The effect of lifelong musicianship on age-related changes in auditory processing

by

Benjamin Rich Zendel

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Department of Psychology
University of Toronto

© Copyright by Benjamin Rich Zendel 2011
The effect of lifelong musicianship on age-related changes in auditory processing

Benjamin Rich Zendel
Doctor of Philosophy
Department of Psychology
University of Toronto
2011

Abstract

Age-related declines in hearing abilities are common and can be attributed to changes in the peripheral and central levels of the auditory system. Although central auditory processing is enhanced in younger musicians, the influence of lifelong musicianship on age-related decline in central auditory processing has not yet been investigated.

Therefore, the purpose of this dissertation was to investigate whether lifelong musicianship can mitigate age-related decline in central auditory processing. In the first experiment, age-related declines on four hearing assessments were compared between musicians and non-musicians. Speech-in-noise and gap-detection thresholds were found to decline at a slower rate in musicians, providing an increasing advantage with age. Furthermore, musicians had a lifelong advantage in detecting a mistuned harmonic, although the rate of age-related decline was similar for both musicians and non-musicians.

Importantly, there was no significant effect of musicianship on pure-tone thresholds, suggesting that lifelong musicianship can mitigate age-related decline in central but not peripheral auditory processing. To test this hypothesis, a second experiment compared auditory evoked responses (AERs) between groups of older and younger musicians and
non-musicians. Results indicated that exogenous neural activity was enhanced in musicians, but that age-related changes were similar between musicians and non-musicians. Furthermore, endogenous, attention-dependent neural activity was enhanced in older adults, suggesting a compensatory cognitive strategy. Importantly, endogenous activity was preferentially enhanced in older musicians, suggesting that lifelong musicianship enhanced cognitive processes related to auditory perception. In the final experiment, the ability to segregate simultaneous sounds was tested in older and younger musicians and non-musicians by using a mistuned harmonic paradigm, where AERs to harmonic complexes were compared to AERs when one of the harmonics was mistuned. Results indicated that musical training in older adults has little effect on early automatic registration of the mistuned harmonic. In contrast, late attention-dependent activity, associated with the perception of the mistuned harmonic as a separate sound, was influenced by musical training in older adults, suggesting that lifelong musicianship preserves or enhances cognitive components of concurrent sound segregation. In summary, musical training was found to reduce age-related decline in hearing abilities due to enhanced central processing of auditory information.
Acknowledgments

People do not exist in isolation. Working together is what allows humans to achieve greatness. Many people worked directly with me on the ideas presented here; others supported me through the years of work it took to complete this dissertation. While only my name appears at the top of this dissertation, a number of people deserve special mention for the help they provided me throughout this process.

First and foremost I would like to thank the love of my life, Karen Ann Willoughby. She is the most amazing person I have ever met, and without her this project would not have been the same. Her caring and loving attitude helped me through tough times, and, her intelligence and knowledge helped me become a better scientist. I can’t say enough nice things about Karen, but I know that I am the luckiest person alive because I get to spend my life with her.

Next I would like to thank my family. First, I want to thank my mom, Natalie Rich. Although she lives in Calgary, her near daily phone calls just to make sure I was doing well were something I looked forward to (even though I sometimes sounded less than enthusiastic). These conversations helped close the 3224 km gap between us. More importantly, they helped me understand my project better because my mom always asked what I was working on, and explaining the minutia of my daily work forced me to understand each detail in a broader context. Next, I want to thank my dad, Ivan Zendel. Obviously my dad’s educational path had a clear impact on mine because I am now the second Dr. Zendel with a PhD in psychology, although we clearly have different career trajectories. Like my mom, most of my regular interactions with my dad were over the phone. Our conversations spanned politics to travel, my mental health to my research, and I always appreciated the unique insight my dad could provide on just about any issue. My sister, brother, Sarah and Adam, also deserve a special mention. For me it has been inspiring watching both of you become interested in academics. Sarah’s interest in Nepal planted the travel bug in the entire family, and more importantly our discussions about qualitative research inspired me to step outside of my quantitative approach, and improved the interpretation of my data. My discussions with Adam about sound engineering always forced me to think outside my academic box, and his work on the development and
function of cities is something I find fascinating. Next I would like to thank my mother-in-law, Teena Willoughby. It has been wonderful getting to know you over the last few years. I always enjoy our conversations about the politics of academia, and all the inside information you gave me has helped me get through grad school. You have also welcomed me into your home and cottage, which provided a refuge from downtown Toronto. Most critical, I have thank you and Kim for raising such an amazing daughter (see above). Finally, I would like to thank all my aunts and uncles for all the help you have given me during my time in Toronto. From home-cooked meals to home furnishings you really made me feel at home. Thanks to Linda & Jacques Lallouz, Sheila & Mark Pinkus, Julia Rich, Lillie Zendel & Jean-Paul Picard, and Orvin & Lynn Zendel.

I can’t thank my supervisor, Claude Alain enough. His easy-going attitude significantly reduced my stress related to completing this project, but he always seemed to know when to push me just bit so that the project would move forward. Moreover, his deep understanding of all things related to EEG and hearing led to some great discussions about my data, and about hearing in general. He is a truly inspiring supervisor, something all academics should strive to achieve. I would also like to thank the rest of my PhD committee Bernhard Ross, and Kathy Pichora-Fuller. Both of you provided a unique perspective on my project, and your input significantly improved the quality of work that I was able to produce. Finally I would like to thank all the students and staff in the ERP/MEG labs at Rotman. I have been lucky to work with such a wonderful group of people.

In addition to these people, countless other friends, co-workers, administrators and colleagues helped me finish this project. Although too numerous to name here, I can’t ignore the important contributions these people made to my social, physical, mental, and academic well-being, so thanks!
# Table of Contents

Acknowledgments................................................................................................................ iv
Table of Contents .................................................................................................................. i
List of Figures ......................................................................................................................... ix
List of Tables ......................................................................................................................... viii

Chapter 1: ............................................................................................................................. 1
1.1 Age-related changes in auditory perception .......................................................... 1
1.2 Hearing from the cochlea to the brain .................................................................... 2
1.3 Electroencephalography and auditory evoked responses .................................. 4
1.4 Auditory Scene Analysis ......................................................................................... 6
1.5 Age-related changes in hearing abilities ............................................................... 7
   1.5.1 Spectral Resolving Power .................................................................................. 11
   1.5.2 Temporal Resolving Power ............................................................................. 12
   1.5.3 Age-related decline in speech perception .................................................. 14
1.6 Cognitive reserve ...................................................................................................... 16
1.7 Musicians vs. Non-musicians .................................................................................. 21
1.8 The current study ..................................................................................................... 24

Chapter 2: ........................................................................................................................... 26
2.1 The influence of age and lifelong musicianship on behavioral measures of auditory processing .................................................................................................................. 26
2.2 Methods ...................................................................................................................... 29
   2.2.1 Participants ..................................................................................................... 29
   2.2.2 Stimuli and Task ............................................................................................ 30
   2.2.3 Data analysis ................................................................................................. 34
2.3 Results ......................................................................................................................... 35
   2.3.1 Pure-tone thresholds ...................................................................................... 35
   2.3.2 Gap-detection thresholds .............................................................................. 36
   2.3.3 Mistuned harmonic detection thresholds .................................................. 36
   2.3.4 Speech-in-noise thresholds .......................................................................... 37
2.4 Discussion ................................................................................................................... 37

Chapter 3: ............................................................................................................................. 42
3.1 The influence of age and musical training on the auditory evoked response ......... 42
3.2 Methods ....................................................................................................................... 47
   3.2.1 Participants ..................................................................................................... 47
   3.2.2 Stimuli ............................................................................................................ 48
   3.2.3 Procedure ...................................................................................................... 49
   3.2.4 Recording of electrical brain activity ......................................................... 50
   3.2.5 Data Analysis ................................................................................................. 51
3.3 Results ......................................................................................................................... 53
   3.3.1 Effects of age and musical training on exogenous scalp recorded auditory evoked responses .......................................................... 54
   3.3.2 Effects of age and musical training on endogenous auditory evoked responses .......................................................... 55
3.4 Discussion ................................................................................................................... 59

Chapter 4: ............................................................................................................................. 65
4.1 The influence of age and musical training on concurrent sound segregation ....... 65
   4.1.1 Aging and concurrent sound segregation .................................................... 70
List of Tables

Table 2.1: Regression table for pure-tone thresholds.
Table 2.2: Regression table for gap-detection thresholds.
Table 2.3: Regression table for mistuned harmonic thresholds.
Table 2.4 Regression table for speech in noise thresholds.
List of Figures

Figure 2.1: Group mean audiograms in older (i.e., > 50 yrs) and younger (i.e., < 50 yrs) participants, plotted separately for musicians and non-musicians. Error bars represent one standard error.

Figure 2.2: Pure-tone thresholds in musicians and non-musicians as a function of age. The data shown in the figure represent individual participants’ averaged pure-tone thresholds across octave frequencies from 250 to 8000 Hz for both ears (i.e., the all-tone average). The black and grey lines reflect the best fit linear trend for the age-related changes in the thresholds of musicians and non-musicians, respectively.

Figure 2.3: Gap-detection thresholds in musicians and non-musicians as a function of age. The black and grey lines reflect the best fit linear trend for the age-related changes in the thresholds of musicians and non-musicians, respectively.

Figure 2.4: Mistuned harmonic thresholds in musicians and non-musicians as a function of age. The black and grey lines reflect the best fit linear trend for the age-related changes in the thresholds of musicians and non-musicians, respectively.

Figure 2.5: Signal-to-noise ratio (SNR) loss for understanding speech-in-noise using the QuickSIN test in musicians and non-musicians as a function of age. SNR loss is a standardized measure that represents the needed increase in SNR to understand 50% of the words in a given sentence. The black and grey lines reflect the best fit linear trend for the age-related changes the thresholds of musicians and non-musicians, respectively.
Figure 3.1: The full electrode montage used for all EEG recording. The nose can be seen at the top of the head-plot.

Figure 3.2: The auditory surrogate model. Five midline sources are shown on the sagittal MRI slice on the left. Six bilateral sources are shown on the transverse MRI slice on the right (only sources on the right are labeled).

Figure 3.3: On the left, AER butterfly plots are shown for each group in the passive listening condition. Each line represents the electrical activity from one of the 65 scalp electrodes. On the right, topographical maps for the P1, N1, P2 and N2 peaks are shown. Peak latency for P1 was 50 ms, for N1 was 100 ms, for P2 was 210 ms in older adults and 189 ms in younger adults, and for N2 was 364 ms in older adults and 323 ms in younger adults.

Figure 3.4: The exogenous auditory evoked responses shown at electrode Cz during passive listening, (low-pass filtered at 30 Hz). The P1, N1, P2 and N2 peaks are labeled.

Figure 3.5: Amplitudes for the P1 (a), N1 (b), P2 (c) and N2 (d) waves, each averaged across nine central electrodes. Amplitudes of the N1 and N2 waves have been multiplied by -1 so that the measurement is a positive value. Error bars represent one standard error.

Figure 3.6: Endogenous neural activity related to making a judgment about an auditory stimulus. At each of the eleven source locations, the source activity during passive
listening was subtracted from the source activity during active listening; the difference wave is plotted. The source model (Figure 3.2) can be seen in the upper right corner.

Figure 3.7: Source activity from each of the eleven regional sources in the surrogate model recorded during active listening.

Figure 3.8: Source activity from each of the eleven regional sources in the surrogate model recorded during passive listening.

Figure 3.9: Mean amplitude during the 550-650 ms epoch from the source in the right auditory cortex, plotted separately for active (A) and passive listening (B). Error bars represent one standard error.

Figure 4.1: Topographical map of the ORN (during both active and passive listening) and the P400. Each column illustrates the head from the left, top and right locations. Latencies for the ORN were chosen based on peak latency from the fronto-central electrode array, while latencies for the P400 were chosen based on peak latency from the right frontal electrode array. For younger musicians the ORN latency was 131 ms, and the P400 latency was 395 ms. For younger non-musicians, the ORN latency was 156 ms and the P400 latency was 395 ms. For older musicians, the ORN latency was 151 ms and the P400 latency was 395 ms. For older non-musicians, the ORN latency was 143 ms and P400 was 395 ms.

Figure 4.2: Perceptual Judgment: Percentage of responses perceived as two sounds as a function of tuning. Error bars represent one standard error.
Figure 4.3: Judgment Consistency: The consistency of responses, plotted as a function of tuning. A 1 represents perfect consistency (i.e., all responses to that stimulus type were the same), while 0 represents chance responses (i.e., half of the stimuli were perceived as a single sound, the other half perceived as concurrent sounds). Error bars represent one standard error.

Figure 4.4: Reaction time to a harmonic complex in milliseconds (ms) as a function of tuning. Error bars represent one standard error.

Figure 4.5: Auditory-evoked response to the tuned stimulus (solid line) and the 16% mistuned stimulus (dashed line) at electrode F2. The thick black line is the difference between the two responses. Responses from active listening are shown in (A), and passive listening in (B). The ORN can be seen as a negative peak ~150 ms in both active and passive listening, while the P400 can be seen as a positive peak around ~400 ms only in active listening.

Figure 4.6: The ORN amplitude, shown during active and passive listening at fronto-central electrodes. ORN amplitude was calculated as the difference in mean amplitude between the tuned and 16% mistuned conditions during the 100-190 ms epoch. ORN amplitude at fronto-central sites was multiplied by (-1) so the ORN amplitude was positive at both fronto-central and mastoid sites. Error bars represent one standard error. To calculate standard error for this graph, the difference in mean amplitude across the nine fronto-central electrodes was calculated between the tuned and 16% mistuned condition. The standard deviation of the difference was then calculated for each group. The standard
deviation was then divided by the square root of N, separately for each group, yielding a standard error.

Figure 4.7: The difference wave (tuned – 16% mistuned) shown at electrode C4. The P400 can be seen as a large positive wave in the active listening condition only that peaks around 400 ms. Note that in passive listening there is no P400 wave.

Figure 4.8: The P400 amplitude shown during active listening at right frontal electrodes during the 250-350 ms epoch and the 350-450 ms epoch. P400 amplitude was calculated as the difference in mean amplitude between the tuned and 16% mistuned stimulus. Error bars represent one standard error. To calculate standard error for this graph, the difference in mean amplitude across the seven right frontal electrodes was calculated between the tuned and 16% mistuned conditions. The standard deviation of the difference was then calculated for each group. The standard deviation was then divided by the square root of N, separately for each group, yielding a standard error.

Figure 4.9: Sensory evoked response shown at electrode POz. Tuned, 2%, 4% and 16% mistuned conditions are shown. The 2% and 4% mistuned conditions are presented as dashed black and solid grey lines to highlight their difference from the tuned (solid black) and 16% mistuned (dotted grey) lines.

Figure 4.10: LPC amplitude plotted as a function of Tuning from nine parito-occipital electrodes. LPC amplitude was the mean between 417-617 ms for younger adults, and 492-692 ms for older adults because LPC latency differed between these two groups. Error bars represent one standard error.
Chapter 1:

1.1 Age-related changes in auditory perception

Age-related decline in auditory processing is a particularly important topic given that most older adults experience difficulties with hearing as they age (Committee on Hearing Bioacoustics and Biomechanics, 1988; Mathers, 2000). Estimates of the incidence for clinically significant age-related hearing loss is around 10-20% for adults at the age of fifty (Cruickshanks, Zhan, & Zhong, 2010; Roth, in press). Importantly, it is rare to find a person over age seventy that has not experienced some degree of hearing loss, even if the loss does not meet diagnostic criteria (Gates & Mills, 2005). Difficulties with hearing can negatively impact quality of life through reduced social interactions, increased confusion, greater financial insecurity and depression (Betlejewski, 2006; Salomon, 1986). This issue is further exacerbated because advances in medicine have increased the average life expectancy, meaning that age-related hearing difficulties will affect an increasing number of people for a longer portion of their life.

Interestingly, many auditory processing tasks that decline in age, such as understanding speech in noise (Frisina & Frisina, 1997) or detecting a small silent gap (Schneider, Pichora-Fuller, Kowalchuk, & Lamb, 1994) are enhanced in younger musicians (e.g., Parbery-Clark, Skoe, Lam, & Kraus, 2009; Rammsayer & Altenmüller, 2006). Furthermore, the cortical processing of acoustic information, assessed via scalp recorded electrical potentials is also known to change with age (e.g., Aine et al., 2005; Ceponiene, Westerfield, Torki, & Townsend, 2008; Chao & Knight, 1997b), and many of these scalp recorded electrical potentials are enhanced in musicians (e.g., Pantev et al., 1998; Shahin, Bosnyak, Trainor, & Roberts, 2003). Therefore, the purpose of this dissertation was to examine if older musicians experience less age-related decline in
auditory processing abilities, as measured via psychoacoustic and electrophysiological assessments.

This introductory chapter will include a summary of the basic anatomy and physiology of the auditory system as well as an introduction to using scalp-recorded auditory evoked potentials to assess the auditory system. Understanding the anatomy and physiology of the auditory system is important for understanding age-related changes in hearing abilities. Scalp-recorded auditory evoked potentials are used to assess central auditory function in experiments reported in Chapters 3 and 4. Following these introductory sections, a review of the literature related to the influence of both aging and musical training on auditory processing is presented.

1.2 Hearing from the cochlea to the brain

Hearing is a complex process that relies on multiple stages of processing throughout the auditory system. Acoustic energy reaches the ear as a summation of all acoustic energy in the environment. The resonant frequencies of the outer ear cause a non-linear gain in amplitude for frequencies between 200-7000 Hz, with the largest gain being for frequencies between 2000-3000 Hz, when sound sources are in front of the listener (Shaw, 1998). The leverage action of the ossicles in the middle ear causes a gain of up to 30 dB for frequencies between 500 and 2500 Hz (Yost, 2000). Transduction from physical energy to neural code occurs in the cochlea. Acoustic energy creates a standing wave along the basilar membrane that causes a maximal displacement at a location that is unique for each frequency. The movement of the basilar membrane causes the tectorial membrane to shear across the surface of the stereocilia of the inner and outer hair cells (Yost, 2000). This shearing action causes the hair cell to depolarize, and to fire an action potential. Therefore, due to the structure of the basilar membrane, each hair cell will fire
maximally to a unique frequency. Importantly, this creates a tonotopic (i.e., frequency) organization within the auditory nerve that is maintained and multiplied through the brainstem, thalamus and into the auditory cortex (Yost, 2000). In addition to this tonotopic ‘place’ coding, spectral information is also encoded through phase locking of a hair cells’ firing rate to the frequency of the acoustic energy (e.g., a hair cell will fire at a rate of 100 Hz to a sound with a frequency of 100 Hz). This phase locking is limited by the refractory period of the hair cell, but by alternating with neighboring hair cells, frequencies of up to 4000 Hz can be encoded by phase locking (Yost, 2000). Importantly, the cochlea acts as an amplifier, where the outer hair cells physically amplify the standing wave along the basilar membrane (Schmiedt, 2010). This amplification process is driven by the endolymphatic potential (i.e., the electrical potential in the endolymph within the cochlea), and the gain is non-linear, such that low amplitude and high frequency sounds are amplified more than higher amplitude and low frequency sounds (Schmiedt, 2010).

Once the auditory signal is transduced, the signal ascends the brainstem to the thalamus, and then to the auditory cortex. The nuclei along the brainstem begin to organize the acoustic input based on spectral and temporal information found within the acoustic stream, although top-down efferent processing can influence this activity (Musacchia, Sams, Nicol, & Kraus, 2006). The thalamus acts as a relay station for all sensory input (except smell), and passes information on to primary cortical areas. Acoustic information is passed on to the primary auditory cortex located in the transverse temporal gyrus. The transverse temporal gyrus is also commonly referred to as Heschl’s gyrus, the primary auditory cortex or simply ‘A1’. In addition to sending afferent acoustic information to the primary auditory cortex, the thalamus can also relay information to other cortical or subcortical areas for fast reflexive actions (e.g., orienting your head to a
loud sound). Once the sound reaches the auditory cortex, it can be organized into perceptual units based on spectro-temporal features and higher order schemata, such as knowledge of a language. Our knowledge of the structure and function of the auditory system comes from myriad neuro-scientific and psychoacoustic techniques, one of which is electroencephalography.

1.3 Electroencephalography and auditory evoked responses

In humans, scalp recorded event-related potentials (ERPs) provide a powerful way to investigate the neural mechanisms involved in auditory processing. An ERP is a series of positive and negative deflections that represent phasic neural activity from large neuronal ensembles, derived from a continuous recording of electrical brain activity known as electroencephalography (EEG). EEG activity is a result of a summation of post-synaptic potentials from bundles of pyramidal neurons that have parallel geometric orientations. This electrical activity is calculated as the voltage difference between two locations, one that is considered to be a reference and another where activity is measured. For a more detailed description of EEG recording see Luck (2005). An ERP is an average of the continuous electroencephalogram time locked to the onset of an event which can include a stimulus presentation (e.g. a short sound) or a stimulus response (e.g. pressing a button). The averaging procedure reduces the noise in the EEG signal. That is, on any given trial, the EEG signal will contain a phase-locked response to the evoking ‘event’, in addition to other ‘random’ neural activity not related to the onset of the ‘event’. By averaging multiple trials, the signal remains stable, while the noise is ‘averaged out’ because it is random from trial to trial; accordingly with an infinite number of trials the noise would be reduced to zero. ERPs time locked to the onset of a transient auditory stimulus are also known as auditory evoked responses (AER). A typical AER is
comprised of several positive and negative deflections (i.e., waves) and this sequence of waves tracks neuronal processing of acoustic information as it travels through the auditory system. The first responses arise from neuronal bundles in the brainstem; this stage of processing is known as the auditory brainstem response (ABR) and occurs 1-10 ms after sound onset (Jewett & Williston, 1971). The next stage of processing is known as the mid-latency auditory evoked response (MLAER), is thought to arise from the transmission of the afferent auditory signal from the thalamus to the primary auditory cortex, and occurs 10-50 ms after the onset of a sound (Cohen, 1982; Ribary et al., 1991). The final stages of the auditory evoked response are the long-latency responses that arise from activity in secondary auditory areas along the superior temporal plane (Näätänen & Picton, 1987; Picton et al., 1999). Long-latency evoked potentials begin 50 ms after the onset of a stimulus, and are described in terms of their polarity (P or N) and sequence of occurrence.

The typical AER to a short transient sound is a P1-N1-P2-N2 complex with each response occurring at latencies of around 50, 100, 180 and 250 ms respectively in a healthy adult with a normal auditory system. The P1-N1-P2-N2 complex is an obligatory response related to auditory signal detection and is observed if the signal is audible (Hillyard, Squires, Bauer, & Lindsay, 1971; Martin, Sigal, Kurtzberg, & Stapells, 1997). Furthermore, AERs can be recorded while participants are not attending to the auditory environment. For example, by asking listeners to read a book, or watch a subtitled (silent) movie, while presenting listeners’ with a series of transient auditory stimuli, an experimenter can reduce the attentional resources available for auditory processing (Pettigrew et al., 2004). AERs recorded in these situations can be compared to AERs recorded in a situation where the listener is asked to focus their attention towards the auditory environment by making a judgment about the auditory stimulus (Alain, Arnott, & Picton, 2001; Hansen & Hillyard, 1980; Näätänen, Gaillard, & Mantysalo, 1978).
type of comparison can separate exogenous and endogenous neural activity, which is critical for separating automatic (exogenous) and cognitively based (endogenous) neural activity and more importantly how they contribute to auditory perception.

1.4 Auditory Scene Analysis

The perceptual organization of auditory information into meaningful perceptual units is known as auditory scene analysis (Bregman, 1990). Auditory scene analysis entails organizing the incoming acoustic signal into distinct constituents (i.e., sources or objects), and then tracking those objects over time to form independent auditory streams (Bregman, 1990). Therefore, the number of active sound sources in the environment increases the difficulty of perceptually parsing the auditory scene. For example, understanding speech in a noisy acoustic environment is more difficult compared to understanding speech in a quiet environment.

The first stage in forming acoustic streams is to extract relevant acoustic primitives, which can provide clues as to the structure of the summed acoustic waveform. The tonotopic nature of the afferent auditory signal allows for integration or segregation of information across the frequency spectrum. This is particularly important because natural vibrating bodies (i.e., sound sources) produce acoustic energy across multiple frequency bands, known as harmonics or overtones, that are normally integer multiples of the fundamental frequency (i.e., they are periodic). Therefore, the auditory system will tend to perceptually group harmonically related components, and segregate those that differ in periodicity (Moore, Glasberg, & Peters, 1986). Modulations of these harmonic components are then tracked over time and space to form a perceptual scene that resembles the acoustic scene. Based on this spectro-temporal analysis, acoustic energy
that has similar onsets, intensities, locations and/or periodic frequencies are more likely to be perceived as coming from the same source compared to acoustic energy that differs in terms of onset, intensity, location, and/or frequency (Bregman, 1990). Many of these grouping processes are considered automatic, primitive, or low-level since they can be observed in infants (Winkler et al., 2003) and animals such as birds (Hulse, MacDougall-Shackleton, & Wisniewski, 1997; MacDougall-Shackleton, Hulse, Gentner, & White, 1998), monkeys (Fishman, Reser, Arezzo, & Steinschneider, 2001) or chinchillas (Sinex, 2008; Sinex, Guzik, Li, Henderson & Sabes, 2003; Sinex & Li, 2002). The primitive stage of processing may also be subjected to top-down (also: schema-driven or controlled) processes, which apply learned schemata (e.g., language), previous information from the auditory environment (Mondor & Bregman, 1994), or information from other sensory modalities (Musacchia et al., 2006; Sach, Hill, & Bailey, 2000) to further refine the perceptual auditory scene. Successful hearing therefore requires peripheral encoding of the acoustic environment, early processing of acoustic primitives, and interpretation of the neural signal into a perceptually meaningful auditory scene.

1.5 Age-related changes in hearing abilities

Age-related difficulties in hearing are common, but highly variable between individuals. In general, age-related hearing difficulties are due to changes in the peripheral encoding of acoustic information as well as the central processing of that information. Age-related decline in peripheral sensitivity can be assessed using pure-tone thresholds (for more information on the collection of pure tone thresholds see: Chapter 2, Mazelova, Popelar, & Syka, 2003 or Nelson & Hinojosa, 2006). Current normative data reveal an accelerated decrease in audibility starting in early middle age of about 0.0025 dB
HL/year² for frequencies from 125-1000 Hz, and up to 0.015 dB HL/year² for frequencies at 8000 Hz (unit is given in dB HL/years² because pure-tone thresholds decline at an accelerating rate with age; ISO 7029, 2000). A World Health Organization review, that used pure-tone audiometry as a metric, suggests that more than four-hundred million people world-wide have mild age-related hearing impairment, with an additional two hundred forty million people meeting the criteria for moderate, severe or profound hearing loss (Mathers, Smith & Concha, 2000). In a more recent review of the literature, Roth (in press) reported that by age 70, 20% of females and 30% of males met audiometric criteria for at least mild to moderate age-related hearing loss, and by age 80 the prevalence rises to 45% of females and 55% of males. In another review the incidence rate was also shown to increase with age, with about 10% of females and 20% of males meeting clinical criteria for age-related hearing impairment at age 50, while by age 80 nearly 100% of both men and women met the clinical criteria (Cruickshanks et al., 2010). In these reviews, epidemiological information was computed from multiple smaller studies; thus while there may have been some inconsistencies in the procedures between the individual studies, the overall pattern of age-related decline in hearing abilities are clear.

Age-related change in hearing abilities are known as presbycusis. Schuknecht (1955) originally described four types of presbycusis, sensory, metabolic, neural and mechanical. The mechanical type was believed to be due to a stiffening of the basilar membrane, however, evidence has not emerged to support this phenomena (Gates & Mills, 2005; Schmiedt, 2010). The second form of presbycusis is the neural type. This type is characterized by a reduction in the density and volume of ganglion cells at the cochlear terminus of the auditory nerve; however, despite the anatomical changes there are no known functional changes related to this type of presbycusis (Schmiedt, 2010). A third
type of presbycusis is sensory presbycusis, and is characterized by loss of high frequency sensitivity, in addition to a loss in sensitivity around 4 kHz (Gates & Mills, 2005). This type of presbycusis is thought to be related to noise exposure where the outer hair cells are damaged or missing (Gates & Mills, 2005; Schmiedt, 2010). The final type of presbycusis is the metabolic type, and is characterized by a decrease in sensitivity to higher frequencies, that progresses to lower frequencies at an accelerating rate with age (Gates & Mills, 2005). This form of presbycusis is due to an age-related decrease in vascular support to the stria vascularis, which in turn results in a decrease to the endolymphatic potential in the cochlea (Gates & Mills, 2005). This change weakens the cochlear amplifier, and thus decreases the ability of the inner hair cells to transduce acoustic energy into a neural code, especially for low amplitude and high frequency sounds (Schmiedt, 2010).

Interestingly, age-related hearing loss is influenced by life-long exposure to high amplitude noise. It is well known that the inner and outer hair cells are damaged by long-term exposure to high amplitude noise (Lawner, Harding, & Bohne, 1997; Spoendlin, 1971). In terms of aging, Gates et al. (2000) found that age-related changes in peripheral sensitivity were exacerbated when there was earlier evidence of noise exposure. Experimental data from mice confirmed this finding, and demonstrated that there was a critical period, early in the lifespan (8-16 weeks in mice) in which the noise induced hearing loss interacted with age-related loss (Kujawa & Liberman, 2006). Surprisingly, only the noise exposed mice showed post-mortem histological pathology (Kujawa & Liberman, 2006). Importantly, the damage induced by the noise had no measurable impact on the inner hair cells, while there was significant damage to the spiral ganglion cells that innervate the inner hair cells (Kujawa & Liberman, 2009). These findings
suggest that while high amplitude noise can damage the hair cells, long term exposure to lower levels of noise can damage the parts of the auditory nerve and exacerbate age-related hearing problems.

While age- and noise-related changes to the cochlea result in a decreased peripheral sensitivity to acoustic energy, there are also age-related changes to the central processing of acoustic information, which has also been described as central presbycusis (Gates & Mills, 2005). Across numerous animal studies age-related changes have been observed in both the structure and function of the nuclei along the auditory pathway (Canlon, Illing, & Walton, 2010). In humans these age-related changes probably contribute to a decreased ability to make acute perceptual judgments. For instance, older adults have a decreased ability to discriminate between two frequencies (Raz, Millman, & Moberg, 1989), make a sound location judgment (Abel, Giguère, Consoli, & Papsin, 2000), understand speech (especially in noise; Pichora-Fuller, 1995; Frisina & Frisina, 1997), discriminate between durations (Fitzgibbons & Gordon-Salant, 1994, 1995) and to detect a silent gap between audible acoustic markers (Heinrich & Schneider, 2006). Few aspects of central auditory processing are preserved in aging; however, evidence suggests that streaming (i.e., tracking separate auditory objects over time) is not negatively impacted by age (Snyder & Alain, 2007).

It is possible that age-related changes to central auditory processing are actually due to changes at the peripheral level. Support for this claim comes from the finding that in some studies speech perception was highly correlated with sensori-neural (peripheral) hearing loss (Humes & Roberts, 1990; van Rooij & Plomp, 1990; van Rooij, Plomp, & Orlebeke, 1989). However, there is also evidence that decline observed at the central level is independent of peripheral decline. For example pure-tone thresholds are not well
correlated with detecting a small silent gap (e.g., Heinrich & Schneider, 2006; Schneider & Hamstra, 1999), detecting a mistuned harmonic (Alain, McDonald, Ostroff, & Schneider, 2001), or speech recognition (Frisina & Frisina, 1997; Marshall, 1981). It is important to note that in most cases older adults have elevated pure-tone thresholds compared to young adults, making the distinction between age-related changes to peripheral and central processing difficult to disentangle.

Decline in central auditory processing can be considered along two dimensions, spectral and temporal. Interestingly older adults have more difficulty resolving fine spectral (Alain, McDonald et al., 2001; Grube, von Cramon, & Rubsamen, 2003) and temporal (Moore, Peters, & Glasberg, 1992; Schneider & Hamstra, 1999; Snell, 1997) details in an acoustic signal. Accurate spectral and temporal coding is paramount to auditory scene analysis because spectral information is used to segregate concurrently occurring sounds (Alain, Arnott, & Picton, 2001) and deterioration of temporal processing is related to decreased speech recognition (Mazelova et al., 2003).

1.5.1 Spectral Resolving Power

One way to quantify the spectral resolving power of the auditory system is to assess the perception of complex (i.e. containing multiple harmonics) periodic acoustic stimuli, where one harmonic component can be mistuned from its’ original value. This process is known as periodicity coding. Moore, Peters and Glassberg (1985) found that the threshold for detecting mistuning in a harmonic complex (a sound made up of

---

1 Auditory information is integrated over time, thus resolving spectral details requires the encoding of timing information to determine the frequency of the acoustic input. Alternatively, spectral resolution could be described as resolving the relationship of phase information between bands of acoustic energy, based on temporal fine structure. However, for the purpose of clarity we will refer to this as spectral resolution throughout.
harmonically related pure tones), was anywhere from 0.1% to 5% depending on stimulus characteristics including fundamental frequency ($f_0$), duration, the frequency and/or amplitude of the mistuned harmonic. Detecting violations of periodicity is important because it is thought to underlie concurrent sound segregation, which is an integral part of auditory scene analysis (Bregman, 1990). In support of the relationship between periodicity coding and concurrent sound segregation, Moore, Glassberg and Peters (1986) also found that once the mistuning levels were sufficient, most listeners reported hearing a second, simultaneous sound. The thresholds for perception of a second auditory object were slightly higher than simple detection of the mistuning (1%-6%; Moore, 1986). Older adults have higher thresholds for detecting mistuning (Alain, McDonald et al., 2001; Grube et al., 2003), and this increase in thresholds is paralleled by a decline in the ability to perceptually segregate the mistuned harmonic as a second auditory object (Alain & McDonald, 2007). Furthermore, neural activity related to segregating concurrently occurring sounds declines with age (Alain & McDonald, 2007). Importantly, these age-related changes were not related to changes in pure-tone sensitivity, suggesting that the age-related decline in the ability to detect and segregate acoustic features based their spectral properties is due to changes at the central level.

1.5.2 Temporal Resolving Power

One way to evaluate the temporal resolving power of the auditory system is to measure the smallest perceivable silent gap, known as gap-detection. The stimuli used for gap-detection tasks usually involve the presentation of a leading and lagging acoustic marker, separated by a silent interval. The ability to detect small silent gaps in a stimulus declines with age, and in many cases is independent of pure-tone thresholds (Heinrich &
Importantly, neuro-physiological evidence supports a decline in gap-detection abilities with age. At the earliest stages of auditory processing in the brainstem, a gap is encoded by detecting the offset of the leading marker, and the onset of the lagging marker. When the gap is very small, the onset of the lagging marker is not encoded (Werner, Folsom, Mancl, & Syapin, 2001). For older adults a larger gap is required for successful and consistent encoding of the gap at the level of the brainstem (Poth, Boettcher, Mills, & Dubno, 2001). Surprisingly, even when gaps are successfully encoded in the brainstem, a listener may not actually perceive the gap, suggesting a more complex system for gap detection (Werner et al., 2001). When measuring the response of the auditory cortex to gaps, a similar pattern emerges; older adults require a larger gap compared to younger adults to evoke a gap related response in the auditory cortex (Alain, McDonald, Ostroff, & Schneider, 2004; Bertoli, Smurzynski, & Probst, 2002). In another study, Ross et al.
(2010) found that neural activity related to the perceptual grouping of a gap stimulus into two auditory objects (i.e., the leading and lagging marker) was reduced in older adults, despite similar encoding of the onsets and offsets of the gap in older and younger adults, suggesting that gap detection also relies on higher cognitive processes. Age-related changes in temporal acuity occur at multiple stages of auditory processing, providing further support for the hypothesis that changes in central auditory processing contribute to age-related difficulties with hearing.

1.5.3 Age-related decline in speech perception

Given that aging negatively impacts nearly every process related to auditory perception, it is not surprising that older adults often report difficulty understanding speech (Gates & Mills, 2005). Understanding speech is a complex process that relies on precise transduction of the acoustic signal, in addition to acute processing of spectro-temporal features. Supporting this claim are studies demonstrating that speech understanding is related to peripheral sensitivity (Frisina & Frisina, 1997; van Rooij et al., 1989), temporal acuity (Mazelova et al., 2003), and spectral acuity (Drennan, Won, Dasika, & Rubinstein, 2007). In addition, understanding speech in a multi-talker conversation requires comprehension of the meaning of each utterance, as well remembering what has been said and by whom. Thus, understanding speech also requires higher order cognitive functions, such as working memory and attention (Schneider, Pichora-Fuller, & Daneman, 2010). Therefore, age-related changes in the peripheral encoding and central processing of acoustic information and more general cognitive factors all influence speech understanding in older adults (Humes & Dubno, 2010). Most importantly, when assessing speech understanding, all of these factors interact with each
other, and it can be difficult to determine if age-related decline on a speech understanding task is caused by age-related decline in peripheral encoding, central processing, or general cognitive factors; however, Humes and Dubno (2010) considers peripheral decline to be the main contributing factor in speech understanding. On the other hand, Wingfield, Tun, and McCoy (2005) suggests that age-related decline in both peripheral and central auditory processing as well as cognitive factors all contribute to a reduction in speech understanding, and that we should be investigating how these factors interact to influence speech understanding in older adults.

One way to observe interactions between peripheral, central and cognitive factors is by manipulating the complexity of the acoustic environment. These types of experiments suggest that central and cognitive factors play a role in age-related decline in speech understanding. As the acoustic environment becomes more complex (i.e., more active sound sources), age-related differences in speech understanding increase. For example, older adults have more difficulty understanding speech in noise, especially when the context of the speech cannot be used to aid in understanding (Pichora-Fuller, Schneider, & Daneman, 1995). It is likely that part of this age-related decline is due to a decrease in the ability to detect salient spectro-temporal features within the acoustic stream. Supporting this hypothesis are experiments demonstrating that older adults have more difficulty discriminating between speech sounds that have minute temporal differences (Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Barrett, 2006), and other experiments demonstrating that older adults have more difficulty segregating simultaneously occurring vowel sounds (Snyder & Alain, 2005; Vongpaisal & Pichora-Fuller, 2007). Importantly, the participants in these studies were older adults with normal audiometric thresholds, suggesting that the age-related differences were due to changes at
the central level. Interestingly, when older adults can utilize the context of a sentence, their ability to understand speech in noise improves to a greater extent compared to younger adults, suggesting that older adults may automatically engage higher-level ‘controlled’, ‘schema-driven’ or cognitive processing to overcome impoverished primitive processing (Pichora-Fuller et al., 1995; Pichora-Fuller, 2008; Wingfield & Tun, 2007). Furthermore, neuro-physiological evidence supports the findings that older adults do in fact engage cognitive processes to partially overcome impoverished primitive processing (Alain et al., 2004). The ability of older adults to use controlled processing to overcome perceptual deficits is however limited by an age-related slowing of processing speed that reduces working memory abilities (Salthouse, 1993; Van der Linden et al., 1999). Not surprisingly, the benefit of using controlled processes in older adults’ breaks down when cognitive abilities are over-taxed (Humes, Lee, & Coughlin, 2006). These studies suggest that age-related decline in speech understanding is related to age-related changes in peripheral sensitivity, age-related changes in the central processing of acoustic information, and age-related changes in cognitive abilities. Given the ubiquity of age-related hearing problems, finding ways to mitigate age-related decline at any stage of speech understanding are of utmost importance.

1.6 Cognitive reserve

There is increasing evidence that life style choices can have a significant impact on successful aging. For instance, older adults who engage in cognitively stimulating activities later in life show slower rates of cognitive decline, independent of early education levels (Ghisletta, Bickel, & Lovden, 2006; Valenzuela & Sachdev, 2006). Other research suggests that higher educational and occupational achievements contribute
to delayed onset of age-related cognitive decline (Qiu, Backman, Winblad, Aguero-Torres, & Fratiglioni, 2001; Stern et al., 1994). In a recent review of the literature, Middleton and Yaffe (2009) reported that engaging in cognitive activities (e.g., reading, learning, or game playing), being physically or socially active can delay or prevent dementia. These findings suggest that early educational or occupational achievement create a ‘cognitive reserve’ that can delay age-related cognitive decline, and continued engagement in stimulating activities can maintain or enhance this reserve of enhanced cognitive abilities.

Importantly, staying mentally active does not prevent all aspects of age-related cognitive decline. For example, Jastrzembiski, Charness, & Vasyukova, (2006) demonstrated that older expert chess players were just as skilled at identifying threatening chess positions; however the older adults reaction times to detect this threat was significantly longer. Therefore, while the expertise related to identifying chess positions remained with age, the processing speed of that information declined. In another study that examined a large sample of expert chess players, it was shown that the overall chess rating reached a peak during middle age (~40’s) and declined steadily thereafter (Roring & Charness, 2007). Interestingly, the decline was slower when the peak performance level in middle age was higher, suggesting that high levels of expertise can slow age-related decline (Roring & Charness, 2007). Again, this trend is likely related to the distinction between age related decline in chess speed and knowledge (Jastrzembiski et al., 2006). In aging, numerous studies have shown that older adults demonstrate significant decline in speeded, process-based, or fluid tasks, while performance on tasks that rely on crystallized knowledge stay relatively stable (Salthouse, 2010). In terms of expertise in chess players, the outcomes are clear; expertise results in an enhancement or preservation
of crystallized knowledge in age, but has little influence on age-related change to speeded tasks. Thus the concept of cognitive reserve is likely limited to certain cognitive domains.

The concept of cognitive reserve naturally extends to the brain, known as the brain reserve hypothesis. In general this hypothesis states that larger brains result in higher cognitive functioning, which persists in age (Christensen, Anstey, Leach, & Mackinnon, 2008). In young, healthy adults, there are correlations between brain volume and overall intelligence (as measured by IQ; Willerman, Schultz, Rutledge, & Bigler, 1991). A more detailed analysis, using both older and younger adults found that IQ scores were only related to volume in specific grey matter brain regions, including frontal, temporal and parietal areas (Haier, Jung, Yeo, Head, & Alkire, 2004). In terms of age, having greater brain reserve doesn’t alter the rate of age-related change in cognitive abilities; however, this means that individuals with greater brain reserve (and hence higher cognitive abilities) will maintain higher cognitive abilities throughout the lifespan (Christensen et al., 2008).

In a longitudinal sample, where participants were measured in childhood, and again in old age, Gow et al. (2010) demonstrated that those who had higher cognitive functioning in childhood tended to also have higher cognitive functioning in old age. Importantly, higher cognitive abilities did not protect against age-related cognitive decline, rather, cognitive abilities were found to be relatively stable, and thus those with greater abilities in childhood had greater abilities in old age (Gow et al., 2010). At the same time, when there was an age-related pathology (e.g. dementia, Alzheimer’s), those with greater brain reserve did show a slower rate of cognitive decline (Christensen et al., 2008).

Interestingly, the volume of some neural regions that are positively correlated with general cognitive abilities are also related to enhanced abilities to understand speech in noisy environments in older adults, including the left pars triangularis and the left
superior frontal gyrus (Broadman areas 45 and 8, respectively; Wong, Ettlinger, Sheppard, Gunasekera, & Dhar, 2010; Haier et al., 2004). Furthermore, older adults (compared to younger adults) tend to engage a more widespread neural network to understand speech in noise that includes pre-frontal areas related to attention and working memory, while they show reduced activation in primary and secondary auditory areas in the temporal lobe (Wong et al., 2009). These findings suggest that older adults use a more cognitive strategy to overcome age-related decline in lower-level perceptual processing. In a review of multiple studies Christensen et al. (2008) found that brain reserve was related more to lifestyle choices compared to intelligence or education. It is therefore possible that lifestyle choices that require an acute auditory system (e.g. being a musician) may result in a decrease in age-related decline in auditory processing abilities. Musicians may therefore be an excellent model to study cognitive reserve, especially in terms of age-related decline in processing auditory information.

Most musicians undergo rigorous training in childhood, and continue to practice over the lifespan. Importantly, auditory scene analysis entails a cognitive stage of processing in addition to basic perceptual processing, both of which may be preserved in older musicians. Musical training engenders advanced development of auditory processing abilities (e.g., Beauvois & Meddis, 1997; Koelsch, Schroger, & Tervaniemi, 1999; Parbery-Clark, Skoe, Lam et al., 2009; Rammsayer & Altenmuller, 2006), motor skills (e.g., Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995; Krampe & Ericsson, 1996; Repp & Doggett, 2007), attention (e.g., Baumann, Meyer, & Jancke, 2008; Nager, Kohlmetz, Altenmuller, Rodriguez-Fornells, & Munte, 2003), working memory (e.g., Y. Lee, Lu, & Ko, 2007), long-term memory (e.g., Jakobson, Lewycky, Kilgour, & Stoesz, 2008), and integration of multiple sensory modalities including visual and auditory (e.g., Musacchia, Sams, Skoe, & Kraus, 2007). Musicians are therefore an excellent population
to examine potential cognitive reserve, but surprisingly few studies have examined lifelong effects of musical training on cognitive tasks (Andrews, Dowling, Bartlett, & Halpern, 1998; Krampe & Ericsson, 1996; Meinz, 2000). Krampe et al. (1996) found that expert musicians, compared to amateur musicians experienced less age-related decline on speeded motor tasks related to music performance, but that general processing speed was not influenced by musical expertise. Andrews et al. (1998) found that the ability to recognize speeded or slowed melodies declined with age, and that musicians were better than non-musicians, but that these factors did not interact. Meinz et al. (2000) found that memory and perceptual speed in musical situations declined with age in pianists, and that more experienced pianists performed better than non-musicians, but that all levels of pianists declined at the same rate. In terms of brain structure, age-related decline in grey matter volume in Broca’s area is mitigated in musicians (Sluming et al. 2002). Broca’s area is located within the left inferior frontal gyrus, and is thought to be involved in numerous auditory processing tasks important for understanding and producing both music and language (see Sluming et al. 2002). Despite these findings, no one has investigated the possibility that continued engagement in music performance over the lifespan may prevent or delay age-related decline in auditory processing abilities. When examining how life-long expertise influences age-related changes in cognitive, perceptual, or neurophysiological measures, two patterns of data may emerge.

The first pattern is that of differential preservation, which is characterized by interactions between age and expertise, where the rate of age-related decline on a given task is slower in experts (Krampe & Ericsson, 1996; Sluming et al. 2002; Salthouse, 2006). The second pattern is that of a preserved differentiation which is characterized by main effects of expertise without interactions with age (Andrews et al., 1998; Meinz, 2000; Salthouse, 2006). Here experts perform better than non-experts, but the difference
is equal across the lifespan. Importantly, only differential preservation can be indicative of a protective effect of expertise because it indicates an increasing effect of the expertise over time (Salthouse, 2006). Differential preservation suggests an accumulation of cognitive reserve over the life span. This hypothesis was supported through computational models that simulated aging. Using a neural network model, Mireles et al. (2002) demonstrated that aging could be simulated by progressively adding noise to the network, while increasing knowledge of the problem (in the network) improved the performance of the network to a greater degree with age (demonstrating differential preservation). This model suggests that expertise and continued training could offset age-related declines in auditory perception. Not surprisingly, numerous studies have shown that younger musicians have advantages in auditory processing abilities (Beauvois & Meddis, 1997; Koelsch et al., 1999; Micheyl, Delhommeau, Perrot, & Oxenham, 2006; Parbery-Clark, Skoe, Lam et al., 2009; Rammsayer & Altenmuller, 2006; Schellenberg & Moreno, 2010). However, the influence of being a musician on age-related changes to auditory processing abilities has not been investigated, and is the purpose of the current dissertation.

1.7 Musicians vs. Non-musicians

Younger musicians present with enhanced auditory abilities on many of the same tasks that decline in older adults. For example, musicians are better able to detect violations of periodicity (Koelsch et al., 1999), have enhanced abilities in detecting a silent gap in a sound (Rammsayer & Altenmuller, 2006), are better at discriminating between the duration of two sounds (Rammsayer & Altenmuller, 2006) and detecting violations of isochrony (Rammsayer & Altenmuller, 2006). In addition, younger musicians are able to maintain auditory streams in memory longer (Beauvois & Meddis,
1997), have enhanced abilities to detect small frequency differences between two tones (Besson, Schon, Moreno, Santos, & Magne, 2007; Micheyl et al., 2006; Schellenberg & Moreno, 2010), have reduced effects of auditory informational masking (Oxenham, Fligor, Mason, & Kidd, 2003), are better able to discriminate between sounds with minute duration differences (Jeon & Fricke, 1997) and are less susceptible to timbral influences on pitch perception (Pitt, 1994). In addition, musicians are better able to automatically process more complex forms of auditory material. Automatic processing of acoustic information can be assessed using scalp-recorded electrical potentials (or magnetic fields), while a listener is attending away from the auditory environment. For example, van Zuijen et al. (2005) presented musicians and non-musicians with repeating sequences of tones that were of the same pitch and numerical regularity (e.g., four tones at 349 Hz, then four tones at 369 Hz, then four tones at 329 Hz, etc…). A rare deviant sequence would violate the numerical regularity (i.e., five tones instead of four). Van Zuijen (2005) found that this violation was detected automatically in musicians, but not in non-musicians. Along similar lines, Fujioka et al. (2004) presented listeners with repeating five note melodies all with an ascending melodic contour. Rare deviant five note melodies had either a final note that descended (violating the regularity of the melodic contour) or had a final note that was significantly higher in pitch than the standard final note. While these violations were automatically detected by both musicians and non-musicians the neural representation of these violations was more robust in musicians (Fujioka et al. 2004). Given that musicians have enhanced auditory processing abilities for both simple and complex stimuli it is not surprising that younger musicians are better able to understand speech embedded in multi-talker babble (Parbery-Clark, Skoe, Lam et al., 2009). None of the studies reported above found any evidence of enhanced peripheral encoding of acoustic information, suggesting that the auditory processing enhancements found in
musicians are due to enhanced central processing abilities. This suggests that neurophysiological measures of the structure and function of auditory processing areas in the central nervous system should be enhanced in musicians.

There is increasing evidence that musical training involves neuro-plastic changes in the auditory cortex. In musicians, musical aptitude is correlated with an increase in grey matter volume of Heschl’s Gyrus (Gaser & Schlaug, 2003; Schneider et al., 2002). In addition, studies have shown that musicians have enhanced grey matter volume around the central sulcus, including motor and somatosensory areas, the inferior temporal gyrus, corpus collosum and hippocampus (Gaser & Schlaug, 2003; Herdener et al., 2010; Schlaug, Jancke, Huang, & Steinmetz, 1995). These findings are paralleled by studies that have found enhancements in the functional properties of the musicians’ brain.

Accordingly, numerous studies have shown that phasic neuro-electric or neuro-magnetic brain activity generated in and around the primary auditory cortex (i.e., the AER) is enhanced in musicians (Besson et al., 2007; Kuriki, Kanda, & Hirata, 2006; Pantev et al., 1998; Pantev, Roberts, Schulz, Engelien, & Ross, 2001; Pantev et al., 2003; Shahin et al., 2003). Other research has found enhanced activity in the auditory brainstem (Lee, Skoe, Kraus, & Ashley, 2009; Musacchia et al., 2007; Parbery-Clark, Skoe, & Kraus, 2009).

While these data are suggestive that musical training alters brain structure and function, it is possible that musicians self-select based on predispositions.

The only way to prove a causal relationship between musical training and differences in brain structure or function is to conduct a controlled experiment, utilizing random assignment. Experimentally there are significant difficulties with employing this type of paradigm. Musicians often train for many years, and spend many hours each day honing their abilities. Therefore it is difficult to track and control the practice over the timeframe that is necessary for an individual to become a musician. Despite these
difficulties, a few studies have provided musical training to non-musicians, and observed neuro-physiological changes. Specifically, enhancements to the structure (Hyde et al., 2009), and functional properties (Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Lappe, Herholz, Trainor, & Pantev, 2008) of the auditory cortex were observed in participants who received music training compared to control groups. In addition, neural coding in the auditory brainstem was enhanced after auditory training (Russo, Nicol, Zecker, Hayes, & Kraus, 2005). These longitudinal studies lend support to a causal interpretation of the cross-sectional studies.

1.8 The current study

Given that age-related decline in central auditory processing is common, and given that musicians have enhanced auditory processing abilities, the goal of this dissertation was to examine whether musicians experience less age-related decline on auditory processing abilities. In the first experiment, a large sample of musicians and non-musicians ranging in age from 18 to 91 were given a series of hearing assessments. Here it was discovered that, for hearing tasks that rely on central processing mechanisms, musicians demonstrate less age-related decline compared to non-musicians (Chapter 2). Findings from this experiment suggested a central source for the benefit in older musicians. To assess central auditory processing, a second experiment examined auditory evoked responses in four groups of older and younger musicians and non-musicians. Here it was found that there were age-related changes to both exogenous and endogenous neural activity related to the processing of a complex sound. While some exogenous processing of acoustic information was enhanced in musicians, the age-related decline was similar in both musicians and non-musicians. On the other hand, endogenous neural activity was largest in older musicians, suggesting that lifelong musicianship creates and maintains a
cognitive advantage in processing acoustic information (Chapter 3). The final experiment investigated how lifelong musicianship and aging interact to influence how acoustic information is organized into a perceptual scene. Here it was found that the ability to segregate simultaneously occurring sounds was enhanced in both older and younger musicians. This enhancement was related to attention-dependent neural activity, which again suggests that older musicians engage a cognitive strategy to aid in perceptual organization (Chapter 4). The relevance of these results is considered in terms of current theories of cognitive and neurological aging, and is discussed in Chapter 5.
Chapter 2:

2.1 The influence of age and lifelong musicianship on behavioral measures of auditory processing

Auditory-based communication problems are prevalent in the elderly. One of the most commonly reported age-related auditory difficulty is in understanding speech, especially in adverse listening situations (Bergman, 1980; Committee on Hearing Bioacoustics and Biomechanics, 1988; Frisina & Frisina, 1997; Pichora-Fuller et al., 1995). Age-related decline in speech perception is thought to reflect bilateral sensorineural hearing loss due to physical changes in the cochlea (Gates & Mills, 2005; Stenklev & Laukli, 2004). In addition there are age-related changes to central stages of auditory processing where spectro-temporal features of the auditory signal are extracted and used to group and segregate acoustic information into a perceptual scene (Bregman, 1990; Alain, Dyson, & Snyder, 2006; 2010; Alain, McDonald et al., 2001; Grube et al., 2003; Schneider et al., 1994). Importantly, acute detection of spectro-temporal features is important for speech understanding (Mazelova et al., 2003; Drennan, et al. 2007). For example detecting violations of periodicity is important for segregating simultaneously occurring sounds (Moore, Glasberg, & Peters, 1986) and for understanding speech (Drennan et al. 2007). Not surprisingly, older adults have increased thresholds for detecting violations of periodicity in a complex sound (Alain, McDonald et al., 2001; Grube et al., 2003). In the temporal domain, the ability to detect brief silent gaps can be important for distinguishing speech phonemes (Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Barrett, 2006; Strouse, Ashmead, Ohde, & Grantham, 1998). Importantly, older adults have a reduced ability to detect brief gaps in an otherwise continuous sound (Gordon-Salant et al., 2006; Heinrich & Schneider, 2006; Schneider et al., 1994; Strouse
et al., 1998). Therefore, some of the age-related changes in the ability to understand speech, especially in adverse listening situations, are likely due to a decline in the ability to detect salient spectro-temporal features within the auditory stream. Given the prevalence, and negative outcomes of age-related decline in hearing abilities, finding ways to prevent, mitigate or delay these changes is of utmost importance.

Musicians are one group of people that have highly developed auditory abilities. Compared to non-musicians, they are better at detecting mistuning in a harmonic complex (Koelsch et al., 1999), detecting a short silent gap (Rammsayer & Altenmuller, 2006), and discriminating pure-tone frequencies (Micheyl et al., 2006; Schellenberg & Moreno, 2010). Musicians also show superior performance in general cognitive domains including auditory working memory (Lee et al., 2007), IQ as measured by the WISC-III (Schellenberg, 2004), and the identification of speech embedded in multi-talker babble (Parbery-Clark, Skoe, Lam et al., 2009). Moreover, cross-sectional studies using neuro-imaging techniques have shown that musicians have enhanced electrophysiological activity in the auditory brainstem (Bidelman & Krishnan, 2010; Musacchia et al., 2007; Parbery-Clark, Skoe, & Kraus, 2009; Wong, Skoe, Russo, Dees, & Kraus, 2007) and auditory cortex (Hyde et al., 2009; Koelsch et al., 1999; Pantev et al., 1998; Schneider et al., 2002; Shahin et al., 2003). Longitudinal studies have yielded similar findings (i.e., enhanced electrophysiological and structural measures of the auditory cortex after training), suggesting that the differences between musicians and non-musicians are due to experience-dependent plastic brain changes rather than self-selection due to pre-existing genetic differences (Fujioka et al., 2006; Hyde et al., 2009). It is therefore likely that any influence of musicianship on age-related changes in auditory processing abilities are due to both enhancements in the perceptual processing of acoustic features, as well as the higher-order cognitive interpretations of those features. Hence, musicians are one group
of experts that could demonstrate an age by expertise interaction on tasks that tap into central auditory processes.

This experiment was designed to test the hypothesis that playing a musical instrument throughout the lifespan mitigates age-related decline in auditory perception. Musicians and non-musicians ranging in age from 18 to 91 completed four hearing assessments: pure-tone thresholds, mistuned harmonic detection thresholds, gap detection thresholds, and the QuickSIN test (Speech-In-Noise; Etymotic Research, Version 1.3, 2001; Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004). These four tests were chosen because performance on all these tasks declines with age (Alain, McDonald et al., 2001; Frisina & Frisina, 1997; Gordon-Salant et al., 2006; Grube et al., 2003; Heinrich & Schneider, 2006; Pichora-Fuller et al., 1995; Schneider et al., 1994; Stenklev & Laukli, 2004; Strouse et al., 1998). At the same time performance on gap detection, mistuned harmonic detection, and speech in noise perception are enhanced in young adult musicians (Koelsch et al., 1999; Parbery-Clark, Skoe, Lam et al., 2009; Rammsayer & Altenmuller, 2006). More importantly, these tests assess disparate aspects of auditory processing, with pure-tone thresholds reflecting the functioning of the peripheral auditory system (Mazelova et al., 2003; Nelson & Hinojosa, 2006), gap-detection reflecting the ability to detect fine temporal details, mistuned harmonic detection reflecting the ability to detect fine spectral details, and speech-in-noise detection reflecting realistic auditory situations. These distinctions are important because it is likely that the rate of age-related decline on tasks that require central processing will be slower in musicians. On the other hand there is some evidence that life-long musicianship can cause noise induced hearing loss due to long-term exposure to high amplitude sounds (Jansen, Helleman, Dreschler, & de Laat, 2009). Therefore age-related decline in pure-tone thresholds will likely be the same or greater in musicians.
2.2 Methods

2.2.1 Participants

A total of 163 native English speakers took part in the study and provided formal informed consent in accordance with the joint Baycrest Centre-University of Toronto Research Ethics Committee. Seventy-four participants were musicians (age 19-91, 35 females), while 89 were non-musicians (18-86, 51 female). Here we defined a musician as someone who started musical training on any musical instrument by the age of 16, continued practicing music until the day of testing, and had an equivalent of at least six years of formal music lessons. The sample of musicians included both professional and amateur musicians; however, the distinction between professional and amateur musician was not directly quantified because of the difficulty in distinguishing them. For example, some “amateur” musicians had other (non-musical) careers, but occasionally taught students part-time or had performed for payment during their lifetime; others earned their main income through music for a few years, but then went on to a different career; others had retired, and no longer performed professionally, but continued to practice and perform for their own enjoyment. Clearly the distinction between amateur and professional is difficult to pin down, and therefore the sample of musicians was not quantified into ‘amateur’ and ‘professional’ subgroups. Non-musicians had no more than two years of formal or self-directed lessons throughout their lifetime, and did not ever play a musical instrument. These criteria excluded many people with intermediate levels of musical achievement (e.g., took piano lessons in childhood and stopped playing, or learned guitar in college and played in a band for a few years). Participants for this study were initially recruited for other studies; however, preliminary analyses of data from these participants’ revealed trends that musicians had age-related advantages on some auditory assessments.
This led to the recruitment of life-long musicians and non-musicians to confirm that these trends were real. In the final sample, musicians and non-musicians were matched in terms of years of education (16.9 vs. 16.2, respectively, \( t(161) = 1.36, p > 0.05 \)), and age (45.3 vs. 49.3, \( t(161) = 1.21, p > 0.05 \)). In addition, for the musicians, we collected the age of onset of musical training, and the average hours per week of practicing with their instrument (in the year the participant was tested). These self-reported demographics were used to correlate with performance on the experimental tasks. Older participants were screened over the phone for age-related psychological disorders including dementia before they participated. No participants wore hearing aids during the testing, and no participants used hearing aids on a regular basis. Finally, all participants spoke English as their first language.

2.2.2 Stimuli and Task

All stimuli were presented through ER 3A insert earphones (Etymotic Research, Elk Grove, USA), while participants were seated in a sound-attenuating room. Before the experiment began, participants were familiarized with the stimuli and response methods for each task. Furthermore, stimuli for the gap detection, mistuned harmonic detection and the speech in noise tasks were presented well above hearing thresholds for all participants as no participant reported any difficulty hearing the stimuli.

2.2.2.1 Pure-tone thresholds.

Pure-tone thresholds were collected for each octave between frequencies of 250 to 8000 Hz binaurally using an audiometer (model GSI 61). Participants were instructed to press a button if they detected a tone. The procedure began by presenting a 1000 Hz tone to the left ear at 40 dB HL. If the participant indicated that they heard the tone, the amplitude was reduced by 5 dB HL. This continued until the participant no longer
responded to the tone. The amplitude of the tone was then increased by 5 dB HL; if a response was recorded the amplitude was lowered again by 5 dB HL; if there was no response the amplitude of the tone was increased by 5 dB HL, and that amplitude was recorded as the pure-tone threshold for the frequency and ear of presentation. If the participant could not hear the initial tone at 40 dB HL, the amplitude was increased in 5 dB steps until the amplitude was 10 dB HL above where the participant first detected the tone. The procedure then continued as stated above. Each tone was presented at an irregular time interval to prevent false positive responses. The same procedure was repeated for each frequency octave in the following order: 2000, 4000, 8000, 250 and 500 Hz. The 1000 Hz stimulus was then tested again to ensure that the initial test at 1000 Hz was not impaired by a lack of familiarity with the procedure. This entire procedure was then repeated in the right ear.

2.2.2.2 Gap Detection.

Stimuli were tone pips produced by multiplying a 1-kHz pure tone by a temporal window created by summing a series of Gaussian envelopes spaced 0.5 ms apart (Schneider et al., 1994). Two of these 10-ms tones marked the beginning and end of the gap (i.e., leading and lagging marker). The duration of the gap, $\Delta t$, was defined as the time between the last Gaussian in the leading marker and the first Gaussian of the lagging marker. The comparison stimulus (a tone whose duration and energy were equal to that of the two markers defining the gap) was created by filling in the missing Gaussians between the two markers. Stimuli were generated by a Tucker-Davis real time processor (RPvdsEx version 6.2), with a 500 ms interval between the first and second stimulus. The initial gap size was 31 ms, and stimuli were presented in a two-alternative forced-choice (2AFC) paradigm, with a 3 down 1 up tracking procedure to determine the 79.4% accuracy point.
on the psychoacoustic curve (Levitt, 1971). All stimuli were presented binaurally at 75 dB SPL. In each trial, two stimuli were presented sequentially in a random order, and the listener identified which of the two sounds contained a gap by pressing a button on a response box. For the first reversal, the size of the gap was reduced by 8 ms after three correct responses, or increased by 8 ms after one incorrect response. The amount of change in gap size was reduced by 50% after each reversal to a minimum of 0.5 ms. Each block of trials lasted until there were 12 reversals. The threshold was determined by averaging the last 8 reversals. This procedure was repeated three times, and the final gap threshold was the average of these three blocks. Feedback about the correctness of the response was given to participants after each trial to ensure that each participant achieved his/her lowest possible threshold.

2.2.2.3 Mistuned Harmonic Detection.

Stimuli were harmonic complexes made up of twelve pure tones at equal intensity levels with a fundamental frequency \( f_0 \) of 200 Hz and were presented at a combined amplitude of 75 dB SPL. Stimulus duration was 200 ms including a 10-ms rise and fall time. The second harmonic (600 Hz) of the complex was mistuned from its original value, with a starting value of 696 Hz (16% above its original value). Stimuli were generated by a Tucker-Davis real time processor (RPvdsEx version 6.2), with a 900 ms interval between the first and second stimulus. As in the gap detection task, a 2AFC procedure was used to estimate the thresholds in detecting a mistuned harmonic. In each trial, two stimuli were presented sequentially, one of which contained the mistuned harmonic component, in a random order. Participants were asked to identify which sound had the mistuned component by pressing a button on a response box. Feedback about the correctness of the response was given to participants after each trial to ensure that each
participant achieved his/her lowest possible threshold. The amount of mistuning was reduced by 50% after three correct responses (i.e., from 16% of the original value to 8% of the original value); a single incorrect response resulted in an increase of mistuning by 32% on the next trial. After the first two reversals, the amount of mistuning decreased by 24% after three correct responses. Thresholds were an average of the last eight reversals. This procedure was repeated across three blocks, and the final threshold was the average of the three blocks.

2.2.2.4 Understanding speech-in-noise.

The effects of age and musical training on listeners’ ability to process speech in noise were assessed using the QuickSIN test (Speech-In-Noise; Etymotic Research, Version 1.3, 2001; Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004). Participants were presented with five lists of six sentences each with five key words per sentence embedded in four-talker babble noise, made up of meaningful English utterances. The sentences were presented at a combined amplitude of 70 dB SPL using pre-recorded signal-to-noise ratios (SNRs) which decreased in 5-dB steps from 25 dB SNR (very easy) to 0 dB SNR (very difficult). Participants were asked to repeat back the target sentence, and were given a single point for each of the five “key words” in each sentence, for a possible 30 points on each list. Performance was assessed by subtracting the total number of correct words from 25.5, which yields the SNR loss score (Killion, 1997). SNR loss is a standardized measure that indicates a difference in the SNR needed to identify 50% of the words in a sentence, where 0 dB represents standard performance. Positive values represent a SNR loss, while negative values represent a SNR gain. The standardization procedure of this test can be found in (Killion et al., 2004).
2.2.3 Data analysis

Each hearing assessment was analyzed separately using multiple linear regressions. Due to experimenter error, data loss, or participants’ fatigue, SIN thresholds were not available for five participants, while gap-detection and mistuned harmonic thresholds were not available for nine participants. Three factors were used in each analysis: age, musical training (‘dummy coded’ as 0.5 and -0.5 for musicians and non-musicians, respectively), and the age by musical training interaction (i.e. age multiplied by the dummy coded variable). A regression analysis was used because it allowed for an analysis of the age-related linear trends in the data, as well as for main effects (i.e., differences in intercepts of the regression lines) or interactions (i.e., differences in the slope of the regression lines) related to being a musician on these hearing tests. More importantly, these two comparisons differentiated whether advantages for musicians, if any, were due to preserved differentiation (characterized by main effects of being a musician) or differential preservation (characterized by age by musician interactions). For each analysis, eighteen years were subtracted from every participant’s age so the comparison between musicians and non-musicians would compare intercepts in the regression lines that matched with the youngest participants in the study, as opposed to extrapolated intercepts at age zero. Note that all figures (i.e., Figures 2.1, 2.2, 2.3 & 2.4) show participants’ real age.

In addition to the multiple regression analyses, correlations in musicians with age of musical training onset, and the number of hours per week of music practice (in the year the participant was tested) while controlling for the effect of age were calculated. Age was included as a covariate because performance on each of the hearing tests significantly declined with age, and the purpose of calculating these correlations was to determine if
hours of practice or age of training onset was related to task performance independent of age.

Finally, correlations were calculated between the pure-tone threshold at the stimulus frequencies used to measure the gap-detection thresholds (i.e. pure-tone threshold at 1000 Hz), and mistuned harmonic detection threshold (i.e. pure-tone threshold at 250, 500 & 1000 Hz) to determine if peripheral sensitivity to the stimulus frequency was related to task performance. Again age was included as a covariate because the influence of age on all hearing assessments was significant, and the purpose of this analysis was to determine if there was a relationship independent of the effect of age. Only significant correlations are reported (i.e., \( \alpha > 0.05 \)).

### 2.3 Results

#### 2.3.1 Pure-tone thresholds

Group mean audiograms for older and younger adults can be seen in Figure 2.1. The mean response for both ears and all pure-tone frequencies (all-tone average) was analyzed in a single regression. The overall regression model was significant for the all-tone average (see table 2.1, Figure 2.2). Importantly, the only significant factor in this analysis was age. No significant effect of musical training or it’s interaction with age were found. In separate analyses, not reported here, this pattern was stable for each octave tested (250, 500, 1000, 2000, 4000 and 8000 Hz). The only difference between each frequency octave was the rate of age-related change (i.e., slope of the regression line increased with frequency). Sensitivity to lower frequencies exhibited less age-related change compared to higher frequencies (see Figure 2.1). Furthermore, the left- and right-ear thresholds were compared at each frequency octave in each participant to ensure participants were free of retrocochlear and conductive pathology. Two participants
exceeded the 15 dB difference between the left and right ears at one or more octave bands. All further analyses (below) were run with and without these participants, and no difference in the pattern of results was found, thus they were included in all analyses.

2.3.2 Gap-detection thresholds

The overall regression model for gap detection thresholds was significant (see table 2.2, Figure 2.3). The beta weights for the age and the age by musical training factors were significant, but the beta weight for the musical training factor was not significant. That is, gap-detection thresholds increased with age; however, the rate (slope of the regression line) at which gap detection thresholds increased with age was slower in musicians compared to non-musicians, as indicated by a significant beta weight for the age by musical training interaction. Finally, gap-detection thresholds were not correlated significantly with 1000 Hz pure-tone thresholds, while controlling for age ($r(157) = -0.061$), indicating that sensitivity at the tested frequency band was not related to the ability to detect a silent gap.

2.3.3 Mistuned harmonic detection thresholds

The overall regression model for mistuned harmonic threshold was significant (see table 2.3, Figure 2.4). In this regression, the beta weights for both the age and the musical training factors were significant. That is, mistuned harmonic thresholds increased with age; however, musicians had lower thresholds compared to non-musicians, as indicated by a significant beta weight for the musical training factor. In addition, it was also found that mistuned harmonic threshold was correlated with pure-tone thresholds at 250, 500 and 1000 Hz, while controlling for age, ($r(157) = 0.163, 0.245, 0.182$, respectively, $p < 0.05$ in all cases) suggesting that pure-tone thresholds at the frequencies used in the mistuned harmonic task may contribute to the ability to detect a mistuned harmonic. At the same
time, in musicians, mistuned harmonic thresholds were negatively correlated with hours per week of musical practice while controlling for age ($r(60) = -0.313, p < 0.05$). That is, the more time a musician spent engaged in musical activities, the lower his/her mistuned harmonic threshold.

2.3.4 Speech-in-noise thresholds

The overall regression model for speech in noise threshold was significant (see table 2.4, Figure 2.5). In this regression, the beta weights for both age and the age by musical training factors were significant. That is, speech-in-noise thresholds increased with age; however, the rate (slope of the regression line) at which speech in noise thresholds increased with age was slower in musicians compared to non-musicians. One participant exceeded the pure-tone threshold of 45 dB HL that in the clinical version of the assessment, would require an increase in the amplitude of the stimulus level; however, in this case the level was not adjusted in order to maintain experimental control. For musicians, speech-in-noise thresholds were negatively correlated with hours per week of musical activities while controlling for age ($r(61) = -0.267, p < 0.05$). That is, the more time a musician spent engaged in musical activities, the lower his/her speech-in-noise threshold.

2.4 Discussion

The primary objective of this study was to examine whether age-related changes in peripheral and central auditory processing were mitigated in musicians. It was observed that musicians experienced less age-related decline for both gap-detection and speech-in-noise thresholds. Moreover, musicians showed a lifelong advantage in detecting a mistuned harmonic compared to non-musicians. Performance on the mistuned harmonic and the speech-in-noise tasks were correlated with hours of music practice, that is the
more musicians practiced, the better they were at detecting a mistuned harmonic and understanding speech in noise. For speech in noise thresholds, the relationship between practice and performance suggests that the accumulation of practice over many years may result in preservation of this ability in musicians, as demonstrated by the interaction between age and musicianship for this test. Most importantly, no influence of being a musician was found for pure-tone thresholds, suggesting that aging can differentially influence central and peripheral stages of auditory processing, and that the differences in musicians and non-musicians are due to differences at the central level. Thus, being a musician may contribute to better hearing in old age by delaying some of the age-related changes in central auditory processing. Alternatively, it is also possible that musicians are self-selected based on in-born characteristics that endow them with enhanced auditory processing abilities.

These findings are consistent with a growing body of literature demonstrating that staying mentally and physically active can prevent or delay age-related cognitive decline (Bialystok, Craik, & Freedman, 2007; Middleton & Yaffe, 2009; Christensen et al., 2008). In the present study, it was found that in musicians, amount of practice was correlated with the ability to detect mistuning and to understand speech in multi-talker babble. This suggests that continued practice throughout life may alleviate some of the age-related decline in speech perception often experienced by older adults.

One possibility is that continued practice of a musical instrument may enhance cognitive reserve, thereby freeing perceptual, attentional and cognitive resources which could then be dedicated to the processing of auditory stimuli. For example, it is known that older adults use a more cognitive strategy to overcome age-related decline in sensory processing (Schneider et al., 2010), and this change is related to an increased engagement
of brain areas related to attention and working memory in complex listening situations (Wong et al., 2009). Therefore, enhanced cognitive and perceptual abilities in older adult musicians may allow for more effective and flexible strategies during complex auditory perceptual tasks. Increased cognitive reserve would be particularly important in older adults who often experience audiometric threshold elevations, even when hearing is within a clinically normal range. For example, almost all older adults demonstrate some age-related decline in hearing sensitivity, even though many do not meet clinical criteria for age-related hearing loss (Gates & Mills, 2005). Any improvement in the processing of this impoverished signal will enhance overall hearing abilities. Therefore, if lifelong musicianship prevents the decline of low-level perceptual processing, then cognitive and other attentional resources would be freed to provide greater compensation for declining peripheral sensitivity. If, on the other hand, lifelong musicianship slows age-related cognitive decline, then older musicians would simply have more resources to deal with complex auditory processing tasks. The continued practice of a musical instrument may therefore result in greater neural efficiency, greater neural capacity, or the ability for compensation via the recruitment of additional brain regions during auditory processing, in addition to enhanced perceptual processing of acoustic features.

Indeed, evidence from cross-sectional and longitudinal neuro-imaging studies in children and young adults have shown that musical training can have long lasting effects on the central auditory system. For instance, using a cross-sectional design, Schneider et al. (2002) demonstrated that young musicians have enhanced grey matter density in the auditory cortex that correlated with musical aptitude. Utilizing longitudinal designs, Hyde et al. (2009) showed that children who received 15 months of music lessons had structural brain changes that correlated with performance in musically-related tasks. Fujioka et al.
(2006) compared neuro-magnetic brain activity in children before and after participating in music lessons over a one year period, and found enhanced activity in musically trained children compared to matched controls who did not receive music lessons. Another study examining neuro-magnetic brain activity demonstrated that in young adults, two weeks of musical training can yield reliable changes in central auditory processing (Lappe et al., 2008). Together, these studies provide converging evidence that musical training is associated with neuroplastic changes in the auditory cortex that is related to an enhanced ability to process auditory material. Enhanced perceptual processing of auditory material should result in improved performance on any auditory task, including speech processing, because, although understanding speech requires a broad network of brain areas (Wong et al., 2009), early perceptual processing of all auditory information is similar.

Indeed, the benefit of musical training spans from simple perceptual tasks that require detecting a single acoustic feature (e.g., Micheyl et al., 2006; Rammsayer & Altenmuller, 2006; Schellenberg & Moreno, 2010), to higher order cognitive abilities that tap into both perceptual and cognitive functions such as language and memory (Lee et al., 2007; Parbery-Clark, Skoe, Lam et al., 2009; Schellenberg, 2004). Any task that requires processing of an acoustic input will engage the auditory cortex, including detection of a silent gap or a mistuned harmonic (Alain, 2007; Ross et al., 2010). More complex auditory tasks, such as processing speech in the presence of background noise will engage a more complex network that comprises the auditory, parietal, and prefrontal cortices (Wong et al., 2009). Interestingly, when processing speech in background noise, older adults, compared to younger adults, show increased activity in the fronto-parietal network, and decreased activity in the auditory cortex, suggesting that older adults automatically compensate for declining perceptual abilities with a more cognitive strategy (Wong et al., 2009).
It is therefore possible that differential age-related decline in musicians for the speech-in-noise test observed in the current study indicates enhanced engagement of this fronto-parietal network. On the other hand, younger musicians demonstrate a more robust subcortical encoding of speech embedded in noise (Parbery-Clark, Skoe, & Kraus, 2009). This suggests that the differential age-related decline observed for the speech-in-noise task in the current study may reflect preserved early processing and encoding of speech sounds in the ascending auditory pathway, thereby providing older musicians with a more robust signal at the auditory cortex. Furthermore, the benefit musicians have in detecting a gap or mistuned harmonic may also be due to enhanced subcortical processing and/or enhanced cognitive strategies. To clarify these findings, an experiment comparing the functional properties of the auditory cortex was carried out in older and younger musicians and non-musicians (Chapter 3).

Understanding the neural locus of training-induced plasticity in musicians may be useful for designing specific rehabilitation programs to improve the auditory processing abilities of older adults. These data demonstrate that musicians have enhanced auditory processing abilities and suggest that continued musical practice throughout life contributes to slower decline of auditory processing tasks that rely on central mechanisms. In the next two chapters some of these central mechanisms will be investigated via analysis of scalp-recorded brain potentials.
Chapter 3:

3.1 The influence of age and musical training on the auditory evoked response

One reason older adults have difficulty with hearing is because of age-related changes to central auditory processing. At the same time, age-related decline on auditory assessments that rely on central processing are mitigated in lifelong musicians (see Chapter 2). Therefore, one important question that remains unanswered is how age-related changes to the central auditory system are mitigated by lifelong musical training.

Central processing of auditory information can be separated into two sub-processes, each of which may be differentially enhanced in older musicians. The first stage of auditory processing can be described as automatic, bottom-up, or obligatory and occurs regardless of where a listener's attention is focused (e.g., Alain, Arnott, Hevenor, Graham, & Grady, 2001; Bregman, 1990; Schneider et al., 2010). Appropriately, neural activity related to this processing stage is considered exogenous. The second stage of auditory processing has been described as controlled, top-down, attention-dependent processing that requires a listener to focus attention towards the auditory environment (e.g., Alain, Arnott, Hevenor et al., 2001; Bregman, 1990; Schneider et al., 2010). Neural activity related to this processing stage is therefore endogenous. The findings from Chapter 2 suggested a central benefit for older musicians; however, because psychoacoustic testing necessitates a response, it is impossible to determine if the benefit was due to enhanced exogenous or enhanced endogenous processing of acoustic information, or both. One way to separate these two stages of auditory processing is to compare scalp recorded auditory evoked responses (AERs) in situations where listeners
ignore an auditory signal to AERs when listeners attend to and make a judgment about the same auditory signal (see Chapter 1 for more information about AERs). In both situations automatic exogenous processing of acoustic features will be present, and reflected in the AER (Näätänen, 1992). On the other hand, additional endogenous neural activity, which is dependent on attending to, and, making a judgment about the auditory stimulus are only present in the latter situation. This endogenous neural activity may modify components of the AER, and engender entirely attention-dependent components that may be influenced by aging and/or musical experience.

Numerous studies have reported endogenous neural activity by comparing AERs to stimuli that are attended to versus stimuli that are ignored (Hansen & Hillyard, 1980; Hillyard, Hink, Schwent, & Picton, 1973; Näätänen et al., 1978). Endogenous neural activity becomes increasingly prominent as the latency from the onset of the acoustic stimulus increases (Coles & Rugg, 1995). This provides support for a two-stage model of auditory processing, where primitive acoustic features are first automatically extracted from the acoustic signal, and second where conscious recognition and cognitive processing of those features occurs. Some evidence suggests that endogenous neural activity can be observed as early as 20-50 ms post-stimulus onset, where selective attention results in an increase in positivity for the AER, described as a P20-50 (Woldorff & Hillyard, 1991). There is some debate as to whether this early influence of attention is truly a separate endogenous component, or simply an enhancement of the exogenous processing that precedes the formation of a complete stimulus representation (Näätänen, 1992). Following the P20-50 is a much larger negative wave that is first observed during the N1-P2 epoch, and can persist for several hundred milliseconds. This endogenous activity has been described as a negative difference wave (Nd; Hansen & Hillyard, 1980) or a processing negativity (PN; Näätänen et al., 1978). While early stages of the Nd/PN
likely represent an endogenous enhancement of obligatory auditory processing, later stages of the Nd/PN most likely represent the maintenance of an acoustic stimulus in memory so that its’ features can be compared to stored schemata (Näätänen, Kujala, & Winkler, 2010). Comparing endogenous and exogenous activity is particularly important because it is well known that exogenous processing of acoustic features declines in older adults, and that older adults tend to use controlled, and thus endogenous neural processes to overcome these deficits (Schneider et al., 2010).

Not surprisingly, aging has an impact on the long-latency exogenous AER; however, the results have been somewhat inconsistent, likely due to increased population variability in hearing and cognitive abilities for older adults. Many studies have shown that the amplitude of the P1 (Aine et al., 2005; Smith, Michalewski, Brent, & Thompson, 1980), N1 (Alain & Woods, 1999; Anderer, Semlitsch, & Saletu, 1996; Chao & Knight, 1997a) and P2 (Amenedo & Diaz, 1998; Anderer et al., 1996; Ceponiene et al., 2008; Chao & Knight, 1997a) waves are larger in older adults. On the other hand, some studies have failed to replicate these age-related findings for P1 (Tremblay, Piskosz, & Souza, 2003), N1 (Ceponiene et al., 2008) and P2 (Smith et al., 1980). It has also been shown that the N2 response decreases in amplitude with age (Anderer et al., 1996; Bertoli & Probst, 2005; Ceponiene et al., 2008). In addition, the P2 (Anderer et al., 1996) and N2 (Ceponiene et al., 2008) latencies increase with age, while the P1 (Aine et al., 2005; Ceponiene et al., 2008; Smith et al., 1980; Tremblay et al., 2003) and N1 (Ceponiene et al., 2008; Smith et al., 1980) waves have similar latencies in older and younger adults. It is unlikely that these changes are due to age-related decreases in peripheral sensitivity, as decreasing the amplitude of a stimulus (which can be analogous to a decrease in sensitivity) results in a decrease in amplitude of the AER (Crowley & Colrain, 2004; Näätänen & Picton, 1987), while in older adults the P1-N1-P2 tend to be enlarged.
One theory explaining this pattern of results is that of pre-frontal cortical inhibitory decline. This theory postulates that aging results in a decrease in the effectiveness of top-down modulation of auditory areas by the pre-frontal cortex (Knight, Staines, Swick, & Chao, 1999; Kok, 1999; Bertoli & Probst, 2005). In younger adults these top-down modulations inhibit irrelevant incoming acoustic information from entering the cortex from the thalamus, thus reducing the amplitude of the AERs. In older adults the pre-frontal inhibitory activity declines, which results in an age-related amplitude increase to the AER. Supporting this possibility was the finding that patients with damage to the pre-frontal cortex also exhibited enlarged AERs (Knight, Scabini, & Woods, 1989). Furthermore, in older adults reduced pre-frontal activity was related to an increase in the amplitude of mid-latency AERs, suggesting a connection between pre-frontal areas and early cortical processing of acoustic information (Chao & Knight, 1997b). This decline in inhibitory function allows a greater amount of irrelevant acoustic information to enter the cortex resulting in a larger, but noisier exogenous auditory signal. Moreover, this irrelevant acoustic information would require greater endogenous processing of acoustic information in order for the older listener to focus on relevant acoustic cues.

In addition to age-related changes in the AER, there have also been observations of enhancements to the AER in musicians. For instance the N1 (Baumann et al., 2008; Pantev et al., 1998; Pantev et al., 2001), and P2 (Kuriki et al., 2006; Shahin et al., 2003; Shahin, Roberts, Pantev, Trainor, & Ross, 2005) waves are enhanced in musicians. Some evidence suggests that the N1 enhancements may be specific to musical stimuli (Pantev et al., 1998), or even the instrument the musician trained on (Pantev et al., 2001). That is, the enhancements to N1 were largest when violinists were presented with violin tones or when trumpeters were presented with trumpet tones. On the other hand, P2 was enhanced
in musicians, regardless of whether the stimulus was a pure-tone or had a musical timbre (Kuriki et al., 2006; Shahin et al., 2003; Shahin et al., 2005). Interestingly, P2 amplitude was also related to the spectral complexity of the sounds, with the difference in P2 amplitude increasing as the spectral complexity (i.e., more harmonics) of the evoking stimulus increased (Baumann et al., 2008; Shahin et al., 2005).

Musician-related modulations of the AER suggests neuro-plastic changes in the underlying neural generators, which could be expressed as: (1) an increase in the size of cortical area (2) higher degrees of synchronization within a particular neural ensemble; (3) sharpening the tuning of neurons and/or (4) changes in cortical maps. Schneider et al. (2002) found enhanced grey matter volume for Heschl’s gyrus in musicians that was correlated with musical aptitude, supporting the idea of an increased cortical area related to processing auditory information. In addition, Baumann et al. (2008) found an Nd-like component during a selective auditory attention task that was larger in musicians compared to non-musicians. Source analysis of this Nd-like component revealed that it originated in the anterior cingulate cortex (Baumann et al., 2008), which is a structure that is important for attending to complex cognitive tasks (Posner & Dehaene, 1994). These findings suggest that, in addition to enhanced cortical processing of basic acoustic parameters, musicians also have an advantage in focusing attention to salient acoustic features (Baumann et al., 2008). Therefore, advantages in auditory processing observed in younger musicians are likely related to enhanced exogenous and enhanced endogenous neural activity that underlie the processing of acoustic information.

One question that remains unanswered is how age, attention and musical training interact to influence the auditory evoked response. This is a particularly important question because age-related decline in central auditory processing is common, and
finding ways to mitigate this decline could improve the quality of life for many older adults. To test this possibility sounds were presented to four groups of participants (younger and older, musicians and non-musicians) in two listening conditions: one that required attentive listening (active listening) and another placed the focus of attention elsewhere (passive listening). The attentional manipulation was important because it allows for the separation of endogenous and exogenous stages of auditory processing.

3.2 Methods

3.2.1 Participants

Fifty-seven participants were recruited for the study and provided formal informed consent in accordance with the joint Baycrest Centre-University of Toronto Research Ethics Committee. These participants were made up of four groups: older musicians (range: 58-91 yrs, M = 69 yrs, SD = 9.24), older non-musicians (range: 61-84 yrs, M = 69.2 yrs, SD = 6.69), younger musicians (range: 23-33 yrs, M = 28.1 yrs, SD = 3.17) and younger non-musicians (range: 23-39 yrs, M = 29.9 yrs, SD = 5.97). Musicians were defined as having advanced musical training (e.g., university degree, Royal Conservatory grade 8, college diploma, or equivalent) and continued practice on a regular basis until the day of testing, while non-musicians had no more than two years of formal training throughout life, and did not currently play a musical instrument. All participants were screened using a telephone interview for neurological or psychiatric illness, and for hearing loss using pure-tone audiometry. Noise-induced hearing loss is a common problem for older musicians because of life-long exposure to high amplitude sounds (Jansen, Helleman, Dreschler, & de Laat, 2009). Not surprisingly, some participants in the older musician group met the threshold for mild hearing loss, based on a pure-tone
threshold assessment (i.e., 25-35 dB HL; for more information about pure-tone thresholds see Chapter 2). To compensate for this, older non-musicians with mild-hearing loss were recruited so that pure-tone thresholds in older non-musicians did not differ from older musicians. To confirm this, a 2 (Musical training: musician, non-musician) by 6 (Pure Tone Frequency: 250, 500, 1000, 2000, 4000 & 8000 Hz) repeated measures ANOVA was calculated for the older adults. Neither the main effect of Musical training, nor the interaction between Musical training and Pure-tone Frequency were significant (p > 0.5 for both). All younger adults had pure-thresholds within the normal range (i.e., below 25 dB HL at all frequency octaves). Some participants also completed the psychoacoustic tests reported in Chapter 1, as the assessments reported in Chapter 1 have become part of the standard hearing assessment for all participants in our lab.

3.2.2 Stimuli
Stimuli consisted of six complex sounds that were each created by adding together six pure tones of equal intensity (i.e., 220, 440, 660, 880, 1100, and 1320 Hz). The fundamental frequency ($f_0$) was 220 Hz and the third tonal element was either tuned (i.e., 660 Hz) or mistuned by 1% (666.6 Hz), 2% (673.2 Hz), 4% (686.4 Hz), 8% (712.8 Hz) or 16% (675.6 Hz) of its original value, yielding six complex sounds. In this chapter only responses to the tuned stimuli are considered; the influence of mistuning is considered in Chapter 4. Importantly, the mistuned harmonic was limited to five possible values in this study, while the mistuned harmonic in the detection task reported in Chapter 2 was adjusted based on the responses of the participant.

The pure tones were generated at a sampling rate of 22 050 Hz using Sig-Gen software (Tucker-Davis Technology, Alachua, FL), and were combined into a harmonic
complex using Cubase SX (Steinberg, V.3.0, Las Vegas, NV). All six harmonic complex
tones had durations of 150 ms with 10 ms rise/fall times. They were presented binaurally
at 80 decibels sound pressure level (dB SPL) using a GSI 61 Clinical Audiometer via ER-
3A transducers (Etymotic Research, Elk Grove, IL, USA). The intensity of the stimuli
and the noise were measured using a Larson-Davis sound pressure level meter that
measured the amplitude of the stimuli presented from the left insert earphone.

3.2.3 Procedure
All participants were presented with the same stimuli in two listening conditions,
active and passive. In the active listening condition, participants were asked to indicate
whether the incoming stimulus was perceived as a single complex sound (i.e., a buzz) or
two concurrently occurring sounds (i.e., a buzz plus another sound with a pure tone
quality; see Alain et al., 2001, Moore et al., 1986, or Chapter 4). Responses were
registered using a multi-button response box, and no feedback related to the responses was
given. In the passive condition participants were instructed to relax, and to ignore the
sounds, while watching a muted subtitled movie of their choice. This design allowed for
the examination of the effects of age and musical training on exogenous cortical activity
elicited by stimuli while minimizing the influence of top-down processes on AER
amplitude. The use of muted subtitled movies has been shown to effectively capture
attention without interfering with auditory processing (Pettigrew et al., 2004).

In both listening conditions, 720 stimuli were presented in each listening condition
(120 exemplars of each stimulus type). All participants completed six blocks of trials.
The first and last blocks were passive and each included 360 stimulus presentations (60 in-
tune harmonic complexes), the middle four blocks were active and each included 180
stimulus presentations (30 in-tune harmonic complexes). The probability of hearing an in-
tune harmonic complex on any given stimulus presentation was 1 in 6. The stimuli were
presented at an inter-stimulus onset interval (ISI) that was randomly varied according to a
rectangular distribution between 1200 and 2000 ms during passive trials and 2000-3000
ms during active trials to allow time for a response.

Increasing the mistuning is known to increase the likelihood of perceiving the
mistuned component as a separate, concurrently occurring sound (Moore et al., 1986).
Therefore, to assess the basic functional properties of the auditory system only AERs to
the in-tune stimuli are analyzed here; the influence of mistuning is considered in Chapter
4.

3.2.4 Recording of electrical brain activity

Neuroelectric brain activity was digitized continuously from 64 scalp locations
(Figure 3.1) with a band-pass filter of 0.05–100 Hz and a sampling rate of 500 Hz per
channel using SynAmps2 amplifiers (Compumedics Neuroscan, El Paso, TX) and stored
for offline analysis. Electrodes on the outer canthi and at the superior and inferior orbit
monitored ocular activity (see Figure 3.1: IO1, IO2, LO1, LO2, FP9, FP10). During
recording, all electrodes were referenced to the midline central electrode (i.e., Cz);
however, for data analysis, the ERPs were re-referenced to an average reference, and
electrode Cz was re-instated. All averages were computed using BESA software (version
5.2). The analysis epoch included 100 ms of pre-stimulus activity and 1000 ms of post-
stimulus activity. Trials containing excessive noise (±130 µV) at electrodes not adjacent
to the eyes (i.e., IO1, IO2, LO1, LO2, FP1, FP2, FPz, FP9, FP10) were rejected before
averaging. ERPs were then averaged separately for each condition, stimulus type, and
electrode site.
For each participant, a set of ocular movements was obtained before and after the experiment (Picton et al., 2000). From this recording, averaged eye movements were calculated both for lateral and vertical eye movements as well as for eye blinks. A principal component analysis (PCA) of these averaged recordings provided a set of components that best explained the eye movements. The scalp projections of these components were then subtracted from the experimental ERPs to minimize ocular contamination such as blinks, saccades, and lateral eye movements for each individual average. ERPs were then digitally low-pass filtered to attenuate frequencies above 30 Hz.

3.2.5 Data Analysis

To determine if there were any differences in exogenous neural activity, peak amplitude and latency during the 40-80 ms (P1), 50-150 ms (N1), 140-260 ms (P2), and 220-400 ms (N2) epochs’ were compared during passive listening only. To further examine early exogenous cortical responses, the Pa & Nb waves of the mid-latency response were also analyzed. To isolate this activity, filter settings were changed to a band-pass set at 10-100 Hz. Moreover, peaks for these waves were not clear for each participant, therefore the mean amplitude during the 30-45 ms & 40-55 ms epoch were quantified as the Pa and Nb waves respectively. The effects of age, musical status, and listening condition on passive AER latency and amplitude were quantified at nine central electrodes (FC1, FCz, FC2, C1, Cz, C2, CP1, CPz & CP2; see Figure 3.1). The amplitude and latency for each of the waves in the AER were analyzed using a mixed design repeated measures ANOVA with Age group and Musical training as the between-subject factors and Electrode site as the within-subject factor. When appropriate, the degrees of freedom were adjusted with the Greenhouse-Geisser epsilon (ε) and all reported
probability estimates were based on the reduced degrees of freedom, although the original
degrees of freedom are reported. The nine electrode montage was used to ensure a stable
and reliable estimate for the AERs, and therefore interactions with the Electrode factor are
not reported.

The influence of endogenous neural activity on the scalp topography was different
in each group, therefore, picking an electrode montage may have biased the results. To
eliminate this potential bias, a discrete source model was used to reduce the scalp data into
regional sources. The analysis assumed a 4-shell ellipsoidal head model with relative
conductivities of 0.33, 0.33, 0.0042, and 1 for the head, scalp, bone, and cerebrospinal
fluid, respectively, and sizes of 85 mm (radius), 6 mm (thickness), 7 mm (thickness), and
1 mm (thickness). As an initial step, we used a surrogate model from BESA software
(version 5.2) designed to model auditory evoked potentials. This model was comprised of
11 regional sources, each containing 3 orthogonal dipoles to account for all directions of
current flow at the source location (Frontal pole, frontal left, frontal midline, frontal right,
central midline, parietal left, parietal midline, parietal right, occipital, auditory cortex left
& auditory cortex right; see Figure 3.2). In each participant, the resulting model was held
fixed and used as a spatial filter to derive source waveforms for both listening conditions
(active and passive). To limit the number of statistical analyses, a visual inspection of the
endogenous activity (i.e., the difference between active and passive listening for the group
averaged data) was used to pick epochs and sources where group differences in
endogenous neural activity likely existed. The largest group differences in endogenous
activity were from 50-850 ms in the left and right auditory cortical sources, in addition to
shorter epochs from sources at the at the frontal pole, occipital cortex, and parietal cortex.
To quantify these differences, neural activity was averaged across 100 ms epochs (i.e., 50-
Therefore, when an effect was significant across multiple epochs, multiple F values are reported, sequentially. Probability estimates for these multiple comparisons are the sum of each of the p values for each of the F values reported (i.e., the Bonferroni correction). The sources in the left and right auditory cortex were analyzed together using a mixed design ANOVA that included Musical training, Age group, Listening condition, and Hemisphere in the analysis. The source activity from other sources was analyzed using the same factors above, excluding Hemisphere. The focus of these analyses was on endogenous neural activity, therefore only group (Age group and/or Musical training) by Listening condition interactions are reported.

3.3 Results
In all participants, the harmonic complex tone evoked P1, N1, P2, and N2 waves that peaked around 50 ms, 100 ms, 200 ms and 320 ms, respectively. In addition, band-pass filtering between 10-100 Hz revealed an Pa-Nb complex that overlapped the P1 response. All waves peaked in amplitude around central electrode sites, and had inversions (i.e., the opposite polarity) around mastoid sites, suggesting sources along the superior temporal plane. The response from each of the 64 electrodes, and related scalp topography for each of the AERs can be seen in Figure 3.3. Figure 3.4 shows the AER recorded at electrode Cz, while Figure 3.5 illustrates the amplitude for each of the waves averaged across nine central electrodes, separately for each of the four groups.
3.3.1 Effects of age and musical training on exogenous scalp recorded auditory evoked responses.

For P1 latency, the main effects of Age group and Musical training and their interaction was not significant ($p > 0.1$).

P1 amplitude was larger in older adults compared to younger adults ($F(1, 53) = 12.36, p < 0.01$), and was smaller in Musicians compared to non-musicians ($F(1, 53) = 10.87, p < 0.01$). Importantly, the Age group and Musical training factors did not interact ($p > 0.1$).

For N1 latency the main effects of Age group and Musical training and their interaction were not significant ($p > 0.1$).

N1 amplitude was larger in older adults compared to younger adults ($F(1, 53) = 8.50, p < 0.01$). The main effect of Musical training and its’ interaction with Age group were not significant ($p > 0.1$).

P2 latency was longer in older adults (210 ms, S.E. = 3.9) compared to younger adults (187 ms, S.E. = 3.8; $F(1, 53) = 18.03, p < 0.001$). The main effect of Musical training and it’s interaction with Age group were not significant ($p > 0.1$).

For P2 amplitude, the main effect of Musical training, Age group and their interaction were not significant ($p > 0.1$).

N2 latency was longer in older adults (364 ms, S.E. = 6.7) compared to younger adults (323 ms, S.E. = 6.5; $F(1, 53) = 19.25, p < 0.001$). The main effect of Musical training and its’ interaction with age group were not significant ($p > 0.1$).

N2 amplitude was smaller in older adults compared to younger adults ($F(1, 53) = 14.11, p < 0.001$). The main effect of Musical training and it’s interaction with Age group were not significant ($p > 0.1$).
Pa amplitude was not influenced by Age group or Musical training. On the other hand the Nb was larger in musicians compared to non-musicians ($F(1,53) = 11.10, p < 0.01$), and larger in younger adults compared to older adults ($F(1,53) = 22.01, p < 0.001$). The interaction between Musical training and Age group was not significant.

These data demonstrate that there are significant age-related changes to exogenous auditory evoked responses, including the P1, N1, P2 and N2. On the other hand, being a musician only resulted in a reduction in P1 amplitude, and this change did not interact with age. Furthermore, the P1 effect was paralleled by similar changes in the Nb wave. Given that these two waves overlap with each other, it is highly probable that the influence of Age group and Musical training on P1 amplitude was caused by differences in the Nb wave. Thus, exogenous processing, as indexed by Nb amplitude, demonstrates a preserved enhancement throughout the lifespan.

### 3.3.2 Effects of age and musical training on endogenous auditory evoked responses.

To investigate the influence of aging and musical training on endogenous auditory evoked responses, a surrogate source model containing eleven regional sources was applied as a spatial filter to the data. To examine endogenous responses from each of these sources, the statistical analyses focused on group interactions with Listening condition. Importantly, analyses were done separately for each epoch at each source location, and F values for each epoch are listed sequentially. The main findings are highlighted in Figure 3.6.

#### 3.3.2.1 Auditory Cortex

Although the focus of these analyses was on group interactions with listening condition, it is worth noting that the main effect of hemisphere was significant during
every epoch analyzed from 50-850 ms, as overall source activity was larger in the right auditory cortex ($F(1, 53) = 23.59, 25.07, 25.34, 25.06, 23.41, 14.74, 12.88, 5.36, p < 0.05$ combined). In addition, Musicians had greater overall source activity compared to non-musicians from 150-650 ms ($F(1, 53) = 8.46, 12.80, 11.88, 9.08, 5.05, p < 0.05$ combined) indicating enhanced exogenous processing of acoustic material beyond the epochs analyzed in the AER (see above). Most importantly, there was significant endogenous activity from 50-850 ms as the main effect of Listening Condition was significant ($F(1, 53) = 58.40, 10.08, 25.37, 50.37, 61.54, 57.29 & 62.69, p < 0.05$ combined; except the 150-250 epoch, where the effect was marginal, $F(1, 53) = 2.87, p = 0.09$). Surprisingly, the main effect of Age group on the overall source strength was not significant during any of the epochs examined ($p > 0.1$). The effect of Listening condition can be observed in Figure 3.6. In this figure, when the difference waves are above the horizontal grey line, there was greater activity during active listening, thus indicating endogenous neural activity.

To illustrate endogenous neural activity, difference waves (active listening minus passive listening) are plotted in Figure 3.6. The actual source waveforms recorded during active listening are shown in Figure 3.7, and the source waves recorded during passive listening are shown in Figure 3.8. The Age group by Listening condition interaction was significant during the 350-850 ms epoch ($F(1, 53) = 4.63, 10.84, 5.76, 5.86, 5.15, p < 0.05$ each). During this epoch, older adults had greater endogenous activity compared to younger adults as there was a larger difference between active and passive listening in older adults ($t(27) = 5.09, 6.97, 6.69, 6.58, 6.24, p < 0.01$ combined) compared to younger adults, ($t(28) = 2.06, 2.65, 4.10, 3.87, 4.90, p > 0.05$ combined).

The Musical training by Listening condition interaction was significant during the 450-550 ms epoch ($F(1, 53) = 6.12, p < 0.05$), and was marginally significant during the
550-750 ms epoch \((F(1, 53) = 3.69, p = 0.06 & 3.21, p = 0.07)\). During this epoch, musicians had greater endogenous activity compared to non-musicians as there was a larger difference between active and passive listening in musicians \((t(28) = 5.31, p < 0.01)\) compared to non-musicians \((t(27) = 3.97, p < 0.05)\).

During the 550-650 ms epoch there was a significant Age group by Musical training by Listening condition by Hemisphere interaction \((F(1, 53) = 5.43, p < 0.05; \text{ see Figure 3.6 & 3.9})\). To determine the source of this interaction, planned t-tests revealed that the endogenous activity (active listening minus passive listening) was largest in the right auditory cortex for older musicians \((t(14) = 5.59, p < 0.001)\), followed by younger musicians \((t(13) = 4.29, p < 0.01)\), then older non-musicians \((t(12) = 2.67, p < 0.05)\). Interestingly younger non-musicians did not demonstrate significant endogenous activity during this epoch \((t(14) = 1.92, p > 0.05)\). This four way interaction continued at a trend level through the 650-750 ms epoch \((F(1, 53) = 2.25, p = 0.14)\).

3.3.2.2 Occipital midline source

There was significant endogenous neural activity during the 250-450 ms epoch at the occipital source, as the main effect of Listening condition was significant \((F(1, 53) = 25.97 & 36.07, p < 0.01 \text{ combined})\). During this epoch endogenous neural activity was larger in older adults \((F(1, 53) = 7.30 & 5.21, p < 0.05 \text{ combined})\). The effect of Musical training and its’ interactions with Age group were not significant during this epoch.

3.3.2.3 Parietal midline source

There was significant endogenous neural activity during the 150-850 ms epoch, as the main effect of Listening condition was significant \((F(1, 53) = 25.97, 36.07, 40.94)\).
43.84, 21.53, 23.00, 34.01, \( p < 0.01 \) combined). More importantly, there was greater endogenous neural activity in musicians compared to non-musicians, as the Listening condition by Musical training interaction was significant \( (F(1, 53) = 5.03, 8.83, 11.83, 12.42, 7.10, 8.31, 10.70, p < 0.05 \) combined). The effect of Age group, and its’ interaction with musical training was not significant \( (p > 0.1) \).

3.3.2.4 Central midline source

There was significant endogenous neural activity during the 550-750 ms epoch as the main effect of listening condition was significant \( (F(1, 53) = 48.84 \& 47.41, p < 0.01 \) combined). Although it appears in Figure 3.6 that there is greater endogenous activity in older musicians, none of the higher order interactions were significant \( (p > 0.1) \).

3.3.2.5 Frontal midline source

There was significant endogenous neural activity during the 450-750 ms epoch as the main effect of listening condition was significant \( (F(1, 53) = 22.76, 30.99 \& 31.98 p < 0.01 \) combined). Although it appears in Figure 3.6 that there is reduced endogenous activity in older non-musicians, none of the higher order interactions were significant \( (p > 0.1) \).

3.3.2.6 Frontal pole midline source

There was significant endogenous neural activity during the 250-850 ms epoch as the main effect of listening condition was significant \( (F(1, 53) = 6.69, 15.56, 24.16, 30.34, 30.95 \& 26.45, p < 0.01 \) combined). Surprisingly, the differences between groups that are
visible on Figure 3.6 were not real, as none of the higher order interactions were significant \((p > 0.1)\).

### 3.3.2.7 Left and right frontal and parietal sources

Although there appear to be some group differences in these sources (see Figure 3.6), no group by Listening condition interactions were significant \((p > 0.1)\).

### 3.4 Discussion

Results from the current study replicate previously reported age-related changes to obligatory components of the auditory evoked response. The current study demonstrated that the P1 and N1 were larger in older adults (the P1 effect may be due to a reduced Nb amplitude), P2 amplitude was delayed in older adults, and the N2 was smaller and delayed in older adults. Surprisingly, the only influence of musical training on the AER was that musicians demonstrated a reduction in P1 (or increase in Nb) amplitude. Other than P1, scalp recorded exogenous sensory evoked responses were little influenced by musical training; however, source activity from the auditory cortex was enhanced in both active and passive listening in musicians during the 150-650 ms epoch, suggesting enhanced exogenous processing of acoustic material beyond the P1-N1-P2-N2 complex. At the same time, attention-dependent endogenous responses were significantly influenced by both musical training and aging. Endogenous responses from the occipital source were larger in older adults during the 250-450 ms epoch, while endogenous responses from the parietal source during the 150-850 ms and from bilateral auditory cortices during the 450-750 ms epoch were larger in musicians. Most importantly, endogenous activity was largest for older musicians in the right auditory cortex during the 550-750 ms epoch.
The first step of the current analysis was done on the scalp recorded auditory evoked potentials. This ensured that the current data was consistent with previous research that has reported age or music training related changes to the AER. The P1 and N1 waves were larger in older adults, which is consistent with previous research demonstrating an age-related increase in P1 (Aine et al., 2005; Smith et al., 1980) and N1 amplitude (Alain & Woods, 1999; Anderer et al., 1996; Chao & Knight, 1997a). In addition the P2 wave was delayed, and the N2 wave was smaller in older adults, both of which are consistent with previously reported age-related changes to the AER (Anderer et al., 1996; Bertoli & Probst, 2005; Ceponiene et al., 2008). This pattern of results has been attributed to an age-related decline in automatic pre-frontal inhibition of irrelevant auditory information in older adults (Knight et al., 1999; Kok, 1999; Bertoli & Probst, 2005). The current data found no age by musical training interactions for the scalp recorded AERs; however, the P1 in older musicians resembled that of younger non-musicians, which may reflect a preserved differentiation of enhanced inhibitory function in older musicians.

Musicians had a reduced P1 amplitude compared to non-musicians, and increased exogenous activity in the auditory cortex during the 150-650 ms epoch. Interestingly, these changes did not interact with age, suggesting that processing of exogenous acoustic information is enhanced in musicians. More importantly it suggests that age-related changes that occur in the exogenous processing of acoustic information declines at the same rate in both musicians and non-musicians. The difference in P1 amplitude in musicians may be due to a more robust signal entering the auditory cortex from the thalamus and brainstem. Previous work demonstrated that younger musicians more robustly encode the higher harmonics of an acoustic stimulus in the brainstem (Bidelman & Krishnan, 2010; Parbery-Clark, Skoe, & Kraus, 2009; Wong, Gauthier, Woroch,
DeBuse, & Curran, 2005). Another study demonstrated larger mid-latency AERs in musicians (Schneider et al., 2002). Given the similarities between the influence of musical training on P1 and Nb, and that they occur during overlapping epochs, it is likely that the reduction in P1 amplitude was actually due to an enhanced negativity in the Nb wave. Furthermore, the evoking stimulus in the current study was complex, and contained multiple harmonics, therefore it is also possible that the changes to P1/Nb amplitude were related to enhanced brainstem encoding of multiple harmonics. Interestingly, the P1/Nb in older musicians was similar to the P1/Nb in the younger non-musicians. Thus, while the age-related change in P1/Nb was similar for both musicians and non-musicians, lifelong musicianship seems to maintain the P1/Nb at or above the amplitude observed for young adults.

The increased activity in the auditory cortex between 150-650 ms likely represents enhanced cortical processing of harmonically complex acoustic material. A similar enhancement has been observed in younger adults for the N1c and P2 waves, which occur around 150 and 190 ms respectively, and have radial sources that originate in the auditory cortex (Shahin et al., 2003, 2007). Data from the current study suggest that these enhancements may actually be due to an increase in neural activity from the radially oriented sources in the auditory cortex, and not an enhancement to tangentially oriented sources in the auditory cortex, which are known to underlie scalp recorded AERs measured at fronto-central electrodes. Given that each component of the AER has multiple generators that sum together, this proposal is highly probable (Nääätänen & Picton, 1987; Crowley & Colrain, 2004; Shahin et al., 2007). Importantly, these findings also demonstrate that the exogenous processing enhancement observed in musicians persists beyond the N1-P2 complex. While these findings suggest that processing exogenous auditory information is enhanced in musicians, none of these findings point to
the moderating factor that slows age-related changes in processing acoustic information (as shown in Chapter 2), as none of the exogenous processes demonstrated age by musical training interactions.

It is therefore likely that there are some endogenous processes that become increasingly enhanced via lifelong musical training. Data from the current study found endogenous brain activity at each of the 11 regional sources, but was clearly most prominent at sources in the auditory cortex (see Figure 3.6). Group differences were noted at three sources: bilateral auditory cortices, the midline parietal source, and the midline occipital source. At the midline occipital source and in bilateral auditory cortical sources, older adults exhibited greater endogenous activity compared to younger adults. Although the occipital lobe contains the primary visual cortex, it has been implicated as part of a neural network used to understand speech (Gonzalez-Castillo & Talavage, 2011; Wong, Uppunda, Parrish, & Dhar, 2008; Zekveld, Heslenfeld, Festen, & Schoonhoven, 2006). Given the latency of the age-related enhancement to endogenous neural activity in the auditory and occipital sources this activity may be due to the processing negativity that is related to matching a stimulus to a stored schema (Näätänen et al., 2010). Thus older adults may engage this process to a greater extent to compensate for age-related decline in the peripheral encoding of auditory information, so that an accurate judgment about the incoming stimulus can be made. This finding further supports the idea that older adults use a more cognitive strategy to process acoustic information (Schneider et al., 2010).

Most importantly, these endogenous processes were largest in older musicians during the 550-750 ms epoch in the right auditory cortex. It is well established that the right auditory cortex is more sensitive to processing spectral aspects of the acoustic stimulus (Warrier et al., 2009; Zatorre, 1988). This specialization for spectral processing in the right auditory cortex is in contrast to the left auditory cortex, which has been shown
to be more sensitive to rapid temporal changes within an acoustic stimulus (Warrier et al., 2009; Liegeois-Chauvel et al. 1999). Interestingly, tonal patterns in music tend to change slowly, compared to language which involves much more rapid temporal fluctuations. At the same time, music perception requires much more complex spectral analysis compared to speech perception. Accordingly, this hemispheric distinction has led researchers to describe the right hemisphere as being important for music perception, and the left hemisphere as being important for language perception (Zatorre, Belin & Penhune, 2002).

In the current study, the experimental task during active listening required the listener to make a judgment about the spectral relationships in the acoustic stimulus. Results here demonstrate that older non-musicians engage endogenous activity during this epoch, but older musicians engage it to a greater extent. The pattern of results suggests that older adults partially compensate for decreased peripheral processing of spectral information by engaging additional endogenous processes, and that this additional processing is occurring in a brain area known to be important for processing spectral information. Furthermore, this implies that musical training accumulates over the lifespan to further compensate for age-related decline in the peripheral encoding of acoustic information. Importantly this age-related, musical training benefit is for endogenous, attention-dependent neural activity related to processing acoustic information.

The final group difference in endogenous activity was in musicians compared to non-musicians. Greater endogenous activity was observed in musicians during the 150-850 ms epoch at the parietal source, and from 450-750 ms in bilateral auditory cortical sources. The activity from the parietal source, which is near the posterior cingulate is likely related to selecting important features within the acoustic stream (Posner & Dehaene, 1994). In this specific case it may be related to focusing attention to the frequency region where the mistuned harmonic component was. As with the older adults,
the additional endogenous activity in the auditory cortex observed in musicians likely reflects the enhanced processing of acoustic information required in complex listening situations that are common in musical performance.

Data from the current study suggest that being a musician influences age-related changes in auditory processing. Exogenous processing of acoustic information is likely enhanced in musicians, but declined at an equal rate in musicians and non-musicians, demonstrating a pattern of preserved differentiation (Salthouse, 2006). While some endogenous processes were enhanced in age, activity in the right auditory cortex was further enhanced in older musicians suggesting a pattern of differential preservation (or enhancement; Salthouse, 2006). This implies that the benefit of musical training accumulates over the lifespan, which further suggests that musical training alters endogenous processes related to processing complex acoustic stimuli. More importantly, it may explain why older adults show less age-related decline on psychoacoustic auditory tasks. The next chapter will examine the influence of aging and musical training on the perceptual organization of acoustic information based on spectral information in an auditory stimulus.
Chapter 4:

4.1 The influence of age and musical training on concurrent sound segregation

Most auditory scenes are complex, in that there are multiple active sound sources at any given time. Thus, one important process in auditory scene analysis is the ability to segregate concurrently occurring sounds (Bregman, 1990). Sounds that are perceptually segregated can then be tracked as separate auditory streams over time to form a dynamic perceptual auditory scene (Alain & Bernstein, 2008; Carlyon, 2004). There are multiple ways the auditory system can detect the presence of concurrently occurring sounds, including: onset asynchrony, spatial location, differences in fundamental frequency ($f_0$) and periodicity (Bregman, 1990). Using periodicity to group or segregate acoustic energy into single or multiple precepts is a particularly important process because it relies mainly on the relationship between spectral acoustic information. When acoustic energy is periodic across the frequency spectrum (i.e., bands of acoustic energy [harmonics or overtones] are multiples of a $f_0$) the separate bands of energy are perceived as a single sound object, while acoustic energy that is not related to the same $f_0$ is segregated into a second auditory precept. This is because natural sound sources (i.e., vibrating bodies) normally produce periodic patterns of acoustic energy.

Given the importance of periodicity in auditory perception, it is not surprising that humans can detect very small amounts of mistuning in an otherwise in-tune harmonic complex (Moore et al., 1985). Mistuning a higher frequency harmonic component adds a roughness to overall timbre of the sound (Moore et al., 1985), while mistuning a lower harmonic results in the perception of two simultaneous sounds, one with a ‘buzz-like’ quality and the other with a pure-tone ‘beep-like’ quality (Alain, 2007; Moore et al.,
Interestingly, the differential effect of mistuning higher or lower harmonics is thought to be related to the ability of auditory nerve fibres to phase lock to acoustic energy (Hartmann, McAdams, & Smith, 1990). Auditory nerve fibres exhibit phase locking for lower frequencies, but not higher frequency, with the cutoff being around 4000 Hz. The phase locking of auditory nerve fibres to lower frequency harmonics allows for an acute comparison of harmonic to the \( f_0 \), while a lack of phase locking at higher frequency harmonics makes such an acute comparison more difficult. Furthermore, acoustic energy in lower frequency bands is more likely to be the \( f_0 \) of a second auditory object because the \( f_0 \) is usually the lowest frequency component in a complex sound.

Mistuning a harmonic component within a harmonic complex is not the only way to induce the perception of concurrently occurring sounds. For example, altering the spatial location of a harmonic component can facilitate the perception of the spatially segregated component as a separate sound (McDonald & Alain, 2005). Assmann and Summerfield (1994) found that the ability to understand simultaneous vowels improved as the difference in \( f_0 \) between the two vowels increased. In another study, Hautus et al. (2005) presented listeners with a broadband noise that had an onset delay for a narrow frequency band within the noise presented to the left ear, and a non-delayed version presented to the right ear. This stimulus resulted in the perception of a low-frequency pitch, related to the onset asynchrony of the narrow band of noise, on the right side of space, known as a dichotic pitch, in addition to the broadband noise which was perceived as being centrally located. Interestingly, Carlyon et al. (1992) demonstrated that incoherent frequency modulation of a mistuned harmonic component can aid in the detection of the mistuned component, but that incoherent frequency modulation on its own does not lead to the perception of a second auditory object. Clearly, there are many ways to manipulate the acoustic parameters of a stimulus to induce the perception of
concurrently occurring sounds. Surprisingly, despite disparate methods used to induce the perception of concurrent sounds, a similar pattern of neural activity is observed during the perception of concurrent sound objects (Alain, 2007).

One way to observe neural activity related to the perception of concurrent sounds is by using the subtraction method of scalp recorded auditory evoked responses (AERs; see Chapter 1). In this method, AERs to stimuli that do not contain physical cues to aid in sound segregation are subtracted from AERs to stimuli that contain such cues. This subtraction yields brain activity related to processing of acoustic features that gives rise to the perception of two sounds, while the common brain activity related to the detection of a stimulus (i.e., the AER) is subtracted out.

Numerous studies have found that the perception of concurrently occurring sounds is paralleled by an increase in negativity, known as an ‘object related negativity’ (ORN) that peaks between the N1 and P2 waves, around 150 ms post stimulus onset (Alain, Arnott, & Picton, 2001). Appropriately, the ORN is best visualized as a difference wave where the AER to a stimulus perceived as being a single sound is subtracted from the AER to a stimulus perceived as two sounds. The ORN has been observed when the perception of the second sound is due to a mistuned harmonic (Alain, Arnott, & Picton, 2001; Alain, Schuler, & McDonald, 2002), a dichotic pitch produced by inter-aural time differences (Hautus & Johnson, 2005), the spatial location of a harmonic component (McDonald & Alain, 2005), or a difference in $f_0$ between concurrently occurring vowels (Alain, Reinke, He, Wang, & Lobaugh, 2005). Importantly, the ORN is correlated with the likelihood of reporting the presence of concurrently occurring sounds, when listeners are asked to make a perceptual judgment about the incoming acoustic stimulus (Alain, Arnott, & Picton, 2001).
More, importantly, the ORN is an automatic response because it is observed regardless of where a listeners’ attention is focused. For example the ORN is evoked when a listener is making a judgment about the incoming acoustic stimulus (i.e., active listening), or when a listener is ignoring the incoming acoustic stimulus by watching a silent, subtitled movie or by reading a book (i.e., passive listening; Alain, Arnott, & Picton, 2001; Alain et al., 2005; Alain et al., 2002; Hautus & Johnson, 2005; McDonald & Alain, 2005). Furthermore, the ORN is little influenced by task demands, or selective attention. Alain & Izenberg (2003) presented listeners with tuned and mistuned harmonic complexes separately, alternating between the left and right ear. Importantly, listeners were instructed to only attend to sounds presented to one of the ears, and ignore sounds presented to the other ear. The ORN evoked by stimuli presented to the unattended ear was not influenced by the difficulty of the auditory judgment at the attended ear (Alain & Izenberg, 2003). These results suggest that the ORN reflects automatic neural activity related to the organization of acoustic primitives into separate auditory objects.

Scalp-recording of AERs have also revealed attention-dependent, endogenous neural activity related to the perception of a second, concurrently occurring auditory object. When listeners were asked to make a judgment about an incoming acoustic stimulus, and the stimulus was perceived as being two concurrently occurring sounds, the ORN was followed by a positive wave that peaked at around 400 ms post-stimulus onset, appropriately termed the P400 (Alain, Arnott, & Picton, 2001; Hautus & Johnson, 2005). Given that the P400 was present only when participants were required to make a response about the stimulus, it is thought to index the conscious registration of concurrent sound objects. Like the ORN, the P400 can be visualized as a difference wave (i.e., AERs evoked by stimuli perceived as a single sound are subtracted from AERs evoked by stimuli perceived as two sounds). Moreover, the P400 amplitude was correlated with the
perception of concurrently occurring sounds, and was larger when participants were more likely to report hearing two concurrent sound objects (Alain, Arnott, & Picton, 2001; Alain et al., 2002). It is therefore likely that concurrent sound segregation occurs in two stages. In the first stage acoustic features are organized automatically, regardless of a listeners’ attentional state; this stage is reflected in the ORN. In the second stage, there is a conscious registration of the automatically segregated mistuned harmonic as a second auditory object. This stage of processing requires a listeners’ focused attention, and is reflected in the P400.

Another attention dependent endogenous wave has also been observed in relation to the perception of concurrently occurring sounds, and has been described as a late-positive complex (LPC; Zendel & Alain, 2009). Unlike the P400 and ORN this wave is related to the confidence in the categorization of the incoming acoustic stimulus and is not related to the perception of simultaneous auditory objects. Accordingly it is largest when the judgment is easier (e.g. when a harmonic complex is tuned or mistuned by 16%), and decreases in amplitude as the judgment becomes more difficult (e.g. when a harmonic complex is only mistuned by 2% or 4%; Zendel & Alain, 2009). It is likely that this late-positive activity is related other late-positivities that arise when participants are asked to make acoustic judgments and are influenced by their confidence in making that judgment. For example, when participants were asked to recognize if a word had been presented previously, a LPC was enhanced when the word was correctly recognized, compared to when the participant was less sure if the word had been presented previously (Rugg et al., 1998; Curran, 2000). The LPC can therefore be considered an attention-dependent wave that is reflective of confidence in making an acoustic judgment.
4.1.1 Aging and concurrent sound segregation

Given that older adults have difficulty segregating speech from background noise (e.g., Duquesnoy, 1983; Pichora-Fuller et al., 1995), it is likely that aging negatively impacts concurrent sound segregation. A number of studies have confirmed this. Older adults have more difficulty detecting inharmonicity within a harmonic complex (Alain, McDonald et al., 2001; Grube et al., 2003; Chapter 2). Not surprisingly, older adults also have difficulty segregating an inharmonic component of a harmonic complex into a separate auditory object; however, this age-related decline is related to the length of the stimulus. When the harmonic complex is short (i.e., 40 ms), older adults were less likely to hear the mistuned harmonic as a separate sound object and this age-related decline coincided with a reduction in ORN amplitude recorded during passive listening (Alain & McDonald, 2007). On the other hand, when the stimulus was longer (i.e., 200 ms) age-related differences were not observed in the likelihood of hearing the mistuned harmonic as a separate sound nor were age-related differences observed in the ORN amplitude (Alain, McDonald, & Van Roon, submitted). When concurrent vowel sounds were presented, the ORN associated with segregating and identifying two vowels presented simultaneously was smaller in older adults (Snyder & Alain, 2005). Interestingly, later activity (reported as an N2b) related to the conscious detection of concurrently occurring vowels, was comparable between older and younger adults (Snyder & Alain, 2005). This pattern of results suggests that aging negatively impacts automatic processing of acoustic features when there is less acoustic information available, while attention-dependent, endogenous processing of acoustic information is relatively preserved. This finding should however be interpreted with caution, as it still remains unknown how the P400 is influenced by age when using a mistuned harmonic paradigm. Alain & McDonald (2007) and Alain et al. (Submitted) recorded AERs only during passive listening, and thus no
P400 was evoked. It is possible that the N2b reported in Snyder & Alain (2005) was related to the conscious perception of simultaneous vowels; however, the use of vowel sounds likely engaged schema-driven processes because of the over-learned nature of speech stimuli, thus, this negativity is likely different from the P400 wave observed in young adults while using a mistuned harmonic paradigm.

4.1.2 Musicians and concurrent sound segregation

While aging has a deleterious effect on the ability to detect and segregate a mistuned harmonic, young musicians have an enhanced ability to detect a mistuned harmonic component (Koelsch et al., 1999). This advantage for musicians was maintained in older adults (Chapter 2). More importantly, musicians are more likely to hear a mistuned harmonic as a separate auditory object, and accordingly, the ORN, P400, and LPC are enhanced in younger musicians (Zendel & Alain, 2009). Furthermore, participants trained to segregate simultaneous vowels showed significant improvement in their ability to correctly identify both vowels (Reinke, He, Wang, & Alain, 2003), which suggests that the advantage for musicians in concurrent sound segregation may be learned and not inborn. While concurrent sound segregation is enhanced in musicians, and declines in older adults, it remains unknown if age-related decline in concurrent sound segregation is mitigated in musicians.

4.1.3 Summary

The purpose of this study was to investigate if being a musician can mitigate age-related changes to concurrent sound segregation. It is known that the ORN and P400 are enhanced in musicians. On the other hand only the ORN has been shown to decline in older adults, and this age-related change is limited to short duration stimuli. Furthermore, the influence of aging on the P400 is not well known. Given that only endogenous neural
activity related to the processing of complex sounds was enhanced in older musicians (Chapter 3), it is likely that only endogenous activity related to the processing of concurrent sound objects will be preserved or enhanced in older musicians. Therefore, the hypothesis for the current study is that the ORN will be enhanced in musicians, but not differentially influenced by age, while the P400 and/or the LPC will be enhanced in musicians, and these enhancements will be differentially preserved with age.

4.2 Methods

Data reported in this Chapter and in Chapter 3 were collected at the same time. Accordingly, Participants, Stimuli, Procedure and Recording of electrical brain activity were identical to Chapter 3 (i.e. sections 3.1.1 - 3.1.4). In this chapter the influence of mistuning a harmonic in a complex sound is considered.

4.2.1 Data analysis (Behavioral)

For the behavioral task participants were asked to perceptually categorize the harmonic complex as either a single buzz, or a buzz with an addition pure-tone (beep-like) component. This response data was analyzed in two ways. The first was utilized a percentage calculation, where the dependent measure was the percentage of trials reported as two sounds at each level of Tuning. For the tuned stimulus, this measure approaches zero percent (i.e., most trials were perceived as a single sound) while the 16% mistuned condition approaches 100% (i.e., most trials were reported as two sounds). This analysis was termed ‘perceptual judgment.’

A second measure of stimulus ‘judgment consistency’ was calculated from the same data. To calculate judgment consistency, perceptual judgment scores for each individual participant were converted using the following equations, where X is equal to the perceptual judgment score in percentage, and C is judgment consistency measure.
If $X > 0.5$ then $C = (2X)-1$

If $X < 0.5$ then $C = 2(1-X)-1$

This calculation yields a score that ranges between 0 and 1, with 0 representing chance responses for that stimulus type (each response was selected 50% of the time), and 1 representing perfectly consistent responses for that stimulus type (100% of responses were the same). The purpose of this measurement was to distinguish between participants that were making different perceptual judgments for the same stimuli (especially in the 2 & 4% mistuned conditions), and participants that were consistently making the same perceptual judgment for each stimulus, regardless of what the perceptual judgment was. Importantly, this calculation reveals information about a group of participants that is not directly available in the perceptual judgment measure. For example, in a group of highly consistent responders, some participants will report hearing two sounds on almost all trials, while other participants will report hearing only a single sound on almost all trials. Despite each individual participant’s response consistency, the overall perceptual judgment score for that group would be around 50%. Now imagine a second group where each individual participant was variable in their response for each stimulus, such that they reported hearing a single sound on 50% of the trials, and reported hearing two sounds on the other 50% of the trials. In this case the perceptual judgment score would also be around 50%. Thus, despite different individual response patterns between these two groups, the perceptual judgment scores are identical. When the perceptual judgment scores are converted to the consistency measure described above, the score for the group of consistent responders would be 1, while the score for the variable responders would be
0. Stimulus response consistency therefore describes the within-subject consistency of perception given an identical stimulus.

In addition to these two analyses of the response data, reaction time was calculated from the onset of the stimulus to the button press indicating a response, and is reported in milliseconds (ms). Each of the three behavioral measures were statistically analyzed with a 6 (Tuning condition) by 2 (Age group) by 2 (Musical training) mixed design repeated measures ANOVA.

4.2.2 Data analysis (electrophysiological)

ORN amplitude was quantified as the mean amplitude during the 100-190 ms epoch, over nine fronto-central electrodes (F1, Fz, F2, FC1, FCz, FC2, C1, Cz & C2; see Figure 3.1). These sites were chosen because previous studies have found that the ORN is largest at fronto-central sites (Alain, Arnott, & Picton, 2001). Importantly the ORN is a difference wave (i.e., AER from a mistuned stimuli is subtracted from the tuned stimuli), and is therefore measured statistically as a main effect of Tuning. Specifically, the ORN is due to an increase in negativity during the 100-190 ms epoch, related to an increasing amount of mistuning. During active listening this increase in negativity is associated with the likelihood of hearing concurrently occurring sounds; however the ORN is also observed when the listener attending away from the auditory environment. Therefore, only the main effect of Tuning, and interactions with Tuning are indicative of an ORN.

To quantify the change in mean amplitude related to Tuning, orthogonal polynomial decompositions were calculated, with a focus on the linear or quadratic trends. Before analysis, activity from each of the nine fronto-central electrodes was re-referenced to the linked mastoid. That is, the average amplitude of electrodes TP9 & TP10 was subtracted from the amplitude of each of the fronto-central electrodes (Luck, 2005). The purpose of
this re-referencing was to maximize voltage potentials at fronto-central sites. Previously, source analysis of the ORN (and P400) revealed generators along the superior temporal plane, that was orientated towards the vertex (i.e. electrode Cz; Alain, Arnott, & Picton, 2001). This source configuration results in a polarity reversal around mastoid sites. Visual analysis of the ORN scalp topographies from the current data set confirm that the ORN was maximal around electrode Cz (see Figure 4.1, top views), and was reversed in polarity around mastoid sites (see Figure 4.1, side views) Thus by using a linked mastoid reference the polarity reversal was included in the analysis of the fronto-central electrodes, which increases the reliability of the ORN amplitude. The analysis was carried out using a mixed design ANOVA that included Age group, and Musical training as between subject factors, and Listening condition and Tuning as within-subject factors.

P400 amplitude was quantified as the mean amplitude during the 250-350 & 350-450 ms epoch, over a frontal-right electrode montage (FC2, C2, CP2, C4, FC6, CP6 & C6; see Figure 3.1). These electrodes were chosen based on a visual inspection of the data that revealed the P400 peak to be around electrode C4 for all subjects (see Figure 4.1). Importantly the P400 is a difference wave (i.e., AER from a mistuned stimuli is subtracted from the tuned stimuli), and is therefore measured statistically as a main effect of Tuning. Specifically, the P400 is due to an increase in positivity during the 250-450 ms epoch, related to an increasing amount of mistuning. During active listening this increase in positivity is associated with the likelihood of hearing concurrently occurring sounds. Importantly, the P400 was not present during passive listening. Therefore, only the main effect of Tuning, and interactions with Tuning are indicative of the P400. To quantify the change in mean amplitude related to Tuning, orthogonal polynomial decompositions were calculated, with a focus on the linear or quadratic trends. Like the ORN data, the P400 data was re-referenced to the linked mastoid. Statistical analyses were the same as the
ORN analysis, except it did not include Listening condition as a factor, because P400 was not present during passive listening.

While the amplitude of the ORN and P400 was quantified by comparing the mean amplitude between tuned and mistuned conditions, the latencies for the ORN and P400 were determined by calculating a difference wave between the tuned and 16% mistuned condition for each participant. This limits the measure of ORN and P400 latency to the 16% mistuned condition. This was done because the 16% mistuned condition resulted in a clear ORN and P400 in all participants, while the ORN and P400 became increasingly difficult to observe at smaller levels of mistuning. ORN latency was defined as the largest negative value in the difference wave between 100-200 ms post-stimulus onset at fronto-central electrodes, while P400 latency was calculated as the largest positive value in the difference wave between 250-500 ms post stimulus onset at frontal-right electrodes. P400 latency was only calculated for active listening because there was no P400 during passive listening. Therefore the final analysis for ORN and P400 Latency was a 2 (Musical training) by 2 (Age Group) by 2 (Listening condition [ORN only]) ANOVA.

During active listening, the participants’ response on every trial was accompanied by a late positive complex (LPC) that peaked at about 550 ms post stimulus onset at parietal and parieto-occipital sites. Given that the LPC was present only during the active listening condition, the effects of mistuning, age and musical training was only examined during active listening. The LPC latency was quantified as the largest positive peak during the 300-800 ms epoch, averaged across a nine electrode montage over parietal regions (CP1, CPz, CP2, P1, Pz, P2, PO3, POz & PO4; see Figure 3.1). To account for the different latency between older and younger adults (see results), LPC amplitude was quantified as the mean amplitude for the 200 ms epoch surrounding the peak, yielding an analysis window of 417-617 ms for younger adults and 492-692 ms for older adults.
Importantly, the LPC is not related to the perception of a second auditory object, but rather related to making an auditory judgment, and is present at all levels of Tuning.

To determine the relationship between behavioral measures, and electrophysiological measures, within subject correlations were calculated between the amplitude of ORN (during active listening), P400 (350-450 ms epoch), and LPC (the amplitude was averaged across the electrode montage used, yielding a single amplitude measure for each level of Tuning), with the three behavioral measures: perceptual judgment, response consistency and reaction time. That is, a correlation co-efficient was calculated for each participant between the electrophysiological data and each behavioral measure across the six levels of Tuning. The mean of these correlations is reported, and is indicative of the relationship between behavior and electrophysiology in each participant. Significance was assessed using a one-sample t-test that compared the value of the correlation co-efficient to zero (α > 0.05).

4.3 Results

4.3.1 Perceptual Judgment

Figure 4.2 shows the group mean perceptual judgment in younger and older musicians and non-musicians. Overall, the likelihood of reporting the perception of two concurrent sound objects increased with mistuning ($F(5, 265) = 282.75, p < 0.001$; linear trend $F(1, 53) = 837.55, p < 0.001$). The main effect of Musical training was significant ($F(1, 53) = 5.86, p < 0.05$). Moreover, the interaction between Tuning and Musical training was also significant ($F(5, 265) = 4.21, p < 0.01$; linear tend ($F(1, 53) = 7.98, p < 0.01$). Follow-up pairwise comparisons indicated that musicians were more likely to report hearing two sounds when the harmonic was mistuned by 4%, 8% and 16% ($t(55) =...$
2.05, 2.56, and 3.65, respectively, \( p < 0.05 \) in all cases). In addition, musicians report hearing two sounds more often when the harmonic was mistuned by 2%; however, results were only significant at a trend level (\( t(55) = 1.87, p = 0.066 \)). The main effect of age was not significant nor was the interaction between Age group and Tuning (\( p > 0.05 \)). Although the influence of musical training appears to be smaller in older adults compared to younger adults in Figure 4.2, the interaction between Age group and Musical training was not significant nor was the three way interaction between Age group, Musical status and Tuning (\( p > 0.5 \)), suggesting that the benefit observed in musicians was similar for older and younger adults.

4.3.2 Judgment consistency

Figure 4.3 shows the group mean judgment consistency scores. The consistency of participants’ responses varied as a function of Tuning (\( F(5, 265) = 28.66, p < 0.01 \); quadratic trend \( F(1, 53) = 108.3, p < 0.001 \)). Pairwise comparisons revealed that participants were the least consistent when the harmonic was mistuned by 2% and 4% (\( p < 0.001 \)). Overall, musicians were more consistent in their responses than non-musicians (\( F(1, 53) = 17.72, p < 0.01 \)). However, Musical training did not interact with Tuning (\( p < 0.1 \)), suggesting that musicians were more consistent at all levels of Tuning. The main effect of Age group was not significant (\( p > 0.1 \)); however, there was a significant interaction between Age group and Tuning (\( F(5, 265) = 2.61, p < 0.05 \); linear trend \( F(1, 53) = 5.12, p < 0.05 \); quadratic trend \( F(1, 53) = 4.37, p < 0.05 \)). Follow up t-tests revealed that older adults were more consistent in the tuned and 1% mistuned condition compared to younger adults (\( t(55) = 2.83 \) & 2.01, respectively, \( p < 0.05 \) in both cases). No differences were found at the other levels of mistuning. Finally, the Tuning by Age group by Musical training interaction was not significant (\( p > 0.1 \)), suggesting that the effects of
Musical training were similar in both Age groups, and that the effects of aging were similar for musicians and non-musicians.

### 4.3.3 Reaction Time

Figure 4.4 shows the group mean reaction times. Overall, mistuning influenced participants reaction time, as the main effect of Tuning was significant \( F(5, 265) = 35.52, p < 0.01 \); quadratic trend \( F(1, 53) = 94.1, p < 0.001 \). Pairwise comparisons revealed that reaction time was longest for the 2% and 4% conditions \( (p < 0.001 \) in all cases). Moreover, musicians responded more quickly than non-musicians \( F(1, 53) = 8.25, p < 0.01 \). The main effect of Age group was not significant \( (p > 0.1) \). Although in Figure 4.4 there appears to be a reduction in the influence of musical training in older adults, and that older non-musicians were faster than the younger non-musicians, Age group did not interact with Tuning or Musical training \( (p > 0.1) \).

### 4.3.4 Electrophysiological Data

Figure 4.5 shows the group mean AERs elicited by the tuned and the 16% mistuned conditions recorded at a representative fronto-central scalp site (i.e., FC2) in young and older musicians and non-musicians. To illustrate the effect of mistuning the 2\(^{nd}\) harmonic on the AER, the difference between the AER to the tuned stimulus and 16% mistuned stimulus is shown in a thick black line. During active listening (Figure 4.5a), a clear ORN can be seen overlapping the N1-P2 complex, in addition to a P400 that can be seen peaking around 400 ms. During passive listening (Figure 4.5b) a clear ORN can be seen overlapping the N1-P2 complex, while the P400 is non-existent. The ORN and P400 are labeled on the plot for Younger Musicians. The scalp topographies for these responses
are illustrated in Figure 4.1 separately at three angles (top, left & right) for each wave and each group of participants. The ORN has a fronto-central orientation that can be seen from the top angle, while the P400 is lateralized slightly to the right, hence the use of a different electrode montage to quantify this wave. The inversion of the ORN and P400 activity can be seen around mastoid sites on both the left and right sides. Importantly, the overall topographies are fairly similar between each of the groups, suggesting that the underlying sources are also similar.

4.3.5 Object-related negativity

ORN latency was 135 ms (S.E. = 4.07) for younger musicians, 156 ms (S.E. = 3.93) for younger non-musicians, 150 ms (S.E. = 3.93) for older musicians and 145 ms (S.E. = 4.22) for older non-musicians. The interaction between Age group and Musical training was significant \((F(1, 53) = 13.09, p < 0.01)\). To determine the source of this interaction, follow up t-tests were calculated to compare the influence of aging in musicians and non-musicians. The ORN latency was shorter in younger musicians compared to older musicians \((t(27) = 3.81, p < 0.01)\), but was not different between younger and older non-musicians \((p > 0.1)\).

At fronto-central sites, the main effect of Tuning was significant, which was indicative of an ORN, as the amount of negativity increased from the tuned to the 16% mistuned condition \((F(5, 265) = 45.76, p < 0.001; \text{linear trend } F(1, 53) = 88.98, p < 0.001;\) Figure 4.5 & 4.6). The Tuning by Age group and the Tuning by Musical training interactions were not significant \((p > 0.1)\). However, the interaction between Tuning, Age group and Musical training was significant \((F(5, 265) = 2.89, p < 0.05; \text{linear trend; } F(1, 53) = 5.55, p < 0.05)\). Importantly, while in Figure 4.6 it appears as if there are group
differences in ORN amplitude between the listening conditions (i.e., older non-musicians had a larger ORN compared to the older musicians during active listening, but not during passive listening), the four-way interaction involving Listening condition, Tuning, Age group and Musical training was not significant ($p > 0.1$). Therefore, follow-up tests for the Tuning by Age group by Musical training interaction were based on the average ORN amplitude during active and passive listening. To determine the influence of Age group on the ORN, follow-up simple two-way interactions were calculated separately for musicians and non-musicians that revealed a larger ORN in older non-musicians compared to younger non-musicians ($F(5, 130) = 3.37, p < 0.05$; linear trend $F(1, 26) = 7.65, p < 0.01$), but no age-related change in ORN amplitude for musicians ($p > 0.05$). In a second follow-up analysis, to determine the influence of Musical training, simple two-way interactions confirmed that the ORN was larger in younger musicians compared to younger non-musicians ($F(5, 135) = 3.94, p < 0.01$; linear trend $F(1, 27) = 7.35, p < 0.05$), but that the ORN was similar between older musicians and non-musicians ($p > 0.1$).

Finally, the ORN was smaller in active compared to passive listening, as the interaction between Listening condition and Tuning was significant ($F(15, 795) = 2.45, p < 0.05$; linear trend $F(1, 53) = 6.17, p < 0.05$), but as mentioned above, the Listening condition by Tuning interaction was not influenced by group factors.

### 4.3.6 P400

The P400 peaked around 395 ms in all participants. The main effects of Age group, Musical training, and their interactions were not significant ($p > 0.1$).

During the 250-350 ms epoch over right frontal sites, the main effect of Tuning was significant, which was indicative of the P400 ($F(5, 265) = 19.82, p < 0.001$; linear trend $F(1, 53) = 31.67 p < 0.001$; Figure 4.7 & 4.8). The P400 was larger in musicians...
compared to non-musicians ($F(1, 53) = 5.48, p < 0.01$; linear trend $F(1, 53) = 6.66, p < 0.05$). The Age group by Tuning and the Age group by Musical training by Tuning interactions were not significant ($p > 0.1$). P400 topography is illustrated in Figure 4.1.

During the 350-450 ms epoch over the right frontal sites, the main effect of Tuning was significant which was indicative of a P400 ($F(5, 265) = 35.84, p < 0.001$; linear trend $F(1, 53) = 57.24, p < 0.001$; Figure 4.7 & 4.8). The P400 was larger in musicians compared to non-musicians, but was only significant at a trend level ($F(5, 265) = 2.05, p = 0.07$; quadratic trend $F(1, 53) = 2.04, p > 0.05$). The Tuning by Age group by Musical training interaction and the Tuning by Age group interactions were not significant ($p > 0.1$).

### 4.3.7 Late-positive complex (LPC)

The LPC peaked earlier in younger adults (517 ms, S.E. = 15.64) compared to older adults (592 ms, S.E. = 15.95; $F(1, 53) = 11.15, p < 0.01$). The main effect of musical training was not significant nor was the interaction between age and musical status ($p > 0.1$). The LPC latency decreased as mistuning increased ($F(5, 265) = 4.68, p < 0.01$; linear trend $F(1, 53) = 12.63, p < 0.01$). LPC latency decreased from 565 ms (S.E. = 11.82) in the tuned condition to 523 ms (S.E. = 10.99) in the 16% mistuned condition. The effects of Tuning on the LPC latency did not interact with other factors.

The LPC amplitude was comparable in young and older adults ($p > 0.1$). On the other hand, the LPC was larger in musicians compared to non-musicians ($F(1, 53) = 5.41, p < 0.05$; see Figure 4.9 & 4.10). The interaction between Age group and Musical training was not significant ($p > 0.1$). The main effect of tuning was significant for the LPC amplitude ($F(5, 265) = 11.06, p < 0.001$; quadratic trend $F(1, 53) = 39.65, p < 0.001$).
LPC amplitude was smallest in the 2% and 4% mistuned conditions, as pairwise comparisons revealed that these two conditions were smaller than the tuned and 16% mistuned conditions ($p < 0.01$ combined). Furthermore, the Tuned and 16% conditions did not differ ($p > 0.1$). The Age group by Tuning by Musical training interaction was not significant ($p > 0.1$), despite the appearance on Figure 4.9 & 4.10 that the LPC was largest in older musicians.

### 4.3.8 Correlations between behavior and electrophysiology

Within-subject correlations were calculated between electrophysiological measures, and behavioral measures.

The ORN amplitude was correlated with perceptual judgment ($r = -.32$, $t(56) = -5.37$, $p < 0.001$), judgment consistency ($r = -.23$, $t(56) = -4.39$, $p < 0.001$), and reaction time ($r = .24$, $t(56) = 4.49$, $p < 0.001$).

The P400 amplitude was correlated with perceptual judgment ($r = .66$, $t(56) = 11.76$, $p < 0.001$), judgment consistency ($r = .27$, $t(56) = 4.56$, $p < 0.001$), and reaction time ($r = -.36$, $t(56) = -5.8$, $p < 0.001$).

The LPC was correlated with judgment consistency ($r = .22$, $t(56) = 3.35$, $p < 0.01$) and reaction time ($r = -.33$, $t(56) = -4.96$, $p < 0.001$), but did not correlate with perceptual judgment ($p > 0.1$).

In addition, the ORN was correlated with the P400 ($r = -.16$, $t(56) = -2.60$, $p < 0.05$), and the P400 was correlated with the LPC ($r = .18$, $t(56) = 2.47$, $p < 0.05$). The ORN and LPC were not correlated ($r = .002$).
4.4 Discussion

There are four main findings from this study. The first was that although the ORN was larger in younger musicians (Zendel & Alain, 2009), it was similar between older musicians and non-musicians. Second, the P400 was larger, and began earlier in musicians. Moreover, the P400 was reduced in the older non-musicians; however this effect did not reach significance. Third, the LPC was larger in musicians and earlier in younger adults. Fourth, musicians had faster reaction times, were more likely to report hearing two sounds when mistuning was above 2%, and made more consistent responses overall. Surprisingly, no age-related effects were found for any of the behavioral measures; however, across all three measures, the benefit of musical training appeared to be smaller for older adults, although this effect did not reach statistical significance. The next section will consider each of the results in more detail, which will be followed by a broader interpretation of the overall pattern of results in terms of how they relate to previous research.

The three electrophysiological measures (i.e., the ORN, P400 & LPC) reported in this study are likely indicative of three separate but related processes. The ORN likely indicates the automatic detection the presence of concurrent sound objects, as it occurs regardless of attention; however, in situations where listeners do respond to the stimulus the ORN amplitude was correlated with the likelihood of reporting the presence two simultaneous sounds (Alain, Arnott, & Picton, 2001; Alain et al., 2005; Alain et al., 2002; Hautus & Johnson, 2005; McDonald & Alain, 2005). The P400 likely represents the conscious detection of concurrently occurring sounds, as it’s amplitude was also correlated with the likelihood of reporting two sounds, and only occurs when the participant is asked to make a perceptual judgment about the stimulus (Alain, Arnott, & Picton, 2001; Alain, 2007). While the P400 is related to the conscious detection of a
second auditory object, the LPC is most likely related to confidence in matching an incoming acoustic stimulus, and was therefore largest for both the tuned and 16% mistuned stimuli, as these were the most easily categorized. Appropriately, the LPC was only correlated with the judgment consistency and reaction time measures, but not the perceptual judgment score, supporting the idea that the LPC was related to confidence in matching the acoustic stimulus to a representative acoustic schema, as increased confidence would increase consistency in responses, and shorten reaction times. The next section will consider the influence of age and musical training on these waves.

While some previous work demonstrated that the ORN was smaller in older adults (Alain & McDonald, 2007), more recent work suggests that this age-related difference was due to the length of the stimulus (Alain et al., Submitted). The length of the stimulus in the current study was 150 ms; shorter than 200 ms, (Alain et al. Submitted), but longer than 40 ms, (Alain & McDonald, 2007). Furthermore, it has been shown that age-related decline in detection of a mistuned harmonic is smaller when the stimulus length is longer (Alain, McDonald et al., 2001). The current data support these findings, by demonstrating that the ORN, and the ability to segregate a mistuned harmonic as a separate auditory object is similar between older and younger adults when the stimulus length is 150 ms.

Based on previous findings that the ORN was enhanced in younger adults (Zendel & Alain, 2009), it was somewhat surprising that there was little influence of musical training on the ORN wave for older adults. Like previous work (e.g., Alain et al., 2001) the ORN amplitude was positively correlated with the likelihood of reporting the perception of a second auditory object. Furthermore, the likelihood of reporting two concurrently occurring sounds in the 16% mistuned condition was much greater for younger musicians compared to younger non-musicians, while this difference was smaller between older musicians and older non-musicians. Even though the electrophysiological
and behavioral data were related, suggesting good internal consistency, it is still surprising that the musical training effect was reduced in older adults. One possible explanation for this was the motivation of the participants. Older adults, compared to younger adults, tend to be more interested in research participation. This enhanced motivation may have influenced the ORN selectively in older non-musicians, especially considering that ORN amplitude in older non-musicians appears largest in the active listening condition (i.e. when the participants are responding to stimuli).

Another possible way to interpret the larger ORN in older non-musicians is that it overlaps with an oddball response, known as the mismatch negativity (MMN). The MMN is an electrophysiological response to an oddball sound in an otherwise identical stream of acoustic information, and is observed during a similar epoch and at a similar scalp location to the ORN (Näätänen et al. 2004). Previous work has shown the ORN and MMN to be distinct from each other, and more importantly that the MMN response is more susceptible to changes in attention level (Alain & Izenberg, 2003). It is therefore possible that the larger ORN in older non-musicians was due to an overlapping MMN. The MMN may have been selectively evoked in the older non-musicians because of higher thresholds for detecting a mistuned harmonic, resulting in the 16% mistuned stimulus being detected as an oddball. In addition, then enlarged ORN in this group was limited to the active listening condition, where increased attention may have selectively enhanced the MMN response.

The other wave related to the perceptual segregation of a mistuned harmonic was the P400. Unlike the ORN, the P400 is indicative of the conscious registration of the perception of two sounds (Alain, Arnott, & Picton, 2001; Alain, 2007). Positivities around this latency have also been attributed to the conscious detection of an oddball in a stream of otherwise identical acoustic stimuli (i.e. P300 or P3b; Picton, 1992).
unlikely that the P400 is related to the detection of an oddball because in the current study, the number of stimuli that were clearly perceived as a one sound or two sounds was approximately equal. Nearly all the tuned and 1% mistuned stimuli were perceived as a single sounds, while nearly all the 8% and 16% mistuned stimuli were perceived as two sounds. The 2% and 4% mistuned conditions were more ambiguous, and during different trials were reported as being a single or a concurrently occurring sound. Therefore there were no ‘oddballs’ in the current study. Moreover, while the stimuli were presented in a random order, it was possible that a series of stimuli perceived as a single sound would be followed by a stimulus perceived as two sounds. In this situation the stimulus perceived as two sounds would be an oddball, and may result in a P300 wave. Importantly, the reverse is also possible, where a series of stimuli perceived as two sounds was followed by a stimuli perceived as a single sound. In this case the stimulus perceived as a single would be the oddball, and may result in a P300 wave. The P400 was only present when the stimuli were perceived as two concurrently occurring sounds, was correlated with the likelihood of reporting the presence of concurrently occurring sounds and thus is not a P300 type wave.

There has been no previous reports of a P400 wave in older adults using a mistuned harmonic paradigm, as previous work that has examined age-related changes to the ORN used exclusively passive paradigms (Alain & McDonald, 2007; Alain et al., Submitted). Thus, one important finding was that there is a P400 in older adults, and that it was correlated with behavioral performance of a concurrent sound segregation task. A non-significant age-related change in behavioral performance and in the P400 indicates that actively segregating sounds in the environment may not decline with age, especially when the stimulus is long (see above), and when the mistuning is large. Sixteen percent mistuning is well above the perceptual threshold for both older and younger adults (see
Chapter 1 & 2), and should therefore be perceivable by both groups. Furthermore, the stimulus was longer which has also been shown to reduce age-related differences in concurrent sound perception (Alain et al., Submitted).

Despite the mistuning being perceivable by both older and younger adults, the P400 was larger in both older and younger musicians, and while the peak of the P400 occurred at the same time in all participants, P400 activity started much earlier in musicians compared to non-musicians. In addition, the P400 was most strongly correlated with perceptual judgment, supporting the hypothesis that the differences in P400 amplitude were related to perceptually segregating concurrently occurring sounds. The effect of musical training on the P400 and the perceptual judgment suggests that musicians are better able to actively organize the auditory scene into separate auditory precepts, which is reflected in a larger P400 response. Previous work that examined how musicians spatially organize the auditory environment demonstrated that musicians, and conductors specifically were more sensitive to peripheral components of the auditory scene, and this was reflected in late positive electrophysiological activity (Nager et al., 2003). In both the current study and in Nager et al. (2003) participants had to assign acoustic input to at least two perceptual objects (or streams). Thus musicians are likely better at organizing the auditory environment, and in the current study this is related to an enhanced P400.

Importantly, the pattern of results when considering the ORN suggests a differential enhancement to the P400 in older musicians. The ORN and P400 were correlated, suggesting that some of the variance in P400 amplitude was related to the ORN amplitude. Although the ORN was similar in older musicians and non-musicians, the P400 was larger in older musicians compared to older non-musicians. This suggests that there was enhanced endogenous activity in the older musicians compared to the older non-musicians that cannot be explained by a larger ORN. In younger adults, the enhancement
to the P400 can be attributed to a larger ORN because both waves were enhanced in younger musicians. Therefore this overall pattern of results suggests that older musicians preferentially engage endogenous neural activity that may be related to using a cognitive strategy to help organize the auditory environment. This possibility is further supported by the observation that the P400 was similar in amplitude between the younger adults and older musicians, and appeared to be reduced in older non-musicians, although this effect did not reach statistical significance.

While the P400 is related to the perceptual segregation of concurrent auditory objects, the LPC is most likely related to the overall confidence in making a perceptual judgment, regardless of what that judgment is. For example, late-positivities that have similar scalp topographies, and latencies have been described as an index of word recognition memory (Rugg & Curran, 2007). That is, an LPC arises when a participant is asked if a word was presented previously. In terms of recognition, the amplitude of the LPC tends to increase as the participant correctly recognizes the stimulus (Rugg et al., 1998; Curran, 2000). These studies provide support for the hypothesis that the late positive activity is in the current study was related to identification of the auditory stimulus into one of two categories. The LPC was largest when the stimuli were easily categorized into the perception of a single sound, or the perception of two sounds, while it was smallest when that categorization was ambiguous (i.e. 2 & 4% mistuning). The LPC was delayed in older adults and larger in musicians. The next section will consider these two findings.

In older adults, increased latency of the LPC suggests that aging may slow the ability to match an acoustic stimulus to memory. Surprisingly, this age-related latency increase was not reflected in the reaction time data, perhaps because pressing a response button introduces another level of variance into the reaction time data (i.e., the motor
system). Therefore, the LPC may be a more accurate measure of stimulus recognition than reaction time. Age-related slowing of cognitive processes are common findings throughout the aging literature (Salthouse, 1996). At the neural level these changes are thought to be due to a decrease in white matter integrity (Head et al., 2004) or an increase in white matter hyperintensities, which may represent demyelination, trauma or another neural disease (Wen & Sachdev, 2004). An age-related change to white matter could slow the propagation of neural information throughout the cortex, because white matter is made up of myelinated axons, and damage to the myelination slows the transmission of a neural signal. Given that the LPC was delayed equally in older musicians and non-musicians, it is unlikely that being a musician prevents age-related changes to neural processing speed. On the other hand, the enhancement to the amplitude of the LPC was preserved in older musicians. This enhancement likely represents a preserved ability to make judgments about the auditory environment, and was reflected in the older musicians preserved ability to consistently, and quickly make perceptual judgments compared to older non-musicians.

The correlation between LPC amplitude and both judgment consistency and reaction time suggests that the enhanced LPC in musicians represents enhanced confidence in matching an acoustic stimulus to a stored acoustic schema. This enhanced recognition process would be a significant benefit for older musicians, because of age-related decline to earlier sensory processing. For example, older adults have a reduced ability to detect many small acoustic cues that can be used to aid auditory scene analysis (e.g. periodicity relationships [Alain, McDonald et al., 2001] or small silent gaps [Schneider et al., 1994]). Given the current pattern of data, the advantage for older musicians seems to be in a preserved ability to organize acoustic information into a conscious precept. This advantage would provide numerous benefits in older adults, because the availability of acoustic cues diminishes with age, thus enhanced ability to
interpret an impoverished signal and make confident judgments about the auditory scene could help older musicians overcome age-related decline in parsing complex auditory scenes.

4.4.1 Summary

Concurrent sound segregation is enhanced in musicians, and this ability is preserved with age. The most likely reason for this advantage is an enhanced ability to recognize a mistuned harmonic as a separate auditory object. It is unlikely that the advantage for musicians is in the detection of mistuning, because, although musicians have a lifelong advantage in detecting a mistuned harmonic (Chapter 2), the stimuli used in the current study were well above those perceptual thresholds. Furthermore, neural activity related to the conscious detection of a mistuned harmonic was enhanced in both older and younger musicians, while the neural activity related to the automatic detection of a mistuned harmonic was only enhanced in younger adults. Thus the advantage for older musicians is most likely related to enhanced endogenous processing of acoustic features related to the segregation of simultaneous sound objects.
Chapter 5:

5.1 Summary, Limitations, and Future directions

The purpose of this dissertation was to explore the possibility that lifelong musicianship can prevent, delay, or slow age-related decline in auditory perception. The findings demonstrate that older musicians experience less age-related decline in auditory processing abilities compared to non-musicians. This was demonstrated by a preservation of musician-related enhancements to early automatic processing of acoustic information, and more importantly, an increasing ability to use controlled auditory processing to overcome age-related decline in hearing abilities.

The benefit of musical training was clear for gap detection, mistuned harmonic detection, and speech in noise thresholds. The rate of age-related decline in both gap detection and speech in noise thresholds were slower in musicians, while musicians had a life-long advantage for detecting mistuned harmonics. Most critical was that there was no difference between musicians and non-musicians for pure-tone thresholds, which is highly indicative of enhanced central auditory processing in musicians compared to non-musicians.

Scalp recorded AERs provided further support for this claim. Age related change to the P1/Nb was similar in both musicians and non-musicians; however, the P1/Nb in older musicians resembled that of younger non-musicians. Most importantly, there were no age by musical training interactions to the scalp recorded AER, suggesting that, while musical training may alter the exogenous auditory evoked response, it does not modify age-related changes to the AER. On the other hand, attention-dependent, endogenous neural activity related to the processing of a complex sound was enhanced in older adults,
and further enhanced in older musicians. This finding suggests that older adults engage
greater amounts of endogenous neural activity to overcome age-related decline in the
exogenous processing of acoustic information. Most importantly, this age-related change
was greatest in musicians, suggesting that lifelong musical training accumulates to further
enhance endogenous neural processes that aid in auditory perception.

One of the main difficulties older adults have is in the perceptual organization of
the auditory scene (e.g. separating speech from background noise). Therefore, the next
step was to investigate neural activity in older and younger musicians and non-musicians
in terms of the ability to segregate a mistuned harmonic as a separate auditory object. The
behavioral responses of musicians were faster and more consistent than non-musicians,
but surprisingly, behavioral performance was not impacted by aging in this sample.
Furthermore, automatic processing of the mistuned harmonic (i.e., the ORN) was only
enhanced in younger musicians, while the endogenous neural activity related to the
perception of a second auditory object (i.e., the P400) was enhanced in musicians across
the lifespan. Finally, the LPC, which was related to confidence and consistency in
perception was enhanced in musicians, and delayed in older adults. This pattern of results
is highly suggestive of a cognitive advantage for older musicians because the early
automatic processing of concurrent sounds was similar in older musicians and non-
musicians, while later endogenous neural activity was enhanced in older musicians.

5.2 Main effects of musical training to exogenous auditory
processing

Exogenous, attention-independent processing of acoustic information is enhanced
in musicians, and this enhancement is maintained in older musicians across the lifespan.

Previous work demonstrated that automatic processing of acoustic information, as
measured by electrophysiological responses from the brainstem, are enhanced in younger
musicians (Parbery-Clark, Skoe, & Kraus, 2009; Wong et al., 2007; Bidelman & Krishnan, 2010). Wong et al., (2007) demonstrated a more robust encoding of pitch change in the brainstem of musicians, while both Parbery-Clark, Skoe, & Kraus (2009) and Bidelman & Krishnan (2010) demonstrated that higher harmonics of a complex sound are more robustly encoded in the brainstem of musicians compared to non-musicians in the presence of background noise. Supporting these findings, Koelsch et al. (1999) demonstrated that small amounts of mistuning in a complex sound were represented automatically in the cortex of musicians, while there was no automatic registration of the mistuning in non-musicians. Clearly, musicians have an advantage in the automatic processing of acoustic information. The current data demonstrated that musicians have an advantage in detecting a mistuned harmonic component, and that this advantage is consistent across the lifespan (Chapter 2). Based on previous work, this advantage is likely due to a more robust encoding of the higher frequency harmonics in the brainstem (Parbery-Clark, Skoe, & Kraus, 2009; Wong et al., 2007; Bidelman & Krishnan, 2010).

Providing support for this claim was the finding that the P1 wave was smaller (or Nb was larger) in musicians in response to a complex sound (Chapter 3). The P1 wave is the boundary between the mid-latency and long-latency AERs, and may be a reflection of afferent and efferent information transfer between the thalamus and cortex (Cohen, 1982; Ribary et al., 1991). It is likely that part of the change to the P1/Nb wave was due to enhanced encoding of acoustic primitives related to spectral relationships in the brainstem, especially considering the stimulus was made up of multiple harmonics.

Interestingly, these data demonstrate a preserved advantage for musicians across the lifespan that seems to be related to processing spectral relationships in the acoustic scene. Although the ORN was similar in older musicians and non-musicians, it remains possible that early periodicity encoding is enhanced in older musicians, particularly
because the perception of higher harmonics is important for musical performance. Musicians are highly sensitive to timbre, evidenced by the finding that early cortical responses distinguish between different instruments in musicians, but not in non-musicians (Pantev et al., 2001). Importantly, the perception of timbre is influenced by the relative amplitude and number of higher harmonics. The current data suggest that musical training, and continued practice, plastically alter the brain to maintain a high level of spectral acuity, as the ability to detect mistuning of a single harmonic was enhanced in musicians, and correlated with hours of music practice. Therefore, continued musical practice helps maintain a high level of spectral acuity, due to neuro-plastic modulation of early automatic spectral processing at the level of the brainstem, likely due to efferent feedback from the cortex (Tzounopoulou & Kraus, 2009). A lack of accumulation of this ability over the lifespan (i.e. not differential preservation), suggests that greater amounts of musical training may increase spectral acuity, but that this ability still declines with age. Future research could examine spectral abilities in people who took music lessons for many years, but are no longer musicians. It is likely that despite the early training, these former musicians would no longer demonstrate a benefit in spectral acuity because an enhanced ability in processing spectral information is not required for everyday hearing. Therefore the neurons in the brainstem that encode higher harmonics no longer receive feedback from higher cortical centres, and their functional connections in the brainstem nuclei may return to a baseline level.

5.3 Age by Musical training interactions for endogenous auditory processing

The lifelong advantage in automatic processing of spectral information cannot explain interactions between aging and musical training on speech in noise and gap detection thresholds. Speech-in-noise and Gap detection thresholds declined at a slower
rate in musicians. Given that early exogenous neural processes evoked by a complex sound were not differentially preserved in older musicians, it is likely that the benefit for older musicians was due to enhanced endogenous processing of acoustic information.

Evidence from the current studies supports this possibility. That is, endogenous neural activity, that represents attention dependent controlled processing of acoustic information, is differentially preserved in older musicians. Endogenous neural activity related to the processing of a complex sound was enhanced in musicians, and further enhanced in older musicians. Interestingly this enhancement peaked around 600 ms, and was best modeled by sources in the right auditory cortex. The right auditory cortex is believed to be specialized for processing spectral information (Zatorre, 1988), and late endogenous activity is believed to be related to the conscious matching of an acoustic stimulus to a stored representation (Näätänen et al., 2010). Therefore the advantage observed in older musicians may derive from an enhanced ability to match a spectrally rich acoustic signal to a stored representation. Critically, there was no exogenous or automatic neural activity that was preferentially enhanced in older musicians, suggesting that controlled, attention dependent processes are the best explanation to why older musicians show less age-related decline on some psychoacoustic measures.

This possibility was further supported by examining endogenous activity in terms of processing a mistuned harmonic. Two separate indices of endogenous neural processing were enhanced in older musicians, the P400, and the LPC. The P400, which is thought to index the conscious registration of simultaneous sounds (Alain, Arnott, & Picton, 2001), was enhanced in musicians. The ORN and P400 were correlated with each other, and there were asymmetrical group differences in the amplitudes of both waves. A larger P400 in younger musicians may be related to an enhanced ORN (i.e. automatic neural activity related to detecting the presence of concurrently occurring sounds). On the
other hand, it is unlikely that the difference in P400 amplitude in older musicians and non-musicians was related to the ORN, because the ORN was comparable in older musicians and non-musicians, suggesting that in older musicians the enhanced P400 was related to enhanced endogenous neural processes. The LPC is thought to be related to matching a complex auditory stimulus to a memory (i.e. a word; Rugg & Curran, 2007). In the current study the LPC was correlated with both reaction time and judgment consistency, and was largest when the perceptual judgment was easiest, supporting the idea that the LPC represents a match between stimulus and memory. This endogenous activity was delayed in older adults, and enhanced in musicians, which further suggests that the propagation of neural information is slowed in older adults, but that life-long musicianship compensates for this decline through enhanced cognitive abilities.

Interactions between aging and musical training on endogenous neural activity may be indicative of cognitive reserve in older musicians. Previous work has shown that older adults who engage in cognitively stimulating activities later in life show slower rates of cognitive decline, known as a cognitive reserve (Christensen et al., 2008; Ghisletta et al., 2006; Valenzuela & Sachdev, 2006). In the current study, cognitive abilities related to auditory perception were preferentially enhanced in older musicians, suggesting that lifelong musicianship enhanced cognitive components of auditory perception. However, it remains unknown if these enhancements are due to a general cognitive reserve, or if it the enhancement is domain specific to auditory material.

While it is likely that attention dependent, controlled auditory processing is differentially preserved in older musicians, an alternative possibility to explain the age by musical training interactions observed for the psychoacoustic data in Chapter 2 is that automatic processing of temporal information is differentially preserved in older musicians. Supporting this possibility was the finding that gap detection thresholds
declined at a slower rate in musicians. Thus, this alternative interpretation suggests that
the ability to detect spectral cues to help organize the auditory scene declines at an equal
rate in musicians and non-musicians, while the ability to detect and use temporal details to
organize the auditory scene is relatively preserved. The current dissertation did not
investigate temporal processing beyond gap detection abilities, but future research should
investigate this question. This interpretation leaves open the possibility that both
primitive processing of acoustic features on the temporal domain and controlled, attention
dependent auditory processing of all acoustic information contribute to the differential
preservation in understanding speech in noisy environments.

5.4 Non-significant effects

First, differences in pure-tone thresholds were not observed between musicians and
non-musicians in Chapter 2 (groups of older adults were matched on pure-tone abilities for
the studies presented in Chapters 3 & 4). This lack of difference is important because it
suggests that the age-related benefits found in musicians for central aspects of auditory
perception are due to differences in the structure and function of the brain, and not due to
differences in the functioning of ear. Therefore, the benefit musicians have in auditory
processing is most likely due to experience dependent neural plasticity or in-born genetic
predispositions. Alternatively, it is possible that there was cochlear pathology that could
not be detected by pure-tone sensitivity, and this pathology selectively influenced older
non-musicians. While this possibility seems unlikely, without a more detailed assessment
of cochlear pathology, it cannot be ruled out.

Another surprising result was that younger musicians and non-musicians had
similar speech in noise and gap detection thresholds. Previous work had demonstrated
that both speech in noise thresholds (Parbery-Clark, Skoe, Lam et al., 2009), and gap
detection thresholds (Rammsayer & Altenmuller, 2006) were lower in younger musicians
compared to non-musicians. Data in this dissertation found an advantage for older
musicians on these tasks, but not for younger musicians. The null finding of musical
training for younger adults on gap detection abilities may have been due to the stimuli.
Rammsayer & Altenmuller (2006) used broadband noise markers, while the current study
used pure-tone markers. The broadband noise may have provided musicians with the
advantage over non-musicians, because musicians encode higher harmonics more robustly
then non-musicians (Bidelman & Krishnan, 2010; Parbery-Clark, Skoe, & Kraus, 2009).
While broadband noise does not contain harmonics, it does contain acoustic energy at
higher frequencies, which may also be encoded more robustly in musicians. The more
robust encoding of higher frequencies would ease detection of the onset or offset of a
marker, and thus young musicians would have an advantage for detecting the gap. The
lack of benefit for speech in noise thresholds in younger musicians may have been due to
the criteria used to define a ‘musician’. Parbery-Clark, Skoe, Lam et al. (2009) used the
same assessment, and had similar participant demographics, except they only selected
musicians that had started musical training before age 7. In the current study musicians
were recruited based on having achieved a minimum amount of training before age 18.
Therefore, the young musicians in Parbery-Clark et al. (2009) had likely been musicians
for more years compared to the young musicians in the current study. This suggests that
the benefit musical training confers on speech-in-noise thresholds takes many years to
accumulate.

In the electrophysiological data, there were no enhancements related to being a
musician for the N1 and P2 waves. This was surprising, as these enhancements have been
previously reported in younger musicians (Pantev et al., 1998; Shahin et al., 2003; Shahin
et al., 2005). The lack of effect for the N1 wave was not surprising as follow-up research demonstrated that the N1 enhancement in musicians was timbre specific to the instrument of training (Pantev et al., 2001). Furthermore other research has demonstrated that the enhancement to the P2 wave in musicians was limited to spectrally rich sounds with musical timbres (Shahin et al., 2003). The current study did not use musical type timbres to elicit an auditory evoked response, thus the lack of effect of musical training may be related to the stimuli used, and suggest that the plasticity of the N1/P2 response is specific to sounds with musical timbres.

5.5 Limitations and future directions

The data reported in this dissertation are the first to demonstrate that older musicians have better auditory perceptual abilities then age-matched non-musicians. Some of these advantages are lifelong, while others accumulate over the lifespan. These two patterns of results point to an important limitation of the current study. Central auditory processing and age-related changes to central auditory perception are not unitary phenomena. This was demonstrated by the different patterns of age-related change on psychoacoustic, and electrophysiological measures of auditory perception in this dissertation. Specifically, the current data suggest that musicians experience a lifelong advantage in using spectral information to parse the auditory scene. At the same time musicians show less age-related decline on temporal processing abilities. Therefore, one of the limitations of this research was that temporal processing was solely assessed by gap detection thresholds. Future research should therefore investigate how older musicians process temporal information in more detail.
Manipulating stimulus characteristics will also help uncover the neural source of the auditory processing advantage for older musicians. The current data support a cognitive advantage for musicians. That is, based on AER data, older musicians engage additional endogenous neural activity that likely reflects a cognitive strategy to overcome perceptual deficits, which suggests that musical training and continued practice creates and maintains a cognitive reserve for processing auditory information. Given the wide reaching benefits of musical training in terms of cognitive abilities, it is possible that musical training may prevent other aspects of cognitive decline. Examining how lifelong musicianship interacts with general cognitive abilities is therefore an important question.

There was, however, some evidence that early exogenous processing of acoustic information was enhanced in older musicians. Unfortunately the current studies were not designed to examine the earliest stage of exogenous auditory processing in the brainstem, because assessing brainstem function requires a different set of stimuli and procedures. Therefore, another important direction is to examine the early brainstem responses from older musicians and non-musicians to determine if automatic auditory processing is differentially preserved in older musicians.

Regardless of the stage of auditory processing where older musicians experience differential preservation, comparing musicians and non-musicians does not allow determination of the direction of causality. That is, it is possible that lifelong engagement in musical training plastically alters the neurophysiological substrates of auditory perception; however, it is also possible that people who become and stay musicians have inborn genetic characteristics that result in an enhanced auditory system. To overcome this limitation, one must conduct controlled experiments where young adults are randomly assigned to be or not to be a musician. Ideally, the auditory abilities of these two groups would be followed for a lifetime. If the group of musicians demonstrates an auditory
advantage over the non-musician group, a causal link between musical training and enhancements to auditory processing over the lifespan would be proven. For many reasons, this type of experiment is unfeasible (e.g., cost, time commitment for participants, etc…). However, some researchers have examined how short-term musical training (over the course of a year or two; Fujioka et al., 2006; Hyde et al., 2009) or laboratory based musical type training (Lappe et al., 2008) influenced auditory processing. These studies demonstrated that musical training causes auditory processing enhancements, and neuroplastic change to the brains of children and young adults.

In older adults, controlled administration of musical training has been shown to increase performance on tasks related to working memory. Bugos et al (2007) randomly assigned older adults to one of two groups, one of which received six-months of music lessons, the other which served as a control. They administered a series of cognitive assessments, and found that performance on tasks related to working memory were enhanced in the group that received music lessons, suggesting that musical training could mitigate aspects of age-related cognitive decline (Bugos et al. 2007). While this study did not measure brain plasticity, the findings are highly suggestive of it, and supported by another training study that observed structural brain changes after training. In this study older adults learned to juggle for three months, and this learning resulted in increased grey matter volume in areas related to processing movement in the visual cortex (V5), and in areas related to memory (hippocampus; Boyke et al. 2008). These results suggest that neural plasticity is possible in older adults, even when using short-term training paradigms (compared to lifelong musical training). While no study has found that musical training in older adults can improve auditory processing, results from these training studies, and the current dissertation suggest that it is worth trying.
5.6 Summary

The results of this dissertation demonstrate that life-long musicianship can mitigate age-related decline in some auditory perceptual abilities. These benefits are due to life-long enhancements related to the automatic processing of acoustic information, and more importantly preferentially enhanced cognitive processing of acoustic information. This is a particularly important finding because age-related decline in auditory perception is very common among older adults, and can result in social isolation and even depression (Betlejewski, 2006; Salomon, 1986). Therefore reducing age-related decline in auditory perception is of utmost importance. In the long run, specific auditory rehabilitation programs may work better than musical training to prevent age-related decline in auditory perception; however music lessons are already widely available, and can be quite fun. So, go learn a musical instrument, your auditory system will thank you.
6.1 References


**Table 2.1**

*Regression analysis for pure-tone thresholds.*

<table>
<thead>
<tr>
<th>R Statistic</th>
<th>R Square</th>
<th>df</th>
<th>F value</th>
</tr>
</thead>
<tbody>
<tr>
<td>.831</td>
<td>.691</td>
<td>3, 159</td>
<td>118.51**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstandardized</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>.581</td>
<td>.947</td>
</tr>
<tr>
<td>Age</td>
<td>.481</td>
<td>.026</td>
</tr>
<tr>
<td>Musician</td>
<td>-1.19</td>
<td>1.89</td>
</tr>
<tr>
<td>Musician x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-.023</td>
<td>.053</td>
</tr>
</tbody>
</table>

*Note: **p < 0.01; df: degrees of freedom; S.E.: standard error.*
Table 2.2

*Regression analysis for gap-detection thresholds.*

<table>
<thead>
<tr>
<th>R Statistic</th>
<th>R Square</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.356</td>
<td>.126</td>
<td>3, 150</td>
<td>7.239**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstandardized</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>S.E.</td>
</tr>
<tr>
<td>(Constant)</td>
<td>.732</td>
<td>.374</td>
</tr>
<tr>
<td>Age</td>
<td>.031</td>
<td>.011</td>
</tr>
<tr>
<td>Musician</td>
<td>.43</td>
<td>.748</td>
</tr>
<tr>
<td>Musician x Age</td>
<td>.042</td>
<td>.021</td>
</tr>
</tbody>
</table>

*Note: *p < 0.05 **p < 0.01; df: degrees of freedom; S.E.: standard error.*
Table 2.3

*Regression analysis for mistuned harmonic detection thresholds.*

<table>
<thead>
<tr>
<th>R Statistic</th>
<th>R Square</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>.328</td>
<td>.108</td>
<td>3, 150</td>
<td>6.032**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstandardized</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>S.E.</td>
</tr>
<tr>
<td>(Constant)</td>
<td>3.583</td>
<td>.507</td>
</tr>
<tr>
<td>Age</td>
<td>.036</td>
<td>.014</td>
</tr>
<tr>
<td>Musician</td>
<td>-2.191</td>
<td>1.014</td>
</tr>
<tr>
<td>Musician x Age</td>
<td>.007</td>
<td>.029</td>
</tr>
</tbody>
</table>

*Note: *p < 0.05 **p < 0.01; df: degrees of freedom; S.E.: standard error.*
Table 2.4.

Regression analysis for speech-in-noise thresholds.

<table>
<thead>
<tr>
<th>R Statistic</th>
<th>R Square</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>.595</td>
<td>.353</td>
<td>3, 154</td>
<td>28.061**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstandardized</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>S.E.</td>
</tr>
<tr>
<td>(Constant)</td>
<td>-.318</td>
<td>.239</td>
</tr>
<tr>
<td>Age</td>
<td>.052</td>
<td>.007</td>
</tr>
<tr>
<td>Musician</td>
<td>.346</td>
<td>.478</td>
</tr>
<tr>
<td>Musician x Age</td>
<td>-.033</td>
<td>.013</td>
</tr>
</tbody>
</table>

Notes: *p < 0.05 **p < 0.01; df: degrees of freedom; S.E.: standard error.
Figure 2.1: Group mean audiogram.
Figure 2.2: All-tone average threshold.
Figure 2.3: Gap-detection thresholds.
Figure 2.4: Mistuned harmonic detection thresholds.
Figure 2.5: Speech-in-noise thresholds.
Figure 3.1: Electrode montage.
Figure 3.2: Auditory surrogate source model.
Figure 3.3: Butterfly plots and scalp topography for the P1, N1, P2 & N2 waves.
Figure 3.4: Auditory evoked response at electrode Cz.
Figure 3.5: Amplitude of the P1, N1, P2 and N2 waves.
Figure 3.6: Endogenous source activity.
Figure 3.7: Source activity during active listening.
Figure 3.8: Source activity during passive listening.
Figure 3.9: Mean amplitude of source activity during the 550-650 ms epoch from the right auditory cortex
Figure 4.1: Topographic maps for the ORN and P400.
Figure 4.2: Perceptual judgment.

The figure shows the percentage of stimuli reported as two sounds across different levels of mistuning of the second harmonic. The x-axis represents the mistuning of the second harmonic (1%, 2%, 4%, 8%, 16%) and the y-axis represents the percentage of stimuli reported as two sounds. Four groups are compared: Younger Musician, Younger Non-musician, Older Musician, and Older Non-musician. The trends indicate that younger musicians and younger non-musicians are more likely to perceive the stimuli as two sounds compared to older musicians and non-musicians at higher mistuning levels.
Figure 4.3: Response consistency.

![Graph showing response consistency across mistuning of 2nd harmonic for different age and musical groups.]
Figure 4.5: Deriving the ORN and P400 from the auditory evoked response

A. Active listening

Younger Musicians

Younger Non-musicians

Older Musicians

Older Non-musicians

B. Passive listening

Younger Musicians

Younger Non-musicians

Older Musicians

Older Non-musicians
Figure 4.6: ORN amplitude in active and passive listening.

![Graph showing ORN amplitude comparison between younger and older musicians and non-musicians in active and passive listening.](image-url)
Figure 4.7: ORN and P400 at electrode C4
Figure 4.8: P400 amplitude during two epochs.
Figure 4.9: Late-positive complex at electrode POz.

The diagram shows the late-positive complex (LPC) at electrode POz for different groups: younger and older musicians, as well as younger and older non-musicians. The graphs compare responses under different conditions: Tuned, 2% mistuned, 4% mistuned, and 16% mistuned.
Figure 4.10: Amplitude of the LPC.

The graph illustrates the amplitude of the LPC (Linear Predictive Coding) for different mistuning conditions of the second harmonic. The data is categorized by musical experience and age:
- Younger Musician
- Younger Non-musician
- Older Musician
- Older Non-musician

The x-axis represents the mistuning of the second harmonic, ranging from Tuned to 16%. The y-axis shows the LPC amplitude in microvolts (µV), ranging from 0 to 5. The bars indicate the range of values.

The graph shows how mistuning affects the LPC amplitude differently across the categories, with younger musicians generally showing higher amplitudes compared to non-musicians, especially at higher mistuning values.