you ever been exposed to loud noise?

Appendix Figure 1. Subject background form (Page 1)

Military_____ Occupation/Job_____ Recreational_____

For checked item, please briefly describe the type of noise and the hearing protection device (HPD) used if any.

Military: ______________________________________ HPD: ______________
Occupation/Job: _______________________________ HPD: ______________
Recreational: __________________________________ HPD: ______________

9. **Have you been diagnosed with neurological, neuropsychological or neuropsychiatrical conditions?** No_____ Yes_____

10. **Other medical problems (check all that apply):**

Infectious disease _____ Diabetes _____ Heart problems _____
Head injury _____ High blood pressure _____ Headache _____
Kidney failure _____ Pacemaker/Defibrillator _____
Other (please explain): ____

11. **Do you have any other concerns over your hearing?** No_____ Yes_____
If yes, please explain: ____

Comments or questions for the Principle Investigator:

____
____
____

===== Please continue to noise exposure questionnaire on next page. Thank you. ====

Appendix Figure 2. Subject background form (Page 2)

Appendix B Noise exposure questionnaires

Exposed Noise and Hearing Disorders of Conscripts (revised)

Subject ID: ______________ Date: ______________

The purpose of this questionnaire is to determine your personal noise exposure except for the major-related noise. For example, if you are a dentist, your major-related noise will be the noise from equipment used in your dental practice. If you are a violinist, your major-related noise will be the sound practicing with your violin alone and in a group. Since there is a separate questionnaire for major-related noise exposure, we would like you to exclude the major-related noise exposure when completing this questionnaire.

Please estimate the sound level of your leisure-time activities and non-major occupational noise exposure using a rating of 1 to 5. Please also mark the exposure time per time, how many times a week, how many weeks over the period of last 1 year, last 3 to 2 years, and before 3 years.

To estimate the sound level of each item, you write:

1 to refer to the sound level of normal conversation,
2 to refer to the sound level of loud conversation,
3 to refer to the sound level at which you must shout over a distance of one meter (over table) to be heard,
4 to refer to the normal sound level of pub music that makes you shout to be heard over a near distance,
5 to refer to loud pub music that makes communication almost impossible.

Activity		estimation of loudness (fill in 1 to 5)	hours/week	#weeks in the past 52 weeks (1 year)	#weeks from 3 years ago to 1 year ago (total 104 weeks)	weeks/year x # years from birth till 3 years ago	% time using hearing protection device
Leisure time noise part 1	Listening to a home stereo/radio via loudspeakers						
	Listening to a home stereo/radio via headphones						
	Watching television						
	Playing computer games						
	Watching movies or the theatre						
	Going to nightclubs or pubs						
	attending classical music concerts and festivals						
	attending rock/pop music concerts and festivals						
	Attending motor sports						
	Using tools indoors (name machine)						
	Using tools outdoors (name machine)						
	Attending or participating indoor sports events (e.g. ice-hockeys)						
	Attending or participating outdoor sports events (e.g. football)						
	Exercising to music						
	Attend car/truck races						
	Attend any other events with amplified public announcement (PA)/music systems						
	Ride in or pilot small aircraft/private airplanes						
	Other, name:						
	Other, name:						
Leisure time noise part 2 (music major student, do not answer)	Playing in a band or orchestra						
	Practicing a musical instrument (name instrument: _________________)						
Occupational noise	Any work involving power tools, chainsaws, or other shop tools						
	Any work using drive heavy equipment or loud machinery (such as tractors, trucks, or farming or lawn equipment like mowers/leaf blowers)						
Activity		#shots/session	sessions/week	#weeks in the past 52 weeks (1 year)	#weeks from 3 to 1 year ago (total 104 weeks)	# years before 3 years ago	% time using hearing protection device
Impulse noise	Fireworks						
	Use of firearms, such as rifles, pistols, shotguns, etc. for hunting, sports shooting						

Appendix Figure 3. General noise exposure questionnaire

Dental equipment use questionnaire

Subject ID: ______________ Date: ______________

Category	Equipment	1st year			2nd year			3rd year			4th year		
		# hours/week using this device	# weeks	% time HPD	# hours/week using this device	# weeks	% time HPD	# hours/week using this device	# weeks	% time HPD	# hours/week using this device	# weeks	% time HPD
Student handpieces	Turbine (high speed)												
	Contra-angle hand-piece												
	Straight hand-piece												
Clinic handpieces (not used until 3rd year)	Ultrasonic scaler												
	Turbine												
	Contra-angle hand-piece												
	Straight hand-piece												
Lab-and-clinic equipment	Polishing equipment												
	Vibrating equipment												
	Lathe equipment 3000												
	Stone trimmer												
	Suction pump (low volume)												
	Suction pump (high volume)												
	Air-water syringe												
	Sandblaster												

Appendix Figure 4. **Dental noise exposure questionnaire**

Appendix C Diagnostic criteria and statistics of the hierarchical multiple linear regression model

This section includes the diagnostic criteria and the resulting statistics for building the HMLR model in section 5.2.

Appendix Table 1 listed the statistical tests for these assumptions, the proposed criteria of the assumptions being met, and the solution when an assumption was violated (Cohen, Cohen, West & Aiken, 2003).

Appendix Table 1. Assumptions of multiple linear regression

Assumptions	Procedures and Measures	Proposed criteria of assumptions being met	Solution when an assumption is violated
Linearity	1. Plot the standardized residuals against the predicted values 2. Plot the standardized residuals against each IV	The data points should scatter randomly around a horizontal line at error = 0	Exclude or transform IVs Polynomial regression
Normality of residuals	1. Histogram of standardized residuals	Histogram of standardized residuals show mean, kurtosis and skewness approximately equal to zero;	Examine subset data pattern Examine outlier Transformation of DV
	2. Q-Q plot of standardized residuals	Q-Q plot: data point closely hug around the straight line of expected normal distribution;	
	3. Shapiro-Wilk test	S-W test $P > 0.05$	

Appendix Table 1 continued

Homoscedasticity	1. Plot the standardized residuals against the predicted values	We should observe a uniform thickness on the scatter cloud around the horizontal line at a residual of zero.	Transformation of DV Weighted least squares method
	2. Plot the standardized residuals against each IV	Residuals have constant variance across values of the IV	
	3. Breusch-Pagan test (Breusch & Pagan, 1979)	B-P test P > 0.05	
Multicollinearity	Variance inflation factor (VIF)	VIF above 10 indicates severe multicollinearity	Average data from IVs of the same measure Factor analysis Drop unexplainable IVs
Identifying outliers and influential data point	1. Generate univariate plots including histograms, boxplots and detrended Q-Q plot, as well as bivariate scatterplots of the IV and the DV	Data point away from the trend line of the rest of the data based on visual inspection	Input the correct data for the outlier or report the model with and without the outlier
	2. Centered Leverage for outlying X observations	2k/n for large N 3k/n for small N	
	3. Externally studentized residuals	±3.0 or ±4.0 for large N; ±2.0 for small N	
	4. Compute DFFITS	For large N: > $\left\| \frac{\pm 2}{\sqrt{(k+1)/n}} \right\|$ For small N: > ± 1.0	
	5. Compute DFBETAS	For large N: $\left\| \frac{\pm 2}{\sqrt{n}} \right\|$ For small N: > ±1.0	

Appendix C.1 Assumption check for outcome variables of LPF unvoiced speech tests

Assumption check was conducted for the performance of the LPF unvoiced speech recognition in SMN (LPSMN) as well as of the silence interrupted LPF unvoiced speech recognition at 48 Hz (LP48). Statistics of normality test is listed in Appendix Figure 5. The assumption of normality was not violated for either outcome variables.

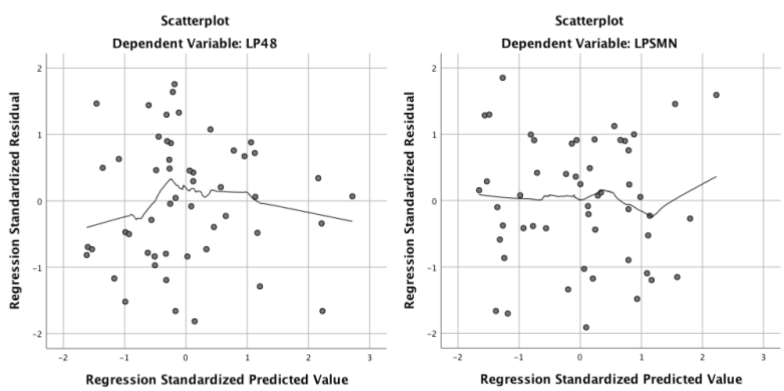
Tests of Normality

	1:ND, 2:D	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
LPSMN	1	.089	27	.200*	.972	27	.648
	2	.117	25	.200*	.986	25	.974
LP48	1	.140	27	.187	.924	27	.050
	2	.192	25	.018	.935	25	.114

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Appendix Figure 5. SPSS output for test of normality for LPSMN and LP48



Appendix Figure 6. SPSS plot of standardized residuals against standardized predicted value for homoscedasticity. Left: LP48. Right: LPSMN

Statistics of homoscedasticity are displayed in Appendix Figure 6, which shows that for both outcome variables, homoscedasticity was not violated.

Statistics of multicollinearity are displayed in Appendix Figure 7, which shows that for both outcome variables, there was no severe issue of multicollinearity.

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	Age	.988	1.012
	Years of Music	.988	1.012
2	Age	.902	1.109
	Years of Music	.957	1.045
	LifetimeDNE	.946	1.057
	LifeNDNE	.927	1.079
3	Age	.898	1.114
	Years of Music	.871	1.148
	LifetimeDNE	.913	1.095
	LifeNDNE	.874	1.144
	LEipsiAR	.835	1.198
4	Age	.795	1.258
	Years of Music	.783	1.278
	LifetimeDNE	.888	1.126
	LifeNDNE	.853	1.172
	LEipsiAR	.744	1.344
	LEDP0.5_2.8K	.752	1.330
	LEDP3.5_8.8K	.904	1.106

a. Dependent Variable: LP48

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	Age	.988	1.012
	Years of Music	.988	1.012
2	Age	.902	1.109
	Years of Music	.957	1.045
	LifetimeDNE	.946	1.057
	LifeNDNE	.927	1.079
3	Age	.898	1.114
	Years of Music	.871	1.148
	LifetimeDNE	.913	1.095
	LifeNDNE	.874	1.144
	LEipsiAR	.835	1.198
4	Age	.795	1.258
	Years of Music	.783	1.278
	LifetimeDNE	.888	1.126
	LifeNDNE	.853	1.172
	LEipsiAR	.744	1.344
	LEDP0.5_2.8K	.752	1.330
	LEDP3.5_8.8K	.904	1.106

a. Dependent Variable: LPSMN

Appendix Figure 7. SPSS output for multicollinearity. Left: LP48. Right: LPSMN. LifetimeDNE: dental noise exposure. LifeNDNE: non-dental noise exposure. LEipsiAR: left ear ipsilateral acoustic reflex. LEDP0.5_2.8K: low frequency DPOAE. LEDP3.5_8.8K: high frquency DPOAE.

Appendix D Summary of findings and potentially affected auditory processing from animal models of NTNE

Appendix Table 2. Summary of findings and potentially affected basic auditory processing and temporal processing from animal models of NTNE

Source study *	Paradigm	Frequency region	Threshold	Supra-threshold intensity	System noise	Temporal processing		
						Onset phase locking (to pulse train)	Excitation duration	Envelope phase locking (temporal envelope processing)
1	Moderate-level prolonged sharp-edged broadband noise (BBN)	At NB	◎ no change	◎ no change	× increased SFR, assume enhanced		× missing	× assume abnormal
		Outside NB	◎ no change		× increased SFR, assume enhanced		× lengthened	× assume abnormal
3	Moderate-level prolonged intermittent sharp-edged BBN	At NB	◎ no change					
		Outside NB	◎ no change					
5, 6	Moderate-level prolonged factory noise	At NB	◎ no change		× increased SFR, assume enhanced		◎ no change	◎ assume normal
		Outside NB	◎ no change		× increased SFR, assume enhanced		◎ no change	◎ assume normal
9, 21, 22	Moderate-level prolonged flat spectrum BBN	Low-frequency end of NB	◎ no change		× reduced inhibitory neurons, assume enhanced	× AC not follow fast-rate pulses		
		High-frequency end of NB						
2, 5, 6, 7, 8	Moderate-level prolonged sharp-edged octave-band noise (OBN)	At NB	◎ no change	◎ AC × ANF	× increased SFR, assume enhanced		× missing	× assume abnormal
		Outside NB	◎ no change	◎ AC × ANF	◎ no change			
6	Moderate-level prolonged sloped OBN	At NB	◎ no change	◎ no change				
		Outside NB	◎ no change	◎ no change				

Appendix Table 2 continued

2, 4, 5	Moderate-level prolonged sharp-edged narrow band noise pairs	At & between NBs	◎ no change	◎ no change	◎ no change			
		Outside NBs	◎ no change	◎ no change	× increased SFR, assume enhanced			
13 – 20	Intense short-term OBN (chronic exam)	At NB	◎ no change	× ANF × IC	× increased SFR, assume enhanced		◎ no change	◎ no change
		Above NB	◎ no change	× ANF ◎ IC	× increased SFR, assume enhanced			× reduced
10 - 12, 21	Intense short-term BBN (chronic exam)	At NB	◎ no change	× ANF		× ANF not follow fast-rate clicks		× assume abnormal
		Above NB	◎ no change	× ANF				× assume normal

* Numeric codes of source studies: 1, Pienkowski & Eggermont, 2009; 2, Pienkowski & Eggermont, 2010a; 3, Pienkowski & Eggermont, 2010b; 4, Pienkowski, Munguia & Eggermont, 2011; 5, Munguia, Pienkowski & Eggermont, 2013; 6, Pienkowski, Munguia & Eggermont, 2013; 7, Sheppard et al., 2017; 8, Pienkowski, 2018; 9, Zhou & Merzenich, 2012; 10, Liu, Wang, Shi, Almuklass, He, et al. 2012; 11, Shi, Liu, He, Guo et al., 2013; 12, Shi, Chang, Li., Aiken, Liu, & Wang, 2016; 13, Hickox & Liberman, 2014; 14, Liberman, Suzuki & Liberman, 2015; 15, Kujawa & Liberman, 2009; 16, Wang & Ren, 2012; 17, Lin, Furman, Kujawa & Liberman, 2011; 18, Furman, Kujawa & Liberman, 2013; 19, Lobarinas, Spankovich & Le Prell, 2017; 20, Shaheen, Valero & Liberman, 2015; 21, Zhou et al., 2011; 22, Kamal et al., 2013; 21, Hesse et al., 2016

Appendix Table 3. Summary of findings and potentially affected spectral processing from animal models of NTNE (see Table 20 for source study)

Source study	Paradigm	Frequency region	Spectral processing			
			Place coding/tonotopic map	TFS coding (pitch)	Spectral receptive field	Spectral shape processing
1	Moderate-level prolonged sharp-edged broadband noise (BBN)	At NB	× missing	× assume abnormal	× suppressed	× assume abnormal
		Outside NB	× enhanced		× enhanced	× assume abnormal
3	Moderate-level prolonged intermittent sharp-edged BBN	At NB	× missing	× assume abnormal	× weakened	× assume abnormal
		Outside NB			× enhanced	× assume abnormal
5, 6	Moderate-level prolonged factory noise	At NB	◎ no change		◎ not suppressed	
		Outside NB	◎ no change		× enhanced	× assume abnormal
9, 21, 22	Moderate-level prolonged flat spectrum BBN	Low-frequency end of NB	× missing	× abnormal	× broader FTC	× assume abnormal
		High-frequency end of NB	× enhanced		× broader FTC	× assume abnormal
2, 5, 6, 7, 8	Moderate-level prolonged sharp-edged octave-band noise (OBN)	At NB	× missing	× assume abnormal	× AC, missing IC, enhanced	× assume abnormal
		Outside NB	× suppressed		× AC, missing near NB IC, enhanced	× assume abnormal
6	Moderate-level prolonged sloped OBN	At NB	× weakened			
		Outside NB	× weakened in 2 oct			
2, 4, 5	Moderate-level prolonged sharp-edged narrow band noise pairs	At & between NBs	× missing	× assume abnormal	× missing	× assume abnormal
		Outside NBs	× enhanced		× enhanced	× assume abnormal
13 – 20	Intense short-term OBN (chronic exam)	At NB			× reduced LSR ANF	× assume abnormal
		Above NB				
10 - 12, 21	Intense short-term BBN (chronic exam)	At NB			× assume reduced LSR ANF	× assume abnormal

Bibliography

Ahmed, H. O., & Ali, W. J. (2017). Noise levels, noise annoyance, and hearing-related problems in a dental college. Archives of environmental & occupational health, 72(3), 159-165.

Ai, Z. T., Mak, C. M., & Wong, H. M. (2017). Noise level and its influences on dental professionals in a dental hospital in Hong Kong. Building Services Engineering Research and Technology, 38(5), 522-535.

Al-Dujaili, M., Thomson, W., Meldrum, R., & Al-Ani, A. (2014). Noise levels in dental school clinics. NZ Dent J, 110(3), 105-8.

Albert, D. G., & Decato, S. N. (2017). Acoustic and seismic ambient noise measurements in urban and rural areas. Applied Acoustics, 119, 135-143.

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

ANSI S3.5 (1997). Methods for Calculation of the Speech Intelligibility Index. New York: American National Standards Institute.

Arai, T., Pavel, M., Hermansky, H., & Avendano, C. (1999). Syllable intelligibility for temporally filtered LPC cepstral trajectories. The Journal of the Acoustical Society of America, 105(5), 2783-2791.

Auerbach, B. D., Rodrigues, P. V., & Salvi, R. J. (2014). Central gain control in tinnitus and hyperacusis. Frontiers in neurology, 5, 206.

Bahannan, S., El-Hamid, A. A., & Bahnassy, A. (1993). Noise level of dental handpieces and laboratory engines. The Journal of prosthetic dentistry, 70(4), 356-360.

Bashford Jr, J. A., Meyers, M. D., Brubaker, B. S., & Warren, R. M. (1988). Illusory continuity of interrupted speech: speech rate determines durational limits. The Journal of the Acoustical Society of America, 84(5), 1635-1638.

Bashford, J. A., Riener, K. R., & Warren, R. M. (1992). Increasing the intelligibility of speech through multiple phonemic restorations. Perception & psychophysics, 51(3), 211-217.

Bashford, J. A., Warren, R. M., & Brown, C. A. (1996). Use of speech-modulated noise adds strong “bottom-up” cues for phonemic restoration. Perception & Psychophysics, 58(3), 342-350.

Benkí, J. R. (2001). Place of articulation and first formant transition pattern both affect perception of voicing in English. Journal of Phonetics, 29(1), 1-22.

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. Journal of Neuroscience, 35(5), 2161-2172.

Bhargava, P., & Başkent, D. (2012). Effects of low-pass filtering on intelligibility of periodically interrupted speech. The Journal of the Acoustical Society of America, 131(2), EL87-EL92.

Bhargava, P., Gaudrain, E., & Başkent, D. (2016). The intelligibility of interrupted speech: Cochlear implant users and normal hearing listeners. Journal of the Association for Research in Otolaryngology, 17(5), 475-491.

Bhatt, I. (2017). Increased medial olivocochlear reflex strength in normal-hearing, noise-exposed humans. PloS one, 12(9).

Bhatt, I. S., & Wang, J. (2019). Evaluation of dichotic listening performance in normal-hearing, noise-exposed young females. Hearing research, 380, 10-21.

Bořil, H., Boyraz, P., & Hansen, J. H. (2012). Towards multimodal driver's stress detection. In Digital signal processing for in-vehicle systems and safety (pp. 3-19). Springer, New York, NY.

Bramhall, N. F., Konrad-Martin, D., McMillan, G. P., & Griest, S. E. (2017). Auditory brainstem response altered in humans with noise exposure despite normal outer hair cell function. Ear and Hearing, 38(1), e1-e12.

Bramhall, N. F., Niemczak, C. E., Kampel, S. D., Billings, C. J., & McMillan, G. P. (2020). Evoked Potentials Reveal Noise Exposure–Related Central Auditory Changes Despite Normal Audiograms. American Journal of Audiology, 1-13.

Burk, A., & Neitzel, R. L. (2016). An exploratory study of noise exposures in educational and private dental clinics. Journal of occupational and environmental hygiene, 13(10), 741-749.

Cai, S., Ma, W. L. D., & Young, E. D. (2009). Encoding intensity in ventral cochlear nucleus following acoustic trauma: implications for loudness recruitment. Journal of the Association for Research in Otolaryngology, 10(1), 5-22.

Can, A., Leclercq, L., Lelong, J., & Botteldooren, D. (2010). Traffic noise spectrum analysis: Dynamic modeling vs. experimental observations. Applied Acoustics, 71(8), 764-770.

Chait, M., Greenberg, S., Arai, T., Simon, J. Z., & Poeppel, D. (2015). Multi-time resolution analysis of speech: evidence from psychophysics. Frontiers in neuroscience, 9, 214.

Chatterjee, M., Peredo, F., Nelson, D., & Başkent, D. (2010). Recognition of interrupted sentences under conditions of spectral degradation. The Journal of the Acoustical Society of America, 127(2), EL37-EL41.

Chen, W. L., Chen, C. J., Yeh, C. Y., Lin, C. T., Cheng, H. C., & Chen, R. Y. (2013). Workplace noise exposure and its consequent annoyance to dentists. Journal of Experimental & Clinical Medicine, 5(5), 177-180.

Chen, H., Shi, L., Liu, L., Yin, S., Aiken, S., & Wang, J. (2019). Noise-induced cochlear synaptopathy and signal processing disorders. Neuroscience, 407, 41-52.

Chi, T., Ru, P., & Shamma, S. A. (2005). Multiresolution spectrotemporal analysis of complex sounds. The Journal of the Acoustical Society of America, 118(2), 887-906.

Choosong, T., Kaimook, W., Tantisarasart, R., Sooksamear, P., Chayaphum, S., Kongkamol, C., ... & Phakthongsuk, P. (2011). Noise exposure assessment in a dental school. Safety and health at work, 2(4), 348-354.

Cole, R. A., & Cooper, W. E. (1975). Perception of voicing in English affricates and fricatives. The Journal of the Acoustical Society of America, 58(6), 1280-1287.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., & Gerstman, L. J. (1952). Some experiments on the perception of synthetic speech sounds. The Journal of the Acoustical Society of America, 24(6), 597-606.

Da Cunha, K. F., Dos Santos, R. B., & Júnior, C. A. K. (2017). Assessment of noise intensity in a dental teaching clinic. BDJ open, 3, 17010.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. The Journal of the Acoustical Society of America, 27(4), 769-773.

Dorman, M. F., & Loizou, P. C. (1998). The identification of consonants and vowels by cochlear implant patients using a 6-channel continuous interleaved sampling processor and by normal-hearing subjects using simulations of processors with two to nine channels. Ear and hearing, 19(2), 162-166.

Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. The Journal of the Acoustical Society of America, 102(4), 2403-2411.

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. The Journal of the Acoustical Society of America, 95(2), 1053-1064.

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2003). Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing. The Journal of the Acoustical Society of America, 113(4), 2084-2094.

Eggermont, J. J. (2017). Effects of long-term non-traumatic noise exposure on the adult central auditory system. Hearing problems without hearing loss. Hearing research, 352, 12-22.

Elliott, T. M., & Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. PLoS computational biology, 5(3), e1000302.

Elmehdi, H. M. (2010). Assessing acoustic noise levels in dental clinics and its link to dental anxiety and fear among UAE population. Compress. Air, 87, 14. Fernandes et al., 2006;

Fant, G. (1970). Acoustic theory of speech production (No. 2). Walter de Gruyter.

Fernandes, J. C., Carvalho, A. P. O., Gallas, M., Vaz, P., & Matos, P. A. (2006). Noise levels in dental schools. European Journal of Dental Education, 10(1), 32-37.

Fernandez, K. A., Jeffers, P. W., Lall, K., Liberman, M. C., & Kujawa, S. G. (2015). Aging after noise exposure: acceleration of cochlear synaptopathy in “recovered” ears. Journal of Neuroscience, 35(19), 7509-7520.

Fernandez, K. A., Guo, D., Micucci, S., De Gruttola, V., Liberman, M. C., & Kujawa, S. G. (2020). Noise-induced Cochlear Synaptopathy with and Without Sensory Cell Loss. Neuroscience, 427, 43-57.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech‐reception threshold for impaired and normal hearing. The Journal of the Acoustical Society of America, 88(4), 1725-1736.

Fogerty, D., Xu, J., & Gibbs, B. E. (2016). Modulation masking and glimpsing of natural and vocoded speech during single-talker modulated noise: Effect of the modulation spectrum. The Journal of the Acoustical Society of America, 140(3), 1800-1816.

Formby, C., & Muir, K. (1988). Modulation and gap detection for broadband and filtered noise signals. The Journal of the Acoustical Society of America, 84(2), 545-550.

Frisina, R. D., Smith, R. L., & Chamberlain, S. C. (1990). Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. Hearing research, 44(2-3), 99-122.

Füllgrabe, C., Moody, M., & Moore, B. C. (2020). No evidence for a link between noise exposure and auditory temporal processing for young adults with normal audiograms. The Journal of the Acoustical Society of America, 147(6), EL465-EL470.

Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. Journal of neurophysiology, 110(3), 577-586.

George, E. L., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America, 120(4), 2295-2311.

Gilbert, G., & Lorenzi, C. (2010). Role of spectral and temporal cues in restoring missing speech information. The Journal of the Acoustical Society of America, 128(5), EL294-EL299.

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. Hearing research, 47(1-2), 103-138.

Goswami, M., Singh, D., Vashist, B., & Marwaha, S. (2017). Noise levels and sound pollution associated with various operative procedures and equipments in a pediatric dental environment—A clinical study. Journal of oral biology and craniofacial research, 7(3), 182-187.

Grose, J. H., Mamo, S. K., & Hall III, J. W. (2009). Age effects in temporal envelope processing: Speech unmasking and auditory steady state responses. Ear and hearing, 30(5), 568.

Grossberg, S., & Kazerounian, S. (2011). Laminar cortical dynamics of conscious speech perception: Neural model of phonemic restoration using subsequent context in noise. The Journal of the Acoustical Society of America, 130(1), 440-460.

Guest, H., Munro, K. J., Prendergast, G., Millman, R. E., & Plack, C. J. (2018). Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. Hearing research, 364, 142-151.

Hall, J. W. (2000). Handbook of otoacoustic emissions. Cengage Learning.

Hall, J. W., Buss, E., Grose, J. H., & Roush, P. A. (2012). Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. Ear and Hearing, 33(3), 340.

Henry, J. A., Griest, S., Zaugg, T. L., Thielman, E., Kaelin, C., Galvez, G., & Carlson, K. F. (2015). Tinnitus and hearing survey: a screening tool to differentiate bothersome tinnitus from hearing difficulties. American journal of audiology, 24(1), 66-77.

Hesse, L. L., Bakay, W., Ong, H. C., Anderson, L., Ashmore, J., McAlpine, D., ... & Schaette, R. (2016). Non-monotonic relation between noise exposure severity and neuronal hyperactivity in the auditory midbrain. Frontiers in neurology, 7, 133.

Hickox, A. E., & Liberman, M. C. (2014). Is noise-induced cochlear neuropathy key to the generation of hyperacusis or tinnitus?. Journal of Neurophysiology, 111(3), 552-564.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. The Journal of the Acoustical Society of America, 97(5), 3099-3111.

Hou, Z., & Pavlovic, C. V. (1994). Effects of temporal smearing on temporal resolution, frequency selectivity, and speech intelligibility. The Journal of the Acoustical Society of America, 96(3), 1325-1340.

Houtgast, T., & Steeneken, H. J. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. The Journal of the Acoustical Society of America, 77(3), 1069-1077.

Huggins, A. W. F. (1975). Temporally segmented speech. Perception & Psychophysics, 18(2), 149-157.

Ihlefeld, A., Deeks, J. M., Axon, P. R., & Carlyon, R. P. (2010). Simulations of cochlear-implant speech perception in modulated and unmodulated noise. The Journal of the Acoustical Society of America, 128(2), 870-880.

Irino, T., Aoki, Y., Kawahara, H., & Patterson, R. D. (2012). Comparison of performance with voiced and whispered speech in word recognition and mean-formant-frequency discrimination. Speech Communication, 54(9), 998-1013.

Jin, S. H., & Nelson, P. B. (2006). Speech perception in gated noise: The effects of temporal resolution. The Journal of the Acoustical Society of America, 119(5), 3097-3108.

Jin, S. H., & Nelson, P. B. (2010). Interrupted speech perception: The effects of hearing sensitivity and frequency resolution. The Journal of the Acoustical Society of America, 128(2), 881-889.

Johnson, T. A., Cooper, S., Stamper, G. C., & Chertoff, M. (2017). Noise Exposure Questionnaire: A tool for quantifying annual noise exposure. Journal of the American Academy of Audiology, 28(1), 14-35.

Jokitulppo, J., Toivonen, M., & Bjoörk, E. (2006). Estimated leisure-time noise exposure, hearing thresholds, and hearing symptoms of Finnish conscripts. Military medicine, 171(2), 112-116.

Kadanakuppe, S., Bhat, P. K., Jyothi, C., & Ramegowda, C. (2011). Assessment of noise levels of the equipments used in the dental teaching institution, Bangalore. Indian Journal of Dental Research, 22(3), 424.

Kaltenbach, J. A., Zacharek, M. A., Zhang, J., & Frederick, S. (2004). Activity in the dorsal cochlear nucleus of hamsters previously tested for tinnitus following intense tone exposure. Neuroscience letters, 355(1), 121-125.

Kamal, B. S., Holman, C., & de Villers-Sidani, E. (2013). Shaping the aging brain: role of auditory input patterns in the emergence of auditory cortical impairments. Frontiers in systems neuroscience, 7, 52.

Kawahara, H., & Irino, T. (2005). Underlying principles of a high-quality speech manipulation system STRAIGHT and its application to speech segregation. In Speech separation by humans and machines (pp. 167-180). Springer, Boston, MA.

Kawahara, H., Takahashi, T., Morise, M., & Banno, H. (2009). Development of exploratory research tools based on TANDEM-STRAIGHT. In Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference (pp. 111-120). Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference, International Organizing Committee.

Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: cochlear nerve degeneration after “temporary” noise-induced hearing loss. Journal of Neuroscience, 29(45), 14077-14085.

Kumar, U. A., Ameenudin, S., & Sangamanatha, A. V. (2012). Temporal and speech processing skills in normal hearing individuals exposed to occupational noise. Noise and Health, 14(58), 100.

Lau, C., Zhang, J. W., McPherson, B., Pienkowski, M., & Wu, E. X. (2015). Long-term, passive exposure to non-traumatic acoustic noise induces neural adaptation in the adult rat medial geniculate body and auditory cortex. Neuroimage, 107, 1-9.

Levitt, H. C. C. H. (1971). Transformed up-down methods in psychoacoustics. The Journal of the Acoustical society of America, 49(2B), 467-477.

Li, Y., Ropp, T. J. F., May, B. J., & Young, E. D. (2015). Dorsal cochlear nucleus of the rat: Representation of complex sounds in ears damaged by acoustic trauma. Journal of the Association for Research in Otolaryngology, 16(4), 487-505.

Liberman, M. C., & Dodds, L. W. (1984). Single-neuron labeling and chronic cochlear pathology. II. Stereocilia damage and alterations of spontaneous discharge rates. Hearing research, 16(1), 43-53.

Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H., & Maison, S. F. (2016). Toward a differential diagnosis of hidden hearing loss in humans. PloS one, 11(9), e0162726.

Liberman, A. M., Delattre, P. C., & Cooper, F. S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. Language and speech, 1(3), 153-167.

Liberman, M. C. (1978). Auditory‐nerve response from cats raised in a low‐noise chamber. The Journal of the Acoustical Society of America, 63(2), 442-455.

Liberman, L. D., Suzuki, J., & Liberman, M. C. (2015). Dynamics of cochlear synaptopathy after acoustic overexposure. Journal of the Association for Research in Otolaryngology, 16(2), 205-219.

Lin, H. W., Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2011). Primary neural degeneration in the Guinea pig cochlea after reversible noise-induced threshold shift. Journal of the Association for Research in Otolaryngology, 12(5), 605-616.

Liu, L., Wang, H., Shi, L., Almuklass, A., He, T., Aiken, S., ... & Wang, J. (2012). Silent damage of noise on cochlear afferent innervation in guinea pigs and the impact on temporal processing. PLoS One, 7(11), e49550.

Lobarinas, E., Spankovich, C., & Le Prell, C. G. (2016). Evidence of “hidden hearing loss” following noise exposures that produce robust TTS and ABR wave-I amplitude reductions. Hearing Research.

Loizou, P. C., Dorman, M., & Tu, Z. (1999). On the number of channels needed to understand speech. The Journal of the Acoustical Society of America, 106(4), 2097-2103.

Ma, W. L. D., & Young, E. D. (2006). Dorsal cochlear nucleus response properties following acoustic trauma: response maps and spontaneous activity. Hearing research, 216, 176-188.

May, B. J., Huang, A., Le Prell, G. S., & Hienz, R.D. (1996). Vowel formant frequency discrimination in cats: comparison of auditory nerve representations and psychophysical thresholds. Auditory neuroscience, 3(2), 135.

Miller, G. A., & Licklider, J. C. (1950). The intelligibility of interrupted speech. The Journal of the Acoustical Society of America, 22(2), 167-173.

Mojarad, F., Massum, T., & Samavat, H. A. M. I. D. (2009). Noise levels in dental offices and laboratories in Hamedan, Iran. Journal of Dentistry of Tehran University of Medical Sciences, 181-186.

Moore, B. C., & Gockel, H. E. (2011). Resolvability of components in complex tones and implications for theories of pitch perception. Hearing research, 276(1-2), 88-97.

Moore, B. C., Glasberg, B. R., Plack, C. J., & Biswas, A. K. (1988). The shape of the ear's temporal window. The Journal of the Acoustical Society of America, 83(3), 1102-1116.

Moore, B. C. (2019). The roles of temporal envelope and fine structure information in auditory perception. Acoustical Science and Technology, 40(2), 61-83.

Munguia, R., Pienkowski, M., & Eggermont, J. J. (2013). Spontaneous firing rate changes in cat primary auditory cortex following long-term exposure to non-traumatic noise: Tinnitus without hearing loss?. Neuroscience letters, 546, 46-50.

Myers, J., John, A. B., Kimball, S., & Fruits, T. (2016). Prevalence of tinnitus and noise-induced hearing loss in dentists. Noise & health, 18(85), 347.

Narayan, S. S., Temchin, A. N., Recio, A., & Ruggero, M. A. (1998). Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae. Science, 282(5395), 1882-1884.

Nassiri, P., Golbabai, F., & Mahmoudi, M. (1993). The effect of noise-induced hearing loss on dentists. Medical Journal of The Islamic Republic of Iran (MJIRI), 7(2), 83-86.

Neitzel R, Sexias NS, Camp J, Yost M: An assessment of occupational noise exposure levels in four construction trades. Am Ind Hyg Assoc J 1999;60(6):807-817.

Nelson, P. B., & Jin, S. H. (2004). Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners. The Journal of the Acoustical Society of America, 115(5), 2286-2294.

Nelson, P. B., Jin, S. H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. The Journal of the Acoustical Society of America, 113(2), 961-968.

National Institute for Occupational Safety and Health. (1998). Criteria for a recommended standard: Occupational noise exposure, revised criteria 1998. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention.

Noreña, A. J. (2015). Revisiting the cochlear and central mechanisms of tinnitus and therapeutic approaches. Audiology and Neurotology, 20(Suppl. 1), 53-59.

Oppenheim, A. V., & Schafer, R. W. (2010). Discrete-Time Signal Processing, Third Edition. Pearson Education, Inc.

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. The Journal of the Acoustical Society of America, 59(3), 640-654.

Peters, R. W., Moore, B. C., & Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. The Journal of the Acoustical Society of America, 103(1), 577-587.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. The Journal of the Acoustical Society of America, 24(2), 175-184.

Pickles, J. O. (2013). An introduction to the physiology of hearing. Leiden: Brill

Pienkowski, M., & Eggermont, J. J. (2009). Long-term, partially-reversible reorganization of frequency tuning in mature cat primary auditory cortex can be induced by passive exposure to moderate-level sounds. Hearing research, 257(1), 24-40.

Pienkowski, M., & Eggermont, J. J. (2010a). Passive exposure of adult cats to moderate-level tone pip ensembles differentially decreases AI and AII responsiveness in the exposure frequency range. Hearing research, 268(1), 151-162.

Pienkowski, M., & Eggermont, J. J. (2010b). Intermittent exposure with moderate-level sound impairs central auditory function of mature animals without concomitant hearing loss. Hearing research, 261(1), 30-35.

Pienkowski, M., Munguia, R., & Eggermont, J. J. (2011). Passive exposure of adult cats to bandlimited tone pip ensembles or noise leads to long-term response suppression in auditory cortex. Hearing research, 277(1), 117-126.

Pienkowski, M., Munguia, R., & Eggermont, J. J. (2013). Effects of passive, moderate-level sound exposure on the mature auditory cortex: spectral edges, spectrotemporal density, and real-world noise. Hearing research, 296, 121-130.

Plack, C. J., & Moore, B. C. (1990). Temporal window shape as a function of frequency and level. The Journal of the Acoustical Society of America, 87(5), 2178-2187.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. Speech communication, 41(1), 245-255.

Powers, G. L., & Wilcox, J. C. (1977). Intelligibility of temporally interrupted speech with and without intervening noise. The Journal of the Acoustical Society of America, 61(1), 195-199.

Prendergast, G., Guest, H., Munro, K. J., Kluk, K., Léger, A., Hall, D. A., ... & Plack, C. J. (2017a). Effects of noise exposure on young adults with normal audiograms I: Electrophysiology. Hearing Research, 344, 68-81.

Prendergast, G., Millman, R. E., Guest, H., Munro, K. J., Kluk, K., Dewey, R. S., ... & Plack, C. J. (2017b). Effects of noise exposure on young adults with normal audiograms II: Behavioral measures. Hearing research, 356, 74-86.

Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. The Journal of the Acoustical Society of America, 114(1), 446-454.

Qsaibati, M. L., & Ibrahim, O. (2014). Noise levels of dental equipment used in dental college of Damascus University. Dental research journal, 11(6), 624.

Paul, B. T., Bruce, I. C., & Roberts, L. E. (2018). Envelope following responses, noise exposure, and evidence of cochlear synaptopathy in humans: Correction and comment. The Journal of the Acoustical Society of America, 143(6), EL487-EL489.

Paul, B. T., Waheed, S., Bruce, I. C., & Roberts, L. E. (2017). Subcortical amplitude modulation encoding deficits suggest evidence of cochlear synaptopathy in normal-hearing 18–19 year olds with higher lifetime noise exposure. The Journal of the Acoustical Society of America, 142(5), EL434-EL440.

Rämö, J., Välimäki, V., & Tikander, M. (2013, May). Perceptual headphone equalization for mitigation of ambient noise. In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 724-728). IEEE.

Reiss, L. A., Ramachandran, R., & May, B. J. (2011). Effects of signal level and background noise on spectral representations in the auditory nerve of the domestic cat. Journal of the Association for Research in Otolaryngology, 12(1), 71-88.

Rhode, W. S., & Greenberg, S. (1994). Encoding of amplitude modulation in the cochlear nucleus of the cat. Journal of neurophysiology, 71(5), 1797-1825.

Rhode, W. S., Oertel, D., & Smith, P. H. (1983). Physiological response properties of cells labeled intracellularly with horseradish peroxidase in cat ventral cochlear nucleus. Journal of Comparative Neurology, 213(4), 448-463.

Rhode, W. S. (1999). Vertical cell responses to sound in cat dorsal cochlear nucleus. Journal of neurophysiology, 82(2), 1019-1032.

Rodríguez Valiente, A., Trinidad, A., García Berrocal, J., Górriz, C., & Ramirez Camacho, R. (2014). Extended high-frequency (9–20 kHz) audiometry reference thresholds in 645 healthy subjects. International journal of audiology, 53(8), 531-545.

Hosford-Dunn, H., Roeser, R. J., & Valente, M. (Eds.). (2007). Audiology: diagnosis. Thieme.

Ropp, T. J. F., Tiedemann, K. L., Young, E. D., & May, B. J. (2014). Effects of unilateral acoustic trauma on tinnitus-related spontaneous activity in the inferior colliculus. Journal of the Association for Research in Otolaryngology, 15(6), 1007-1022.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 336(1278), 367-373.

Rothauser, E. H. (1969). IEEE recommended practice for speech quality measurements. IEEE Trans. on Audio and Electroacoustics, 17, 225-246.

Ruggero, M. A., Narayan, S. S., Temchin, A. N., & Recio, A. (2000). Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chinchilla. Proceedings of the National Academy of Sciences, 97(22), 11744-11750.

Sachs, M. B., & Young, E. D. (1979). Encoding of steady‐state vowels in the auditory nerve: representation in terms of discharge rate. The Journal of the Acoustical Society of America, 66(2), 470-479.

Sachs, M. B., & Abbas, P. J. (1974). Rate versus level functions for auditory‐nerve fibers in cats: tone‐burst stimuli. The Journal of the Acoustical Society of America, 56(6), 1835-1847.

Saunders, J. C., Dear, S. P., & Schneider, M. E. (1985). The anatomical consequences of acoustic injury: a review and tutorial. The Journal of the Acoustical Society of America, 78(3), 833-860.

Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. Journal of Neuroscience, 33(34), 13686-13694.

Setcos, J. C., & Mahyuddin, A. (1998). Noise levels encountered in dental clinical and laboratory practice. International journal of prosthodontics, 11(2).

Shafiro, V., Fogerty, D., Smith, K., & Sheft, S. (2018). Perceptual Organization of Interrupted Speech and Text. Journal of Speech, Language, and Hearing Research, 61(10), 2578-2588.

Shafiro, V., Sheft, S., & Risley, R. (2011). Perception of interrupted speech: Effects of dual-rate gating on the intelligibility of words and sentences. The Journal of the Acoustical Society of America, 130(4), 2076-2087.

Shaheen, L. A., Valero, M. D., & Liberman, M. C. (2015). Towards a diagnosis of cochlear neuropathy with envelope following responses. Journal of the Association for Research in Otolaryngology, 16(6), 727-745.

Shailer, M. J., & Moore, B. C. (1983). Gap detection as a function of frequency, bandwidth, and level. The Journal of the Acoustical Society of America, 74(2), 467-473.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. Science, 270(5234), 303-304.

Shehorn, J., Strelcyk, O., & Zahorik, P. (2020). Associations between speech recognition at high levels, the middle ear muscle reflex and noise exposure in individuals with normal audiograms. Hearing Research, 107982.

Sheppard, A. M., Chen, G. D., Manohar, S., Ding, D., Hu, B. H., Sun, W., ... & Salvi, R. (2017). Prolonged low-level noise-induced plasticity in the peripheral and central auditory system of rats. Neuroscience, 359, 159-171.

Shi, L., Liu, L., He, T., Guo, X., Yu, Z., Yin, S., & Wang, J. (2013). Ribbon synapse plasticity in the cochleae of Guinea pigs after noise-induced silent damage. PloS one, 8(12), e81566.

Shi, L., Chang, Y., Li, X., Aiken, S. J., Liu, L., & Wang, J. (2016). Coding deficits in noise-induced hidden hearing loss may stem from incomplete repair of ribbon synapses in the cochlea. Frontiers in neuroscience, 10, 231.

Shinn-Cunningham, B., Varghese, L., Wang, L., & Bharadwaj, H. (2017). Individual differences in temporal perception and their implications for everyday listening. In The Frequency-Following Response (pp. 159-192): Springer.

Sivonen, P., Maess, B., Lattner, S., & Friederici, A. D. (2006). Phonemic restoration in a sentence context: evidence from early and late ERP effects. Brain research, 1121(1), 177-189.

Smith, S. B., Krizman, J., Liu, C., White-Schwoch, T., Nicol, T., & Kraus, N. (2019). Investigating peripheral sources of speech-in-noise variability in listeners with normal audiograms. Hearing research, 371, 66-74.

Souza, P., & Rosen, S. (2009). Effects of envelope bandwidth on the intelligibility of sine-and noise-vocoded speech. The Journal of the Acoustical Society of America, 126(2), 792-805.

Stamper, G. C., & Johnson, T. A. (2015a). Auditory function in normal-hearing, noise-exposed human ears. Ear and hearing, 36(2), 172.

Stamper, G. C., & Johnson, T. A. (2015b). Letter to the Editor: Examination of potential sex influences in Stamper, GC & Johnson, TA (2015). Auditory function in normal-hearing, noise-exposed human ears, Ear Hear, 36, 172-184. Ear and hearing, 36(6), 738.

Stickney, G. S., Zeng, F. G., Litovsky, R., & Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. The Journal of the Acoustical Society of America, 116(2), 1081-1091.

Stone, M. A., & Moore, B. C. (2014). On the near non-existence of “pure” energetic masking release for speech. The Journal of the Acoustical Society of America, 135(4), 1967-1977.

Stone, M. A., Moore, B. C., & Greenish, H. (2008). Discrimination of envelope statistics reveals evidence of sub-clinical hearing damage in a noise-exposed population with ‘normal’ hearing thresholds. International journal of audiology, 47(12), 737-750.

Valero, M. D., Burton, J. A., Hauser, S. N., Hackett, T. A., Ramachandran, R., & Liberman, M. C. (2017). Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta). Hearing research, 353, 213-223.

Valero, M. D., Hancock, K. E., Maison, S. F., & Liberman, M. C. (2018). Effects of cochlear synaptopathy on middle-ear muscle reflexes in unanesthetized mice. Hearing research, 363, 109-118.

van der Horst, R., Leeuw, A. R., & Dreschler, W. A. (1999). Importance of temporal-envelope cues in consonant recognition. The Journal of the Acoustical Society of America, 105(3), 1801-1809.

Verhulst, S., Ernst, F., Garrett, M., & Vasilkov, V. (2018). Suprathreshold psychoacoustics and envelope-following response relations: Normal-hearing, synaptopathy and cochlear gain loss. Acta Acustica united with Acustica, 104(5), 800-803.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. The Journal of the Acoustical Society of America, 66(5), 1364-1380.

Wang, X., & Humes, L. E. (2010). Factors influencing recognition of interrupted speech. The Journal of the Acoustical Society of America, 128(4), 2100-2111.

Wang, Y., & Ren, C. (2012). Effects of repeated “benign” noise exposures in young CBA mice: shedding light on age-related hearing loss. Journal of the Association for Research in Otolaryngology, 13(4), 505-515.

Keppel, G., & Wickens, T. D. (2004). Design and analysis: a researcher's handbook. Upper Saddle River.

Yeend, I., Beach, E. F., Sharma, M., & Dillon, H. (2017). The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. Hearing research, 353, 224-236.

Yost, W. A., & Sheft, S. (1997). Temporal modulation transfer functions for tonal stimuli: Gated versus continuous conditions. Aud. Neurosci, 3, 401-414.

Young, E. D., & Brownell, W. E. (1976). Responses to tones and noise of single cells in dorsal cochlear nucleus of unanesthetized cats. Journal of Neurophysiology, 39(2), 282-300.

Yousuf, A., Ganta, S., Nagaraj, A., Pareek, S., Atri, M., Singh, K., & Sidiq, M. (2014). Acoustic noise levels of dental equipments and its association with fear and annoyance levels among patients attending different dental clinic setups in Jaipur, India. Journal of Clinical and Diagnostic Research: JCDR, 8(4), ZC29.

Zhou, X., & Merzenich, M. M. (2012). Environmental noise exposure degrades normal listening processes. Nature communications, 3, 843.

Zhou, X., Panizzutti, R., de Villers-Sidani, É., Madeira, C., & Merzenich, M. M. (2011). Natural restoration of critical period plasticity in the juvenile and adult primary auditory cortex. Journal of Neuroscience, 31(15), 5625-5634.