Effect of Hearing Loss and Simulated Hearing Aid Amplification on Prosody Perception: From Neural Encoding to Behavioral Response

Maryam Karimi Boroujeni

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Doctorate in Philosophy degree in Rehabilitation Sciences

School of Rehabilitation Sciences
Faculty of Health Sciences
University of Ottawa

Abstract

Speech prosody, an essential component of social interaction for conveying emotional information, has been often overlooked in the assessment and management of hearing-impaired individuals. Furthermore, existing literature remains inconclusive on whether hearing loss impairs emotion perception and to what extent hearing aids can compensate for potential deficits in prosody processing. Previous studies have predominantly examined behavioral performance in emotion perception tasks, with limited insight into the neural processing of prosody in normal-hearing and hearing-impaired populations.

This study investigated whether the speech-evoked frequency-following response (FFR) can track variations in fundamental frequency (F0 contour) in emotional speech, and to examine male/female differences in the neural encoding of emotional prosody. Additionally, we evaluated the impact of hearing loss and simulated hearing aid processing on the neural representation of F0 contours. To complement electrophysiological measures, the behavioral responses of hearing-impaired individuals were also examined through an emotion identification task conducted under cognitively demanding conditions. We further investigated the association between individuals' ability to perceive prosody and their speech-in-noise perception.

First, FFRs were recorded from normal-hearing young adults in response to the word "balloon" spoken by one male and one female talker with sad and happy emotion, and F0 tracking accuracy was calculated. This method was then applied to hearing-impaired individuals (aided and unaided) and age-matched normal-hearing controls using male-spoken "balloon" expressing sad and happy emotions. Afterward, all participants underwent an emotion identification test with a secondary free recall task. The Quick Speech-In-Noise test (QuickSIN) was also conducted to determine participants' SNR loss.

The results showed the FFR can track F0 contours in emotional speech, influenced by emotion type and talker voice characteristics but not by listener sex. Hearing loss disrupted neural representation of F0 contours, although amplification could partially restore it. Hearing loss also adversely affected behavioral performance and increased cognitive demands for prosody perception; however, amplification could not fully compensate for this perceptual deficit. Moreover, SNR loss was negatively correlated with emotion identification accuracy.

This study emphasized the need to consider prosody perception difficulties in the assessment and rehabilitation of hearing-impaired individuals and suggested the FFR's potential for identifying deficits in prosodic processing.

Preface

This thesis is the original work of Maryam Karimi Boroujeni (MK), the doctoral candidate, and was conducted under the supervision of Prof. Christian Giguère and Prof. Hilmi Dajani. All aspects of the research, including study design, data collection, data analysis, interpretation of findings, and manuscript preparation were carried out by MK. In case of collaborations, collaborators/co-authors information are included and acknowledged in the relevant chapters or sections.

This research project was conducted with the approval of the University of Ottawa Research Ethics Board. The first and second studies of this thesis, presented in Chapters 3 and 4, were approved under ethics file number H-03-22-7915. The third study, presented in Chapter 5, was approved under ethics file number H-01-23-8637. Chapters 2, 3, 4, and 5 are formatted according to the submission or publication guidelines of the peer-reviewed journals in which they have been published or are intended to be submitted.

Dedication

To Mehraban Badiee Ardestani,
an undergraduate student I never knew in life, but whose absence reached into mine.

We were both part of the same Faculty of Health Sciences at the University of Ottawa, you at the beginning of your journey, full of promise and light; I, just beginning my PhD. Perhaps we passed each other in the hallway or boarded the same bus, never knowing that someone so young, so talented, so full of dreams was present in the same spaces I passed through.

One day, a small potted flower, yours, was placed in my hands. I was told it had belonged to you. Since then, I've cared for it the way I wish the world had cared for you, with gentleness, with hope, with both hands. It blooms quietly in the corner of my home, growing through seasons, holding onto life. Like you might have.

I look at it often and wonder who you would have become. A researcher? A nurse? A doctor? A friend I might have had? I grieve for the questions left unanswered, for the brilliance never realized. But I also carry you forward, in this work, in this story, in the silent bond between two lives that never touched but still found each other.

This thesis is for you, Mehraban*.
For the future you were supposed to have. For the flower still blooming in your name.

** Mehraban was one of 176 bright souls, each carrying dreams, futures, and untold stories, lost in the heartbreaking tragedy of Flight PS752 in January 2020.*

Acknowledgements

Before any of the pages that follow, before the research, the findings, the discussions, I need to acknowledge the quiet and powerful forces that carried me here. This thesis is the result of countless unseen moments of patience, support, and love that sustained me through doubt, exhaustion, and growth.

To my supervisor, Dr. Christian Giguère, thank you for your calm and continuous support throughout every stage of this process. Your knowledge, valuable feedback, and kind mentorship helped shape not only this thesis but my growth as a researcher and as a person. You have been both my strongest advocate and my most thoughtful challenger, and I am deeply grateful for the commitment you invested in me. You saw my potential even when I struggled to see it myself. I will always remember your patience and belief in my ability to finish this work.

To my co-supervisor, Dr. Hilmi Dajani, thank you for your invaluable guidance and constant encouragement. Your analytical insight and thoughtful questions helped me push beyond surface answers. You always brought clarity when my thoughts felt scattered and uncertain. Your engineering perspective in this project gave me strength and direction, and I'm truly grateful for your dedication and support.

To my committee members, Dr. Amineh Koravand and Dr. Anna Zumbansen, thank you for your time, your trust, and your insightful contributions to my journey. Your questions helped me see things from new perspectives, and your encouragement reminded me that my work matters.

To Dr. Elisabeth Fitzpatrick, Director of the program, thank you for leading with kindness and creating an environment where I could truly learn and grow. Your support behind the scenes made a world of difference.

To my parents, whose sacrifices and unconditional love have laid the very foundation of my life, thank you for believing in this dream even when the path felt impossible. You supported me not just with words, but with your hearts. Every achievement of mine is rooted in your belief in me. You gave me the wings to fly.

To my siblings, Marjan and Mehdi. Not a day passed without your love reaching me from thousands of kilometers away. I felt your care in every check-in and every time you stepped in for the family when I couldn't be there. Thank you for holding the space I left behind with such grace, and for always reminding me that no matter where life takes us, we're in this together.

To Mina, my dearest friend who lives across time zones but whose love always arrived right on time: thank you for your emotional presence and your open hearts. To my friends Fatemeh, Marjan, and Kimia, thank you for showing up in the quiet ways that meant the most, in the random texts and the reminders to eat and sleep.

I also owe my deepest respect and gratitude to the incredible participants of this research. They gave their time, energy, and personal stories with such grace and generosity. Some came in during their days off, some took time away from their works and families, and some even participated around the holidays, yes, even on Christmas Eve and Christmas Day. The fact that you showed up, trusted me, and shared your experiences is something I will carry with me forever. This thesis would simply not exist without you. Thank you for contributing not just to my research, but to the broader understanding of hearing health and rehabilitation sciences.

And finally, to my husband, Ali. No words could ever fully capture what you have meant to me throughout this journey. From the day you met me at the airport, camera in hand and heart wide open, to every sleepless night we spent apart, video calling through time zones, to the moment we stood together and said "I do", you have been next to me. Through pandemic separations, devastating loss, cross-province commutes, and every storm we weathered, you held me, uplifted me, believed in me when I was falling apart, and celebrated every little victory as if it were your own. This thesis may carry my name, but your soul is etched into every page. You were my calm during chaos, my home between two cities, my reason to keep writing when I wanted to stop. Thank you for everything!

To all of you who held me in silence, who cheered for me without needing recognition, who walked beside me even from afar: thank you.

Table of Contents

List of Tables and Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Abbreviations

CI	cochlear implant
dB	decibel
dB SL	decibels-sensation level
dB SPL	decibels-sound pressure level
ESD	emotional speech database
FFR	frequency following response
F0	fundamental frequency
F1	formant 1
HAS	harmonic Amplitude Summation
HA	hearing aid
HI	hearing impaired
Hz	hertz
ms	milliseconds
NH	normal hearing
PS	pleasant surprise
QuickSIN	Quick Speech-in-Noise
RT	reaction time
RMSE	Root Mean Squared Error
SNHL	sensorineural hearing loss
SNR	signal-to-noise ratio
TESS	Toronto Emotional Speech Set

Chapter 1. Introduction

Introduction

Interactive communication relies on speech being transmitted in a clear, dynamic, and responsive flow [1]. This requires the speaker to produce speech components accurately and the listener to effectively perceive and interpret the relevant cues to make sense of the message. Some of these cues are at the consonant/vowel segmental level and the others are suprasegmentals spread across multiple segments [2, 3]. Despite suprasegmental features contributing an important layer of meaning to speech [4], most speech perception studies have focused on the segmental aspects of speech, leaving suprasegmental features largely unexamined.

Suprasegmental information, also known as prosody, conveys a wide range of linguistic and affective information [5, 6]. Linguistic prosody delivers grammatical and pragmatic cues, while affective prosody expresses the attitude and the emotional state of the speaker [5, 6]. In emotional expression, prosodic cues are acoustically characterized by variations in fundamental frequency (F0 contour), level (intensity contour), and duration [7, 8]. Each emotion has a specific acoustic profile that the auditory system must faithfully receive and accurately encode. At higher levels of processing, the auditory system integrates these cues with broader neural networks involved in language, cognition, and memory to extract contextual meaning and pick out relevant information [9-12].

When hearing impairment occurs, the auditory system has limited access to the acoustic features of speech not only due to reduced audibility, but also because of degraded spectro-temporal resolution and disrupted integrity of peripheral structures and the central auditory pathway [13-15]. Even when hearing aids provide audibility, they often do not fully restore speech perception [16, 17]. The degree to which hearing loss and hearing aids affect prosody perception remains elusive. Some behavioral studies have indicated that hearing loss leads to deficits in emotion perception [18, 19], whereas others have argued that residual hearing may be adequate to preserve emotional cues [20, 21]. Studies on hearing aids have also shown inconsistent results, with some evidence supporting benefits for emotional speech perception [22], while other findings suggest that frequency and amplitude compression may distort the intensity and spectral cues essential for distinguishing emotions [23, 24]. Despite the importance of neural networks involved in prosody processing [9, 25], these studies have focused on the behavioral aspects of emotion perception, and there is limited understanding of the subcortical and cortical processing in both normal-hearing individuals and those with hearing impairment.

Electrophysiological approaches make it possible to investigate the neural mechanisms underlying speech processing in the brain, and the Speech-evoked Frequency Following Response (sFFR) is one such non-invasive neural measure. Whereas the traditional FFR relied on simple non-speech stimuli such as pure tones or clicks to reflect basic temporal encoding, the Speech-evoked FFR (sFFR) uses acoustically rich speech signals to reflect how the central auditory system encodes the dynamic temporal and spectral features of speech [26, 27]. F0-related information is one of the key speech features encoded in the sFFR, based on the timing pattern of neural discharges and interspike intervals [28]. Using synthetic or natural isolated vowels and simple syllables, some sFFR studies have demonstrated that the phase-locked neural activity responsible for the FFR is dynamic enough to track small variations in speech periodicity over time [29-31]. However, the ability of the sFFR to reliably track F0 contours beyond segmental information remains unclear. Also, as prior studies have indicated that the sFFR is affected by the acoustic characteristics of male and female voice [32], and that F0 encoding may vary across male and female listeners [33, 34], further investigation is needed to assess the effect of talker and listener on neural encoding of prosody. This study aimed to examine whether the sFFR can capture F0 contours associated with sad and happy emotions in young adults with normal hearing, and to assess sex-related differences in F0 tracking ability.

Additionally, previous studies have provided strong evidence supporting the use of the sFFR in hearing loss assessment and hearing aid fitting using isolated vowel sounds or monosyllabic stimuli. Specifically focusing on F0 encoding, several studies have reported both reduced and enhanced F0 amplitude, averaged over brief vowels/monosyllables, in individuals with hearing loss, as well as in those using hearing aids [32, 35-38]. However, the extent to which hearing loss and hearing aids influence the neural encoding of time-varying F0 information within the context of prosody remains to be investigated. Accordingly, another aim of this study was to evaluate F0 contour tracking accuracy in response to sad and happy emotions in hearing-impaired individuals under both unaided and aided conditions, and to compare their performance to that of age-matched normal-hearing individuals.

To complete the test battery and provide more comprehensive assessment, this study additionally constructed and employed an emotion identification test. Given that verbal prosody perception engages cognitive functions [9, 39], a free recall task was added to the emotion identification test to evaluate the impact of hearing loss and hearing aid on emotional prosody perception in a more cognitively demanding condition. Moreover, as prosody is not routinely assessed in standard audiologic evaluations, we aimed to determine whether common clinical tests could help predict performance on emotional prosody tasks among hearing-impaired individuals. Given that Quick Speech-in-Noise test (QuickSIN) relies on auditory and cognitive skills also essential for emotion perception, such as pitch perception, temporal resolution, and cognitive functions, we examined the extent to which speech-in-noise performance predicts emotion recognition ability. Identifying a correlation between these two tests may shed light for audiologists on potential prosodic deficits when patients perform poorly on speech perception tasks in noisy environment.

Thesis objectives

This project focused on the prosody perception abilities of hearing-impaired individuals, both with and without amplification. To this end, we first aimed to provide an overview of the different functions and forms of prosody and reviewed existing research on prosody perception in hearing-impaired individuals and the role of hearing assistive technologies, with the goal of identifying current gaps and limitations in literature.

Secondly, to address some of these gaps, we investigated the feasibility of using the sFFR to evaluate prosody processing in normal-hearing young adults by assessing the accuracy of F0 contour tracking across different emotions. This included examining the effects of talker and listener sex on neural encoding of prosodic features.

After confirming the applicability of the FFR to prosody, our third objective was to investigate how hearing loss and hearing aid processing affect the neural tracking of F0 contours.

Lastly, we aimed to examine the perceptual performance of the same hearing-impaired individuals on a behavioral emotion identification test, and how cognitively demanding the task was for them compared to normal-hearing participants. We also investigated whether their QuickSIN scores were correlated with their performance on the emotion identification test.

Thesis Outline

This thesis consists of six chapters and is presented as a series of manuscripts, some of which have been submitted or published, and others are ready for submission to peer-reviewed journals, as outlined below.

Chapter 2 provides a background on perception of prosody in hearing-impaired individuals and users of hearing assistive devices. This narrative review has been published in Journal of Speech, Language, and Hearing Research.

Chapter 3 includes the result of the sFFR recording in normal-hearing individuals and indicates the sFFR capability to capture emotional cues and the variables that modulate it. This chapter has been submitted to a peer-review journal for publication as original research.

Chapter 4 reports the findings on how hearing loss and hearing aid influence the neural encoding of acoustic-related emotional cues. The chapter is complete and ready for journal submission.

Chapter 5 presents the behavioral performance of hearing-impaired individuals, both with and without hearing aid, in response to emotional speech. It also addresses the relationship between speech perception in noise and emotion perception in hearing impaired individuals. This chapter is ready to be submitted.

Finally, Chapter 6 provides a general discussion that summarizes the key findings of this thesis and draws connections between behavioral and electrophysiological results. It concludes by proposing new approaches for assessing prosody perception in individuals with hearing loss and suggests new avenues for enhancing clinical evaluation and rehabilitation.

References

1. Ng, S.H. and J.J. Bradac, *Power in Language: Verbal Communication and Social Influence*. 1993, Newbury Park, CA: SAGE Publications, Inc.
2. Bolinger, D. and D.L.M. Bolinger, *Intonation and Its Parts: Melody in Spoken English*. 1986: Stanford University Press.
3. Gaikwad, S.K., B.W. Gawali, and P. Yannawar, *A review on speech recognition technique.* International Journal of Computer Applications, 2010. **10**(3): p. 16-24.
4. Nygaard, L.C., D.S. Herold, and L.L. Namy, *The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning.* Cognitive science, 2009. **33**(1): p. 127-146.
5. Raithel, V. and M. Hielscher-Fastabend, *Emotional and linguistic perception of prosody: Reception of prosody.* Folia Phoniatrica et Logopaedica, 2004. **56**(1): p. 7-13.
6. Nooteboom, S., *The prosody of speech: Melody and rhythm. The handbook of phonetic sciences. Hardcastle WJ, Laver J*. 1997, Oxford: Basil Blackwell.
7. Pickett, J. and S.R. Morris, *The acoustics of speech communication: Fundamentals, speech perception theory, and technology*. 2000, Acoustical Society of America.
8. Banse, R. and K.R. Scherer, *Acoustic profiles in vocal emotion expression.* Journal of personality and social psychology, 1996. **70**(3): p. 614.
9. Grandjean, D., *Brain networks of emotional prosody processing.* Emotion Review, 2021. **13**(1): p. 34-43.
10. Fries, P., *A mechanism for cognitive dynamics: neuronal communication through neuronal coherence.* Trends in cognitive sciences, 2005. **9**(10): p. 474-480.
11. Ceravolo, L., et al., *Basal ganglia and cerebellum contributions to vocal emotion processing as revealed by high-resolution fMRI.* Scientific reports, 2021. **11**(1): p. 10645.
12. Pannese, A., D. Grandjean, and S. Frühholz, *Subcortical processing in auditory communication.* Hearing research, 2015. **328**: p. 67-77.
13. Moore, B.C., *Cochlear hearing loss: physiological, psychological and technical issues*. 2007: John Wiley & Sons.
14. Ponsot, E., et al., *Mechanisms of spectrotemporal modulation detection for normal-and hearing-impaired listeners.* Trends in hearing, 2021. **25**: p. 2331216520978029.
15. Syka, J., *Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning.* Physiological reviews, 2002. **82**(3): p. 601-636.
16. Ng, J.H.-Y. and A.Y. Loke, *Determinants of hearing-aid adoption and use among the elderly: A systematic review.* International journal of audiology, 2015. **54**(5): p. 291-300.
17. Lesica, N.A., *Why do hearing aids fail to restore normal auditory perception?* Trends in neurosciences, 2018. **41**(4): p. 174-185.
18. Most, T. and C. Aviner, *Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing.* Journal of deaf studies and deaf education, 2009. **14**(4): p. 449-464.
19. Goy, H., et al., *Perception of emotional speech by listeners with hearing aids.* Canadian Acoustics, 2016. **44**(3).
20. Irino, T., et al., *Effects of age and hearing loss on speech emotion discrimination.* Scientific Reports, 2024. **14**(1): p. 18328.
21. Boothroyd, A., *Auditory perception of speech contrasts by subjects with sensorineural hearing loss.* Journal of Speech, Language, and Hearing Research, 1984. **27**(1): p. 134-144.

22. Schmidt, J., et al. *Do hearing aids improve affect perception?* in *Physiology, psychoacoustics and cognition in normal and impaired hearing*. 2016. Springer International Publishing.
23. Goy, H., et al., *Acoustic analysis of emotional speech processed by hearing aids.* Canadian Acoustics, 2017. **45**(3): p. 178-179.
24. Arefi, H.N., et al., *Effect of hearing aid amplitude compression on emotional speech recognition.* Auditory and Vestibular Research, 2017. **26**(4): p. 223-230.
25. Karimi-Boroujeni, M., H.R. Dajani, and C. Giguère, *Perception of prosody in hearing-impaired individuals and users of hearing assistive devices: An overview of recent advances.* Journal of Speech, Language, and Hearing Research, 2023. **66**(2): p. 775-789.
26. Krizman, J. and N. Kraus, *Analyzing the FFR: A tutorial for decoding the richness of auditory function.* Hearing research, 2019. **382**: p. 107779.
27. Skoe, E. and N. Kraus, *Auditory brain stem response to complex sounds: a tutorial.* Ear and hearing, 2010. **31**(3): p. 302-324.
28. Coffey, E.B., et al., *Evolving perspectives on the sources of the frequency-following response.* Nature Communications, 2019. **10**(1): p. 5036.
29. Krishnan, A. and J. Parkinson, *Human frequency-following response: representation of tonal sweeps.* Audiology and Neurotology, 2000. **5**(6): p. 312-321.
30. Krishnan, A., et al., *Human frequency-following response: representation of pitch contours in Chinese tones.* Hearing research, 2004. **189**(1-2): p. 1-12.
31. Dajani, H.R., et al., *Recording human evoked potentials that follow the pitch contour of a natural vowel.* IEEE Transactions on Biomedical Engineering, 2005. **52**(9): p. 1614-1618.
32. Easwar, V., et al., *Evaluation of speech-evoked envelope following responses as an objective aided outcome measure: Effect of stimulus level, bandwidth, and amplification in adults with hearing loss.* Ear and Hearing, 2015. **36**(6): p. 635-652.
33. Ahadi, M., et al., *Gender disparity in subcortical encoding of binaurally presented speech stimuli: an auditory evoked potentials study.* Auris Nasus Larynx, 2014. **41**(3): p. 239-243.
34. Krizman, J., et al., *Athleticism and sex impact neural processing of sound.* Scientific reports, 2022. **12**(1): p. 15181.
35. Ananthakrishnan, S., A. Krishnan, and E. Bartlett, *Human frequency following response: neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss.* Ear and hearing, 2016. **37**(2): p. e91-e103.
36. Koravand, A., et al., *Speech-evoked auditory brainstem responses in children with hearing loss.* International Journal of Pediatric Otorhinolaryngology, 2017. **99**: p. 24-29.
37. BinKhamis, G., et al., *Speech auditory brainstem responses in adult hearing aid users: Effects of aiding and background noise, and prediction of behavioral measures.* Trends in Hearing, 2019. **23**: p. 2331216519848297.
38. Jenkins, K.A., et al., *Effects of amplification on neural phase locking, amplitude, and latency to a speech syllable.* Ear and hearing, 2018. **39**(4): p. 810-824.
39. Husain, F.T., J.R. Carpenter-Thompson, and S.A. Schmidt, *The effect of mild-to-moderate hearing loss on auditory and emotion processing networks.* Frontiers in systems neuroscience, 2014. **8**: p. 10.

Chapter 2. Perception of prosody in hearing-impaired individuals and users of hearing assistive devices: An overview of recent advances

Perception of prosody in hearing-impaired individuals and users of hearing assistive devices: An overview of recent advances

This chapter is based on the following peer-reviewed article, which has been published:

Karimi-Boroujeni, M., Dajani, H. R. & Giguère, C. Perception of prosody in hearing-impaired individuals and users of hearing assistive devices: An overview of recent advances. Journal of Speech, Language, and Hearing Research. 66, 775-789 (2023). https://doi.org/10.1044/2022_JSLHR-22-00125

Abstract

Purpose: Prosody perception is an essential component of speech communication and social interaction through which both linguistic and emotional information are conveyed. Considering the importance of the auditory system in processing prosody-related acoustic features, the aim of this article is to review the effects of hearing impairment on prosody perception in children and adults. It also assesses the performance of hearing assistive devices in restoring prosodic perception.

Method: Following a comprehensive online database search, two lines of inquiry were targeted. The first summarizes recent attempts towards determining the effects of hearing loss and interacting factors such as age and cognitive resources on prosody perception. The second analyzes studies reporting beneficial or detrimental impacts of hearing aids, cochlear implants, and bimodal stimulation on prosodic abilities in people with hearing loss.

Results: The reviewed studies indicate that hearing-impaired individuals vary widely in perceiving affective and linguistic prosody, depending on factors such as hearing loss severity, chronological age, and cognitive status. Also, most of the emerging information points to limitations of hearing assistive devices in processing and transmitting the acoustic features of prosody.

Conclusion: The existing literature is incomplete in several respects, including the lack of a consensus on how and to what extent hearing prostheses affect prosody perception, especially the linguistic function of prosody, and a gap in assessing prosody under challenging listening situation such as noise. This review article proposes directions that future research could follow to provide a better understanding of prosody processing in the hearing impaired, which may help healthcare professionals and designers of assistive technology to develop innovative diagnostic and rehabilitation tools.

Keywords: Prosody perception; Hearing loss; Hearing aid; Cochlear implant

1. Introduction

Interactive communication requires information to be clearly transmitted in the form of dynamic and two-way flow of the speech signal (Ng & Bradac, 1993). To make this possible, all speech components must be accurately produced by the speaker, and the listener must be capable of perceiving and interpreting the relevant features to extract the message (Brown, 1995). Some of these features are at the consonant/vowel segmental level and the others are suprasegmentals lying across multiple segments (Bolinger & Bolinger, 1986). The suprasegmental information in speech is neglected in the vast majority of speech perception studies; however, it is an important contributor of information that complements the properties of segmental sounds like vowels and consonants (Miller, 1978; Nygaard et al., 2009).

The suprasegmental information, also called prosody, refers to the melody of speech and is comprised of three main elements, namely intonation, stress, and rhythm (Shattuck-Hufnagel & Turk, 1996). These prosodic elements are defined acoustically as variations in fundamental frequency (F0), intensity, and duration of the utterance (Banse & Scherer, 1996; Cruttenden, 1997). At the perceptual level, these acoustic features correspond to pitch, loudness, and length, respectively (Pickett & Morris, 2000). Intonation is mainly characterized by F0 variations over time, also referred to as F0 contour. However, changes in intensity and duration are also associated with intonation patterns (Lieberman, 1965; Raphael et al., 2007). Stress refers to the intensity variations (namely level contour) and duration, along with F0 (Gay, 1978; van der Hulst, 2006). Rhythm is primarily determined by changes in duration (e.g., syllabic rate) and intensity as well as F0 (Nooteboom, 1997; Raphael et al., 2007). Figure 1 shows an example of how these acoustic parameters vary in different prosodic patterns.

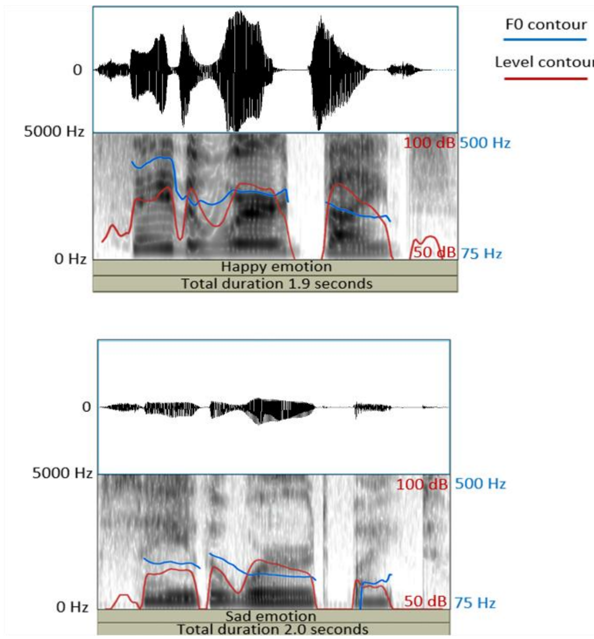


Figure 1. Waveform, spectrogram and F0/level contours for the sentence "Say the word bite" from the Toronto Emotional Speech Set (Dupuis & Pichora-Fuller, 2011) spoken by a female in two simulated emotional states: (a) Happy (duration: 1.90s, F0: mean of 294 Hz, range: 204-419 Hz, Level: mean of 73 dB, range 24-79 dB) and (b) Sad (duration: 2.04s, F0: mean of 196 Hz, range: 75-250 Hz, Level: mean of 61 dB, range 28-68 dB)

Prosody conveys a wide range of information during speech communication, including linguistic and affective information. Linguistic prosody provides valuable markers through which the listener can parse continuous speech into meaningful segments and track information that lies beyond the words (Fischer, 2018; Shattuck-Hufnagel & Turk, 1996; Vaissière, 1983). This function improves speech recognition, especially in noisy environments, by directing attention to particularly important syllables, words, or phrases in a sentence (Binns & Culling, 2007; Chen et al., 2018). Not only does prosody directly reflect the speaker's intended content and meaning but affective prosody also contributes to express the emotional state of speaker (Banse & Scherer, 1996; Mozziconacci, 1998). There is also a growing body of data indicating that prosody perception in children is a prerequisite for healthy language acquisition, reading development, and social development (Barth-Weingarten et al., 2010; Calet et al., 2019; Myers et al., 2018; Paul et al., 2005; Peppé & McCann, 2003).

How well prosodic phenomena are perceived depends on several factors, one of which is the hearing function (Kalathottukaren et al., 2017; Picou et al., 2018). The healthy auditory system is able to encode the spectro-temporal features of speech accurately (Delgutte, 1997; Katz et al., 2015). In the case of hearing impairment, the salience of the prosodic patterns can diminish and lead to impaired perception of prosody, which is associated with poor speech recognition (Grant, 1987; Gürses et al., 2020; Niu et al., 2019). Among hearing-impaired individuals, difficulties in prosody perception may be different from adults to children as they depend on age-related changes in cognition, as well as type and prescription time of hearing assistive devices (Jiam et al., 2017; Picou et al., 2018). Despite the significant impact of prosody perception deficit in day-to-day communication of hearing-impaired individuals, this is rarely considered in clinical evaluation or hearing aid fitting. That modern digital hearing aids (HAs) and cochlear implants (CIs) have not been able to fully satisfy all users lends credence to the importance of this issue (Goy et al., 2016; Meister et al., 2009, 2011; Most & Aviner, 2009; Most & Peled, 2007; WysocKa & Mackiewicz, 2017). However, it remains unclear to what extent hearing assistive devices can restore affective and linguistic prosody.

This review paper provides an overview of recent findings on different functions of prosody and their neurophysiological basis. Also, our purpose is to address the current gaps concerning the prosody perception abilities of people with hearing loss in order to inform the reader of the current state of knowledge on this topic and outline possible avenues for further research. More specifically, we aim to present the range of prosodic perception deficits and their implications in people with hearing loss. Also, the efficacy of hearing assistive devices at restoring prosodic cues under quiet and noise conditions is described.

2. Types of Prosody

Classification of prosody can be useful to distinguish the various functions of this phenomenon in human communication. Given the broad range of prosody perception deficits, a classification system also enables experts to develop specific assessment tools for evaluating different prosodic skills. Accordingly, there is a general agreement that prosody can be functionally categorized into "linguistic prosody" and "affective prosody" (Belyk & Brown, 2014; Raithel & Hielscher-Fastabend, 2004). Alongside these two functions, some studies have also related the indexical information, such as age, identity, and gender of the speaker to prosody.

2.1. Linguistic prosody

The prosodic structure of an utterance signals linguistic contrasts by providing grammatical and pragmatic distinctions (Hirschberg, 2002). The grammatical function of prosody comprises lexical stress, phrasal prominence, syntactic chunking, and the identification of a sentence as declarative or interrogative (Field, 2005; Marshall et al., 2009; Pickett & Morris, 2000; Steffman, 2021). Lexical stress is the emphasis placed on a given syllable in the word and is associated with increased intensity, longer duration, and higher pitch. Which syllable of the word is stressed can produce differences in meaning or word function (e.g., noun "PERmit” vs. verb “perMIT") (Nakatani & Schaffer, 1978; Zuraiq & Sereno, 2007). If stress is put on words at the phrase or sentence level, it could carry prominence that represents the speaker's intended meaning (e.g., "KARA likes cats" communicates that it is Kara, not Lisa, who likes cats) (Pickett & Morris, 2000). Syntactic chunking refers to the condition in which the speaker disambiguates syntactically equivocal strings of words (e.g., “That is just insight” vs. “That is just in sight”) by using pauses and lengthening at the word level. Falling and rising intonation patterns can also differentiate declarative sentences from interrogative sentences (e.g., “That is just insight.” vs. “That is just insight?”) (Pickett & Morris, 2000).

The pragmatic function of prosody includes turn-taking and sarcasm (Hjalmarsson, 2011; Vaissière, 1983). Turn-taking allows the speaker to show the end of conversation by lengthening the final syllable and decreasing intensity and frequency (Ohsuga et al., 2006; Wells & Peppé, 1996). Sarcasm is another phenomenon in verbal communication, expressing the condition where an utterance is used as its opposite meaning (Anolli et al., 2000; Colston, 1997). In the sarcastic condition, the key words usually have longer duration, flatter pitch contours, and shallower rising level contours irrespective of sentence type and speaker gender (Chen & Boves, 2018; Dumouchel et al., 2009).

Perception of these prosodic functions can be assessed by a few published behavioral tests in English, including Perception of Prosody Assessment Tool (PPAT), Profiling Elements of Prosody in Speech-Communication (PEPS-C), and Test of Prosody via Syllable Emphasis (TOPsy) (Klieve & Jeanes, 2001; Nayak et al., 2022; Peppé & McCann, 2003). However, these tests do not comprehensively cover all subsets of the aforementioned functions, which may indicate a need for the development of novel and specific practical approaches. For further details on available prosody perception instruments, see Table S1 in Appendix J.

2.2. Affective prosody

A successful social interaction depends on the extent to which the listener can understand what others feel and think (Busso et al., 2009). This affective communication has a prominent place in human interactions, which aids in social development, social cognition, and emotional well-being (Chatterjee et al., 2019). Limited access to emotional cues could impede communications, resulting in low morale, confliction, poor psychosocial function, and reduced quality of life (Rivers et al., 2012; Schorr et al., 2009). Emotional states, according to the categorical system, can be described with a few basic descriptors such as fear, sadness, joy, happiness, and so on (Calvo & D'Mello, 2010; Calvo & Mac Kim, 2013; Izard, 1977). Also, the varieties of emotions can be categorized based on the dimensional system in which emotions are characterized with two dimensions: arousal (exciting vs. calming) and valence (pleasant vs. unpleasant) (Bradley et al., 2000; Picou et al., 2018). Much of this emotional information is conveyed verbally and through prosodic features of speech (Watson et al., 2020; Scherer, 1995). Generally, those emotions which lead to listener arousal are usually related to high-pitch and high-intensity sounds (Picou et al., 2018; Weninger et al., 2013). Although the acoustical features of different emotions may overlap, each emotion has its own acoustic profile through which the auditory system can distinguish that emotion from others (Banse & Scherer, 1996; Goy et al., 2017; Picou et al., 2018).

Perceiving and interpreting the emotional patterns in speech prosody involves a multi-step process in which the peripheral auditory system first extracts and encodes the acoustic cues and transmits them to the different nuclei in the brainstem for further processing (Fries, 2005; Grandjean, 2021). Thereafter, the emotion-related cues are processed along the ascending auditory pathway including the inferior colliculus in the midbrain, the medial geniculate nucleus in the thalamus, and cortical areas (Ceravolo et al., 2021; Pannese et al., 2015). Alongside the feedforward information flow, feedback information processing is also activated in cortical areas to infer contextual meaning and pick out relevant information through attentional top-down mechanism (Dupuis & Pichora-Fuller, 2014; Fries, 2005). In addition, visual cues (facial expressions and gestures) are required to integrate with speech prosody in order to optimally recognize a speaker's emotions (de Boer et al., 2021; Ernst & Banks, 2002).

As is the case for linguistic prosody, there are some validated but limited number of assessment tools that can be utilized for evaluating the perception of emotional prosody, such as the Emotional Communication in Hearing Questionnaire (EMO-CHeQ), Minnesota Tests of Affective Processing (MNTAP), Aprosodia battery, and Advanced Clinical Solution (ACS) (Lai et al., 1991; Pearson, 2009; Ross et al., 1997; Singh et al., 2019), as described in Table S1 in Appendix J. Furthermore, in auditory emotion research, the Toronto Emotional Speech Set (TESS) has been used for assessing emotion identification in adults, which includes 200 words embedded in the carrier phrase "say the word". These sentences were recorded using one older and one young female voice to express seven different emotions, including angry, disgust, fear, happy, pleasant surprise, sad, and neutral (Dupuis & Pichora-Fuller, 2014).

2.3. Brain networks of linguistic and affective prosody perception

The processing of affective and linguistic prosody is often lateralized to right and left hemispheres, respectively (Sheppard et al., 2020; Witteman et al., 2011). Nevertheless, processing speech prosody involves multiple brain regions consisting of a large-scale network that spreads spatially across both hemispheres, with greater lateralization in temporal-lobe auditory areas (Belyk & Brown, 2014). Some neuroimaging studies have indicated that these two aspects of prosody engage both distinct and shared brain areas (Grandjean, 2021). Affective prosody can be exclusively localized in different regions, including the orbital part (BA 47) of left inferior frontal gyrus (IFG), thalamus, right posterior superior temporal gyrus (pSTG/BA 22) and right middle temporal gyrus (MTG/BA 21) (Kotz et al., 2003; Pichon et al., 2009; Seydell-Greenwald et al., 2020; Wildgruber et al., 2005). Linguistic prosody, in contrast, has been found to strongly stimulate areas like claustrum, left pSTG, opercular part of bilateral IFG (ventral BA 44), bilateral middle frontal gyrus (BA 9) and right supramarginal gyrus (BA 40) (Gandour et al., 2007; Tong et al., 2005). Furthermore, shared overlapping areas between affective and linguistic prosody mostly include but are not limited to, bilateral supramarginal gyrus (SMG/BA 40), right STG (BA 22), right middle frontal gyrus (BA 9), midline supplementary motor area (SMA/BA 6), and left insula (BA 13) (Aziz-Zadeh et al., 2010; Belyk & Brown, 2014; Hesling et al., 2005; Lima et al., 2016). These brain networks are represented in Figure 2.

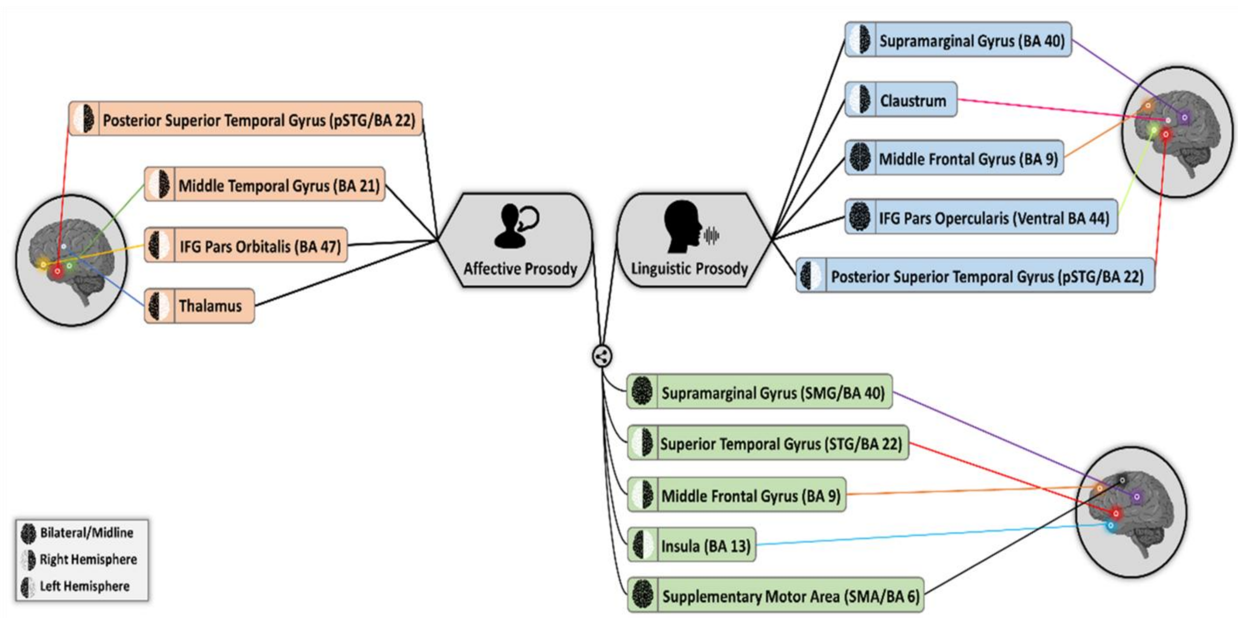


Figure 2. The distinct and shared higher brain areas involved in affective and linguistic prosody

3. Effect of hearing loss and interacting factors on prosody perception

A narrative review of the literature was conducted to summarize peer-reviewed articles on the effect of hearing loss on prosody perception published from 2000 to May 2022. To find pertinent investigations, key words "prosody perception" or "linguistic prosody" or "affective prosody" or "emotion perception" or "emotion" or "emotion identification" combined with "hearing impairment", "hearing loss", "cognition", "children", "adults", "prosody rehabilitation", “auditory training”, “training” were searched on the electronic databases PubMed, Scopus, and Google Scholar. After removing duplications, this search strategy uncovered 94 articles. The title and abstract of each research study were first screened by the first author and then inclusion was validated by the second and third authors. Articles were included if they: (a) had enrolled children or adults who had been diagnosed with hearing loss; (b) had used behavioral or electrophysiological assessment tools to assess speech-embedded vocal prosody; (c) had compared the prosodic abilities of hearing-impaired individuals against normal hearing listeners or other people with hearing loss or under different conditions; and (d) had measured the parameters that appear to mediate prosody perception in hearing impairment, including hearing loss severity, age at onset of hearing impairment, and use of auditory training. Excluded were articles that investigated acoustic variations in tonal languages, visual-only speech prosody perception, prosody production, or emotional perception of music. Also, conference abstracts and review articles were excluded. Subsequently, 25 key studies were selected to screen by full text.

3.1. Hearing loss severity

Once damage to the auditory system occurs, reduced audibility limits access to prosodic cues due to inadequate auditory input. Moreover, even if speech stimuli are presented at a suprathreshold level, degraded temporal and frequency resolution in the auditory system may still restrict access to the tempo-spectral acoustic features carrying speech prosody, such as F0 and level contours (Lai et al., 1991). While an earlier study found that children (14-19 years) with hearing loss, but residual hearing in the low frequencies, did not encounter difficulties in perceiving acoustic information related to emotions (Boothroyd, 1984), recent studies on vocal emotion perception in the hearing-impaired population have shown significant deficit in perceiving auditory-only emotional cues, especially those emotions which are associated with less variation in their acoustic profile such as fear, disgust, and surprise (Most & Aviner, 2009; Most & Michaelis, 2012). While most of these studies involved individuals with severe hearing loss, Dyck et al. (2004) indicated that both younger (6–11 years) and older (12–18 years) children with hearing loss ranging from mild to moderate obtain lower scores in emotion recognition and understanding ability. The authors of this study have attributed such an emotion recognition deficit in hearing-impaired children to their delayed acquisition of language abilities. As children age, increased social and conversational experiences can partly overcome this developmental delay. However, Cannon & Chatterjee (2019) found that voice emotion perception skill in children with mild-to-moderate hearing loss (8 – 14 years) had no significant deficit, and degree of hearing loss did not have impact on their prosodic abilities. In the case of unilateral hearing loss, recent evidence by Gürses et al. (2020) showed that single-sided deafness in adults does not affect emotional prosody perception skills, supporting the claim that audibility in the normal ear is enough to provide the required auditory pattern analysis.

3.2. Age at onset of hearing impairment

If hearing impairment occurs in old age, some interacting factors contribute to increase the chance of a deficit in the perception of prosody (Legris et al., 2021; Paulmann et al., 2008).These factors can be attributed not only to age-related peripheral hearing loss but also to the possible presence of central auditory disorders (Mitchell & Kingston, 2014). The functional magnetic resonance imaging (fMRI) study conducted by Demenescu et al. (2015) showed declining prosody encoding at the level of auditory cortex across the adult lifespan, which was reflected by decreased response of superior temporal gyri (STG) to different verbal emotions in aged individuals compared to their younger peers. This view was later supported by Rasmus & Błachnio (2021) who found that, in older adults, deterioration of affective prosody perception is associated with decrease in spatial hearing ability, frequency pattern recognition, and duration pattern recognition. These authors also showed positive correlations between deficit in the perception of linguistic prosody and age-related declines in pitch discrimination, temporal recognition, and frequency pattern recognition. However, other lines of evidence have revealed different results concerning the effect of age-related central nervous system reorganization on the elderly performance with prosodic content. For instance, using a mismatch negativity (MMN) experiment, Giroud et al. (2019) observed that lower cortical thickness and higher cortical surface area of bilateral auditory-related areas were associated with larger MMN magnitudes for word stress violations among elderly normal-hearing individuals compared to younger adults. The authors pointed out that this observation is likely related to elderly people relying on prosodic cues in order to preserve speech perception that can be compromised by progressive brain atrophy.

Moreover, besides central auditory dysfunction, a growing body of literature has indicated that age-related cognitive decline can negatively affect perception of prosodic content (Zinchenko et al., 2018). Attentional resources that facilitate encoding of speech stimuli into verbal short-term and working memory are known as relating to brain regions that undergo such age-related decline (LaCroix et al., 2020). According to findings by Dupuis & Pichora-Fuller (2015), even older adults with good audiometric thresholds were less accurate in identifying vocal emotions compared to their younger counterparts, suggesting the importance of age-associated alterations in cognition and top-down mechanisms in emotion perception.

In children, on the other hand, impaired prosody perception is linked to more severe outcomes in comparison with young adults and older people, most notably as related to difficulties in language development, reading acquisition, social development, emotion identification, reduced intelligibility of spoken language, inability to interpret sarcasm, and loss of sympathy (Ferre, 2015;

Panzeri et al., 2021; Sobhy et al., 2021; Waaramaa et al., 2018). However, how well children perform in prosody perception depends on individual factors such as chronological age, language age, cognitive status, and duration of prosody perception interventions (Tinnemore et al., 2018; Yeshoda et al., 2020).

3.3. Impact of auditory training

Some studies have investigated the impact of different training plans on hearing-impaired individuals with perceptual prosodic difficulties. Music training is one of the most widely used approach in this context. Review studies have indicated that music impacts tasks beyond music skills like speech prosody perception by placing higher demands on overlapped sensory and cognitive processing mechanisms of speech and music in the brain (Patel, 2014). Using Melodic Contour Training with adult post lingually deafened CI recipients, Lo et al. (2015) demonstrated that the declarative/interrogation prosody perception improves 6 weeks after testing.

In deaf children with CIs (6–15 years), recent evidence by Good et al. (2017) also showed that music training can lead to improved performance on tasks that require emotional prosody perception, supporting the efficacy of such supplementary training techniques following cochlear implantation. Likewise, Lo et al (2020) found that a 12-week music training program could help children (6–9 years) with prelingual sensorineural hearing loss to improve their abilities in identifying question/statement prosody. However, this training had no benefit for emotion perception, which was attributed to the ease of the applied test and ceiling effects. In addition to music training, further evidence was later obtained regarding the efficacy of suprasegmental training on children with hearing loss in improving vocal emotional prosody perception (Yeshoda et al., 2020). Such a finding was also reported for HA users by comparing the pre- and post-training vocal emotional prosody perception scores (Konadath et al., 2021). Following short term speech stress training, higher scores on speech stress perception were achieved in both children (4.1–9.2 years) with hearing impairment using HAs and children (3.2–6.1 years) with normal hearing. However, HA users showed lower performance (mean score of 7.5/10 for stressed and 6.1/10 unstressed condition) compared to their normal-hearing peer group (mean score of 8.5/10 for stressed and 7.2/10 unstressed condition).

4. How hearing assistive devices affect prosodic cues

To identify pertinent studies, a search was performed on peer-reviewed articles in English that were published between 2000 and May 2022 in the electronic databases PubMed, Scopus, and Google Scholar using the following search terms: “hearing aid”, or “cochlear implant”, or “bimodal fitting” or “bimodal simulation” in combination with “linguistic prosody”, “affective prosody”, “prosody perception”, “emotion perception”, “emotion identification”. This search identified 79 articles. The title and abstract of each article study were first screened by the first author and then inclusion was validated by the second and third authors. To be included, the studies had to fulfill the following criteria: (a) enrolling participants who used HA, CI, or bimodal fitting in either children or adults, (b) applying behavioral or electrophysiological methods to assess the efficacy of hearing prostheses in vocal prosody perception, (c) comparing the prosodic abilities of HA users, CI recipients, or bimodal hearing users with normal-hearing listeners, or before and after using hearing prosthesis, and (d) assessing either detrimental or beneficial effects of hearing assistive devices on prosody perception. Articles were excluded if they had investigated emotional perception of music or were focused on tonal languages or non-speech sounds. A total number of 22 articles were identified to be eligible for full text screening, 10 of which investigated prosody perception in HA users, 8 of which assessed the efficacy of CIs in prosody perception, and the rest were about the recipients of bimodal stimulation. These studies are summarized in Table S2 in Appendix K.

4.1. Hearing aids and prosody

While the benefits of HAs on audibility and speech recognition have been amply studied, only a few studies have investigated the effect of HAs on speech prosody recognition (Cannon & Chatterjee, 2019; Chatterjee et al., 2019; Goy et al., 2016). Some evidence demonstrates that increasing audibility in HA technology is associated with altering acoustic cues used in prosody perception. Goy et al. (2016) investigated whether HAs can restore the audibility of suprasegmental cues, and how well HA users can benefit emotional speech and word recognition. Using sentences from the Toronto Emotional Speech Set, they observed that older adults with hearing loss showed improved performance in both word recognition and emotion identification when they use HAs. However, the benefit of HAs was more remarkable for word recognition (38% unaided, 65% aided) than emotion identification (36% unaided, 42% aided), which was attributed to the greater efficacy of HAs in restoring the audibility of speech cues than in transmitting the speech cues necessary for emotion perception. Likewise, Singh et al. (2019) indicated that hearing-impaired older adults with HAs had similar performance to age-matched hearing-impaired peers without HAs on emotion perception based on the result of The Emotional Communication in Hearing Questionnaire (EMO-CHeQ).

HAs can make emotions less distinguishable depending on how and to what extent the amplification methods (linear vs. amplitude compression, dynamic range and speed of compression, single channel vs. multi-channel) change intensity variations and speech spectrum (Goy et al., 2018). As such, the enhanced gain for low-level sounds compared to high-level ones distorts the amplitude envelope of speech by making it less variable. Also, more amplification being applied in the high frequencies in prescription formulas changes the spectral features of speech (Goy et al., 2017), which can presumably impact the perception of intensity variations. Arefi et al. (2017) also described the effect of HA amplitude compression as in the aforementioned study, which indicated reduced recognition of some emotions (fear, sad, angry, and happy) following simulated fast-acting compression, slow-acting compression, and linear amplification. Neither study examined the changes of duration and F0 contours that may result from other speech enhancement algorithms in HAs.

Christensen et al. (2019) also found that HAs could not efficiently compensate for affective prosody perception deficit in English speaking older adults with hearing loss relative to age-matched peers with normal-hearing, suggesting additional contribution of hearing impairment in prosody recognition deficit, separate from age-related prosody perception difficulties. Within the hearing-impaired group, the effect of age and talker gender were significant. This finding was associated with reduced emotional identification accuracy (overall percent correct difference ≃ 13% for male talker and 6% for female talker), increased reaction time (e.g., sad emotion and male talker: minimum reaction time in hearing-impaired group ≃ 0.7 s vs. normal-hearing group ≃ 0.3 s), and lower sensitivity on voice emotion recognition task for adults who had been fitted with bilateral HAs (e.g., d' value for angry emotion and male talker: hearing-impaired group ≃ 2.9 vs. normal-hearing group ≃ 3.7).

Legris et al. (2021) found similar results; emotional identification scores of participants with presbycusis in the Montreal Affective Voice test improved but remained low one month after using HAs compared to age-matched normal peers (mean of 59% before rehabilitation and 69% after rehabilitation vs. 88% for normal-hearing individuals), particularly at high speech levels, where compression was activated to decrease gain and to expand the dynamic range. However, the

Profile of Emotional Competence (PEC) questionnaire revealed improvement in the ability of presbycusis participants to identify their own and others' emotions after HA fitting. Although such an improvement in emotion perception before and after using HAs was not statistically significant, it could nonetheless be a sign of increased quality of life following HA use.

Interestingly, different findings have been found by Cannon & Chatterjee (2019). They noted that the ability of voice emotion recognition by children with mild to moderate hearing loss (8–14 years), some of whom were fitted with HAs, was similar to their normally hearing counterparts (6–17 years) using either adult-directed (less exaggerated prosody) or child-directed (more exaggerated prosody) materials. In both groups, all participants responded better to sentences with more exaggerated prosodic cues, but with different patterns of interactions. The emotion recognition of children with hearing loss showed stronger interaction with developmental age, vocabulary, and materials, while in normal-hearing children, linguistic ability played a prominent role in emotion recognition. Prior to this observation, Most & Michaelis (2012) had indicated that young children (4.0–6.6 years) with HAs do not significantly differ from normal-hearing children in emotion perception under three different conditions: auditory, visual, and auditory-visual modes. However, using within-group comparisons, children with HAs were found to perform better when given both visual and auditory cues (average corrected score of 85.5%) compared to only visual cues (average corrected score of 72.0%), suggesting that additional auditory information could be provided by HA in combined auditory-visual mode. Schmidt et al. (2016) had reported that HAs changed the pattern of acoustic parameters to restore affective prosody so that the emotion perception ability of adult HA users reached the level of age-matched normal-hearing listeners. In their approach, HA users were asked to rate the degree of arousal and valence for some short audio-only utterances. Also, by comparing the affect ratings of HA users with acoustic parameters cueing different emotions, they observed that the HA users were more sensitive to the intensity differences. Laugen et al. (2017) found that not only could preschool children wearing HAs (4–5 years) perform equivalently to typically hearing children on emotion perception tasks, but also that the parents of these children could estimate the child's emotion understanding accurately. Parents' estimation accuracy can reveal whether the parents contribute properly to their children's emotional and cognitive development, which helps in detecting families who may need support in this matter.

4.2. Cochlear implants and prosody

CIs can partially restore hearing in severe to profound hearing loss by stimulating the auditory nerve directly via electrodes (Dhanasingh & Jolly, 2017). However, the limited number of electrodes, limited insertion depth, the position of the electrodes, the spread of neural excitation in electrical hearing, and limitations in processing strategies limit spectral and temporal resolution, especially at low frequencies, which can compromise pitch perception (Canfarotta et al., 2020; Dhanasingh & Jolly, 2017; Fu et al., 2017; Macherey et al., 2011). Given that pitch variations convey essential cues for prosody perception, the tonotopic mismatch between the acoustic input and the cochlear place of stimulation can restrict the user's perception of prosody. Similarly, though to a lesser extent, adult CI users have reduced ability in performing tasks that heavily rely on intensity and duration (Peng et al., 2009). In this regard, the negative effect of electric stimulation was demonstrated by the worse performance of adult CI recipients in perceiving alterations in F0 contour (Meister et al., 2009). Compared to normal listeners, many adult CI users were worse in a sentence stress perception test, especially when varying F0 occurred over a shorter duration, which might limit accurate perception of acoustic cues. In contrast, there were only relatively slight differences between normal and CI participants in perceiving artificial modifications of the F0 contour in statement vs. question test and gender difference test.

More recently, the same research team emphasized the poor performance of adults using CIs in perceiving sentence-based stress by measuring smallest detectable differences or difference limens (i.e., DLs) based on F0, intensity, and duration (Meister et al., 2011). They found that, compared with a control group of normal-hearing listeners, CI recipients had higher DLs for F0 (71.2 Hz vs. 15.7 Hz) and intensity (3.9 dB vs. 1.8 dB) but relatively similar DLs for duration (51.0 ms vs. 40.5 ms). The authors concluded that most CI users have problems discriminating variation in the spectral and intensity domains but less so in the temporal domain. The study by Hegarty & Faulkner (2013) also supported that CI children (7.4–14.6 years) benefited from amplitude and duration cues in the absence of pitch cues, which was reflected in better perception of stress when children were presented with natural speech stimuli compared to sentences with pitch manipulations (approximate mean of correct scores for natural speech of 53 % vs. approximate mean of correct scores for manipulated speech of 47 %). Van Zyl & Hanekom (2013) indicated that even in noise, adult CI users performed better in a certain/hesitant distinction task where durational cues were prominent than in a question/statement distinction task where F0 and intensity cues were prominent. The study by Kalathottukaren et al. (2015), which examined prosody perception in adults using CIs with the Profiling Elements of Prosody in Speech-Communication (PEPS-C) test and Montreal Battery of Evaluation of Amusia (MBEA), also supported the view that CI users were less accurate in the perception of pitch contour, emotion, turn-end, and stress, suggesting poor acoustic transmission of pitch and intensity cues. However, the performance of CI users in the chunking reception subset that relies more on temporal cues was near normal range. An earlier study by Luo et al. (2007) had mentioned that although pitch cues are more salient for vocal emotion identification in adult listeners with CIs, they can perceive less than half of the emotions by using preserved amplitude and durational cues in speech signal. Accordingly, removing the amplitude variations was associated with worsen performance in CI users. This suggests that all spectral, temporal and amplitude cues are required for vocal emotion perception in listeners with CIs.

Agrawal et al. (2012) provided objective evidence of deficit in CI-simulated prosody recognition by recording event-related potentials (ERPs) alongside behavioral measures. In this study, two CI speech-coding strategies, namely the Advanced Combination Encoder (ACE) and Psychoacoustic ACE (PACE), were simulated using vocoded sentences and presented to normal-hearing adults in order to assess the impact of electric hearing on prosody recognition. In line with the behavioral results, vocoded stimuli made recognizing prosody more difficult as indicated by prolonged ERP latencies and reduced amplitudes, especially with the ACE strategy. Indeed, larger N100 and P200 amplitudes with the PACE strategy suggested an advantage over ACE for prosody perception. This finding was attributed to the psychoacoustic-masking model used in the PACE strategy that could simulate masking effect present in a normal-hearing cochlea and provide more natural frequency bands in which the finer acoustic features were preserved and delivered to the implant. Interestingly, among neutral, angry and happy emotions, happiness led to stronger P200 response. The authors attributed this finding to its larger pitch variation, social salience, and its association with widespread brain activity.

In a follow-up investigation, the same authors presented the original sentences to adult CI users instead of simulating CI function with vocoded stimuli (Agrawal et al., 2013). In addition to the N100-P200 complex, time-frequency (TF) analysis was used to reveal non-phased-locked (induced) neural activity that elucidates differences between ACE and PACE coding strategies on vocal emotion recognition. Although all participants showed high accuracy and faster recognition of happy than angry and neutral prosodic sentences, electrophysiological measurements suggested poor emotional voice recognition in CI users compared to normal-hearing subjects. This result was manifested in longer N100 latency and reduced N100 and P200 peak amplitudes in CI users, which reflects the limited speech cues provided by CI devices. Similar to the previous study, a larger amplitude of P200 for happy prosody was found for PACE, confirming the higher efficacy of this coding-strategy in transferring prosodic cues. Based on TF analysis, enhanced activity in the gamma-range lent strong support to the PACE strategy causing higher neural firing rate compared to ACE. The notable aspect of these two studies was that the P200 peak can be considered as an indicator of emotional prosody recognition. Similar perceptual findings have also indicated that electrode configurations and coding strategies can affect intonation perception (Landwehr et al., 2014).

Interestingly, opposite results have been reported by Chatterjee et al. (2015), indicating a remarkable benefit obtained by CIs. Using a speech emotion identification task, both adult (27.5–69.8 years) and school-age (6.0–19.8 years) CI users performed similarly to normal-hearing adults listening to a CI simulation of the same stimuli. The authors suggested that a sufficiently developed cognitive system in adult CI users or learned adaptive strategies in cochlear-implanted children have enabled them to extract emotion from degraded speech. This suggestion is also supported by the finding that normal-hearing children showed significantly poorer performance in response to CI-simulated speech, given their underdeveloped cognitive system compared to adults and not benefiting from adaptation to impoverished speech compared to children using CIs.

4.3. Bimodal stimulation and prosody

In cases where the CI recipients have residual hearing in the non-implanted ear, combined electric stimulation and contralateral HA is often suggested, which is referred to as bimodal hearing (Gantz & Turner, 2004). While using a HA in the opposite ear was found to increase speech recognition, music perception and sound localization over the CI-alone condition (Kong et al., 2004; Turner et al., 2004; Tyler et al., 2002), the benefits of bimodal fitting has not been extensively studied for prosody perception. Using artificially-manipulated F0, Straatman et al. (2010) demonstrated that the percentage correct for distinguishing between questions and affirmations was improved from 56% to 62% when the HA was added in the non-implanted ear of implanted children (6.0–19.8 years). Better performance was also found for the discrimination of pitch movements in the bisyllabic nonsense-word “baba”. However, even in bimodal fitting, children needed exaggerated F0 excursions for doing both tasks. This reinforced the idea that the auditory system can combine the low frequency information from the HA with the high frequency input from the CI. Another study by Most et al (2011) also found that adults using bimodal stimulation scored better compared to adult CI users in perception of intonation (75.0% vs. 55.0%), syllable stress (64.4% vs. 54.6%), and word emphasis (71.9% vs. 63.6%).

Considering that the speech processing strategies in CIs use only the envelope information, not the temporal fine structure, Cullington & Zeng (2011) also compared the affective prosody discrimination performance of bilateral CI users with bimodal users in adults. However, results on the Aprosodia Battery, which is a standardized test for prosody identification, showed no statistically significant difference between bimodal (mean score over all subsets of 67.9 %) and bilateral CI users (61.6 %). Furthermore, three other pitch-related tasks, including Hearing in Noise Test (HINT), Montreal Battery of Evaluation of Amusia, and talker identification were performed in which both groups scored at an approximately equivalent level. Although this study had hypothesized better performance of bimodal fitting on tasks requiring good pitch perception, the obtained results and the lack of correlation between test results led the authors to point out that these tests rely on more diverse cues like stress and timing, which can be adequately extracted from the use of both bimodal fittings and bilateral CIs. Comparing bimodal stimulation with CI alone condition, Hegarty & Faulkner (2013) also showed that there is no bimodal advantage for stress and intonation perception in children (7.4–14.6 years). These children in both the CI and bimodal conditions performed better in response to natural sentences than pitch-manipulated ones, suggesting that duration and amplitude cues may play an important role in perception of prosodic pattern when there is limited pitch information.

5. Gaps in knowledge

Prosody perception in people with hearing loss is a relatively neglected topic. Some investigations have focused on emotional perception in hearing loss and interacting factors such as age and cognitive resources. Such attempts have focused on affective prosody and barely touched on the role of linguistic prosody in speech recognition and efficient communication. Moreover, limited information is available on the extent to which hearing prostheses can restore acoustic-based prosodic cues. So far, only the effect of amplitude compression on affective prosody has been studied to some extent but other speech enhancement technologies in modern HAs have not been investigated at all. In the case of CIs, little is known about the role of different speech coding strategies and electrode array design in restoring prosody perception.

Several studies have provided converging evidence that challenging listening situations, including background noise and reverberation, compromise speech intelligibility by masking the speech signal, especially for hearing-impaired listeners who have temporal and spectral processing deficits (Al Osman et al., 2017; George et al., 2010; Humes, 1991; Xia et al., 2018). Hearing-impaired individuals not only fail to perceive some speech components due to increased susceptibility to masking by ambient sounds, but also face difficulties in understanding audible parts due to severe distortions resulting from hearing loss (Levitt, 2001). Although HAs allow for signal processing strategies to decrease the effect of noise in challenging listening situations while controlling the risk of over-amplification (Brons et al., 2015; Leroux et al., 2018; Nabelek et al., 2004; Shi & Doherty, 2008), the performance of hearing assistive devices in restoring F0 contour, level contour and temporal cues has not been extensively studied in those environments. Quantifying the masking effect of different noises, minimal acceptable signal to noise ratio, and reverberation time in the field of prosody perception can provide supplementary information, potentially paving the path towards more comprehensive rehabilitation plans for people with hearing loss. On the other hand, since prosodic cues can also be conveyed through facial and body movements, the complete response of HA users and CI recipients to prosodic information can only be determined when the visual modality is added to the auditory mode, a topic that has not been adequately addressed.

In addition, as discussed earlier, our understanding of the brain networks involved in prosody perception is incomplete. While there is limited understanding of the cortical networks involved, there has been little or no study on subcortical processing of prosody in normal or hearing-impaired individuals or how hearing prostheses affect this processing. Several auditory evoked potential methodologies may be used to fill this gap to probe an intermediate step in the auditory chain between the acoustic signal and higher brain centers. Some techniques, like the frequency following response, can assess the electrophysiological response to speech stimuli with a high temporal resolution. Although several studies have shown that the frequency following response can extract F0 contour successfully in normal-hearing individuals (Aiken & Picton, 2006; Dajani et al., 2005; Krishnan et al., 2004), this response has not been used in the context of prosodic processing, as most of the research on prosody has been conducted via a variety of behavioral approaches. Filling such gaps could redirect the recent advances in understanding prosody perception of hearing-impaired population towards improving speech recognition and social/cognitive performance in this population.

6. Concluding remarks and future directions

Prosody perception is key to effective communication through which both linguistic and emotional information is conveyed. Since different prosodic patterns are defined by their acoustic profiles, hearing abilities are essential for perceiving prosody in spoken language. Thus, more attention should be given to potential prosody perception deficit in individuals with hearing loss. Based on existing studies, reduced audibility and decreased spectro-temporal resolution make hearing-impaired individuals very vulnerable to a loss of prosodic information. Moreover, the unique characteristics of individuals with hearing loss such as chronological age, cognitive status, degree and configuration of hearing loss, and the presence or absence of clinical intervention can also affect the extent of prosody perception deficit and its consequences.

Current HAs, CI technologies, and bimodal stimulation do not fully restore prosodic cues due to their limitation in extraction, processing, and transmission of acoustic variations. As evidenced by the studies reviewed, hearing-impaired individuals have reduced access to the pitch information, which leads users to rely more on temporal and amplitude cues to perceive prosodic patterns. However, some studies suggested that the performance of individuals with hearing prostheses in perceiving speech prosody is heavily dependent on pitch cues and cannot be modified through using alternative strategies. This may in part be due to prosody processing being an afterthought in the design of hearing assistive devices. With the current rapid advances in technology and speech processing, future investigations should test the benefits of new designs and strategies on the perception of linguistic and affective prosody, thereby helping to reduce communication problems and consequent poor quality of life.

Finally, while the subject of prosody is of high relevance to speech and language pathologists, other specialists, like audiologists, should keep abreast of recent advances in this area. Treating prosody as a multifaceted concept and assessing prosodic deficits in different clinical populations may not only open various opportunities for research but also lead to improved clinical services and innovations in diagnostic and rehabilitation tools.

7. Limitations

We reviewed different aspects of prosody perception in hearing-impaired individuals and users of hearing assistive devices, but some limitations prevented us from covering this subject more comprehensively. First, the prosodic features in speech can have different interpretation or social meaning in different cultures, but such cultural differences were not considered in this review. In addition, current studies have largely or exclusively been conducted under controlled laboratory conditions which can differ markedly from everyday situations. Moreover, we limited our review to studies carried out in spoken language; however, prosodic processing in signed language was not reviewed, but may be covered in future analysis of prosody perception studies in hearing-impaired individuals.

References

Agrawal, D., Thorne, J., Viola, F., Timm, L., Debener, S., Büchner, A., Dengler, R., & Wittfoth, M. (2013). Electrophysiological responses to emotional prosody perception in cochlear implant users. *NeuroImage: Clinical*, *2*, 229–238.

Agrawal, D., Timm, L., Viola, F. C., Debener, S., Büchner, A., Dengler, R., & Wittfoth, M. (2012). ERP evidence for the recognition of emotional prosody through simulated cochlear implant strategies. *BMC Neuroscience*, *13*(1), 1–10.

Aiken, S. J., & Picton, T. W. (2006). Envelope following responses to natural vowels. *Audiology and Neurotology*, *11*(4), 213–232.

Al Osman, R., Dajani, H. R., & Giguére, C. (2017). Self-masking and overlap-masking from reverberation using the speech-evoked auditory brainstem response. *The Journal of the Acoustical Society of America*, *142*(6), EL555–EL560.

Anolli, L., Ciceri, R., & Infantino, M. G. (2000). Irony as a game of implicitness: Acoustic profiles of ironic communication. *Journal of Psycholinguistic Research*, *29*(3), 275–311.

Arefi, H. N., Sameni, S. J., Jalilvand, H., & Kamali, M. (2017). Effect of hearing aid amplitude compression on emotional speech recognition. *Auditory and Vestibular Research*, *26*(4), 223–230.

Aziz-Zadeh, L., Sheng, T., & Gheytanchi, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PloS One*, *5*(1), e8759.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614.

Barth-Weingarten, D., Reber, E., & Selting, M. (2010). *Prosody in interaction* (Vol. 23). John Benjamins Publishing.

Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *9*(9), 1395–1403.

Binns, C., & Culling, J. F. (2007). The role of fundamental frequency contours in the perception of speech against interfering speech. *The Journal of the Acoustical Society of America*, *122*(3), 1765–1776.

Bisgaard, N., & Ruf, S. (2017). Findings from EuroTrak surveys from 2009 to 2015: Hearing loss prevalence, hearing aid adoption, and benefits of hearing aid use. *American Journal of Audiology*, *26*(3S), 451–461.

Bolinger, D., & Bolinger, D. L. M. (1986). *Intonation and its parts: Melody in spoken English*. Stanford University Press.

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *27*(1), 134–144.

Bradley, M., Safi, F., & Lang, P. (2000). *Startle modality and size: Is a big startler always a big startler? 37*, S29–S29.

Brons, I., Houben, R., & Dreschler, W. A. (2015). Acoustical and perceptual comparison of noise reduction and compression in hearing aids. *Journal of Speech, Language, and Hearing Research*, *58*(4), 1363–1376.

Brown, G. (1995). *Speakers, listeners and communication: Explorations in discourse analysis*. Cambridge University Press.

Busso, C., Lee, S., & Narayanan, S. (2009). Analysis of emotionally salient aspects of fundamental frequency for emotion detection. *IEEE Transactions on Audio, Speech, and Language Processing*, *17*(4), 582–596.

Calet, N., Gutiérrez-Palma, N., Defior, S., & Jiménez-Fernández, G. (2019). Linguistic and non-linguistic prosodic skills in Spanish children with developmental dyslexia. *Research in Developmental Disabilities*, *90*, 92–100.

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37.

Calvo, R. A., & Mac Kim, S. (2013). Emotions in text: Dimensional and categorical models. *Computational Intelligence*, *29*(3), 527–543.

Canfarotta, M. W., Dillon, M. T., Buss, E., Pillsbury, H. C., Brown, K. D., & O'Connell, B. P. (2020). Frequency-to-place mismatch: Characterizing variability and the influence on speech perception outcomes in cochlear implant recipients. *Ear and Hearing*, *41*(5), 1349–1361.

Cannon, S. A., & Chatterjee, M. (2019). Voice Emotion Recognition by Children with Mild to Moderate Hearing Loss. *Ear and Hearing*, *40*(3), 477.

Ceravolo, L., Frühholz, S., Pierce, J., Grandjean, D., & Péron, J. (2021). Basal ganglia and cerebellum contributions to vocal emotion processing as revealed by high-resolution fMRI. *Scientific Reports*, *11*(1), 1–15.

Chakravorti, S., Noble, J. H., Gifford, R. H., Dawant, B. M., O'Connell, B., Wang, J., & Labadie, R. F. (2019). Further evidence of the relationship between cochlear implant electrode positioning and hearing outcomes. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *40*(5), 617.

Chatterjee, M., Kulkarni, A. M., Siddiqui, R. M., Christensen, J. A., Hozan, M., Sis, J. L., & Damm, S. A. (2019). Acoustics of emotional prosody produced by prelingually deaf children with cochlear implants. *Frontiers in Psychology*, *10*, 2190.

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., Kulkarni, A. M., & Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing Research*, *322*, 151–162. https://doi.org/10.1016/j.heares.2014.10.003

Chen, A., & Boves, L. (2018). What's in a word: Sounding sarcastic in British English. *Journal of the International Phonetic Association*, *48*(1), 57–76.

Chen, J., Yang, H., Wu, X., & Moore, B. C. (2018). The effect of F0 contour on the intelligibility of speech in the presence of interfering sounds for Mandarin Chinese. *The Journal of the Acoustical Society of America*, *143*(2), 864–877.

Christensen, J. A., Sis, J., Kulkarni, A. M., & Chatterjee, M. (2019). Effects of age and hearing loss on the recognition of emotions in speech. *Ear and Hearing*, *40*(5), 1069.

Colston, H. L. (1997). " I've Never Seen Anything Like It": Overstatement, Understatement, and Irony. *Metaphor and Symbol*, *12*(1), 43–58.

Cruttenden, A. (1997). *Intonation*. Cambridge University Press.

Cullington, H. E., & Zeng, F.-G. (2011). Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination and talker identification. *Ear and Hearing*, *32*(1), 16.

Dajani, H. R., Purcell, D., Wong, W., Kunov, H., & Picton, T. W. (2005). Recording human evoked potentials that follow the pitch contour of a natural vowel. *IEEE Transactions on Biomedical Engineering*, *52*(9), 1614–1618.

de Boer, M. J., Jürgens, T., Cornelissen, F. W., & Başkent, D. (2021). Degraded visual and auditory input individually impair audiovisual emotion recognition from speech-like stimuli, but no evidence for an exacerbated effect from combined degradation. *Vision Research*, *180*, 51–62.

Delgutte, B. (1997). Auditory neural processing of speech. *The Handbook of Phonetic Sciences*, 507–538.

Demenescu, L. R., Kato, Y., & Mathiak, K. (2015). Neural processing of emotional prosody across the adult lifespan. *BioMed Research International*, *2015*.

Dhanasingh, A., & Jolly, C. (2017). An overview of cochlear implant electrode array designs. *Hearing Research*, *356*, 93–103.

Dumouchel, P., Dehak, N., Attabi, Y., Dehak, R., & Boufaden, N. (2009). *Cepstral and long-term features for emotion recognition*. Tenth Annual Conference of the International Speech Communication Association.

Dupuis, K., & Pichora-Fuller, M. K. (2014). Intelligibility of emotional speech in younger and older adults. *Ear and Hearing*, *35*(6), 695–707.

Dupuis, K., & Pichora-Fuller, M. K. (2015). Aging affects identification of vocal emotions in semantically neutral sentences. *Journal of Speech, Language, and Hearing Research*, *58*(3), 1061–1076.

Dyck, M. J., Farrugia, C., Shochet, I. M., & Holmes-Brown, M. (2004). Emotion recognition/understanding ability in hearing or vision-impaired children: Do sounds, sights, or words make the difference? *Journal of Child Psychology and Psychiatry*, *45*(4), 789–800.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

Ferre, J. M. (2015). Auditory dysfunction beyond the 8th nerve: Understanding central auditory processing disorders. *Perspectives on Hearing and Hearing Disorders: Research and Diagnostics*, *19*(1), 4–11.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, *39*(3), 399–423.

Fischer, S. E. (2018). *Speech Language Pathologists and Prosody: Knowledge and Clinical Practices*. University of Mississippi.

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480.

Fu, Q.-J., Galvin, J. J., & Wang, X. (2017). Integration of acoustic and electric hearing is better in the same ear than across ears. *Scientific Reports*, *7*(1), 1–9.

Gandour, J., Tong, Y., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y., Li, X., & Lowe, M. (2007). Neural basis of first and second language processing of sentence-level linguistic prosody. *Human Brain Mapping*, *28*(2), 94–108.

Gantz, B. J., & Turner, C. (2004). Combining acoustic and electrical speech processing: Iowa/Nucleus hybrid implant. *Acta Oto-Laryngologica*, *124*(4), 344–347.

Gay, T. (1978). Physiological and acoustic correlates of perceived stress. *Language and Speech*, *21*(4), 347–353.

George, E. L., Goverts, S. T., Festen, J. M., & Houtgast, T. (2010). *Measuring the effects of reverberation and noise on sentence intelligibility for hearing-impaired listeners*.
Giroud, N., Keller, M., Hirsiger, S., Dellwo, V., & Meyer, M. (2019). Bridging the brain structure—Brain function gap in prosodic speech processing in older adults. *Neurobiology of Aging*, *80*, 116–126.
Gomez, R., & Ferguson, M. (2020). Improving self-efficacy for hearing aid self-management: The early delivery of a multimedia-based education programme in first-time hearing aid users. *International Journal of Audiology*, *59*(4), 272–281.
Good, A., Gordon, K. A., Papsin, B. C., Nespoli, G., Hopyan, T., Peretz, I., & Russo, F. A. (2017). Benefits of music training for perception of emotional speech prosody in deaf children with cochlear implants. *Ear and Hearing*, *38*(4), 455.
Goy, H., Pichora-Fuller, M. K., Singh, G., & Russo, F. A. (2016). Perception of emotional speech by listeners with hearing aids. *Canadian Acoustics*, *44*(3).
Goy, H., Pichora-Fuller, M. K., Singh, G., & Russo, F. A. (2017). Acoustic analysis of emotional speech processed by hearing aids. *Canadian Acoustics*, *45*(3), 178–179.
Goy, H., Pichora-Fuller, M. K., Singh, G., & Russo, F. A. (2018). Hearing aids benefit recognition of words in emotional speech but not emotion identification. *Trends in Hearing*, *22*, 2331216518801736.
Grandjean, D. (2021). Brain networks of emotional prosody processing. *Emotion Review*, *13*(1), 34–43.
Grant, K. W. (1987). Identification of intonation contours by normally hearing and profoundly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *82*(4), 1172–1178.
Gürses, E., Türkyılmaz, M. D., Kalaycıoğlu, C., Karabulut, E., Bajin, M. D., Sennaroğlu, L., & Genç, G. A. (2020). Evaluation of temporal and suprasegmental auditory processing in patients with unilateral hearing loss. *Auris Nasus Larynx*, *47*(5), 785–792.
Hegarty, L., & Faulkner, A. (2013). The perception of stress and intonation in children with a cochlear implant and a hearing aid. *Cochlear Implants International*, *14*(sup4), 35–39. https://doi.org/10.1179/1467010013Z.000000000132
Hesling, I., Clément, S., Bordessoules, M., & Allard, M. (2005). Cerebral mechanisms of prosodic integration: Evidence from connected speech. *Neuroimage*, *24*(4), 937–947.
Hirschberg, J. (2002). Communication and prosody: Functional aspects of prosody. *Speech Communication*, *36*(1–2), 31–43.
Hjalmarsson, A. (2011). The additive effect of turn-taking cues in human and synthetic voice. *Speech Communication*, *53*(1), 23–35.
Humes, L. E. (1991). Understanding the speech-understanding problems of the hearing impaired. *Journal of the American Academy of Audiology*, *2*(2), 59–69.
Izard, C. (1977). *Human emotions New York Plenum*.
Jiam, N., Caldwell, M., Deroche, M., Chatterjee, M., & Limb, C. (2017). Voice emotion perception and production in cochlear implant users. *Hearing Research*, *352*, 30–39.
Kalathottukaren, R. T., Purdy, S. C., & Ballard, E. (2015). Prosody perception and musical pitch discrimination in adults using cochlear implants. *International Journal of Audiology*, *54*(7), 444–452.
Kalathottukaren, R. T., Purdy, S. C., & Ballard, E. (2017). Prosody perception and production in children with hearing loss and age-and gender-matched controls. *Journal of the American Academy of Audiology*, *28*(4), 283–294.

Katz, J., Chasin, M., English, K. M., Hood, L. J., & Tillery, K. L. (2015). *Handbook of clinical audiology* (Vol. 7). Wolters Kluwer Health Philadelphia, PA.
Klieve, S., & Jeanes, R. (2001). Perception of prosodic features by children with cochlear implants: Is it sufficient for understanding meaning differences in language? *Deafness & Education International*, *3*(1), 15–37.
Konadath, S., Raveendran, R., & Yeshoda, K. (2021). Perception of speech stress in children with hearing impairment. *International Journal of Pediatric Otorhinolaryngology*, *140*, 110495.
Kong, Y.-Y., Cruz, R., Jones, J. A., & Zeng, F.-G. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear and Hearing*, *25*(2), 173–185.
Kotz, S. A., Meyer, M., Alter, K., Besson, M., von Cramon, D. Y., & Friederici, A. D. (2003). On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language*, *86*(3), 366–376.
Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequency-following response: Representation of pitch contours in Chinese tones. *Hearing Research*, *189*(1–2), 1–12.
LaCroix, A. N., Blumenstein, N., Tully, M., Baxter, L. C., & Rogalsky, C. (2020). Effects of prosody on the cognitive and neural resources supporting sentence comprehension: A behavioral and lesion-symptom mapping study. *Brain and Language*, *203*, 104756.
Lai, Z., Hughes, S., & Shapiro, E. (1991). Manual for the Minnesota Tests of Affective Processing (MNTAP). *University of Minnesota, MN*.
Landwehr, M., Fürstenberg, D., Walger, M., von Wedel, H., & Meister, H. (2014). Effects of various electrode configurations on music perception, intonation and speaker gender identification. *Cochlear Implants International*, *15*(1), 27–35.
Laugen, N. J., Jacobsen, K. H., Rieffe, C., & Wichstrøm, L. (2017). Emotion Understanding in Preschool Children with Mild-to-Severe Hearing Loss. *The Journal of Deaf Studies and Deaf Education*, *22*(2), 155–163. https://doi.org/10.1093/deafed/enw069
Legris, E., Henriques, J., Aussedat, C., Aoustin, J.-M., Robier, M., & Bakhos, D. (2021). Emotional prosody perception in presbycusis patients after auditory rehabilitation. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, *138*(3), 163–168.
Leroux, T., Laroche, C., Giguère, C., & Voix, J. (2018). *Hearing Aid Use in Noisy Workplaces*.
Levitt, H. (2001). Noise reduction in hearing aids: A review. *Journal of Rehabilitation Research and Development*, *38*(1), 111–122.
Lieberman, P. (1965). On the acoustic basis of the perception of intonation by linguists. *Word*, *21*(1), 40–54.
Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of supplementary motor areas in auditory processing and auditory imagery. *Trends in Neurosciences*, *39*(8), 527–542.
Lo, C. Y., Looi, V., Thompson, W. F., & McMahon, C. M. (2020). Music training for children with sensorineural hearing loss improves speech-in-noise perception. *Journal of Speech, Language, and Hearing Research*, *63*(6), 1990–2015.
Lo, C. Y., McMahon, C. M., Looi, V., & Thompson, W. F. (2015). Melodic contour training and its effect on speech in noise, consonant discrimination, and prosody perception for cochlear implant recipients. *Behavioural Neurology*, *2015*.
Luo, X., Fu, Q.-J., & Galvin, J. J. (2007). Cochlear Implants Special Issue Article: Vocal Emotion Recognition by Normal-Hearing Listeners and Cochlear Implant Users. *Trends in Amplification*, *11*(4), 301–315. https://doi.org/10.1177/1084713807305301

Macherey, O., Deeks, J. M., & Carlyon, R. P. (2011). Extending the limits of place and temporal pitch perception in cochlear implant users. *Journal of the Association for Research in Otolaryngology*, *12*(2), 233–251.

Marshall, C. R., Harcourt-Brown, S., Ramus, F., & Van Der Lely, H. (2009). The link between prosody and language skills in children with specific language impairment (SLI) and/or dyslexia. *International Journal of Language & Communication Disorders*, *44*(4), 466–488.

Meister, H., Landwehr, M., Pyschny, V., Wagner, P., & Walger, M. (2011). The perception of sentence stress in cochlear implant recipients. *Ear and Hearing*, *32*(4), 459–467.

Meister, H., Landwehr, M., Pyschny, V., Walger, M., & Wedel, H. von. (2009). The perception of prosody and speaker gender in normal-hearing listeners and cochlear implant recipients. *International Journal of Audiology*, *48*(1), 38–48.

Mildner, V., & Koska, T. (2014). Recognition and production of emotions in children with cochlear implants. *Clinical Linguistics & Phonetics*, *28*(7–8), 543–554.

Miller, J. L. (1978). Interactions in processing segmental and suprasegmental features of speech. *Perception & Psychophysics*, *24*(2), 175–180.

Mitchell, R. L., & Kingston, R. A. (2014). Age-related decline in emotional prosody discrimination. *Experimental Psychology*.

Moore, J. A., & Teagle, H. F. (2002). *An introduction to cochlear implant technology, activation, and programming*.

Most, T., & Aviner, C. (2009). Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *Journal of Deaf Studies and Deaf Education*, *14*(4), 449–464.

Most, T., Harel, T., Shpak, T., & Luntz, M. (2011). *Perception of suprasegmental speech features via bimodal stimulation: Cochlear implant on one ear and hearing aid on the other*.

Most, T., & Michaelis, H. (2012). *Auditory, visual, and auditory–visual perceptions of emotions by young children with hearing loss versus children with normal hearing*.

Most, T., & Peled, M. (2007). Perception of suprasegmental features of speech by children with cochlear implants and children with hearing aids. *Journal of Deaf Studies and Deaf Education*, *12*(3), 350–361.

Mozziconacci, S. J. L. (1998). *Speech variability and emotion: Production and perception*.

Myers, B., Vaughan, C., Soman, U., Blain, S., Korsnack, K., Camarata, S., & Gordon, R. L. (2018). Prosody perception in children: Studying individual differences with novel brain and behavioral approaches. *BioRxiv*, 281998.

Nabelek, A. K., Tampas, J. W., & Burchfield, S. B. (2004). *Comparison of speech perception in background noise with acceptance of background noise in aided and unaided conditions*.

Nakatani, L. H., & Schaffer, J. A. (1978). Hearing’’words’’without words: Prosodic cues for word perception. *The Journal of the Acoustical Society of America*, *63*(1), 234–245.

Nayak, S., Gustavson, D. E., Wang, Y., Below, J. E., Gordon, R. L., & Magne, C. L. (2022). Test of Prosody via Syllable Emphasis (“TOPsy”): Psychometric Validation of a Brief Scalable Test of Lexical Stress Perception. *Frontiers in Neuroscience*, *16*.

Ng, S. H., & Bradac, J. J. (1993). *Power in language: Verbal communication and social influence*. Sage Publications, Inc.

Niu, Y., Chen, F., & Chen, J. (2019). The effect of F0 contour on the intelligibility of Mandarin Chinese for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *146*(2), EL85–EL91.

Nooteboom, S. (1997). The prosody of speech: Melody and rhythm. *The Handbook of Phonetic Sciences*, *5*, 640–673.

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning. *Cognitive Science*, *33*(1), 127–146.

Ohsuga, T., Horiuchi, Y., Nishida, M., & Ichikawa, A. (2006). Prediction of turn-taking from prosody in spontaneous dialogue. *Transactions of the Japanese Society for Artificial Intelligence*, *21*(1), 1–8.

Pannese, A., Grandjean, D., & Frühholz, S. (2015). Subcortical processing in auditory communication. *Hearing Research*, *328*, 67–77.

Panzeri, F., Cavicchiolo, S., Giustolisi, B., Di Berardino, F., Ajmone, P. F., Vizziello, P., Donnini, V., & Zanetti, D. (2021). Irony Comprehension in Children With Cochlear Implants: The Role of Language Competence, Theory of Mind, and Prosody Recognition. *Journal of Speech, Language, and Hearing Research*, *64*(8), 3212–3229.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, *308*, 98–108. https://doi.org/10.1016/j.heares.2013.08.011

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *35*(2), 205–220.

Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). How aging affects the recognition of emotional speech. *Brain and Language*, *104*(3), 262–269.

Pearson, N. (2009). Advanced clinical solutions for WAIS-IV and WMS-IV: Administration and scoring manual. *San Antonio: The Psychological Corporation*.

Peng, S.-C., Lu, N., & Chatterjee, M. (2009). Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. *Audiology and Neurotology*, *14*(5), 327–337.

Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clinical Linguistics & Phonetics*, *17*(4–5), 345–354.

Pichon, S., de Gelder, B., & Grèzes, J. (2009). Two different faces of threat. Comparing the neural systems for recognizing fear and anger in dynamic body expressions. *Neuroimage*, *47*(4), 1873–1883.

Pickett, J., & Morris, S. R. (2000). *The acoustics of speech communication: Fundamentals, speech perception theory, and technology*.

Picou, E. M., Singh, G., Goy, H., Russo, F., Hickson, L., Oxenham, A. J., Buono, G. H., Ricketts, T. A., & Launer, S. (2018). Hearing, emotion, amplification, research, and training workshop: Current understanding of hearing loss and emotion perception and priorities for future research. *Trends in Hearing*, *22*, 2331216518803215.

Raithel, V., & Hielscher-Fastabend, M. (2004). Emotional and linguistic perception of prosody. *Folia Phoniatrica et Logopaedica*, *56*(1), 7–13.

Raphael, L. J., Borden, G. J., & Harris, K. S. (2007). *Speech science primer: Physiology, acoustics, and perception of speech*. Lippincott Williams & Wilkins.

Rasmus, A., & Błachnio, A. (2021). Auditory Processing Disorders in Elderly Persons vs. Linguistic and Emotional Prosody. *International Journal of Environmental Research and Public Health*, *18*(12), 6427.

Rivers, S. E., Brackett, M. A., Reyes, M. R., Mayer, J. D., Caruso, D. R., & Salovey, P. (2012). Measuring emotional intelligence in early adolescence with the MSCEIT-YV: Psychometric properties and relationship with academic performance and psychosocial functioning. *Journal of Psychoeducational Assessment*, *30*(4), 344–366.

Ross, E. D., Thompson, R. D., & Yenkosky, J. (1997). Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain and Language*, *56*(1), 27–54.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, *9*(3), 235–248.

Schmidt, J., Herzog, D., Scharenborg, O., & Janse, E. (2016). Do hearing aids improve affect perception? In *Physiology, psychoacoustics and cognition in normal and impaired hearing* (pp. 47–55). Springer, Cham.

Schorr, E. A., Roth, F. P., & Fox, N. A. (2009). *Quality of life for children with cochlear implants: Perceived benefits and problems and the perception of single words and emotional sounds*.

Seydell-Greenwald, A., Chambers, C. E., Ferrara, K., & Newport, E. L. (2020). What you say versus how you say it: Comparing sentence comprehension and emotional prosody processing using fMRI. *NeuroImage*, *209*, 116509.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 193–247.

Sheppard, S. M., Keator, L. M., Breining, B. L., Wright, A. E., Saxena, S., Tippett, D. C., & Hillis, A. E. (2020). Right hemisphere ventral stream for emotional prosody identification: Evidence from acute stroke. *Neurology*, *94*(10), e1013–e1020.

Shi, L.-F., & Doherty, K. A. (2008). *Subjective and objective effects of fast and slow compression on the perception of reverberant speech in listeners with hearing loss*.

Singh, G., Liskovoi, L., Launer, S., & Russo, F. (2019). The emotional communication in hearing questionnaire (EMO-CHeQ): Development and evaluation. *Ear and Hearing*, *40*(2), 260.

Sobhy, O. A., Abdou, R. M., Ibrahim, S. M., & Hamouda, N. H. (2021). Effects of a prosody rehabilitation program on expression of affect in preschool children with hearing impairment: A randomized trial. *The Egyptian Journal of Otolaryngology*, *37*(1), 1–13.

Steffman, J. (2021). Prosodic prominence effects in the processing of spectral cues. *Language, Cognition and Neuroscience*, *36*(5), 586–611.

Straatman, L., Rietveld, A., Beijen, J., Mylanus, E., & Mens, L. (2010). Advantage of bimodal fitting in prosody perception for children using a cochlear implant and a hearing aid. *The Journal of the Acoustical Society of America*, *128*(4), 1884–1895.

Tantibundhit, C., Onsuwan, C., Klangpornkun, N., Phienphanich, P., Saimai, T., Saimai, N., Pitathawatchai, P., & Wutiwiwatchai, C. (2013). *Lexical tone perception in Thai normal-hearing adults and those using hearing aids: A case study*. 2262–2266.

Tinnemore, A. R., Zion, D. J., Kulkarni, A. M., & Chatterjee, M. (2018). Children's recognition of emotional prosody in spectrally-degraded speech is predicted by their age and cognitive status. *Ear and Hearing*, *39*(5), 874.

Tong, Y., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y., Li, X., & Lowe, M. (2005). Neural circuitry underlying sentence-level linguistic prosody. *NeuroImage*, *28*(2), 417–428.

Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., & Henry, B. A. (2004). Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing. *The Journal of the Acoustical Society of America*, *115*(4), 1729–1735.

Tyler, R. S., Parkinson, A. J., Wilson, B. S., Witt, S., Preece, J. P., & Noble, W. (2002). Patients utilizing a hearing aid and a cochlear implant: Speech perception and localization. *Ear and Hearing*, *23*(2), 98–105.

Vaissière, J. (1983). Language-independent prosodic features. In *Prosody: Models and measurements* (pp. 53–66). Springer.

van der Hulst, H. (2006). *Word stress*.

Van Zyl, M., & Hanekom, J. J. (2013). Perception of vowels and prosody by cochlear implant recipients in noise. *Journal of Communication Disorders*, *46*(5–6), 449–464.

Waaramaa, T., Kukkonen, T., Mykkänen, S., & Geneid, A. (2018). Vocal emotion identification by children using cochlear implants, relations to voice quality, and musical interests. *Journal of Speech, Language, and Hearing Research*, *61*(4), 973–985.

Watson, D. G., Jacobs, C. L., & Buxó-Lugo, A. (2020). Prosody indexes both competence and performance. *Wiley Interdisciplinary Reviews: Cognitive Science*, *11*(3), e1522.

Wells, B., & Peppé, S. (1996). Ending up in Ulster: Prosody and turn-taking in English dialects. *Studies in Interactional Sociolinguistics*, *12*, 101–130.

Weninger, F., Eyben, F., Schuller, B. W., Mortillaro, M., & Scherer, K. R. (2013). On the acoustics of emotion in audio: What speech, music, and sound have in common. *Frontiers in Psychology*, *4*, 292.

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage*, *24*(4), 1233–1241.

Witteman, J., van IJzendoorn, M. H., van de Velde, D., van Heuven, V. J., & Schiller, N. O. (2011). The nature of hemispheric specialization for linguistic and emotional prosodic perception: A meta-analysis of the lesion literature. *Neuropsychologia*, *49*(13), 3722–3738.

WysoCKa, M., & Mackiewicz, L. (2017). Music and intonation in the Perception of children using Hearing Aids. *Logopedia*, *46*, 105–119.

Xia, J., Xu, B., Pentony, S., Xu, J., & Swaminathan, J. (2018). Effects of reverberation and noise on speech intelligibility in normal-hearing and aided hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *143*(3), 1523–1533.

Yeshoda, K., Raveendran, R., & Konadath, S. (2020). Perception of vocal emotional prosody in children with hearing impairment. *International Journal of Pediatric Otorhinolaryngology*, *137*, 110252.

Zinchenko, A., Kanske, P., Obermeier, C., Schröger, E., Villringer, A., & Kotz, S. A. (2018). Modulation of cognitive and emotional control in age-related mild-to-moderate hearing loss. *Frontiers in Neurology*, *9*, 783.

Zuraiq, W., & Sereno, J. A. (2007). English lexical stress cues in native English and non-native Arabic speakers. *Ratio*, *1*(2), 829–832.

Chapter 3. Evaluating neural encoding of prosody in emotional speech using the speech FFR in normal-hearing adults

Evaluating neural encoding of prosody in emotional speech using the speech FFR in normal-hearing adults

Maryam Karimi-Boroujeni [1], Sajad Sadeghkhani [2], Christian Giguère [1], Saeid R. Seydnejad [2,3], Hilmi R. Dajani [2]

[1] School of Rehabilitation Sciences, University of Ottawa, Ottawa, ON, Canada

[2] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, Canada

[3] Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran.

Abstract

Emotional information in speech is conveyed through prosodic cues, including variations in fundamental frequency (F0 contour), intensity, and duration. Emotion perception has been largely studied behaviorally, but the neural mechanisms underlying it are not well understood. The Speech-evoked Frequency Following Response (FFR) is a non-invasive neural measure reflecting the encoding of speech acoustics in the auditory system. This study investigated the extent to which the FFR can represent prosody-related F0 contours and compared neural responses between male and female listeners. Sixteen normal-hearing adults underwent FFR recording in response to the word "balloon" spoken with sad and happy emotion by a male and a female talker. Using a pitch-tracking algorithm, F0 tracking accuracy was quantified via Root Mean Squared Error (RMSE) and 5% accuracy. The results showed that the FFR can track emotional F0 contours; however, the degree of accuracy is modulated by emotion type and talker voice characteristics. Enhanced F0 tracking was observed for the sad male voice, followed by the happy male, sad female, and happy female voices. On the other hand, listener sex did not influence the results. These findings provide new insights into the use of FFR as a neural measure for prosody assessment.

Keywords: Prosody processing; emotional speech; Frequency Following Response (FFR)

1. Introduction

The speech signal is the primary means of conveying information in interactive human communication [1,2]. Through the vocal apparatus, a sequence of basic linguistic units is produced in acoustical form; syllables build meaningful words, and then words are arranged into phrases to state more complex concepts [3,4]. In natural speech, these segments are not uttered at a fixed tone, loudness, and rate. The segments are threaded together with different melodies and rhythms, which constitute the suprasegmental features of speech [5,6]. Most speech perception studies neglect the suprasegmental information, even though it is an important contributor of information that complements the properties of segmental sounds [7].

Suprasegmental information, also called prosody, contributes to conveying the emotional state of the speaker through which the listener can understand what the speaker feels and thinks [8-10]. The emotional function of prosody is defined acoustically as variations in fundamental frequency (F0), intensity, and duration [11]. To perceive emotional patterns in speech, these acoustic-based prosodic cues are initially received and encoded by the peripheral auditory system [9,12]. Thereafter, the emotion-related cues are transmitted through the ascending auditory pathway, including subcortical and cortical structures, to undergo further processing [13,14]. Additionally, cognitive top-down mechanisms are engaged in a feedback information flow to extract and deduce contextual meaning and pick out pertinent information [12,15]. Despite the vast neural networks involved in prosody processing, most studies have focused on the behavioral aspects of emotion perception, and there is limited understanding of the subcortical and cortical processing in both normal-hearing individuals and those with impairment.

The speech-evoked Frequency Following Response (FFR) is a non-invasive neural measure reflecting the temporal and spectral encoding of speech in the central auditory system

[16,17]. The FFR, recorded via scalp electrodes, arises from synchronized neural activity within the auditory pathway [16,18]. Previous studies using tonal sweeps, synthetic and natural vowels, and isolated syllables have demonstrated that the phase-locked neural activity underlying the FFR is sufficiently dynamic to track small variations in speech periodicity over time [19-21]. However, whether the FFR can successfully capture prosodic features beyond segmental information remains unclear. Since the acoustic characteristics of speech stimuli produced by male and female talkers have been shown to affect FFRs [22], further investigation of talker effects on verbal prosody processing is also needed. In addition, although sex-related differences in the central encoding of speech stimuli have been reported [23-25], their implications for prosody processing are still unclear. In this regard, the FFRs obtained in a suprasegmental context may offer supplementary insight into sex differences in the neural encoding of speech prosody.

This study was designed to investigate the extent to which the FFR can track F0 contours in different emotions using a natural two-syllable word as stimulus. We hypothesized that the FFR can provide a potential diagnostic tool to test subcortical and cortical contributions to the neural encoding of speech prosody. As part of this objective, male- and female-spoken emotional speech stimuli were used to determine if and how the neural representation of F0 contours in young adult listeners with normal hearing can be affected by male and female voices. Additionally, this study aimed to examine whether neural encoding of acoustic features in prosodic speech differs between male and female listeners. Using electrophysiological data, this study contributes to a broader understanding of how efficiently a healthy auditory system can use prosodic features in natural speech processing.

2. Methods and Materials

2.1. Participants

A total of 16 adults (8 males, 8 females) between 18 and 31 years of age participated in this study. All participants had normal hearing defined as pure-tone audiometric thresholds ≤ 25 dB HL at octave frequencies between 250 and 8000 Hz bilaterally. The four-frequency pure-tone average ($PTA_{500\text{-}4000}$) ranged from -4 to 19 dB HL across individuals. Otoscopy confirmed no abnormalities in the external auditory canal and tympanic membrane in any participant. No participant had a history of neurological diseases, brain injuries, or psychiatric problems, as screened through a general development questionnaire (Appendix F). All participants were native or near-native English speakers who scored 9 or 10 on a 10-point scale in self-report questions from the Language Experience and Proficiency Questionnaire (LEAP-Q) [26] about their proficiency in speaking and understanding (Appendix D). Given that speakers of tonal languages can be more sensitive to F0 variations [27,28], only individuals without a tonal language background were recruited. Additionally, since previous studies have shown that musicians exhibit enhanced FFRs [29,30], the Demographic and Music Experience Questionnaire [31] was used to ensure that only non-musicians were included (Appendix E). To exclude the negative effect of cognitive impairment in prosody perception, all participants were screened using the Standardized Mini-Mental State Examination (SMMSE) [32] to assess five cognitive domains, including orientation, registration, attention and calculation, recall, and language, with inclusion based on scoring 29 or 30 on a scale of 30 (Appendix C). Written informed consent was required from all, with experimental protocols approved by the Office of Research Ethics and Integrity at the University of Ottawa (Appendix A).

2.2. Stimuli

The word "balloon" uttered by both male and female speakers with sad and happy emotions was used to elicit the FFR. This word was extracted from the sentence "*No, I burst the balloon!*" in the emotional speech database (ESD) [33]. The ESD database contains 350 parallel utterances spoken by 10 native English and 10 native Chinese speakers across 5 emotions (neutral, happy, angry, sad, and surprise). The word "balloon" was chosen for its fully voiced bisyllabic structure, making it well-suited for tracking suprasegmental variations in F0 across the entire stimulus. Given the limited availability of valid and reliable emotional speech datasets, balloon also demonstrated reliable waveform production during our pilot testing. This study targeted sad and happy emotions produced by one male English speaker (Speaker 11) and one female English speaker (Speaker 15), selected for their distinct F0 contours, including flat, rising, falling, and rising-falling contours. Figure 1 illustrates the F0 variations for the word "balloon" across emotional contexts for both male and female talkers. The duration of the stimuli also varied across conditions, with the male sad stimulus lasting 457 ms, male happy 565 ms, female sad 655 ms, and female happy 450 ms. Table 1 summarizes key F0 parameters for each condition, including mean, maximum, minimum, and range values. Higher mean F0, broader F0 ranges, and greater rates of F0 change were observed for happy compared to sad emotions, with female talker showing mostly higher F0 values than male talker. F0 parameters were extracted using Praat (version 6.2.05; www.praat.org) with the default settings.

Table 1. List of acoustic parameters related to F0 across all the conditions examined.

Word "balloon"		Mean F0 (Hz)	Max F0 (Hz)	Min F0 (Hz)	F0 range (Hz)
Talker	**Emotion**				
Male	Sad	95.5	107.8	88.1	19.7
Male	Happy	150.9	182.5	108.1	74.4
Female	Sad	272.5	332.2	189.6	142.6
Female	Happy	312.9	432.9	179.1	253.8

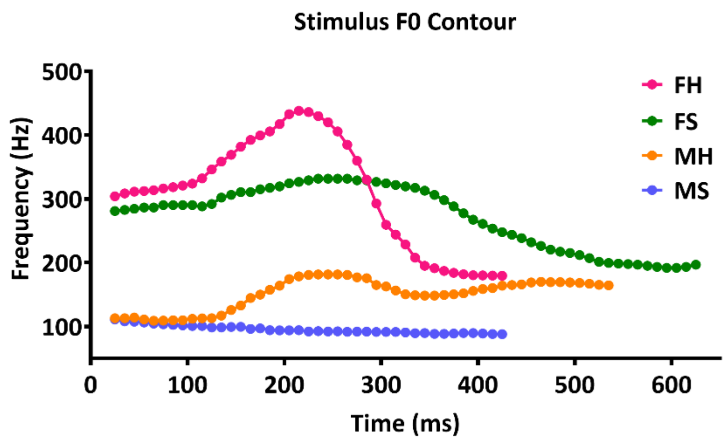


Figure 1. F0 contours for the stimulus word "balloon" spoken by a male and a female with two emotions, sad and happy. The blue curve represents the male talker with sad emotion (MS), the orange curve the male talker with happy emotion (MH), the green curve the female talker with sad emotion (FS), and the pink curve the female talker with happy emotion (FH).

2.3. EEG recording

The FFR responses were recorded with the Duet 2-channel AEP system (Intelligent Hearing Systems, Miami, FL) using four Ag/AgCI disposable disc electrodes. In accordance with the International Electrode System IES 10-20 norms, two inverting electrodes were placed at A1 and A2 (right and left earlobes), a noninverting electrode was positioned at Cz (vertex), and the ground electrode was located at Fz (forehead). Prior to electrode placement, the skin at these sites was prepared using an alcohol wipe and Nuprep Skin Prep Gel. The electrode impedances were kept below 5 kΩ, and inter-electrode impedances did not exceed 3 kΩ.

Two-channel FFR recordings were made over a time window of 793.65 ms (including a 40 ms pre-stimulus time), with an online filter setting of 30–1500 Hz, and sampling period of 175 μs. The four stimulus conditions, consisting of the spoken word /balloon/ by one male and one female speaker expressing happy and sad emotions, were presented binaurally through a shielded insert earphone at 80 dB SPL and a rate of 1.26 sweeps per second with an alternating polarity. Individual sweeps were rejected as artifacts if they exceeded 31 microvolts. This experiment was conducted in two sessions on different days, each lasting a maximum of three hours. Each session included 1500 sweeps (three blocks of 500) of each of the four conditions, resulting in 12 blocks of recording overall. Following each block, participants were given a brief pause to confirm their comfort and reduce the risk of sleepiness. Additionally, after completing 6 blocks in a session, participants were given a 15-minute break to reduce fatigue. In total, 3000 artifact free responses (6 blocks of 500) were acquired per condition for each participant. To prevent order effects, the sequence of emotions and speakers was randomized across participants using a complete counterbalancing approach, ensuring that no emotion or speaker was repeatedly presented in the same order.

The experiment was conducted in an electrically shielded and soundproof booth where participants were seated comfortably in a reclining chair. To minimize artifacts, participants were encouraged to close their eyes, relax, and limit body movements during the recording. They were also asked to try not to fall asleep.

2.4. Data analysis

For each condition by participant, the 3000 response sweeps were coherently averaged. To account for neural delays in the FFR, the stimulus waveform was shifted forward by 10 ms. To extract the F0 contour, both the stimulus and averaged response were segmented into 50 ms frames with a 40ms overlap. The stimulus F0 contour was captured by computing the median of F0 values obtained from each frame using speech pitch estimation methods in MATLAB 2024a specifically the ‘pitch’ and ‘pitchnn’ commands. To extract the response F0 value we followed a Harmonic Amplitude Summation (HAS) approach [34]. A ±50 Hz frequency search range with 1 Hz increments was applied to each frame, centered around the stimulus F0 of the corresponding time-aligned frame. For each candidate frequency within this range, the amplitudes of the candidate and its harmonics in the response frame spectrum were summed. The candidate yielding the most prominent peak was selected as the response F0. The number of harmonics was set to 4 for the sad male condition, and to 2 for the remaining conditions, as responses in the sad male condition exhibited a clearer harmonic structure.

Two evaluation metrics were used to assess the accuracy of F0 estimation: (1) 5% accuracy which quantifies the proportion of response frames where the estimated F0 values fall within 5% of their corresponding stimulus F0 values, and (2) Root-Mean-Square Error (RMSE) which quantifies the average deviation between the estimated response and reference stimulus F0 contours.

2.5. Statistical analysis

Neural encoding of F0 contours in participants' FFR responses, measured through RMSE and 5% accuracy, was assessed across sad and happy stimuli spoken by male and female talkers using a one-way repeated measures analysis of variance (RM-ANOVA) and Tukey's post hoc test, visualized using GraphPad Prism 10.0, and the data are presented as mean ± SDE in Figure 4. To evaluate the sex-related differences in processing acoustic features of emotional speech, an unpaired two tailed t-test and Welch's test were conducted for equal and unequal variance data (Figure 5). Results with *p-values* < 0.05 were considered statistically significant. Statistical significance is shown as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

3. Results

3.1. Spectrographic representation and F0 contour of the FFR

To ensure the robustness of the recorded responses, aggregate spectrograms (6 blocks × 16 participants) were examined for each condition before F0 extraction (Figure 2). The spectrograms were generated with a window length of 50 ms and an overlap of 40 ms. In the response spectrograms for stimuli presented in the male voice (Figure 2A–B), strong energy was observed at F0, 2F0, and 3F0, whereas female-voiced stimuli (Figure 2C–D) showed less prominent F0 and very attenuated harmonics. These observations highlight potential differences in the neural encoding of F0 as a function of stimulus characterisitics.

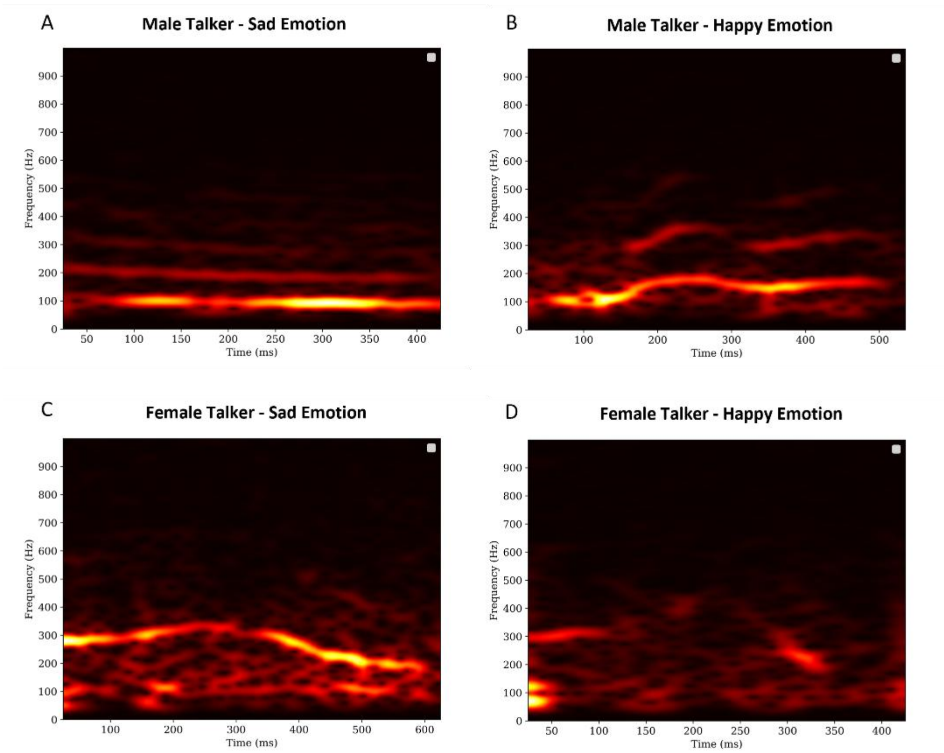


Figure 2. Spectrograms of the average response across all participants (n=16) for the four experimental conditions: male talker and sad emotion (A), male talker and happy emotion (B), female talker and sad emotion (c), and female talker and happy emotion (D).

The HAS algorithm was subsequently applied to each participant's FFR to extract the response F0 contour, and RMSE and 5% accuracy were then computed. Figure 3 displays spectrograms from a selected participant (Participant 13), illustrating that the extracted response F0 contour tracks the stimulus F0 across all four conditions, although its accuracy varies across conditions. According to the 5% accuracy and RMSE metrics for this participant, sad male condition resulted in highest accuracy and lowest error, followed by happy male, sad female, and happy female.

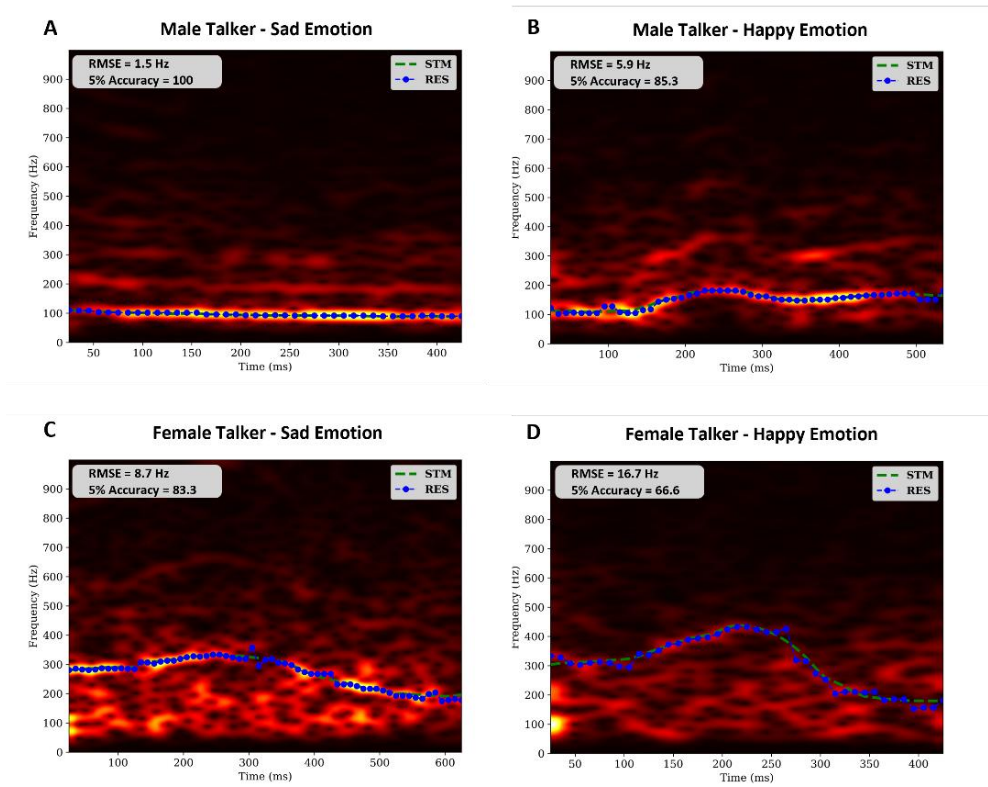


Figure 3. Spectrograms of Participant 13, showing overlaid F0 contours extracted from the stimulus waveform (green curves) and evoked FFR (blue curves) in response to the word "balloon" spoken by male talker and sad emotion (A), male talker and happy emotion (B), female talker and sad emotion (C), and female talker and happy emotion (D).

3.2. Effect of stimulus condition on FFR F0 tracking accuracy

Figure 4 illustrates the results of one-way repeated measures analysis of variance (RM-ANOVA) conducted to statistically examine the impact of stimuli type (sad male, happy male, sad female, happy female) for the word "balloon" on 5% accuracy and RMSE separetely across all participants.

Regarding RMSE, a significant main effect of stimulus was observed ($F(2.74, 41.20)=137.9$, $p < 0.0001$), with an effect size of $R^2=0.90$. As shown in Figure 4A, the lowest F0 tracking error was observed in the sad male condition (Mean RMSE=2.7 Hz), followed by the happy male (Mean RMSE=7.5 Hz), sad female (Mean RMSE=12.0 Hz), and happy female (Mean RMSE=17.5 Hz) conditions. Post-hoc Tukey's mulitiple comparisons test showed significant differences among all conditions ($p < 0.05$), as shown in Figure 4A. Specifically, F0 tracking error for the sad male stimulus was significantly lower than for the happy male (Mean Difference=-5.1 Hz, CI[-7.14, -3.19], $p < 0.0001$), sad female (Mean Difference=-9.6 Hz, CI[-11.84, -7.45], $p < 0.0001$), and happy female conditions (Mean Difference=-15.2 Hz, CI[-17.14, -13.26], $p < 0.0001$). In addition, the happy male condition resulted in significantly lower error compared to both the sad female (Mean Difference=-4.4 Hz, CI[-6.90, -2.04], $p < 0.001$) and happy female conditions (Mean Difference=-10.0 Hz, CI[-12.57, -7.49], $p < 0.0001$). Lastly, F0 tracking for the sad female condition was associated with lower errors compared to the happy female condition (Mean Difference=-5.5 Hz, CI[-7.88, -3.23], $p < 0.0001$).

Similar observations were made for 5% accuracy (Figure 4B), where the RM-ANOVA results showed a significant effect of stimulus type on the FFR's efficiency in accurately encoding the stimulus F0 contour ($F(2.51, 37.67)=84.52$, $P < 0.0001$). The large size effect ($R^2=0.85$) confirmed that a large proportion of the variance in the 5% accuracy values is explained by the experimental conditions and individual differences did not significantly contribute to the obsereved variance ($F(15, 45)=1.84$, $P=0.0580$). As shown in Figure 4B, the highest 5% accuracy was observed for the sad male condition (mean=96.9), followed by happy male (mean=81.8), sad female (mean=81.7), and happy female (mean=63.7). Tukey's multiple comparisons test indicated that this significant difference was present among all conditions ($p < 0.05$), except between the happy male and sad female conditions. Participants demonstrated higher accuracy in response to the sad male stimulus compared to the happy male (Mean Difference=15.1, CI[10.01, 20.18], $p < 0.0001$), sad female (Mean Difference=15.2, CI[8.47, 21.90], $p < 0.0001$), and happy female stimulus (Mean Difference=33.2, CI[28.24, 38.11], $p < 0.0001$). The happy male condition also elicited a significantly more accurate response than the happy female stimulus (Mean Difference=18.1, CI[12.66, 23.50], p < 0.0001), and the sad female condition had significantly higher 5% accuracy than the happy female condition (Mean Difference=17.9, CI[11.81, 24.17], p < 0.0001). Although the happy male condition showed slightly higher 5% accuracy compared to the sad female, this difference did not reach statistical significance (Mean Difference=0.09, CI[-7.282, 7.462], P=0.972), suggesting that FFR is tracking the contour of these two stimuli with relatively similar accuracy (Figure 4B). Both RMSE and 5% accuracy metrics supported that the F0 contours of emotional speech can be extracted from FFR recordings, with greater accuracy and lower error observed for the sad male stimulus, whearas extraction was more challenging for the happy female condition.

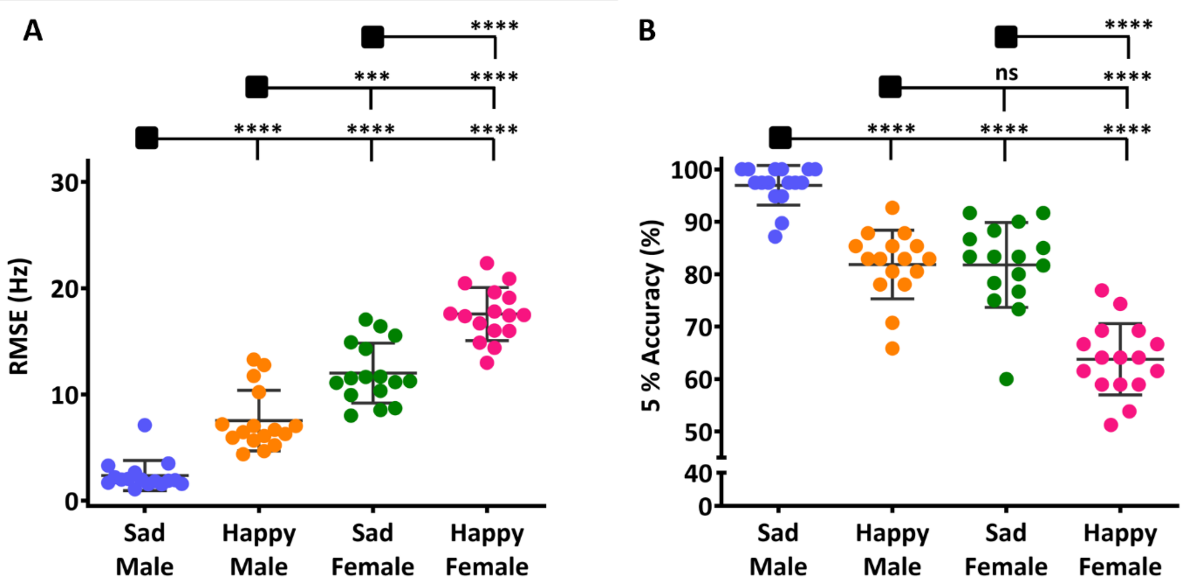


Figure 4. FFR F0 tracking accuracy, measured by either RMSE (A) or 5% accuracy (B) across all participants (n=16), in response to the word “balloon” spoken by male talker and sad emotion (Blue dots), male talker and happy emotion (orange dots), female talker and sad emotion (green dots), and female talker and happy emotion (pink dots).

As depicted in Figure 4, when focusing solely on emotional content, the sad version of word “balloon” was easier to be tracked by FFR at the central level of the auditory system compared to happy vesrion for both male (sad male vs. happy male – RMSE: $p < 0.0001$, 5% Accuracry: $p < 0.0001$) and female talkers (sad female vs. happy female – RMSE: $p < 0.0001$, 5% Accuracry: $p < 0.0001$). Similarly, when considering male and female talkers independently of emotion type, words spoken by male talker resulted in more robust F0 than female talker (sad male vs. sad female – RMSE: $p < 0.0001$, Accuracy: $p < 0.0001$/ happy male vs. happy female – RMSE: $p < 0.0001$, 5% Accuracy: $p < 0.0001$).

3.3. Effect of listener sex on FFR F0 tracking accuracy

We evaluated differences between male and female participants in tracking F0 contours for the sad and happy version of word "balloon". An unpaired two-tailed t-test was used to compare the perfomance of male and female participants in encoding F0 contours, as measured by RMSE and 5% accuracy indices. In cases of unequal variances, Welch's correction was applied.

No noticable difference was seen between the male- and female-averaged FFR F0 contours and response F0 contours, as illustrated in the left panels of Figure 5. Statistical results further supported these initial observations. For the sad male condition, RMSE did not differ significantly between male and female participants ($p = 0.63$, using Welch's correction), nor did 5% accuracy ($p = 0.78$)(Figure 5A, middle and right panels). Similarly, in the happy female condition, no significant differences were found in RMSE ($p = 0.68$) or 5% accuracy ($p = 0.47$), as shown in the middle and right panels of Figures 5D. In the happy male condition, there was no statistically significant difference in 5% accuracy ($p = 0.25$); however, for RMSE, Welch's test revealed a significant difference ($p = 0.0451$) in performance between male (Mean=9.0 Hz) and female (Mean=6.1 Hz) group, with the effect size (R^2) of 0.4076 (Figure 5B). As shown in Figure 5C, similar results were observed for the sad female condition where 5% accuracy did not differ significantly between male and female paeticipants ($p = 0.08$). However, for RMSE, a significant difference (p = 0.0455) was observed between male (Mean= 13.4 Hz) and female (Mean=10.6 Hz) participants.

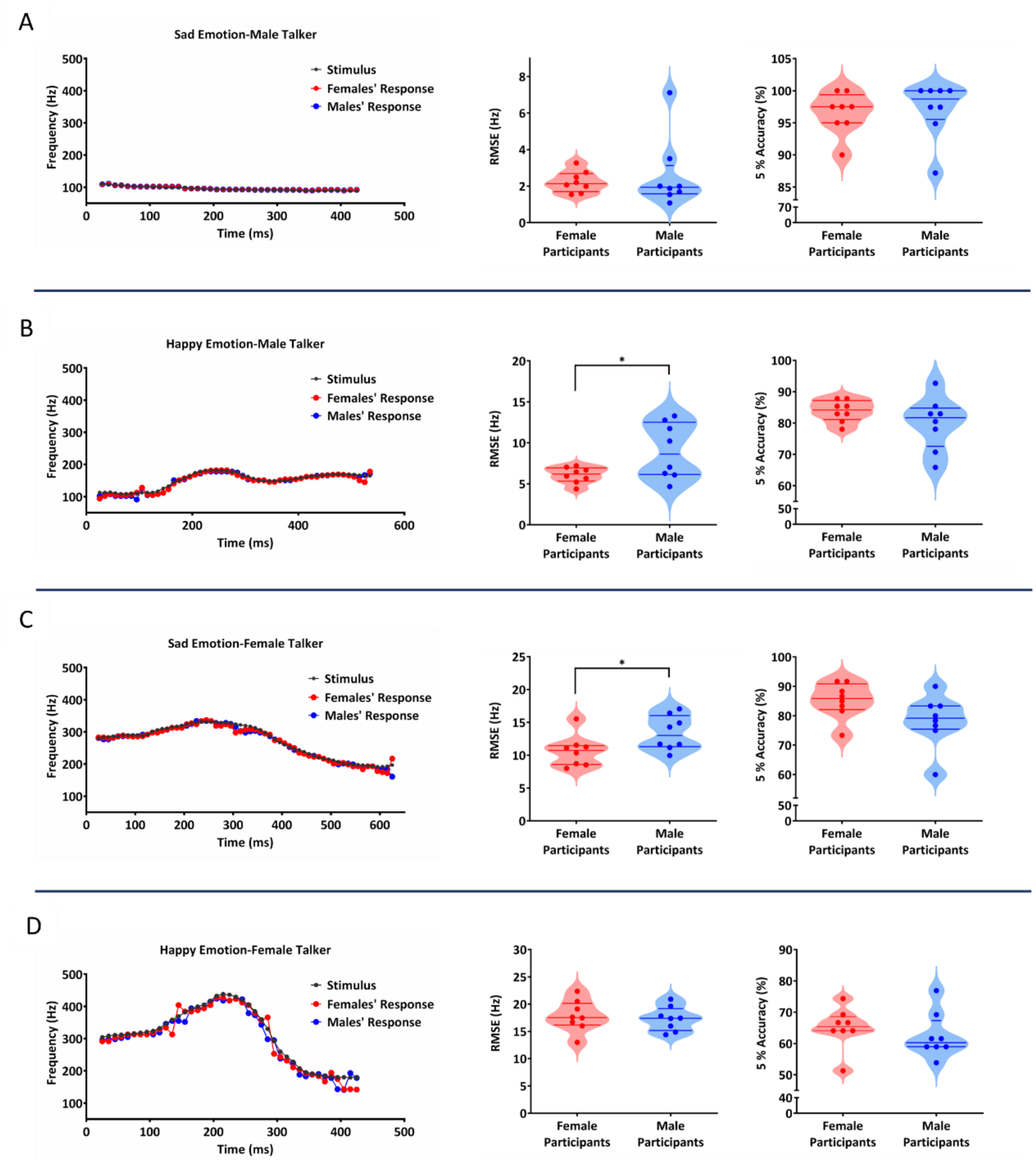


Figure 5. FFR F0 tracking accuracy as a function of listener sex, shown using RMSE (middle panel) and 5% accuracy (right panel). F0 contours extracted (left panel) from the stimulus waveform (black curves) and evoked FFR according to the listener sex, with female participants (red curves) vs. male participants (blue curves) in normal hearing individuals in response to the word “balloon” spoken by male talker and sad emotion (A), male talker and happy emotion (B), female talker and sad emotion (c), and female talker and happy emotion (D).

4. Discussion

F0 variations over time, extending beyond single speech segments, create different prosodic patterns that can convey various emotional states during a conversation. Using FFR recordings from 16 individuals with normal hearing, we investigated the application of the FFR in assessing speech prosody encoding. Although some studies have examined F0 tracking at the segmental level, this is the first study to investigate FFR F0 tracking in response to suprasegmental variations associated with emotional expression.

Our results showed that the FFR can preserve pitch-relevant information in response to all versions of the word "balloon", suggesting that auditory neural circuits can dynamically phase lock to time-variant F0 in complex natural speech stimuli. This aligns with the finding from Krishnan et al. who showed that the FFR can accurately track the dynamic pitch contours of synthesized Mandarin tones, demonstrating the dominant role of discharge periodicities and interspike intervals in variable pitch percepts [19]. By using natural and high-pass filtered vowel, Dajani et al. also reinforced the capability of central auditory processors to encode and track pitch-related variations in the form of synchronized neural firing even in the degraded condition [21]. There is also evidence that this phase-locking ability can even be tolerant to noise, preserving frequency-coding acuity in the presence of both ipsilateral and contralateral noise [35,36].

While our findings demonstrated that the FFR reflects the acoustic properties of prosodic speech in word "balloon", the extent of this representation was influenced by emotion type and acoustic voice characteristics of talkers.

When considering emotion exclusively, we found that the sad stimuli resulted in more robust FFR (higher score in 5% accuracy and lower score in RMSE) compared to the happy versions of "balloon" for both male and female speakers. Sadness, as confirmed by the stimulus characteristics in Table 1 and Figure 1, is typically characterized by a lower average F0 height and a falling F0 trajectory with lower rate of F0 change compared to the happy emotion [37-39]. These acoustic differences may contribute to affecting the FFR's capacity for tracking F0. Previous evidence also confirms that pitch, as a complex perceptual attribute, is shaped by different acoustic characteristics, including average F0 height, direction of F0 movement, and magnitude of F0 slope [40,41]. In the context of F0 encoding through FFR recordings, previous studies have consistently observed that F0 response magnitude decreases as pitch height of the stimulus increases [16,42.43]. Krishnan et al. by focusing on the direction of F0 movement, demonstrated that while FFR pitch contours mirror the pitch trajectory in synthetic Chinese syllables (yi^1, yi^2, yi^3, and yi^4) with varying F0 contours, the pitch strength is greater for stimuli with rising F0 contours compared to ones with a falling contour [19]. This finding implied that the neural elements responsible for generating the FFR exhibit a distinct temporal response pattern for rising versus falling tones. However, in a study examining the effect of pitch direction on pitch strength and pitch tracking accuracy of FFR response, it was found that synthesized Mandarin stimuli of "yi" with linear rising and falling F0 contours can have similar pitch representation, both in native and non-native listeners [44]. These contrasting findings were explained by the idea that a simple linear and unilateral rising or falling pitch contour pattern alone cannot fully characterize how the brainstem neurons are tuned to the complex features of sound. They suggested that using naturalistic and curvilinear pitch contours, like the stimuli used in the present study, can effectively reflect the brainstem's distinct processing abilities, which are strengthened through experience and plasticity. Billings et al. also found that no difference in response strength and temporal coherence was observed between linear rising and falling glides [45]. However, their study added that the rate of frequency change contributes more significantly to FFR strength than glide direction, with shallower slopes notably enhancing response accuracy. Similarly, a few studies also reported that a decrease in the rate of frequency change elicits stronger FFR responses [43,46]. Moreover, other studies have shown that even with absence of F0 in the stimulus, the FFR can represent F0 with the same strength as in the condition where the F0 was physically present [21,47]. This suggests there is a complex relationship between pitch representation and stimulus frequency content and periodicity.

In addition to the type of emotions, our results showed that the FFR at F0 is also influenced by the characteristics of male- and female-spoken voices. In both the sad and happy conditions for the word "balloon", the male talker consistently led to better F0 tracking performance relative to female talker. Since the male voice in our study had an F0 nearly half that of the female voice (Table 1), and previous studies have reported that lower talker F0 leads to larger FFR amplitude [22,48,49], this may account for the increased robustness of the FFR evoked by male voice over female voice. Consistent with our finding, using continuous speech, Van Canneyt et al. calculated the correlation between actual F0 and reconstructed F0 feature from EEG data, and showed that male-narrated stories, which had lower and steadier F0, is associated with stronger response at F0 compared to female-narrated stories [50]. However, they observed that even for high F0 voices, if they create a synthesized version of a female-narrated story with strong higher harmonics and lower rate of F0 change, the strength of the continuous F0 response significantly improves compared to the natural version of the same story. This result was attributed to intrinsic acoustic differences in female and male voices, rather than directly to the talker gender. In agreement with this finding, Polonenko & Maddox also confirmed that while the FFRs to male-narrated speech with an average F0 of 123 Hz was larger and earlier compared to female-narrated speech with an average F0 of 183 Hz, by shifting the F0 of both narrators' voices to an intermediate pitch of 150

Hz, their FFR became very similar in both amplitude and latency [51]. Easwar et al. further highlighted that in addition to F0 properties, formant structure and harmonic distribution of male and female voices can evoke a different FFR [22]. They elaborated that the enhanced F0 FFR elicited by male voices benefits from wider F1 bandwidth and more well-resolved harmonics falling within F1, while FFR to female vowels depends more on high-frequency harmonics, especially in regions above F1 and F2, creating envelope modulation at F0 to generate the F0 FFR. This distinction underscores the role of sex specific vocal characteristics in shaping auditory processing.

Our findings also revealed no sex differences among listeners in encoding of dynamic F0 contours across all conditions. Male and female participants performed similarly in tracking F0 contours for the word “balloon” spoken with both sad and happy emotion by male and female talker, as measured by both RMSE and 5% accuracy. Although a small RMSE difference was observed for the happy male and sad female stimuli, 5% accuracy remained similar between these two conditions, which suggests overall similarity in performance between males and females. Contrary to our results, research by Ahadi et al. found that the FFR to a 40 ms synthesized /da/ syllable differs between male and female listeners [23]. Specifically, females exhibited shorter latencies and larger peak amplitudes for the transient portion of response compared to males. Furthermore, females had a larger representation of F0, first formant (F1), and high frequency components of F1. On the other hand, Krizman et al. showed that dissimilarity between males and females in response to the synthesized syllable /da/ with a time-varying F0 (103–125 Hz) is only limited to the encoding of the fast elements of speech such as having earlier transient peaks and larger spectral magnitude of higher frequency components in females [52]. However, the encoding of F0, either measured through its spectral magnitude or the timing between F0-related peaks, was not affected by sex. This similar F0 phase-locking across sexes implied that, although there are sex differences in the auditory system, both males and females have consistent low-frequency tracking ability. The authors suggested that this difference in how the sexes process fast-varying speech elements compared to slower components may reflect a selective response of the auditory nervous system that is influenced by the stimulus characteristics.

It is also worth noting that our study used a meaningful word with emotional valence, extending beyond the physical acoustic properties of sound, such as frequency, to include paralinguistic features of the speech signal. In a study by Galbraith et al., using a speech stimulus with familiar phonological and prosodic features, it was shown that the FFR has larger amplitude in response to this familiar stimulus compared to its reversed version, which was unfamiliar in sound but with the same phoneme duration and voicing [53]. Presumably, with a real speech stimulus containing both linguistic and paralinguistic information, the neural activity underlying the FFRs would differ compared to when simpler and more isolated stimuli like a synthesized syllable are presented. At the cortical level, Guo et al. also showed that the FFR tracks dynamic F0 more effectively for harmonic-rich speech stimuli than for non-speech stimuli, suggesting a specialized mechanism tuned for complex sound inputs in the brain [54]. It is possible that the distinct FFR characteristics previously reported for male and female listeners in response to isolated syllables may be modulated when more complex and paralinguistically rich speech stimuli are used. Future studies with larger and more diverse samples would allow for a more comprehensive examination of the observed pattern.

In conclusion, this study showed that the prosodic features of speech, such as F0 contour, can be extracted from FFRs with good accuracy. This evidence was obtained using a single stimulus word “balloon”, produced by only two talkers across two emotional conditions. However, the acoustic characteristics of the eliciting stimuli can affect the response's effectiveness in accurately representing prosody. Specifically, when using emotional speech stimuli, the type of emotion and the acoustic characteristics of talker can influence the ability of the FFR to track variations in F0 of the stimulus. Stimuli characterized by a lower F0, smaller F0 slope magnitude, and greater energy at F0 and its harmonics, such as sad emotion or male voices, can result in more robust FFR. Future work is needed to extend this test to individuals with hearing impairment, with or without hearing aids. This would introduce speech prosodic assessment in clinical settings and thus provide a more comprehensive rehabilitation approach.

References

1 Gaikwad, S. K., Gawali, B. W. & Yannawar, P. A review on speech recognition technique. *Int. J. Comput. Appl.* **10**, 16-24 (2010).

2 Ng, S. H. & Bradac, J. J. *Power in Language: Verbal Communication and Social Influence*. (SAGE Publications, Inc., 1993).

3 Tatham, M. & Morton, K. *Speech Production and Perception*. (Springer, 2006).

4 Teoh, E. S., Cappelloni, M. S. & Lalor, E. C. Prosodic pitch processing is represented in delta-band EEG and is dissociable from the cortical tracking of other acoustic and phonetic features. *Eur. J. Neurosci.* **50**, 3831-3842 (2019).

5 Bolinger, D. & Bolinger, D. L. M. *Intonation and Its Parts: Melody in Spoken English*. (Stanford University Press, 1986).

6 Pickett, J. & Morris, S. R. The acoustics of speech communication: Fundamentals, speech perception theory, and technology. *J. Acoust. Soc. Am.* **108**, 1373-1374, doi:10.1121/1.1289701 (2000).

7 Karimi-Boroujeni, M., Dajani, H. R. & Giguère, C. Perception of prosody in hearing-impaired individuals and users of hearing assistive devices: An overview of recent advances. *J. Speech. Lang. Hear. Res.* **66**, 775-789 (2023).

8 Fischer, S. E. *Speech language pathologists and prosody: Knowledge and clinical practices* Honors thesis, university of Mississippi, (2018).

9 Grandjean, D. Brain networks of emotional prosody processing. *Emot. Rev.* **13**, 34-43 (2021).

10 Hammerschmidt, K. & Jürgens, U. Acoustical correlates of affective prosody. *J. Voice* **21**, 531-540 (2007).

11 Xie, X., Buxó-Lugo, A. & Kurumada, C. Encoding and decoding of meaning through structured variability in intonational speech prosody. *Cognition* **211**, 104619 (2021).

12 Fries, P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* **9**, 474-480 (2005).

13 Ceravolo, L., Frühholz, S., Pierce, J., Grandjean, D. & Péron, J. Basal ganglia and cerebellum contributions to vocal emotion processing as revealed by high-resolution fMRI. *Sci. Rep.* **11**, 10645 (2021).

14 Pannese, A., Grandjean, D. & Frühholz, S. Subcortical processing in auditory communication. *Hear. Res.* **328**, 67-77 (2015).

15 Dupuis, K. & Pichora-Fuller, M. K. Intelligibility of emotional speech in younger and older adults. *Ear Hear.* **35**, 695-707 (2014).

16 Skoe, E. & Kraus, N. Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* **31**, 302-324 (2010).

17 Krizman, J. & Kraus, N. Analyzing the FFR: A tutorial for decoding the richness of auditory function. *Hear. Res.* **382**, 107779 (2019).

18 Coffey, E. B. *et al.* Evolving perspectives on the sources of the frequency-following response. *Nat. Commun.* **10**, 5036 (2019).

19 Krishnan, A., Xu, Y., Gandour, J. T. & Cariani, P. A. Human frequency-following response: representation of pitch contours in Chinese tones. *Hear. Res.* **189**, 1-12 (2004).

20 Krishnan, A. & Parkinson, J. Human frequency-following response: representation of tonal sweeps. *Audiol. Neurotol.* **5**, 312-321 (2000).

21 Dajani, H. R., Purcell, D., Wong, W., Kunov, H. & Picton, T. W. Recording human evoked potentials that follow the pitch contour of a natural vowel. *IEEE Trans. Biomed. Eng.* **52**, 1614-1618 (2005).

22 Easwar, V., Purcell, D., Van Eeckhoutte, M. & Aiken, S. J. The influence of male-and female-spoken vowel acoustics on envelope-following responses. *Semin. Hear.* **43**, 223-239 (2022).

23 Ahadi, M., Pourbakht, A., Jafari, A. H., Shirjian, Z. & Jafarpisheh, A. S. Gender disparity in subcortical encoding of binaurally presented speech stimuli: an auditory evoked potentials study. *Auris Nasus Larynx* **41**, 239-243 (2014).

24 Krizman, J. *et al.* Athleticism and sex impact neural processing of sound. *Sci. Rep.* **12**, 15181 (2022).

25 Kırbac, A., Turkyılmaz, M. D. & Yağcıoglu, S. Gender effects on binaural speech auditory brainstem response. *J. Int. Adv. Otol.* **18**, 125 (2022).

26 Marian, V., Blumenfeld, H. K. & Kaushanskaya, M. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *J. Speech. Lang. Hear. Res.* **50**, 940-967 (2007).

27 Krishnan, A., Xu, Y., Gandour, J. & Cariani, P. Encoding of pitch in the human brainstem is sensitive to language experience. *Cogn. Brain Res.* **25**, 161-168 (2005).

28 Krishnan, A., Suresh, C. H. & Gandour, J. T. Tone language experience-dependent advantage in pitch representation in brainstem and auditory cortex is maintained under reverberation. *Hear. Res.* **377**, 61-71 (2019).

29 Bidelman, G. M., Gandour, J. T. & Krishnan, A. Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cogn.* **77**, 1-10 (2011).

30 Xie, X. & Myers, E. The impact of musical training and tone language experience on talker identification. *J. Acoust. Soc. Am.* **137**, 419-432 (2015).

31 Comeau, G., Koravand, A. & Swirp, M. Prevalence of hearing loss among university music students. *Can. Acoust.* **46**, 37-51 (2018).

32 Molloy, D. W. & Standish, T. I. A guide to the standardized Mini-Mental State Examination. *Int. Psychogeriatr.* **9**, 87-94 (1997).

33 Zhou, K., Sisman, B., Liu, R. & Li, H. Emotional voice conversion: Theory, databases and esd. *Speech Commun.* **137**, 1-18 (2022).

34 Sadeghkhani, S., Boroujeni, M. K., Dajani, H. R., Seydnejad, S. R. & Giguère, C. A Robust Method for Pitch Tracking in the Frequency Following Response using Harmonic Amplitude Summation Filterbank. *JASA Express Letters* ***5*** *(2025).*

35 Li, X. & Jeng, F.-C. Noise tolerance in human frequency-following responses to voice pitch. *J. Acoust. Soc. Am.* **129**, EL21-EL26 (2011).

36 Jeng, F.-C., Stehura, K. A., Hart, B. N. & Giordano, A. T. Contralateral noise degrades frequency-coding accuracy in normal-hearing adults. *J. Acoust. Soc. Am.* **150**, A144-A144 (2021).

37 Yoon, T.-J. How Much Does the Dynamic F0 Curve Affect the Expression of Emotion in Utterances? *Appl. Sci.* **14**, 10972 (2024).

38 Pereira, C. & Watson, C. I. Some acoustic characteristics of emotion. in *5th International Conference on Spoken Language Processing.* 927–934(Sydney,1998).

39 Pollermann, B. Z. & Archinard, M. *Acoustic patterns of emotions*. in *Improvements in Speech Synthesis* (ed G. Bailly Keller, A. Monaghan, J. Terken, M. Huckvale) Ch. 23, 237-245 (Wiley, 2002).

40 Gandour, J. Tone perception in Far Eastern languages. *J. Phon.* **11**, 149-175 (1983).

41 Chandrasekaran, B., Gandour, J. T. & Krishnan, A. Neuroplasticity in the processing of pitch dimensions: A multidimensional scaling analysis of the mismatch negativity. *Restor. Neurol. Neurosci.* **25**, 195-210 (2007).

42 Krishnan, A., Suresh, C. H. & Gandour, J. T. Changes in pitch height elicit both language-universal and language-dependent changes in neural representation of pitch in the brainstem and auditory cortex. *Neuroscience* **346**, 52-63 (2017).

43 Van Canneyt, J., Wouters, J. & Francart, T. From modulated noise to natural speech: The effect of stimulus parameters on the envelope following response. *Hear. Res.* **393**, 107993 (2020).

44 Xu, Y., Krishnan, A. & Gandour, J. T. Specificity of experience-dependent pitch representation in the brainstem. *Neuroreport* **17**, 1601-1605 (2006).

45 Billings, C. J., Bologna, W. J., Muralimanohar, R. K., Madsen, B. M. & Molis, M. R. Frequency following responses to tone glides: Effects of frequency extent, direction, and electrode montage. *Hear. Res.* **375**, 25-33 (2019).

46 Clinard, C. G. & Cotter, C. M. Neural representation of dynamic frequency is degraded in older adults. *Hear. Res.* **323**, 91-98 (2015).

47 Coffey, E. B., Colagrosso, E. M., Lehmann, A., Schönwiesner, M. & Zatorre, R. J. Individual differences in the frequency-following response: relation to pitch perception. *PLoS One* **11**, e0152374 (2016).

48 Easwar, V., Bridgwater, E. & Purcell, D. The influence of vowel identity, vowel production variability, and consonant environment on envelope following responses. *Ear Hear.* **42**, 662-672 (2021).

49 Easwar, V., Boothalingam, S. & Flaherty, R. Fundamental frequency-dependent changes in vowel-evoked envelope following responses. *Hear. Res.* **408**, 108297 (2021).

50 Van Canneyt, J., Wouters, J. & Francart, T. Neural tracking of the fundamental frequency of the voice: the effect of voice characteristics. *Eur. J. Neurosci.* **53**, 3640-3653 (2021).

51 Polonenko, M. J. & Maddox, R. K. Fundamental frequency predominantly drives talker differences in auditory brainstem responses to continuous speech. *JASA Express Letters* **4** (2024).

52 Krizman, J., Skoe, E. & Kraus, N. Sex differences in auditory subcortical function. *Clin. Neurophysiol.* **123**, 590-597 (2012).

53 Galbraith, G. C. *et al.* Brain stem evoked response to forward and reversed speech in humans. *Neuroreport* **15**, 2057-2060 (2004).

54 Guo, N. *et al.* Speech frequency-following response in human auditory cortex is more than a simple tracking. *Neuroimage* **226**, 117545 (2021).

Chapter 4. Frequency following response to emotional speech: Effect of hearing loss and simulated hearing aid processing on F0 tracking accuracy

Frequency following response to emotional speech: Effect of hearing loss and simulated hearing aid processing on F0 tracking accuracy

Maryam Karimi-Boroujeni [1], Sajad Sadeghkhani [2], Solenn Ollivier 3, Hilmi R. Dajani [2], Saeid R. Seydnejad [2,4], Christian Giguère [1]

[1] School of Rehabilitation Sciences, University of Ottawa, Ottawa, ON, Canada

[2] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, Canada

[3] Université du Québec, École de technologie supérieure, Montréal, Québec, Canada

[4] Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran.

Abstract

Prosody, a key paralinguistic feature of successful interaction, conveys emotional information and complements the linguistic content of speech. In hearing impairment, reduced audibility and degraded spectro-temporal resolution diminish the salience of prosodic cues. Despite the importance of prosody in daily communications, research on speech perception in hearing-impaired individuals has primarily focused on linguistic elements, with less emphasis on emotional prosody. Inconsistent behavioral evidence leaves unclear whether hearing loss adversely affects emotion perception and to what extent hearing aids can compensate for potential deficits in prosody processing. This study used the speech-evoked frequency-following response (FFR) to investigate the neural encoding of prosody-related acoustic features in the central auditory system. Dynamic variations in fundamental frequency over time, known as the F0 contour, were the focus of our study. The neural representation of F0 contours was assessed using FFR recordings elicited by the word "balloon" spoken with sad and happy emotion, in normal-hearing participants and in hearing-impaired individuals under simulated aided and unaided conditions. Using a pitch-tracking method, F0 tracking accuracy was quantified based on 5% accuracy and Root Mean Square Error (RMSE). Hearing loss disrupted neural encoding of F0 contours in emotional speech, as reflected by lower 5% accuracy and higher RMSE; however, it was partially restored by hearing aid. Additionally, prosodic representation in FFR responses differed by emotion, with higher F0 tracking accuracy for sad stimulus. By providing insight into the neural mechanisms underlying prosody encoding, this study supports considering prosodic deficit in clinical settings to enhance speech comprehension in hearing-impaired individuals.

Keywords: prosody perception, hearing loss, hearing aids, emotion

Significance

Hearing loss is among the most common chronic health conditions worldwide. People with disabling hearing loss often experience communication difficulties. Even with advances in hearing aid technology, many users are dissatisfied with their devices. This may be partly due to fact that hearing assessments and hearing aid fittings prioritize word recognition, often neglecting prosodic cues that convey emotional state of the speaker. This study shows that individuals with hearing loss have difficulty processing emotional speech. The degree to which hearing loss impairs emotion perception differs depending on the emotion type. However, this deficit can be partially improved by hearing aid use. The results suggest including prosodic information in audiological assessment and rehabilitation to enhance speech comprehension in hearing-impaired individuals.

1. Introduction

Effective speech communication extends beyond the perception of linguistic content, relying also on paralinguistic cues that provide critical information to complement the linguistic aspect of speech (1, 2). Prosody is one the key paralinguistic cues through which speakers express their emotions and attitudes (3-5). In emotional expression, prosodic cues are acoustically characterized by the contours of fundamental frequency (F0) and intensity, along with variations in duration (6, 7). Each emotion is typically associated with a characteristic acoustic profile that must be faithfully received, precisely encoded, and effectively processed by a healthy auditory system (6, 8-10).

Once hearing impairment occurs, it not only reduces audibility but also degrades the spectro-temporal resolution and disrupts central auditory processing, contributing to broader deficits in everyday interaction that involve more than an inability to hear words (11-13). Although hearing aids improve audibility, they often do not fully restore speech perception at suprathreshold levels (14, 15). Regarding paralinguistic information, the degree to which hearing loss and hearing aid affect acoustic-based prosodic cues remains elusive. While some behavioral studies suggest that hearing loss leads to deficits in emotion perception (16-18), others argue that residual hearing may be sufficient to preserve emotional information, particularly in cases of mild to moderate hearing loss (19, 20). Similar discrepancies have been reported in hearing aid outcomes, with some evidence supporting benefits for emotional speech perception (21, 22), whereas other findings indicate that frequency and amplitude compression may distort intensity and spectral cues critical for distinguishing emotions (23, 24).

Electrophysiological investigation can provide insight into the neural mechanisms supporting the encoding of prosodic features in the brain. The Frequency-Following Response (FFR) is a non-invasive neural measure that reflects the synchronous activity of neurons across subcortical and cortical auditory pathways (25). It preserves F0-related information by capturing the temporal patterns of neural discharges and interspike intervals (26, 27). Using synthetic or natural isolated vowels and simple syllables, some FFR studies have demonstrated that hearing loss alters phase-locked neural activity, resulting in either reduced or enhanced F0 representation (28-33). Hearing aid use has been reported to compensate for the affected F0 encoding in some cases (30, 34, 35), although no significant improvement has been observed in others (28, 36). However, it remains to be determined how hearing loss and hearing aids may alter the encoding of time-varying F0 information within the context of prosody processing over multisyllabic speech tokens.

Our recent investigation in normal-hearing young adult listeners (Chapter 3) showed that the phase-locked activity underlying the FFR is sufficiently dynamic to track F0 contours across different emotions. The current study aimed to examine F0 tracking accuracy in response to sad and happy emotions in normal-hearing individuals and hearing-impaired individuals in both unaided and aided conditions.

2. Methods and materials

2.1. Participants

Twenty-eight adults aged between 18 and 65 years participated in the study, divided into a normal-hearing group ($n = 14$, mean age = 52.36 years) and an age-matched group with hearing loss ($n = 14$, mean age = 52.07 years). Normal hearing was defined as pure-tone thresholds ≤ 25 dB HL at frequencies from 250 to 8000 Hz. Hearing impairment was characterized by unilateral or bilateral mild to moderately severe sensorineural hearing loss, with air-conduction thresholds > 30 dB HL and ≤ 70 dB HL within the same frequency range. Hearing thresholds for hearing-impaired participants are provided in Table S3 in Appendix L. None of the participants showed abnormalities in the external auditory canal, tympanic membrane, or middle ear, as confirmed by otoscopy and tympanometry. A general development questionnaire was used to ensure that participants had no history of neurological diseases, psychiatric problems, or brain injuries (Appendix F). As musicians and tonal language speakers have demonstrated enhanced neural encoding of acoustic features in FFR studies (37-39), only non-musicians and non-tonal language speakers were included. Accordingly, participants were required to be native or near-native English speakers, with self-assessed scores of 9 or 10 on the 10-point Language Experience and Proficiency Questionnaire (LEAP-Q) (40) for speaking and understanding English (Appendix D). Music background was further assessed using the Demographic and Music Experience Questionnaire (41) to confirm that participants met the non-musician criteria (Appendix E). To control for potential effects of cognitive impairment on prosody perception, all participants were screened using the Standardized Mini-Mental State Examination (SMMSE) (42), which assesses five cognitive domains, including orientation, registration, attention and calculation, recall, and language (Appendix C). Only individuals scoring 29 or 30 out of 30 were included. Once eligibility was confirmed and consent form was signed, participants took part in data collection process. The experimental protocols were approved by the Research Ethics Board at the University of Ottawa (Appendix A).

2.2. Stimulus

To elicit the FFR, the word "balloon" uttered by a male talker with sad and happy emotions was used. This word was taken from the sentence "No, I burst the balloon!" included in the Emotional Speech Database (ESD) (43). In the ESD, 350 utterances were recorded from 20 speakers, 10 native English and 10 native Chinese, each expressing sentences with five different emotions, including neutral, happy, angry, sad, and surprise. Since word "balloon" has more than one syllable and is composed entirely of voiced sounds, it enabled tracking of frequency variations of F0 in the neural response. In this study, sad and happy utterances from a male English talker (Speaker 11) were selected due to their distinct F0 contours, including falling and rising-falling patterns, as shown in Figure 1. F0 parameters were extracted using Praat (version 6.2.05; Boersma & Weenink, www.praat.org) with default settings. Compared to the sad stimulus, the happy "balloon" had a higher mean F0, broader F0 range, and a greater rate of F0 change, as well as a longer duration (565 ms vs. 457 ms).

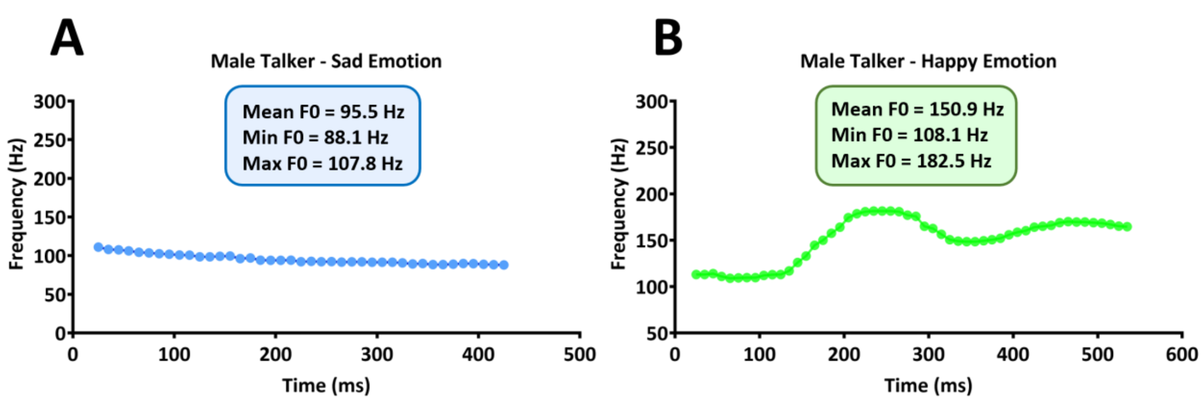


Figure 1. F0 contours of the stimulus word "balloon" spoken with sad (A) and happy (B) emotion by Speaker 11 in ESD, along with corresponding acoustic parameters related to F0 for each condition.

Normal-hearing participants (NH) were tested with the original sound files presented at 65 dB SPL. For the hearing-impaired group (HI), in addition to the original sound files used in the unaided condition (HI-Unaided) at 65 dB SPL, participants were also tested with amplified versions of the word "balloon" representing the aided condition (HI-Aided). Hearing aid processing was simulated using a software-based hearing aid simulator (44) developed in MATLAB (MathWorks, Natick, MA). The simulator applied wide dynamic range compression (WDRC) based on an algorithm combining an NAL-R gain prescription (45) and a 5-channel automatic gain controller (AGC). AGC settings were as follows: threshold = 50 dB SPL, anchor point = 120 dB SPL, attack = 10 ms, release = 50 ms. The sentence "No, I burst the balloon!" at

65 dB SPL were individually processed with the hearing aid simulator according to the audiometric profile of each participant. The word "balloon" was then extracted from the processed sentences and used as the stimulus in the aided condition. Across participants, the aided output level ranged from 66 dB to 85 dB for the sad stimulus and from 66 dB to 87 dB for the happy stimulus.

2.3. EEG recording

The Duet 2-channel AEP system (Intelligent Hearing Systems, Miami, FL) was used to record the FFR responses using four Ag/AgCI disposable disc electrodes. Two inverting electrodes were placed at right and left earlobes (A1 and A2), a noninverting electrode was positioned at vertex (Cz), and the ground electrode was placed at forehead (Fz). To improve electrode contact, the electrode sites were cleansed with an alcohol wipe and prepared using Nuprep Skin Prep Gel. Electrode impedances were maintained below 5 kΩ and inter-electrode differences in impendence were less than 2 kΩ. Two-channel FFR recordings were obtained over a 793.65 ms window, including a 40 ms pre-stimulus baseline. Signals were digitized with a 175 µs sampling period and bandpass filtered online between 30 and 1500 Hz.

The stimuli were presented monaurally through shielded insert earphones with alternating polarity, at a rate of 1.26 sweeps/s. A total of 3,000 sweeps (six blocks of 500) were recorded per emotion. Individual sweeps were rejected as artifacts if they exceeded 31 microvolts. For NH participants, stimuli were presented to the right ear. In participants with unilateral hearing loss, the impaired ear was tested, whereas for those with bilateral hearing loss, the right ear was used. A total of 3,000 (6 blocks of 500) responses per participant were obtained for each emotion condition.

For the NH group, the experiment was conducted in a single session lasting up to three hours. A brief pause was provided after each block to ensure participant comfort and maintain alertness, with a 15-minute break planned after six recording blocks to minimize fatigue. Emotional conditions were randomized across participants using a fully counterbalanced design. HI participants completed two recording sessions on two separate days, beginning with the unaided condition, followed by the aided condition. The break schedule and randomization protocol were consistent across groups. The experiment took place in a sound-attenuated and electromagnetically shielded booth. Participants sat in a reclining chair and were instructed to keep their eyes closed, minimize movement, and stay awake to reduce recording artifacts.

2.4. Data analysis

Both stimulus and response signals were segmented into 50 ms frames with a 40 ms overlap for F0 contour extraction. For the stimulus, the F0 contour was determined by calculating the median F0 within each frame using established speech pitch estimation methods in MATLAB 2024a and specifically the 'pitch' and 'pitchnn' commands. For the FFR response, the Harmonic Amplitude Summation (HAS) method was employed, in which each response frame was assigned a reference F0 based on the stimulus (46). To evaluate candidate F0s, a 100 Hz search range was applied around the reference, and a filter bank computed the summed amplitudes of each candidate and its harmonics in the response spectrogram. The F0 candidate with the most prominent summed amplitude was then selected as the final F0.

To evaluate the accuracy of F0 estimation, two metrics were used: (1) 5% accuracy, measuring the percentage of response frames where the estimated F0 is within 5% of the corresponding stimulus F0; and (2) Root-Mean-Square Error (RMSE), quantifying the mean deviation between the estimated response and reference stimulus F0 contours.

2.5. Statistical analysis

Statistical analyses and data visualization were performed using GraphPad Prism 10.0. Separate two-way mixed-design analysis of variances (ANOVAs) and Tukey's post hoc tests were first applied to compare 5% accuracy and RMSE across hearing conditions (Normal, HI-Unaided, HI-Aided) and emotion types (Happy, Sad), with emotion as a within-subjects factor and hearing condition as a between-subjects factor. To account for within-subject variability in the unaided vs. aided conditions for each evaluation metric, a paired t-test was consequently conducted to enhance statistical sensitivity. All *p-values* were two-sided, and the results with *p-values* < 0.05 were considered statistically significant. Statistical significance is shown as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

3. Results

3.1. Spectrographic representation and F0 contour of the FFR

To perform a preliminary assessment of the robustness of recorded responses, the averaged spectrograms across all participants (6 blocks × 14 participants) were obtained for each condition, as shown in Figure 2. The spectrograms were generated with a window length of 50 ms and an overlap of 40 ms. In the spectrograms of NH participants (Figure 2A-B), greater energy was observed at F0 and even 2F0, particularly in response to the sad stimulus. In contrast, HI participants showed weaker F0 in the unaided condition (Figure 2C-D), which appeared to partially improve with amplification (Figure 2E-F). These observations suggest possible variations in F0 neural encoding depending on stimulus and hearing condition.

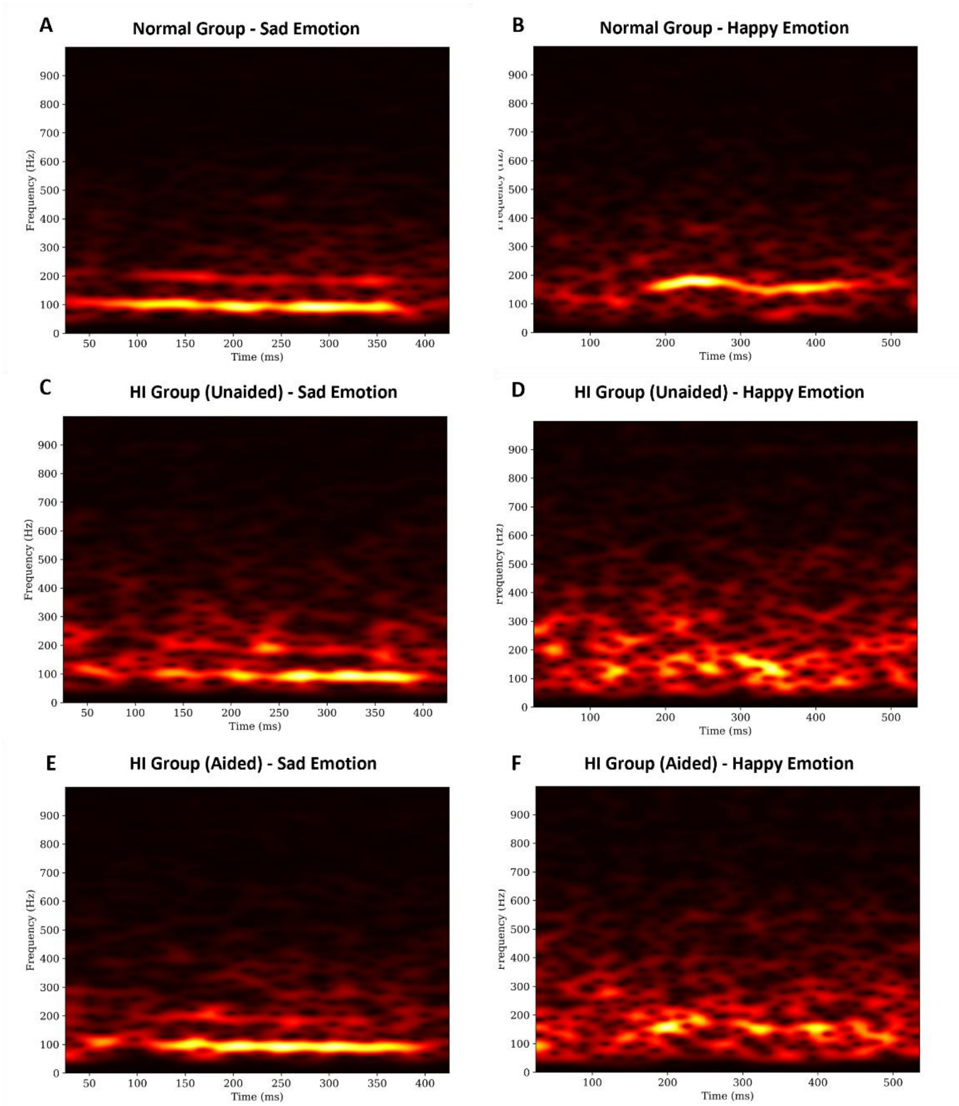


Figure 2. Spectrograms of the averaged responses across all participants (n = 14 per group) for each experimental condition: NH participants and sad emotion (A), NH participants and happy emotion (B), HI participants in the unaided condition and sad emotion (C), HI participants in the unaided condition and happy emotion (D), HI participants in the aided condition and sad emotion (E), and HI participants in the aided condition and happy emotion (F).

As the next step, the HAS algorithm was applied separately to each participant's FFR responses to extract the response F0 (Figure 3). Based on the spectrogram and extracted F0 contours, the response F0 closely tracked the stimulus F0 contour in NH group (Participant 4 shown in Figure 3 as a representative example). F0 tracking accuracy was reduced in the HI group when tested in the unaided condition, but showed improvement under the aided condition (Participant 2 shown as a representative example). Across all hearing conditions, F0 tracking was more accurate for sad stimulus than for happy one. Subsequently, both RMSE and 5% accuracy were calculated.

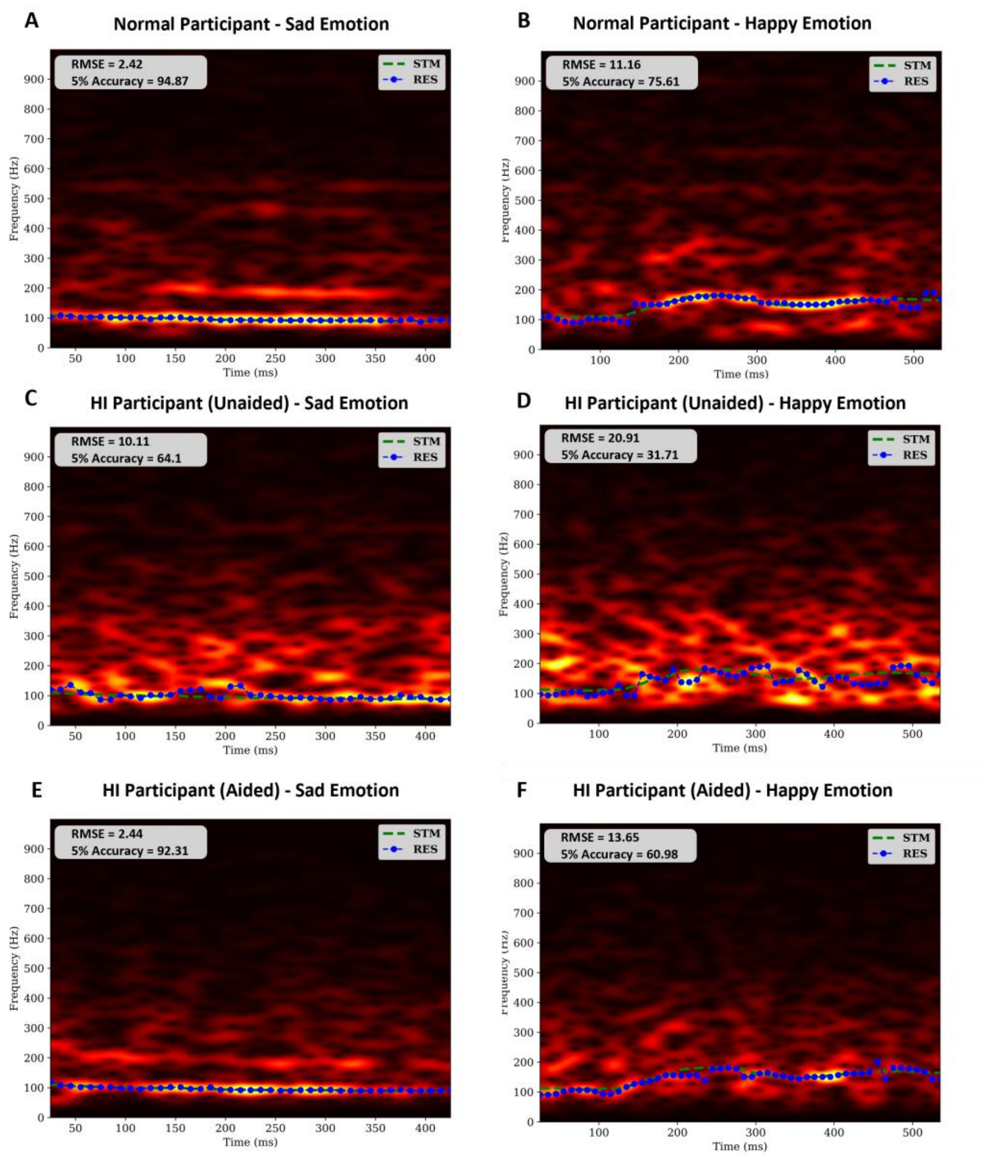


Figure 3. Spectrograms from Participant 4, a NH individual, and Participant 2, a HI individual tested under both unaided and aided conditions. F0 contours extracted from the stimulus waveform (green) and the evoked FFR (blue) are overlaid for the word "balloon". A and B display responses from the NH participant (Normal) to the sad and happy emotions, respectively. C and D correspond to the unaided HI participant's responses (HI-Unaided) to sad and happy stimuli. E and F show the same participant's responses under the aided condition (HI-Aided) for sad and happy emotions, respectively.

3.2. F0 Tracking Accuracy as a Function of Hearing Condition and Emotion Type

Table 1 and Figure 4 provide a descriptive overview of 5% accuracy and RMSE for both sad and happy stmuli across all three hearing conditions. For 5% accuracy, NH participants had the highest mean scores (sad: Mean = 72.9; happy: Mean = 48.2), followed by HI participants in the aided condition (sad: Mean = 66.4; happy: Mean = 38.8), and then in the unaided condition (sad: Mean = 55.5; happy: Mean = 28.2). Mean RMSE values showed a similar trend, with the lowest errors observed in the NH group particularly for the sad stimulus (sad: Mean = 6.3; happy: Mean = 16.7), followed by the HI group in the aided (sad: Mean = 8.0; happy: Mean = 18.6) and unaided (sad: Mean = 9.7; happy: Mean = 20.3) conditions.

Table 1. Mean 5% accuracy and RMSE for sad and happy versions of the word "balloon" across three hearing conditions.

	5% Accuracy (%)		RMSE (Hz)	
Emotion Type / Hearing Condition	Sad stimulus	Happy Stimulus	Sad stimulus	Happy Stimulus
Normal	72.9	48.2	6.3	16.7
HI-Unaided	55.5	28.2	9.7	20.3
HI-Aided	66.4	38.8	8.0	18.6

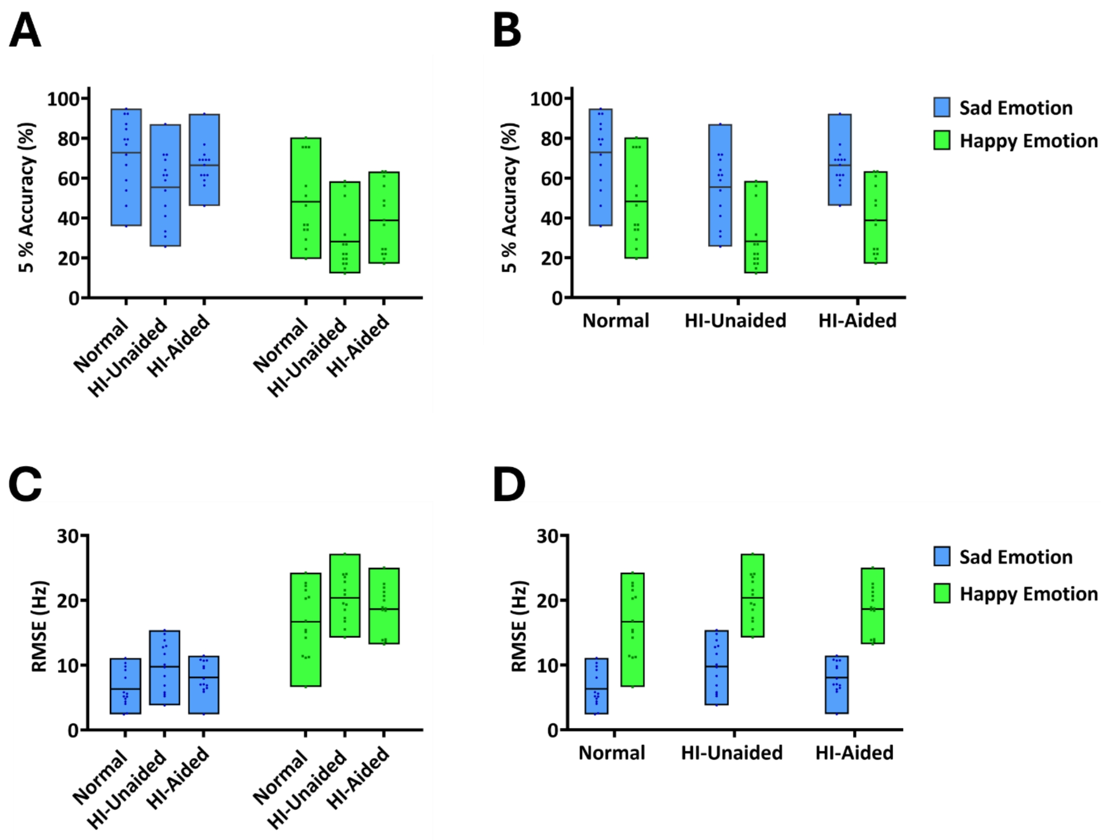


Figure 4. Floating bar plots showing mean, minimum, and maximum values of 5% accuracy (A-B) and RMSE (C-D) in response to sad (blue bar) and happy stimuli (green bar) for the word "balloon"; displayed across and within hearing conditions. Overlaid dots represent individual participant data points.

To statistically assess the influence of hearing condition (Normal, HI-Unaided, HI-Aided) and emotion type (sad, happy) on F0 processing, two-way mixed-design ANOVA revealed a significant main effect of hearing condition ($F (2,39) = 4.984$, $P = 0.0118$) and emotion ($F (1,78) = 50.16$, $p < 0.0001$) on 5% accuracy scores. However, the interaction effect was not significant ($F (2,39) = 0.2026$, $p= 0.8175$), indicating that hearing condition and emotion independently influence the encoding of F0 contour at the central auditory level. A parallel effect was also found for the RMSE metric, with a significant main effect of hearing condition ($F (2, 39) = 4.074$, $p = 0.0247$) and emotion ($F (1, 39) = 339.9$, $P < 0.0001$), but no interaction effect ($F (2, 39) = 0.017$, $p = 0.9825$).

3.3. Effect of hearing condition on F0 tracking accuracy

Tukey's post hoc comparisons demonstrated that FFR F0 tracking accuracy varied significantly across hearing conditions (Figure 5). Based on 5% accuracy results presented in Figure 5A, participants with hearing impairment in unaided condition obtained significantly lower 5% accuracy scores compared to the NH participants (Mean Difference = 18.7, 95% CI [4.2, 33.2], $p = 0.0087$). In contrast, the HI participants in the aided condition did not significantly differ from the NH participants (Mean Difference = 7.9, 95% CI [-6.5, 22.4], $p = 0.388$). Despite higher mean 5% accuracy scores in the aided condition, the initial post hoc comparisons revealed no significant difference between the HI-Aided and HI-Unaided conditions (Mean Difference = -10.8, 95% CI [-25.3, 3.6], $p = 0.1778$). Given that the unaided and aided conditions were assessed within the same participants, a paired t-test was subsequently conducted to eliminate within-subject variability and enhance statistical power (Figures 5B). The analysis revealed that hearing aid use significantly improved prosody processing compared to the unaided condition ($p = 0.0048$), despite the preliminary post hoc comparisons showing no significant difference.

With respect to RMSE results (Figures 5C), Tukey's post hoc comparisons revealed that RMSE values were significantly greater in the HI participants in the unaided condition compared to the NH individuals (Mean Difference = -3.5 Hz, 95% CI [-6.6, -0.5], $p = 0.0185$). In contrast, no significant difference was found between HI individuals in the aided condition and NH participants (Mean Difference = -1.8 Hz, 95% CI [-4.9, 1.2], $p = 0.313$), suggesting improvement in the HI-Aided condition. Similar to 5% accuracy results, Tukey's post hoc comparisons showed no significant differences between the HI-Aided and HI-Unaided conditions (Mean Difference = 1.7 Hz, 95% CI [-1.3, 4.7], $p = 0.3625$). However, a follow-up paired t-test, appropriate for the repeated-measures design within the HI group, showed that F0 tracking error was significantly reduced in the HI-Aided condition compared to the HI-Unaided condition ($p = 0.0368$).

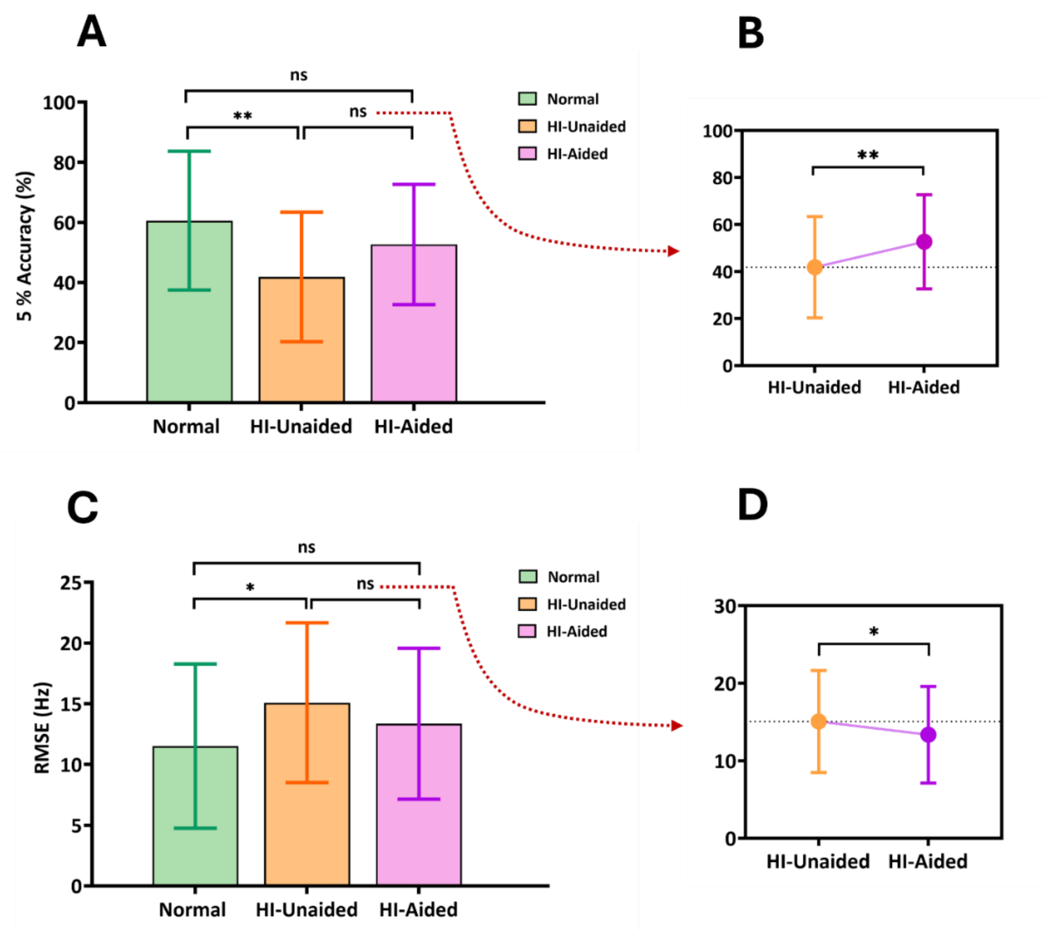


Figure 5. Effect of hearing condition on FFR F0 tracking accuracy in response to the word "balloon" assessed using either 5% accuracy (A-B) or RMSE (C-D) in the NH participants (Normal) and HI participants in both the unaided (HI-Unaided) and aided condition (HI-Aided). (A) and (C) show initial Tukey's post hoc pairwise comparisons across three hearing conditions. (B) and (D) present repeated-measures comparisons between the HI-Unaided and HI-Aided conditions using paired t-tests. The bars desplay the mean ± standard deviation (SD) for each condition. Statistical significance is shown as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

3.4. Effect of emotion on F0 tracking accuracy

Analysis of 5% accuracy revealed a significant main effect of emotion (Mean Difference = 26.5, $p < 0.0001$). The sad stimulus (Mean = 64.9) elicited significantly higher 5% accuracy scores than the happy emotion (Mean = 38.4), as shown in Figure 6A. A similar observation was made for RMSE (Mean Difference = -10.5, $p < 0.0001$), with significantly lower RMSE values for sad stimulus (Mean = 8.0) compared to happy stimulus (Mean = 18.5), showing that the F0 contour associated with the sad version of word “balloon” elicited more robust FFR responses (Figure 6B).

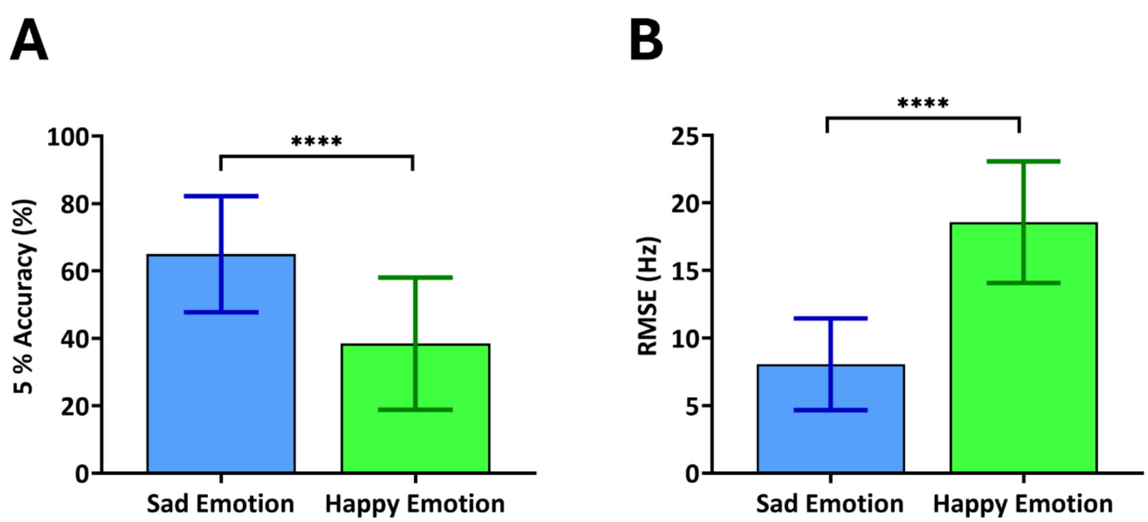


Figure 6. Effect of emotion on FFR F0 tracking accuracy assessed using either 5% accuracy (A) or RMSE (B) spoken by a male talker expressing sad (blue bar) and happy (green bar) emotions. The bars desplay the mean ± standard deviation (SD) for each condition. Statistical significance is shown as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

4. Discussion

This study investigated the potential deficit in prosody processing associated with hearing loss at the central level of the auditory system by examining FFRs to the word “balloon” spoken with sad and happy emotions. It also assessed if hearing aids can restore prosodic cues in emotional speech by comparing FFR tracking accuracy of F0 contours in HI participants under unaided and aided listening conditions.

Hearing loss degrades neural representation of prosodic cues

Our findings showed that hearing loss significantly disrupts neural encoding of F0 contours in emotional speech. Although the FFR was capable of tracking F0 variations over time in response to for both the sad and happy "balloon" stimuli in HI individuals, both 5% accuracy and RMSE were significantly different when compared to the NH condition. One possible explanation for the impaired F0 tracking accuracy observed in individuals with hearing loss is reduced audibility, particularly given that amplification led to improved F0 tracking in our study. Reduced audibility, regardless of the affected frequency range, may limit access to prosodic cues. Although prosody is primarily conveyed through low-frequency components, even high-frequency hearing loss can impact prosodic processing, as previous studies have shown that high-frequency harmonics contribute to the representation of speech periodicity (47). This aligns with previous FFR studies reporting degraded neural representation of F0 following auditory deprivation.

In a study conducted by Ananthakrishnan et al., neural representation of both the envelope (F0) and temporal fine structure (TFS) at first formant (F1) was degraded in adults with mild to moderate hearing loss when stimuli were presented at equal sound pressure levels (SPLs) as those used for NH individuals, in response to a synthetic steady-state vowel /u/ (30). The decreased F0 and F1 magnitudes was attributed to reduced audibility and limited access to high-frequency unresolved harmonics. Similarly, Hao et al. (29) also reported significantly reduced F0 amplitudes and degraded stimulus-response correlation in HI individuals in response to a 170-ms /da/ stimulus presented at 85 dB SPL compared to NH peers, both in quiet and in the presence of white noise. Having ruled out the contribution of age-related neural degeneration, as no significant correlation was found between age and FFR components, the authors attributed the observed impairments to diminished audibility due to peripheral hearing loss.

On the other hand, some other studies have indicated that hearing loss can enhance F0 representation. In a study conducted by Anderson et al., older adults with mild to moderate hearing loss exhibited greater envelop encoding, as reflected by increased amplitude of F0, H1, and H2, in response to synthesized /da/ stimulus compared to age-matched NH controls (28). Relying on animal models, the authors proposed that hearing loss can lead to decreased inhibition and increased excitability throughout the central auditory pathway. As another possibility, they considered that broader tuning curve bandwidths resulting from hearing impairment may lead to enhanced envelope modulation. Similarly, Koravand et al. also found an increase in the amplitude of the fundamental frequency (F0) component of the FFR in children with mild to moderately severe hearing loss (31).

In contrast, Presacco et al. reported no significant differences in either FFR amplitude or stimulus-to-response correlation between older adults with normal hearing and those with hearing impairment in response to a 170 ms synthesized /da/ syllable with a F0 of 100 Hz presented at 75 dB SPL (32). However, when both older groups were compared to younger adults, they exhibited weaker FFR amplitudes, leading the authors to attribute these deficits primarily to age-related temporal processing rather than hearing loss. In contrast, our study revealed that there is a significant difference in F0 tracking accuracy between older adults with and without hearing loss. The difference between our results and this study may be explained by methodological factors. Unlike their use of a synthesized /da/ with steady F0 and RMS amplitude of the FFR signal in the 70–2000 Hz range, we employed a natural speech word with time-varying F0 and focused specifically on the FFR at F0, with the aim of measuring FFR tracking accuracy. The dynamic F0 may have imposed greater temporal demands on the auditory system than the steady F0 used by Presacco et al., which may account for the increased vulnerability to hearing loss observed in our study. Findings by Moris et al. further suggested that both aging and hearing loss contribute to reduced neural synchrony at the brainstem level (33). Using 120-ms tone glides centered around 500 Hz with dynamic changes in F0 (⅓, ⅔, or 1 octave) presented at 80 dB SPL, they demonstrated a decreased likelihood of older and HI individuals producing a detectable FFR, and among those with a detectable response, reduced pitch strength and stimulus–response coherence was also observed.

F0 tracking deficit associated with hearing loss can be partially restored by hearing aid

The present study showed that HI individuals with amplification exhibited a significant improvement in 5% accuracy and a significant reduction in RMSE values, indicating enhanced F0 tracking compared with the unaided condition. These findings suggest that restoring audibility through amplification may enhance the robustness of F0 encoding. Although hearing aids typically apply greater gain to high frequency regions which are more vulnerable to hearing loss, this can still support improved tracking of the F0 contour, a primarily low-frequency feature. Since high-frequency harmonics can represent periodicity (47), hearing aids by amplifying these high-frequency components may indirectly enhance F0 tracking in individuals with hearing loss. However, it is worth noting that, because this study used a hearing aid simulator rather than real devices, some characteristics of real hearing aids, such as processing delays and possible comb filtering effects, may not be reflected in our results. These characteristics could disrupt neural phase locking and distort the speech envelope, which is important for perceiving emotional prosody. Consequently, real hearing aids might produce smaller or less consistent improvements due to the effect of processing delay. Even with this limitation, our results reinforced the view that the FFR can serve as a valuable tool in hearing aid research and fitting, aligning with previous findings that highlight clinical potential of FFR for objective assessment and personalized hearing aid adjustment (48-50).

In line with our results, using a naturally spoken speech token /susaʃi/, Easwar et al. showed that use of amplification in adults with hearing loss led to significantly enhanced representation of F0 (34), as reflected by an increased number of FFR detections and greater response amplitude at F0. Ananthakrishnan et al. also found that when stimuli were presented at equal Sensation Levels (SLs) to both NH and HI participants, F0 representation in HI group closely resembled that observed in NH listeners (30). This observation was taken to imply that amplification can partially restore access to envelope cues. Likewise, relying on the improved audibility and more precise temporal processing provided by hearing aid, BinKhamis et al. reported greater F0 encoding amplitude in response to the amplified 40 ms /da/ stimulus with rising F0, compared to unaided condition (35).

On the other hand, Anderson et al. found that HI individuals exhibited greater envelope encoding than their NH peers when listening to an amplified synthesized /da/ stimulus processed through the NAL-R algorithm, a pattern that was also observed in the unaided condition (28). They interpreted the presence of enhanced envelope encoding, even after restoring audibility through amplification, as indicative of neural alterations arising from central auditory mechanisms rather than peripheral deficits alone in HI individuals. Interestingly, Fu et al. demonstrated significantly reduced pitch strength and stimulus–response coherence in FFRs to both steady tones and linearly frequency-modulated (FM) sweeps in individuals with mild to severe hearing loss, relative to NH participants (36). Using non-speech stimuli presented at individually adjusted levels to equate audibility and match loudness perception, they attributed these deficits to impaired neural synchrony rather than reduced audibility, although the authors mentioned that the effect of sound level could not be fully excluded.

F0 representation in sFFR responses differs by emotion type

To the best of our knowledge, this is the first study to examine the neural correlates involved in encoding emotion-related F0 cues in hearing-impaired individuals, with and without hearing aid amplification, using the FFR. As mentioned in the previous section, a few studies have assessed neural representation of time-varying F0 in individuals with hearing loss; however, no one have explored it within the emotional context of speech. Based on sFFR responses to the word "balloon" spoken by one male talker, our results revealed that F0 representation varies across emotion types, with the sad emotion being more robustly encoded than the happy emotion across all hearing conditions, as reflected by higher 5% accuracy and lower RMSE values.

This result could be linked to the acoustic characteristics of different emotions. In this study, the Sad version of "balloon" was acoustically defined by a lower mean F0, a falling pitch contour, and a slower rate of F0 changes compared to happy emotion (Figure 1). Previous FFR studies in normal-hearing individuals have also shown that the magnitude of the F0 response is enhanced when stimulus pitch is lower and the rate of F0 change is slower (51-53). In our study, with the same talker throughout, the sad stimulus had lower mean F0 of 95.1 Hz (Min F0 = 88.8 Hz, Max F0 = 103.3 Hz) with a gradual falling contour, whereas the happy emotion exhibited a higher mean F0 of 153.4 Hz (Min F0 = 109.5, Max F0 = 182.1 Hz) with a steeper rising and falling slope. Averaged spectrograms across participants also showed greater energy at F0 and even at 2F0 for sad stimuli, particularly in the normal and aided groups (Figure 2). The acoustic profile of the sad stimuli likely facilitated more robust F0 encoding, consistent with previous findings that lower pitch and slower F0 dynamics enhance FFR responses. It should be taken into consideration that these results apply only to happy and sad emotions, and further research with a broader range of emotions, different words, and additional talkers is needed to determine the generalizability of these findings.

Despite insufficient electrophysiological evidence on the effect of hearing loss and hearing aid processing on the neural encoding of emotional cues, a few behavioral studies have shown that hearing loss and hearing aids influence the perception of emotional speech. In a study conducted by Goy et al., older adults with hearing loss performed worse than their NH peers on emotion identification task, and emotion types did not affect their performance (18). Their follow-up study indicated that the use of hearing aids did not significantly improve emotion identification performance in the HI individuals (54). They also reported that different amplification strategies in simulated hearing aid processing could alter the intensity variation and spectral centre-of-gravity in emotional speech (23). These effects were emotion-specific and could make some emotions less distinguishable based on these acoustic cues. Most & Aviner have also found that hearing aid users performed significantly worse than NH individuals when emotions were presented in the auditory-only condition (16). Among the six tested emotions, sadness was the most accurately identified emotion by the hearing aid users, followed by happiness, surprise, fear, anger, and disgust. The authors suggested that acoustic features of sadness such as slower rate, lower average F0, and longer duration could make it more accessible through hearing aids.

In conclusion, the current study expanded our previous findings (Chapter3) demonstrating that phase-locked neural activity underlying the FFR is sufficiently dynamic to encode time-varying acoustic features embedded in speech prosody. It further revealed that hearing loss can compromise the synchronized neural firing necessary for tracking F0 contours, whereas hearing aids can partially preserve this pitch-related information. In addition to hearing status, the type of emotion expressed in the eliciting stimuli can influence the extent to which the sFFR reflects prosodic characteristics, likely through differences in their underlying acoustic features. Considering the importance of prosody in daily communication, these findings highlight the need to identify and rehabilitate prosody deficits in HI individuals. The sFFR can provide an objective tool to assess how prosody-related acoustic cues are processed. However, since perception eventually relies on higher level cortical processing, future studies should incorporate complementary behavioral assessments to provide a more comprehensive understanding of the effect of hearing loss and the benefit of hearing aids on the perception of prosody.

References

1. D. B. Pisoni, R. E. Remez, *The handbook of speech perception* (Wiley Online Library, 2005).
2. L. J. Raphael, G. J. Borden, K. S. Harris, *Speech science primer: Physiology, acoustics, and perception of speech* (Lippincott Williams & Wilkins, 2007).
3. S. J. L. Mozziconacci, Speech variability and emotion: Production and perception. (1998).
4. V. Raithel, M. Hielscher-Fastabend, Emotional and linguistic perception of prosody: Reception of prosody. *Folia Phoniatrica et Logopaedica* **56**, 7-13 (2004).
5. S. Nooteboom (1997) The prosody of speech: Melody and rhythm. The handbook of phonetic sciences. Hardcastle WJ, Laver J. (Oxford: Basil Blackwell).
6. R. Banse, K. R. Scherer, Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* **70**, 614 (1996).
7. J. Pickett, S. R. Morris (2000) The acoustics of speech communication: Fundamentals, speech perception theory, and technology. (Acoustical Society of America).
8. T. Bänziger, K. R. Scherer, The role of intonation in emotional expressions. *Speech Commun.* **46**, 252-267 (2005).
9. D. Grandjean, Brain networks of emotional prosody processing. *Emot. Rev.* **13**, 34-43 (2021).
10. A. Pannese, D. Grandjean, S. Frühholz, Subcortical processing in auditory communication. *Hear. Res.* **328**, 67-77 (2015).
11. D. S. Dalton *et al.*, The impact of hearing loss on quality of life in older adults. *The gerontologist* **43**, 661-668 (2003).
12. J. Hall *et al.* (2015) Handbook of clinical audiology. (Lippincott Williams & Wilkins: Philadelphia, PA, USA).
13. E. M. Picou *et al.*, Hearing, emotion, amplification, research, and training workshop: Current understanding of hearing loss and emotion perception and priorities for future research. *Trends Hear.* **22**, 2331216518803215 (2018).
14. J. H.-Y. Ng, A. Y. Loke, Determinants of hearing-aid adoption and use among the elderly: A systematic review. *Int. J. Audiol.* **54**, 291-300 (2015).
15. N. A. Lesica, Why do hearing aids fail to restore normal auditory perception? *Trends Neurosci.* **41**, 174-185 (2018).
16. T. Most, C. Aviner, Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *JDSDE* **14**, 449-464 (2009).

17. M. Mazaheryazdi, M. Aghasoleimani, M. Karimi, P. Arjmand, Perception of musical emotion in the students with cognitive and acquired hearing loss. *Iran. J. Child Neurol* **12**, 41 (2018).

18. H. Goy, M. K. Pichora-Fuller, G. Singh, F. A. Russo, Perception of emotional speech by listeners with hearing aids. *Can. Acoust.* **44** (2016).

19. A. Boothroyd, Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *J. Speech. Lang. Hear. Res.* **27**, 134-144 (1984).

20. T. Irino, Y. Hanatani, K. Kishida, S. Naito, H. Kawahara, Effects of age and hearing loss on speech emotion discrimination. *Sci. Rep.* **14**, 18328 (2024).

21. J. Schmidt, D. Herzog, O. Scharenborg, E. Janse (2016) Do hearing aids improve affect perception? in *Physiology, psychoacoustics and cognition in normal and impaired hearing* (Springer International Publishing), pp 47-55.

22. M. Karimi-Boroujeni, H. R. Dajani, C. Giguère, Perception of prosody in hearing-impaired individuals and users of hearing assistive devices: An overview of recent advances. *J. Speech. Lang. Hear. Res.* **66**, 775-789 (2023).

23. H. Goy, M. K. Pichora-Fuller, G. Singh, F. A. Russo, Acoustic analysis of emotional speech processed by hearing aids. *Can. Acoust.* **45**, 178-179 (2017).

24. H. N. Arefi, S. J. Sameni, H. Jalilvand, M. Kamali, Effect of hearing aid amplitude compression on emotional speech recognition. *Audit. Vestib. Res.* **26**, 223-230 (2017).

25. E. Skoe, N. Kraus, Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* **31**, 302-324 (2010).

26. E. B. Coffey *et al.*, Evolving perspectives on the sources of the frequency-following response. *Nat. Commun.* **10**, 5036 (2019).

27. J. Krizman, N. Kraus, Analyzing the FFR: A tutorial for decoding the richness of auditory function. *Hear. Res.* **382**, 107779 (2019).

28. S. Anderson, A. Parbery-Clark, T. White-Schwoch, S. Drehobl, N. Kraus, Effects of hearing loss on the subcortical representation of speech cues. *J. Acoust. Soc. Am.* **133**, 3030-3038 (2013).

29. W. Hao *et al.*, Effects of phase-locking deficits on speech recognition in older adults with presbycusis. *Front. Aging Neurosci.* **10**, 397 (2018).

30. S. Ananthakrishnan, A. Krishnan, E. Bartlett, Human frequency following response: neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss. *Ear Hear.* **37**, e91-e103 (2016).

31. A. Koravand, R. Al Osman, V. Rivest, C. Poulin, Speech-evoked auditory brainstem responses in children with hearing loss. *Int. J. Pediatr. Otorhinolaryngol.* **99**, 24-29 (2017).

32. A. Presacco, J. Z. Simon, S. Anderson, Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PLoS One* **14**, e0213899 (2019).

33. M. R. Molis, W. J. Bologna, B. M. Madsen, R. K. Muralimanohar, C. J. Billings, Frequency following responses to tone glides: effects of age and hearing loss. *JARO.* **24**, 429-439 (2023).

34. V. Easwar, D. W. Purcell, S. J. Aiken, V. Parsa, S. D. Scollie, Evaluation of speech-evoked envelope following responses as an objective aided outcome measure: Effect of stimulus level, bandwidth, and amplification in adults with hearing loss. *Ear Hear.* **36**, 635-652 (2015).

35. G. BinKhamis, A. Elia Forte, T. Reichenbach, M. O'Driscoll, K. Kluk, Speech auditory brainstem responses in adult hearing aid users: Effects of aiding and background noise, and prediction of behavioral measures. *Trends Hear.* **23**, 2331216519848297 (2019).

36. Z. Fu, H. Yang, F. Chen, X. Wu, J. Chen, Brainstem encoding of frequency-modulated sweeps is relevant to Mandarin concurrent-vowels identification for normal-hearing and hearing-impaired listeners. *Hear. Res.* **380**, 123-136 (2019).

37. G. M. Bidelman, J. T. Gandour, A. Krishnan, Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cogn.* **77**, 1-10 (2011).

38. A. Krishnan, C. H. Suresh, J. T. Gandour, Tone language experience-dependent advantage in pitch representation in brainstem and auditory cortex is maintained under reverberation. *Hear. Res.* **377**, 61-71 (2019).

39. A. Krishnan, Y. Xu, J. Gandour, P. Cariani, Encoding of pitch in the human brainstem is sensitive to language experience. *Cogn. Brain Res.* **25**, 161-168 (2005).

40. V. Marian, H. K. Blumenfeld, M. Kaushanskaya, The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *J. Speech. Lang. Hear. Res.* **50**, 940-967 (2007).

41. G. Comeau, A. Koravand, M. Swirp, Prevalence of hearing loss among university music students. *Can. Acoust.* **46**, 37-51 (2018).

42. D. W. Molloy, T. I. Standish, A guide to the standardized Mini-Mental State Examination. *Int. Psychogeriatr.* **9**, 87-94 (1997).

43. K. Zhou, B. Sisman, R. Liu, H. Li, Emotional voice conversion: Theory, databases and esd. *Speech Commun.* **137**, 1-18 (2022).

44. S. Ollivier, H. Nélisse, J. Voix, On the use of wide dynamic range compression and other algorithms to improve hearing protection of workers with hearing impairment: A preliminary study on speech intelligibility. *Can. Acoust.* **51**, 110-111 (2023).

45. H. Dillon, *Hearing Aids* (Thieme, 2001).

46. Sadeghkhani, S., Karimi Boroujeni, M., Dajani, H. R., Seydnejad, S. R. & Giguère, C. A method for pitch tracking in the frequency following response using harmonic amplitude summation filterbank. *JASA Express Lett.* **5** (2025).

47. M. Laroche, H. R. Dajani, F. Prévost, A. M. Marcoux, Brainstem auditory responses to resolved and unresolved harmonics of a synthetic vowel in quiet and noise. *Ear Hear.* **34**, 63-74 (2013).

48. H. R. Dajani, B. P. Heffernan, C. Giguère (2013) Improving hearing aid fitting using the speech-evoked auditory brainstem response. in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE), pp 2812-2815.

49. B. Heffernan, H. R. Dajani, C. Giguère, Towards developing a brain-computer interface for automatic hearing aid fitting based on the speech-evoked frequency following response. *Neurotechnix 2017-extended abstracts-neurotechnix*, 3-4 (2017).

50. L. Bellier *et al.*, Speech auditory brainstem response through hearing aid stimulation. *Hear. Res.* **325**, 49-54 (2015).

51. A. Krishnan, C. H. Suresh, J. T. Gandour, Changes in pitch height elicit both language-universal and language-dependent changes in neural representation of pitch in the brainstem and auditory cortex. *Neuroscience* **346**, 52-63 (2017).

52. C. J. Billings, W. J. Bologna, R. K. Muralimanohar, B. M. Madsen, M. R. Molis, Frequency following responses to tone glides: Effects of frequency extent, direction, and electrode montage. *Hear. Res.* **375**, 25-33 (2019).

53. J. Van Canneyt, J. Wouters, T. Francart, From modulated noise to natural speech: The effect of stimulus parameters on the envelope following response. *Hear. Res.* **393**, 107993 (2020).

54. H. Goy, M. K. Pichora-Fuller, G. Singh, F. A. Russo, Hearing aids benefit recognition of words in emotional speech but not emotion identification. *Trends Hear.* **22**, 2331216518801736 (2018).

Chapter 5. Behavioral evidence of prosody perception deficits in hearing-impaired individuals with and without simulated hearing aid processing

Behavioral evidence of prosody perception deficits in hearing-impaired individuals with and without simulated hearing aid processing

Maryam Karimi-Boroujeni [1], Christian Giguère [1], Solenn Ollivier [2], Hilmi R. Dajani [3]

[1] School of Rehabilitation Sciences, University of Ottawa, Ottawa, ON, Canada

[2] Université du Québec, École de technologie supérieure, Montréal, Québec, Canada

[3] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, Canada

Abstract

Background: While prosody perception is a crucial component of effective communication and carries emotional information in speech, it remains largely neglected in the assessment and management of individuals with hearing impairment. Moreover, the limited existing literature has yet to reach a consensus on the effects of hearing loss and hearing aid use on emotional speech perception.

Objectives: This study aimed to examine the prosodic abilities of hearing-impaired individuals under cognitively demanding conditions and to determine whether they engage cognitive resources to the same extent as normal-hearing peers during emotional processing. It also evaluated whether hearing aid amplification can compensate for potential prosodic deficits and reduce the cognitive load required for processing prosodic information. Finally, it assessed whether speech-in-noise performance can predict prosodic processing ability in hearing-impaired individuals.

Method: An emotion identification test paired with a secondary free-recall task was administered to normal-hearing participants and to hearing-impaired individuals under both unaided and aided conditions. Response accuracy, reaction time, and the number of recalled words were recorded. Speech-in-noise perception was also assessed using the QuickSIN test, and SNR loss was calculated.

Results: Significantly lower response accuracy in hearing-impaired individuals indicated that hearing loss degraded emotional prosody processing, with effects varying by emotion type. The reduced number of recalled words among hearing-impaired participants further suggested increased cognitive effort in perceiving emotional speech. Although amplification improved performance, emotional speech perception remained below normal-hearing levels, suggesting that hearing aids may partially restore emotional cues in speech. A significant negative correlation was

also observed between QuickSIN scores and emotion identification accuracy, suggesting that poor speech-in-noise performance may reflect prosodic processing difficulties.

Keywords: emotional prosody, speech perception, hearing loss, hearing aids, speech-in-noise

1. Introduction

Hearing loss, a growing global health concern projected to affect 2.5 billion individuals by 2050 [1], significantly impairs speech recognition and can impact social interactions, cognitive function, and overall quality of life [2-5]. This emphasizes the need to maintain effective speech communication, which requires accurately perceiving all speech components [6]. However, in hearing research and audiology clinical practice, from evaluation to rehabilitation and hearing assistive technologies, the primary focus is on the linguistic aspects of speech such as word and sentence recognition. In contrast, the paralinguistic features, which complement linguistic information, have not been examined as extensively. These features include the prosodic elements of speech such as intonation, stress, and rhythm, which help the listener understand the speaker's emotions and intentions [7-9]. Acoustically, these emotional patterns are produced by variations in fundamental frequency (F0), intensity, and duration [6, 10].

Individuals with hearing impairment often experience reduced audibility, degraded spectro-temporal resolution, and impaired auditory pathway integrity that can disrupt the ascending information flow [11-13], preventing precise encoding of the prosody-related acoustic cues [14-16]. Some evidence has also shown that, through the feedback flow in the descending pathway, hearing impairment can reduce activation in the amygdala and limbic system, key hubs for emotion processing, and alter connectivity between auditory regions and cognitive network [17, 18]. However, research findings on the performance of hearing-impaired individuals in processing speech prosody are inconsistent. While some studies have reported no deficits in the perception of emotional prosody among people with hearing loss [19-22] , other studies have found that hearing-impaired individuals perform worse than normal-hearing counterparts, influenced by the hearing loss severity and the type of emotions being processed [23-27].

This inconsistency is also evident in studies investigating the impact of hearing aids on restoring emotional information. Using emotional speech sentences, some studies have reported that emotion recognition does not improve when hearing-impaired individuals use hearing aids [24, 28, 29]. Other studies have also shown that hearing aids may alter acoustic characteristics critical for emotion perception, such as intensity variation and spectral cues, and thereby weaken the distinction among various emotions [30, 31]. However, some studies have suggested that hearing aid use can significantly compensate for emotional perception deficits in individuals with hearing loss, as reflected by better performance on emotion identification tests and higher arousal ratings in emotion judgment tasks [23, 27].

One possible reason contributing to such mixed results in the aforementioned studies might be that they have solely focused on emotion identification in isolation, without exploring how hearing-impaired individuals interpret prosodic speech in cognitively demanding situations that resemble real-life challenges. This becomes particularly important when considering that effective daily communication relies on cognitive skills that enable listeners to simultaneously process speech signals, retain them, and respond appropriately [32-34]. By introducing a recall component to the emotion identification task, the present study aimed to assess how and to what extent hearing loss can degrade emotional prosody perception under a cognitively demanding condition. Furthermore, this study evaluated the impact of simulated hearing aid amplification on emotion perception ability among hearing-impaired individuals, with a particular focus on whether amplification reduces the cognitive load required for processing prosodic information.

In addition, given the absence of prosody assessment in standard audiologic test batteries, we explored whether any routine audiology tests in clinical settings could serve as a predictor for hearing-impaired individuals' performance on prosody perception. We selected the Quick Speech-in-Noise (QuickSIN) test as our measure of interest. The QuickSIN test is a reliable clinical tool consisting of twelve lists of six sentences, each presented with multi-talker babble noise to simulate real-world listening conditions [35]. It requires listeners to selectively attend to the target talker while suppressing babble talkers [35]. This task involves auditory processing mechanisms such as spectral resolution, temporal processing, and pitch tracking, alongside higher-order cognitive functions, which are also engaged in emotion perception [35-37]. The use of naturally dynamic sentences [35], as opposed to static words in conventional speech audiometry, aligns more with the demands of emotion identification tasks. Adding the recall task to the emotion identification test further increases cognitive load, making it more comparable to the complexity of speech perception in noisy environments. We investigated the degree to which the QuickSIN can predict performance on the emotion identification task. Determining the correlation between these two tests may give audiologists insight about the possibility of deficits in prosody perception when patients perform poorly on speech perception tasks in noisy environments.

2. Methods and materials

2.1. Participants

A total of 28 adults, aged 18–65 years, participated in this study, comprising 14 normal-hearing individuals (mean age = 52.4 years) and 14 age-matched hearing-impaired individuals (mean age = 52.1 years). Otoscopy, tympanometry, and pure-tone audiometry were conducted to assess overall hearing function. Normal hearing (NH) was defined as audiometric pure-tone thresholds $\leq$ 25 dB HL from 250 to 8000 Hz in both ears. The hearing-impaired (HI) group included individuals with unilateral or bilateral mild to moderately severe sensorineural hearing loss (SNHL), defined by air conduction thresholds $>$ 30 dB HL and $\leq$ 70 dB HL at octave frequencies of 250 to 8000 Hz in at least one ear. Their hearing thresholds are shown in Table S3 (Appendix L). Following the administration of the Language Experience and Proficiency Questionnaire (LEAP-Q) [38], the Demographic and Music Experience Questionnaire [39], and the Standardized Mini-Mental State Examination (SMMSE) [40], participants who were native or near-native English speakers, non-musicians, and non-tonal language speakers were recruited for the study (Appendix C-E).

2.2. Data collection

The QuickSIN test was conducted at two presentation levels (55 dB and 70 dB SPL). At each level, participants completed two lists, and the average signal-to-noise ratio (SNR) loss was used to represent their speech-in-noise performance. QuickSIN results are available in Table S4 in Appendix M.

To assess emotion perception, an emotion identification test was structured based on Goy et al. [24] study and further developed to meet the needs of the present study. The test consisted of 105 English sentences selected from the Toronto Emotional Speech Set (TESS) [41], which comprises 200 words embedded within the carrier phrase "Say the word". These sentences were recorded from one older and one younger female speaker, each expressing seven emotions: angry, disgust, fear, happy, pleasant surprise (PS), sad, and neutral. For the present study, only the recordings from the younger female speaker were used. Seven lists were created, each containing 105 sentences, 15 for each of the seven emotions. Although the sentence content was the same across all seven lists, the emotion used to express each sentence varied from one list to another. For each participant, the list was selected in a counterbalanced order.

The participants were asked to listen to sentences and, immediately after each one, choose one of seven emotion options by clicking the corresponding button on a computer screen. Accurate responses and reaction times were recorded both for overall performance and separately for each emotion. A secondary free-recall task was also incorporated into the emotion identification task.

Participants were instructed to remember the final word of each sentence. After every five sentences, they were asked to recall the last word from each of the five sentences without any cues or prompts. The total number of correctly recalled words was then recorded. For NH participants, the right ear was considered as the test ear. For participants with unilateral hearing loss, the test was conducted in the impaired ear; for those with bilateral loss, the right ear was tested.

Sentences were presented at an average level of 65 dB SPL via an insert earphone to the participant's test ear, with the stimuli delivered from a laptop through the audiometer. To maintain the natural variation in vocal intensity across emotions, the sentences were not individually normalized. Participants in the HI group completed the test twice; once with the original and unamplified sentences, and once with amplified sentences representing the aided condition. A software-based hearing aid simulator [42], developed in MATLAB (MathWorks, Natick, MA), was used to amplify the sentences based on each participant's hearing loss. The simulator applied wide dynamic range compression (WDRC) based on an algorithm combining a NAL-R gain prescription[43] and a 5-channel automatic gain controller (AGC). In this study, the AGC was set to a compression threshold of 50 dB SPL, an anchor point of 120 dB SPL, and attack and release times of 10 ms and 50 ms, respectively. Performance with the original stimuli represented the unaided condition (HI–Unaided), and performance with the amplified stimuli by HA simulator represented the aided condition (HI–Aided). Different sentence lists were used for the aided and unaided conditions to avoid repetition for each participant. To minimize learning effects, half of the HI participants began with the aided condition followed by the unaided condition, while the other half completed the conditions in the reverse order. To ensure comparable testing conditions, half of the NH participants were tested twice using the original stimuli with two different lists, and their performance on the second trial was recorded.

This study was conducted in accordance with the guidelines of the Institutional Research Ethics Board at the University of Ottawa, and written informed consent was obtained from all participants (Appendix B).

2.3. Statistical analysis

All statistical analyses and graphical representations were implemented using GraphPad Prism 10.0. Separate one-way analyses of variance (ANOVAs), each followed by Tukey's post hoc test, were first conducted to compare overall emotion identification performance across three hearing conditions (Normal Hearing, HI-Unaided, and HI-Aided) for each outcome: total correct responses, overall reaction time (RT), and the total number of final words recalled. Consequently, to account for repeated measurements within the HI group, repeated measures t-tests were also conducted between the HI-Unaided and HI-Aided conditions as complementary analyses for each outcome.

To evaluate the effects of emotion types and hearing condition on response accuracy, a two-way mixed ANOVA (emotion type x hearing condition) was first conducted with emotion (7 levels) as a within-subjects factor and hearing condition (normal, HI-Unaided, HI-Aided) as a between-subjects factor. Tukey's post hoc test was then applied for pairwise comparisons. To improve sensitivity and account for repeated measurements within the HI group, response accuracy was consequently analyzed using repeated-measures ANOVA across emotions over the HI-Unaided and HI-Aided conditions. To analyze the effects of emotion types and hearing condition on reaction time, a mixed-effects model was applied with Restricted Maximum Likelihood (REML) estimation to address missing values while preserving the within-subjects structure. Tukey's test was then conducted for post hoc pairwise comparisons. A follow-up repeated measures ANOVA was subsequently conducted over the HI group results to more accurately examine the effect of hearing aid on reaction time across the same emotional categories.

Finally, Pearson correlation and simple linear regression analyses were performed within the HI-unaided condition to examine associations between the performance of HI individuals on emotion identification task and Quick SIN scores. All *p-values* were two-sided, and the results with *p-values* < 0.05 were considered statistically significant. Statistical significance is shown as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

3. Results

3.1. Group differences in overall response accuracy, reaction time, and recall performance in the emotion identification test

Overall Response Accuracy. The total number of correct responses (out of max. 105), aggregated across all emotion categories, differed significantly among hearing conditions, as shown by a one-way ANOVA ($F(2, 39) = 10.08$, $p < 0.0003$). Post hoc comparisons using Tukey's correction indicated that participants with normal hearing achieved significantly higher accuracy scores (Mean = 89.6) than HI group in both the unaided (Mean = 57.6, $p = 0.0003$) and aided (Mean = 65.5, $p = 0.0066$) conditions (Figure 1A). No significant difference was observed between the HI-Unaided and HI-Aided conditions ($p = 0.5456$) in preliminary post hoc comparisons. However, since the unaided and aided conditions were assessed within the same HI participants, a complementary repeated measures t-test comparing these two conditions (Figure 1B) was also performed and revealed no statistically significant difference in response accuracy ($p = 0.1677$), despite a higher mean score in the aided condition.

Overall Reaction Time. Total reaction time (RT) did not significantly differ across hearing conditions ($F(2, 39) = 0.02269$, $p = 0.9776$), indicating that the processing speed of emotion identification was similar between the NH group and the HI group in both unaided and aided conditions (Figure 1C). The follow-up repeated measures t-test between the HI-Unaided and HI-Aided conditions also showed no significant difference ($p = 0.7471$), despite a lower mean RT in the HI-aided than the HI-unaided condition (Figure 1D).

Free Recall Performance. The ANOVA revealed a significant main effect of hearing condition ($F(2, 39) = 5.323$, $p = 0.0090$) on free recall performance (out of max. 105). Tukey's test indicated that NH participants (Mean=79.9) obtained higher recall scores compared to their HI peers in both the unaided (Mean = 64.71, $p = 0.0101$) and aided conditions (Mean = 67.9, $p = 0.0483$) (Figure 1E). However, no significant difference was found between the HI-Unaided and HI-Aided conditions ($p = 0.80$) in preliminary post hoc comparisons. This result was further supported by a repeated measures *t*-test over the HI group (Figure 1F), showing no statistical difference in the number of words recalled between the HI-Unaided and HI-Aided conditions ($p = 0.14$)**,** despite higher mean recall scores in the aided condition.

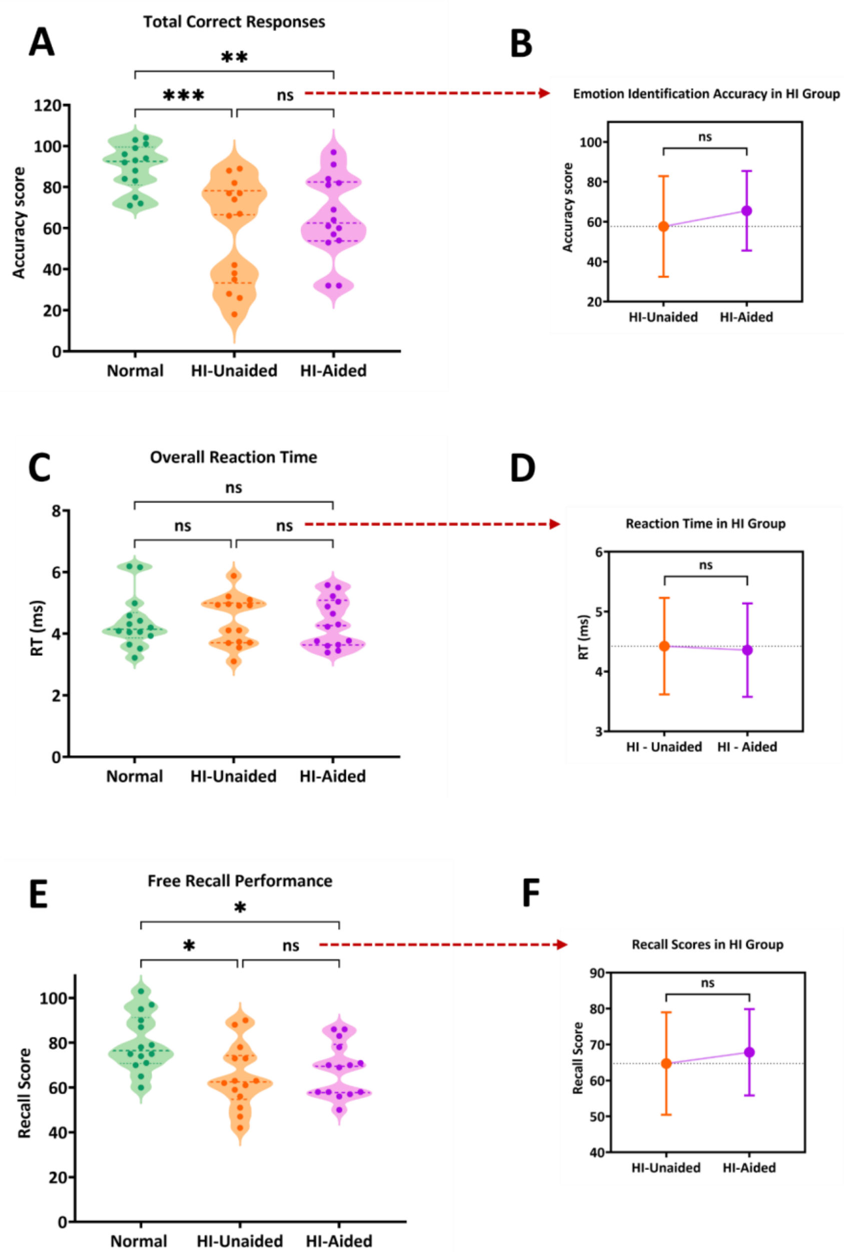


Figure 1. Hearing condition differences (A, C, E) and within-subjects comparisons (B, D, F) for total correct responses (out of 105), overall reaction time (ms), and recall performance (out of 105). **A** illustrates the results of a Tukey's post hoc test for total number of correct responses, comparing NH group (Normal) with HI group in both the unaided (HI-Unaided) and the aided (HI-Aided) conditions. **B** illustrates the result of a repeated measures t-test comparing total correct responses between the HI-Unaided and HI-Aided conditions. **C** and **D** display the results of post hoc test and repeated measures t-test analysis for overall reaction time (RT), respectively. For the recall performance, **E** and **F** present the post hoc and repeated measures t-test results for the total number of recalled words, respectively. Violin plots on the left side illustrate the distribution of individual scores (one dot per participant), with horizontal bars denoting the median and interquartile range. Right-side figures present group means (dots) with vertical bars representing the standard deviation (SD). Statistical significance is shown as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

3.2. Effect of emotion type and hearing condition on response accuracy and RT

Response accuracy. To evaluate how accurately different types of emotion were identified across hearing conditions, a two-way mixed ANOVA was first conducted on accuracy scores, with hearing condition (Normal hearing, HI-Unaided, and HI-Aided) treated as a between-subjects factor and emotion types (7 levels) as a within-subjects factor. Each emotion had a maximum score of 15, representing the total number of correct responses for each emotion. The ANOVA revealed significant main effects of emotion type ($F(4.651, 181.4) = 11.11$, $p < 0.0001$) and hearing condition ($F(2,39) = 10.80$, $p = 0.0002$). However, no significant emotion* hearing condition interaction was observed ($F(12,234) = 1.735$, $p = 0.06$). Although the interaction between emotion type and hearing condition did not reach statistical significance, the near-threshold p-value (p = 0.06) and the study's focus on these factors supported reporting both main effects across the respective factor levels to illustrate potential trends in the data.

As for the effect of emotion type, post hoc analyses revealed that in NH group (Figure 2A, left bars), accuracy scores were highest for fear (Mean = 14.4) and sad (Mean = 13.9), followed by angry (Mean = 13.4), neutral (Mean = 12.7), and disgust (Mean = 12.5), with happy (Mean = 12.3) and PS (Mean = 10.5) being the least accurately identified emotions. No significant differences were found between any pairs of emotions. For HI participants in the unaided condition (Figure 2A, middle bars), accuracy scores were highest for sad (Mean = 11.5) and angry (Mean = 9.8), followed by neutral (Mean = 8.8), fear (Mean = 8.6), happy (Mean = 8.3), PS (Mean= 6.9), and disgust (Mean = 3.5) emotions. For the disgust emotion, accuracy scores differed significantly from those of sad ($p = 0.0011$), angry ($p = 0.0032$), happy ($p = 0.0078$), neutral ($p = 0.0142$), and fear ($p = 0.0376$) emotions. Additionally, there was a significant difference between PS and sad emotions, with PS identified less accurately (p = 0.0176). A similar pattern was observed in the aided condition (Figure 2A, right bars), where accuracy scores were higher for sad (Mean = 12.9) and angry (Mean = 11.2), followed by happy (Mean = 9.6), fear (Mean = 9.5), neutral (Mean = 8.3), PS (Mean= 8.0), and disgust (Mean = 5.7) emotions. Disgust emotion was identified significantly less accurately than both sad (p = 0.0019) and angry (p = 0.0299) emotions.

Regarding the main effect of hearing condition, Tukey's post hoc comparisons showed that the NH group demonstrated significantly higher accuracy than HI participants in both the HI-Unaided (Figure 2B) and HI-Aided groups (Figure 2C) for some emotions, including disgust, fear, happy, and neutral emotions (all $p < 0.04$). However, no significant differences were observed between the HI-Unaided and HI-Aided groups for any emotion (all $p > 0.05$) in the initial analysis, though accuracy scores tended to be numerically higher in the aided condition. Given that the same participants were tested in both aided and unaided conditions, a complementary two-way repeated-measures ANOVA was conducted within the HI group to account for within-subject variability, followed by Tukey's post hoc comparisons (Figure 2E). It showed a significant improvement in identification of disgust emotion in the aided condition relative to the unaided condition (p = 0.0377). No other emotions showed statistically significant improvement (all $p > 0.05$).

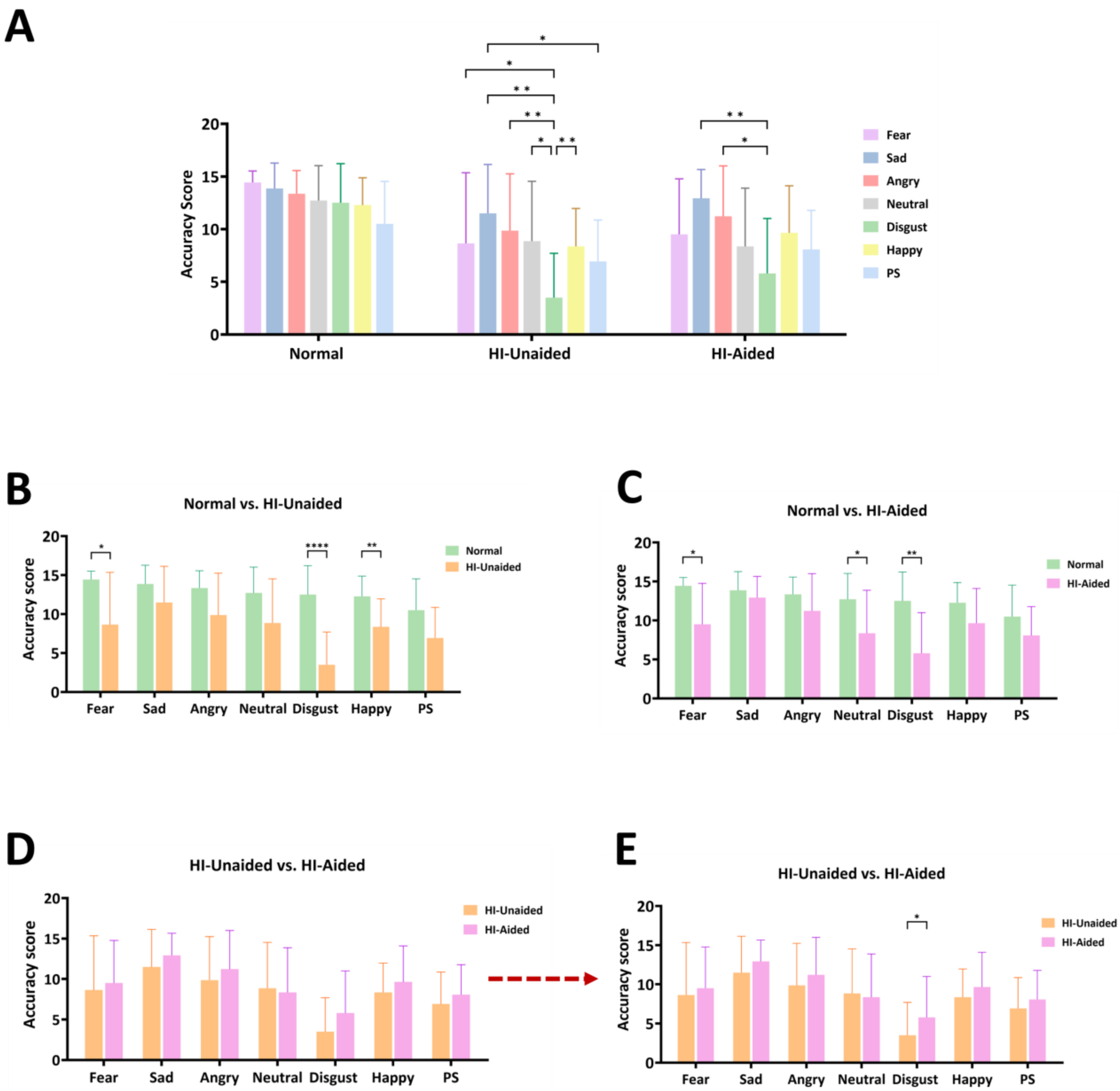


Figure 2. Comparisons of response accuracy across seven emotions and three hearing conditions (Normal Hearing, HI-Unaided, HI-Aided). **A** presents Tukey's post hoc analyses of response accuracy, showing the effect of emotion type across hearing conditions. **B**, **C**, and **D** illustrate results from Tukey's post hoc analyses, highlighting the effect of hearing condition on response accuracy. **E** display repeated-measures comparisons that complement the post hoc analyses by evaluating accuracy differences between the HI-Unaided and HI-Aided conditions. Each bar represents the mean accuracy score, with error bar indicating the standard deviation (SD). Statistical significance is shown as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

Reaction Time. To examine the main effects of emotion type and hearing condition on reaction time (RT), a mixed-effects model was first conducted with hearing condition (Normal hearing, HI-Unaided, and HI-Aided) treated as a between-subjects factor and emotion types (7 levels) as a within-subjects factor. This analysis showed that emotion type had a significant effect on RT ($F(4.801, 175.2) = 16.93$, $p < 0.0001$); however, the main effect of hearing condition was not statistically significant ($F(2,39) = 0.5041$, $p = 0.6079$). No significant interaction was also observed between hearing condition and emotion ($F(12,219) = 0.9757$, $p = 0.4730$).

To further explore the main effect of emotion type, Tukey's analyses revealed that that some emotions required less processing time for accurate identification than others, independent of hearing condition (Figure 3). Mean reaction times (RTs) indicated that participants responded most quickly to sad (Mean = 4.0 ms) and neutral (Mean = 4.23 ms), followed by fear (Mean = 4.29 ms), PS (Mean = 4.4 ms), angry (Mean = 4.5 ms), and happy (Mean = 4.6), with disgust (Mean = 6.0 ms) taking the longest to identify. Pairwise comparisons revealed that disgust elicited significantly slower reaction times than sad, neutral, fear, angry, and happy (all $p < 0.0001$). Reaction times for sad emotion was also significantly different from angry ($p = 0.0346$) and happy ($p = 0.0365$). No other pairwise comparisons reached statistical significance.

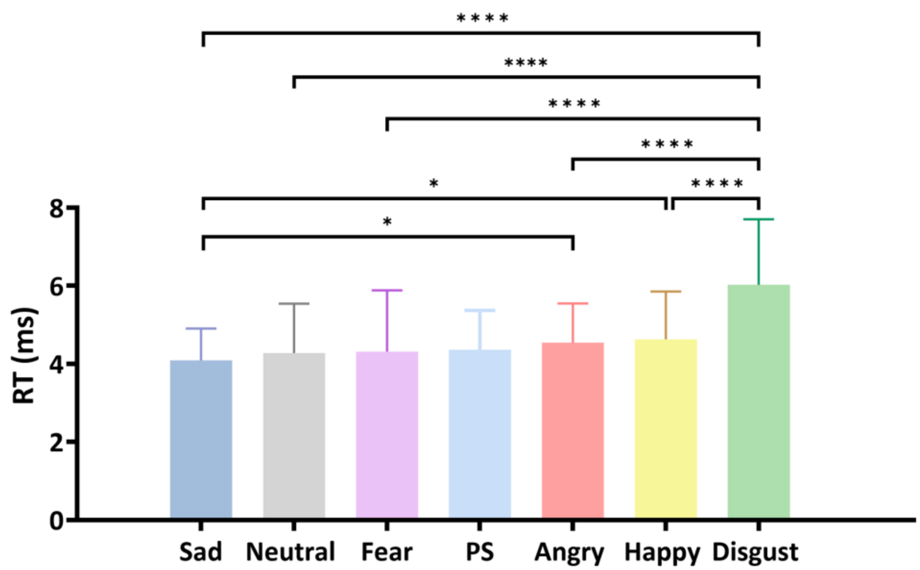


Figure 3. Comparisons of reaction time across seven emotions. Each bar represents the mean accuracy score, with error bar indicating the standard deviation (SD). Statistical significance is shown as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

3.3. Relationship between Speech-In-Noise performance and emotion identification ability

To evaluate whether performance on the QuickSIN test is correlated with and serves as predictor of emotion identification performance, Pearson correlation and simple linear regression analyses were conducted for HI participants in unaided condition at both 55 dB SPL and 70 dB SPL presentation levels for the QuickSIN. Given that only response accuracy, and not reaction time, showed significant group differences between HI and NH participants, analyses were restricted to response accuracy.

At 55 dB SPL, a significant correlation was found between SNR loss and emotion identification accuracy ($r = -0.5624$, $p = 0.036$), indicating that the poorer performance in QuickSIN is associated with lower accuracy scores in emotion identification test (Figure 4, Left panel). The regression analysis confirmed this relationship, showing that SNR loss predicted accuracy scores in the emotion identification test (β = –1.578, 95% CI: –3.038 to –0.1189, R^2 = 0.316, p = 0.036). At 70 dB SPL, a similar pattern was observed where both the Pearson correlation (r = –0.563, p = 0.036) and the linear regression analysis (β = –1.897, 95% CI: –3.65 to –0.145, R^2 = 0.316, p = 0.036) were statistically significant, indicating that as QuickSIN SNR loss increases, emotion accuracy score decreases (Figure 4, Right panel). Based on the regression model, each 1 dB increase in QuickSIN SNR loss was associated with a 1.9-point decrease in accuracy scores. However, individuals who performed well on QuickSIN were not always associated with good prosody perception.

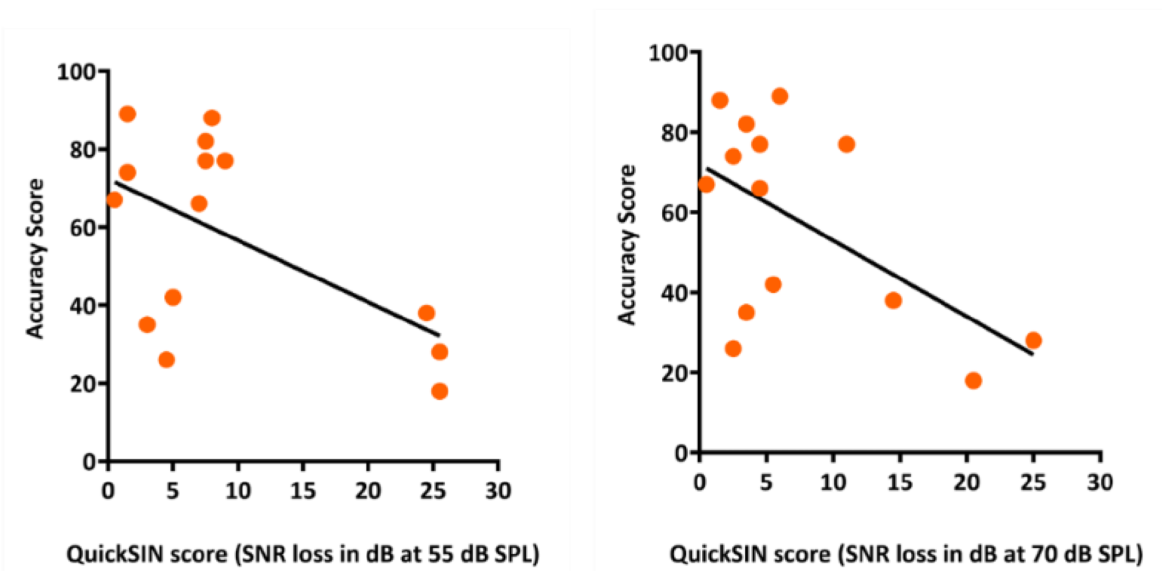


Figure 4. Association between QuickSIN scores (SNR loss) and response accuracy in the emotion identification test in HI individuals tested in unaided condition. Scatterplots show Pearson correlations and fitted regression lines for two presentation levels (55 and 70 dB SPL) of the QuickSIN.

4. Discussion

The effects of hearing loss and hearing aids on emotional speech perception remain unclear, as previous studies have reported inconsistent findings. To further investigate this, we evaluated emotion identification under more ecologically valid conditions by introducing a cognitive component to emotion identification task and age-matching participants to isolate the impact of hearing loss.

Our results showed that the HI group in the unaided condition performed worse than the NH participants, as indicated by significantly lower accuracy scores. With amplification, accuracy scores remained significantly lower than those of the NH group. Compared to the unaided condition, HI participants showed improvement in accuracy scores in the aided condition, although this improvement was not significant. As for reaction time, performance was comparable between groups, suggesting that hearing status and amplification had no influence on processing speed. A breakdown of responses by individual emotions further revealed that both accuracy scores and reaction times were influenced by emotion type. Some emotions were consistently identified with lower accuracy and shorter reaction time than others across all hearing conditions. Furthermore, this study examined the cognitive load associated with emotional speech processing in individuals with hearing impairment, an area that, to our knowledge, has not been previously investigated. Using a recall task, our findings indicated that HI listeners required greater cognitive effort to process prosodic information, leaving fewer resources available for the recall task, as reflected by a lower number of recalled words. While the use of amplified stimuli reduced cognitive load during prosody processing compared to the unaided condition, the difference was not significant. Although a reduced number of correctly recalled words is consistent with increased cognitive effort, it cannot be considered conclusive. Incorporating physiological markers or self-report measures, like pupillometry and workload ratings, would strengthen the assessment of listening and cognitive effort.

Considering all the findings collectively, one possible explanation is that reduced audibility can limit access to emotional cues among individuals with hearing loss. The hearing aid amplification could provide limited benefit and did not significantly improve behavioral emotion identification. This opens up another possibility that beyond peripheral auditory deficits and reduced audibility, higher-order processing may also be affected. Earlier research has shown that individuals with hearing loss may attend to or process emotional cues differently [14]. It is also possible that, given the interaction between the auditory and cognitive systems [44-46], hearing impairment at more advanced stages of auditory processing where cognitive resources are more engaged may heighten cognitive load, contributing to weaker behavioral emotional responses in HI individuals compared to their age-matched NH peers, even with amplification. This may partially explain why hearing aids alone are not sufficient to fully restore emotional processing to the normal level, offering that auditory training may be necessary to enhance their benefit. The emotion-related performance patterns observed in our participants, regardless of hearing status, further suggest that variations in the acoustic properties of each emotion may make some emotions more distinguishable and accessible than others.

Our findings align with those of Goy et al., who also used the same emotional speech database as in our study, albeit without including a free recall task in their experimental design [23]. They reported that older adults with hearing loss performed significantly worse than their NH peers in identifying vocal emotions. In contrast to our findings, however, hearing aids significantly benefited the HI participants in their study. Additionally, no effect of emotion type was observed between aided and unaided conditions, suggesting a general difficulty among HI listeners in emotional prosody recognition. In their later study, consistent with our results, the previously observed differences between aided and unaided conditions were no longer significant, and the type of emotion influenced the HI individuals' performance [24]. Among the emotions, sadness was the most accurately identified, whereas disgust and fear showed the lowest recognition scores. In another study, Most and Aviner also observed emotion-related performance differences among hearing aid users, which were attributed to the acoustic profiles of different emotions, including F0 patterns, intensity, and duration. [25]. Emotions such as sadness, with its slower rate, lower average F0, and longer duration, may be more easily perceived through hearing aids. Accordingly, they explained that sadness was most accurately identified, followed by happiness, surprise, fear, anger, and disgust.

Contrary to our results, Namvar et al. showed that hearing aids, through amplitude compression, can reduce the salience of emotion-specific acoustic cues, making emotion identification more challenging [30]. They showed that fast compression resulted in the greatest degradation of emotional speech recognition, particularly for happiness, sadness, and anger, followed by linear amplification, and slow compression. These effects were attributed to acoustic distortions resulting from frequency-dependent gain variations and differences in amplification strategies. The discrepancy between their findings and ours may stem from the methodological differences; while their study simulated hearing loss in NH listeners, our study involved individuals with actual hearing impairment, whose perceptual processing may differ. Additionally, our emotion identification task included a cognitive component, which may have further influenced performance. Similar to the findings of Arefi et al., Goy et al., through acoustic analysis of simulated hearing aid outputs, demonstrated that all types of hearing aid processing, including linear, slow, and fast compression, significantly alter key acoustic features of emotional speech, such as intensity standard deviation and spectral center of gravity [31]. These acoustic alterations were emotion-specific and resulted in reduced acoustic contrast between emotional categories and made certain emotions, such as anger and disgust, less distinguishable, especially with fast compression. It is important to note that this particular study relied solely on acoustic analysis without examining human behavioral responses, which may account for the discrepancy between their findings and ours. In our study, while performance in the aided condition was still below that of NH participants and not significantly different from their unaided condition, it showed a trend of improvement rather than decline. However, our study used a hearing aid simulator instead of real devices, which may not fully capture the effects of real hearing aid characteristics on emotion perception

Lastly, our study showed that QuickSIN performance was significantly correlated with emotion identification accuracy among HI participants, although individual variability was observed. A moderately severe negative correlation between SNR loss scores and emotion identification accuracy indicated that the poorer performance in QuickSIN is in general associated with lower accuracy scores. However, individuals who performed well on the QuickSIN were not always associated with good prosody perception, suggesting some differences on the roles of auditory and cognitive processing for speech perception in noise versus emotion identification. In this regard, previous studies have also demonstrated that individuals with hearing impairment who perform poorly on speech-in-noise tasks, such as those involving four-talker babble in QuickSIN or speech-shaped noise in the HINT, also exhibit deficits in processing F0-related cues [47, 48]. Since F0 is a critical acoustic feature for emotion identification, this relationship is reasonably expected. On the other hand, speech-in-noise perception requires cognitive resources, including attention, working memory, and short-term memory, in addition to auditory processing skills [36, 37, 49].

Perceiving emotional speech within our paradigm also recruited similar cognitive functions. The overlap in cognitive resources needed for these two tasks may partially account for the observed correlation.

To sum up, prosody perception difficulties should be given more consideration in both the assessment and rehabilitation of individuals with hearing loss. Given that current hearing aid technologies may not fully compensate for prosody perception deficit, exploring innovative algorithms - potentially involving artificial intelligence, brain-computer interfaces, or auditory training - could be beneficial. Future investigations could further examine the role of hearing loss configuration and include individuals with a wider range of cognitive abilities to better clarify how these factors influence prosody perception.

References

1. Organization, W.H., *World report on hearing*. 2021: World Health Organization.

2. Sanchez, V.A., et al., *Effect of hearing intervention on communicative function: A secondary analysis of the ACHIEVE randomized controlled trial.* Journal of the American Geriatrics Society, 2024. **72**(12): p. 3784-3799.

3. Yuan, J., et al., *The risk of cognitive impairment associated with hearing function in older adults: a pooled analysis of data from eleven studies.* Scientific reports, 2018. **8**(1): p. 2137.

4. Chen, H.-L., *Hearing in the elderly: Relation of hearing loss, loneliness, and self-esteem*. 1994, SLACK Incorporated Thorofare, NJ. p. 22-28.

5. Shukla, A., et al., *Hearing loss, loneliness, and social isolation: a systematic review.* Otolaryngology–Head and Neck Surgery, 2020. **162**(5): p. 622-633.

6. Miller, J.L., *Interactions in processing segmental and suprasegmental features of speech.* Perception & Psychophysics, 1978. **24**(2): p. 175-180.

7. Nooteboom, S., *The prosody of speech: Melody and rhythm. The handbook of phonetic sciences. Hardcastle WJ, Laver J*. 1997, Oxford: Basil Blackwell.

8. Lehiste, I., *Suprasegmental features of speech.* Contemporary issues in experimental phonetics, 1976: p. 225-239.

9. Raphael, L.J., G.J. Borden, and K.S. Harris, *Speech science primer: Physiology, acoustics, and perception of speech*. 2007: Lippincott Williams & Wilkins.

10. Nygaard, L.C., D.S. Herold, and L.L. Namy, *The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning.* Cognitive science, 2009. **33**(1): p. 127-146.

11. Moore, B.C., *Cochlear hearing loss: physiological, psychological and technical issues*. 2007: John Wiley & Sons.

12. Ponsot, E., et al., *Mechanisms of spectrotemporal modulation detection for normal-and hearing-impaired listeners.* Trends in hearing, 2021. **25**: p. 2331216520978029.

13. Syka, J., *Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning.* Physiological reviews, 2002. **82**(3): p. 601-636.

14. Picou, E.M., et al., *Hearing, emotion, amplification, research, and training workshop: Current understanding of hearing loss and emotion perception and priorities for future research.* Trends in Hearing, 2018. **22**: p. 2331216518803215.

15. Kalathottukaren, R.T., S.C. Purdy, and E. Ballard, *Prosody perception and production in children with hearing loss and age-and gender-matched controls.* Journal of the American Academy of Audiology, 2017. **28**(04): p. 283-294.

16. Karimi-Boroujeni, M., H.R. Dajani, and C. Giguère, *Perception of prosody in hearing-impaired individuals and users of hearing assistive devices: An overview of recent advances.* Journal of Speech, Language, and Hearing Research, 2023. **66**(2): p. 775-789.

17. Husain, F.T., J.R. Carpenter-Thompson, and S.A. Schmidt, *The effect of mild-to-moderate hearing loss on auditory and emotion processing networks.* Frontiers in systems neuroscience, 2014. **8**: p. 10.

18. Slade, K., C.J. Plack, and H.E. Nuttall, *The effects of age-related hearing loss on the brain and cognitive function.* Trends in neurosciences, 2020. **43**(10): p. 810-821.

19. Boothroyd, A., *Auditory perception of speech contrasts by subjects with sensorineural hearing loss.* Journal of Speech, Language, and Hearing Research, 1984. **27**(1): p. 134-144.

20. Irino, T., et al., *Effects of age and hearing loss on speech emotion discrimination.* Scientific Reports, 2024. **14**(1): p. 18328.

21. Gürses, E., et al., *Evaluation of temporal and suprasegmental auditory processing in patients with unilateral hearing loss.* Auris Nasus Larynx, 2020. **47**(5): p. 785-792.

22. Cannon, S.A. and M. Chatterjee, *Voice emotion recognition by children with mild-to-moderate hearing loss.* Ear and hearing, 2019. **40**(3): p. 477-492.

23. Goy, H., et al., *Perception of emotional speech by listeners with hearing aids.* Canadian Acoustics, 2016. **44**(3).

24. Goy, H., et al., *Hearing aids benefit recognition of words in emotional speech but not emotion identification.* Trends in Hearing, 2018. **22**: p. 2331216518801736.

25. Most, T. and C. Aviner, *Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing.* Journal of deaf studies and deaf education, 2009. **14**(4): p. 449-464.

26. Mazaheryazdi, M., et al., *Perception of musical emotion in the students with cognitive and acquired hearing loss.* Iranian journal of child neurology, 2018. **12**(2): p. 41.

27. Schmidt, J., et al. *Do hearing aids improve affect perception?* in *Physiology, psychoacoustics and cognition in normal and impaired hearing*. 2016. Springer International Publishing.

28. Singh, G., et al., *The emotional communication in hearing questionnaire (EMO-CHeQ): Development and evaluation.* Ear and Hearing, 2019. **40**(2): p. 260-271.

29. Christensen, J.A., et al., *Effects of age and hearing loss on the recognition of emotions in speech.* Ear and hearing, 2019. **40**(5): p. 1069-1083.

30. Arefi, H.N., et al., *Effect of hearing aid amplitude compression on emotional speech recognition.* Auditory and Vestibular Research, 2017. **26**(4): p. 223-230.

31. Goy, H., et al., *Acoustic analysis of emotional speech processed by hearing aids.* Canadian Acoustics, 2017. **45**(3): p. 178-179.

32. Rönnberg, J., et al., *Cognition as a bridge between signal and dialogue: Communication in the hearing impaired and deaf.* Scandinavian Audiology, 1998. **27**(4): p. 101-108.

33. Pichora-Fuller, M.K., P. Mick, and M. Reed. *Hearing, cognition, and healthy aging: Social and public health implications of the links between age-related declines in hearing and cognition*. in *Seminars in hearing*. 2015. Thieme Medical Publishers.

34. Worthington, D.L., G.D. Bodie, and M.E. Fitch-Hauser, *Listening: Processes, functions, and competency*. 2024: Routledge.

35. Killion, M.C., et al., *Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners.* The Journal of the Acoustical Society of America, 2004. **116**(4): p. 2395-2405.

36. Anderson, S., et al., *A dynamic auditory-cognitive system supports speech-in-noise perception in older adults.* Hearing research, 2013. **300**: p. 18-32.

37. Dillon, H. and S. Cameron, *Separating the causes of listening difficulties in children.* Ear and hearing, 2021. **42**(5): p. 1097-1108.

38. Marian, V., H.K. Blumenfeld, and M. Kaushanskaya, *The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals.* Journal of Speech, Language, and Hearing Research, 2007. **50**(4): p. 940-967.

39. Comeau, G., A. Koravand, and M. Swirp, *Prevalence of hearing loss among university music students.* Canadian Acoustics, 2018. **46**(1): p. 37-51.

40. Molloy, D.W. and T.I. Standish, *A guide to the standardized Mini-Mental State Examination.* International psychogeriatrics, 1997. **9**(S1): p. 87-94.

41. Pichora-Fuller, M.K. and K. Dupuis, *Toronto emotional speech set (TESS).* Scholars Portal Dataverse, 2020. **1**: p. 2020.

42. Ollivier, S., H. Nélisse, and J. Voix, *On the use of wide dynamic range compression and other algorithms to improve hearing protection of workers with hearing impairment: A preliminary study on speech intelligibility.* Canadian Acoustics, 2023. **51**(3): p. 110-111.

43. Dillon, H., *Hearing Aids*. 2001: Thieme.

44. Loughrey, D.G., et al., *Association of age-related hearing loss with cognitive function, cognitive impairment, and dementia: a systematic review and meta-analysis.* JAMA otolaryngology–head & neck surgery, 2018. **144**(2): p. 115-126.

45. Taljaard, D.S., et al., *The relationship between hearing impairment and cognitive function: a meta-analysis in adults.* Clinical Otolaryngology, 2016. **41**(6): p. 718-729.

46. Powell, D.S., et al., *Hearing loss and cognition: what we know and where we need to go.* Frontiers in aging neuroscience, 2022. **13**: p. 769405.

47. Presacco, A., J.Z. Simon, and S. Anderson, *Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss.* PloS one, 2019. **14**(3): p. e0213899.

48. Anderson, S., et al., *Brainstem correlates of speech-in-noise perception in children.* Hearing research, 2010. **270**(1-2): p. 151-157.

49. Porto, L., J. Wouters, and A. van Wieringen, *Speech perception in noise, working memory, and attention in children: a scoping review.* Hearing Research, 2023. **439**: p. 108883.

Chapter 6. General discussion

General Discussion

Emotional prosody perception is fundamental to effective speech communication. However, individuals with hearing loss and hearing aid users often face challenges in perceiving and interpreting these verbal prosodic cues [1-3]. This thesis, comprising a narrative review and three experimental studies, investigated the neural and perceptual aspects of prosody processing in individuals with normal hearing and hearing impairment, with and without hearing aid amplification, using both electrophysiological and behavioral paradigms. Across all chapters, the primary focus was on how hearing loss and simulated hearing aid processing influence the encoding and perception of prosodic cues, particularly those conveying emotional information. This general discussion brings together the main findings from each study, highlighting both the overlapping results and the distinct contributions of each.

Chapter 2 presented a narrative overview that laid the theoretical foundation for this thesis, emphasizing the essential role of prosody in speech interaction and the critical gaps in current knowledge regarding the effects of hearing loss and hearing assistive devices on prosody perception [3]. It highlighted that verbal prosody conveys both linguistic and emotional meaning and understanding it engages a wide network of brain regions involved in auditory and cognitive processing. Different prosodic patterns are characterized by distinct acoustic features, particularly the fundamental frequency and intensity contours, along with variations in duration. In individuals with hearing loss, the extraction of these acoustic cues could be challenged not only by reduced audibility, but also by degraded spectro-temporal resolutions and disrupted central auditory processing.

Vulnerability to prosody perception deficits among hearing-impaired individuals was shown to depend on range of factors, such as severity of hearing loss, age at the onset of hearing loss, type of assistive hearing devices, and auditory training experience. Although prosody perception is expected to decline with greater degrees of hearing loss, findings have been inconsistent. Some studies reported deficits even in mild cases, while others observed preserved performance in individuals with mild to moderate hearing loss. In older adults with hearing impairment, age-related deterioration in both auditory and cognitive functions likely contributed to reducing the auditory system's ability to adapt and compensate for prosodic deficits.

A particularly important point raised in this chapter was that current hearing technologies do not fully restore prosodic cues. While hearing aids effectively improve audibility for speech recognition, they are limited in capturing the dynamic acoustic cues essential to emotional communication. This could be due to device algorithms and their compression strategies that prioritize speech intelligibility but do not account for emotion cues. Cochlear implants similarly provided only partial access to F0-related prosodic cues, often shifting reliance to secondary cues such as duration and intensity. Bimodal fittings have been reported to offer added benefit, but this advantage is not consistently supported by the evidence. On the other hand, auditory training demonstrated promising effects in supporting prosody perception in hearing-impaired individuals of all ages.

This chapter also raised attention to a notable gap in understanding the neural mechanisms of prosody perception in individuals with hearing loss. This gap shaped the rationale for our subsequent experimental design, which aimed to integrate behavioral and electrophysiological measures to explore this issue more comprehensively.

In Chapter 3, informed by the background framework outlined in our review, we investigated the feasibility of using speech-evoked FFR to study prosody processing at the central level of the auditory system. As a baseline study, the FFR was recorded from young adults with normal hearing using the two-syllable word "balloon," spoken by both a male and a female talker with sad and happy emotions. After applying the pitch-tracking algorithm on recorded FFRs, F0 tracking accuracy was measured using two metrics: 5% accuracy and RMSE.

First, our findings demonstrated that the FFR can encode the F0 contours embedded in natural emotional speech, indicating that neural fibers can synchronize their firings to time-varying F0 patterns associated with emotional content. Secondly, this study showed that the accuracy of F0 encoding varied depending on the type of emotion and the acoustic characteristics of the talker's voice. Stimuli spoken with sad emotion resulted in significantly more accurate F0 tracking compared to their happy counterparts. The acoustic features of our sad stimuli, including lower average F0, slower changes in F0, and more energy at F0 and its harmonics, align with properties known to enhance phase-locking in the auditory pathway [4], might be a plausible explanation for the stronger F0 representation in response to sad emotion. Voice characteristics of the talker further modulated the FFR response, with male-spoken tokens eliciting stronger F0 tracking than female-spoken ones, regardless of emotion. This finding was consistent with previous reports showing enhanced FFRs to stimuli with lower frequency periodicity [5,6]. In our data, the male voice had lower F0 than the female voice, which may have contributed to enhanced neural synchrony, possibly due to the greater number of harmonics present in lower F0 stimuli.

Our findings also showed that male and female participants exhibited comparable F0 tracking accuracy across all conditions. This finding contrasted with earlier studies using brief and synthetic stimuli, where female listeners were reported to show enhanced FFR amplitudes or shorter response latencies, especially in transient onset responses or high-frequency harmonics rather than for F0 tracking itself [7]. This suggests that sex-related differences in the neural encoding of prosody are minimal, at least at the brainstem level and when processing emotionally rich speech.

Indeed, a key contribution of this study was the use of ecologically valid stimuli. By presenting an emotionally meaningful word "balloon" that likely engaged both linguistic and paralinguistic processing mechanisms, the experimental design more closely resembled real-life communication.

Taken together, this chapter provided insights into how the central auditory system encodes and tracks F0 contour in emotional speech and confirmed that sFFR is a viable and sensitive tool for measuring verbal prosodic cues. However, only happy and sad emotions were tested. By providing baseline data from individuals with normal hearing, this study established the basis for our subsequent study on neural encoding of prosodic patterns in hearing loss.

Extending the findings from normal-hearing individuals, Chapter 4 examined how hearing loss and simulated hearing aid processing affect the neural representation of F0 contours. Based on our previous study showing more accurate FFRs to the male voice, this study used only male-spoken balloon with sad and happy emotion. FFR recordings were obtained from individuals with hearing loss and from age-matched peers with normal hearing. Hearing-impaired participants were tested twice, once in the unaided condition using original stimuli, and once in the aided condition using customized amplification of the stimuli. 5% accuracy and RMSE were then calculated for each condition within the hearing-impaired and normal-hearing groups.

Among the key findings presented in this chapter was that hearing loss was associated with reduced fidelity in F0 tracking. Individuals with hearing impairment exhibited significantly lower 5% accuracy and higher RMSE values in both sad and happy stimuli compared to their normal-hearing peers. As earlier studies have demonstrated reduced F0-related representation in the FFR among HI individuals at sound pressure levels (SPLs) equal to those used for NH individuals [8,9], this study offers further evidence that their ability to track F0 changes over time is also affected. Nonetheless, mixed results are found in the literature regarding the effect of hearing loss on FFRs, with some prior studies reporting either preserved or even enhanced F0 encoding [10-12]. Of note, the present study focused on F0 contour tracking of a word in emotional speech, while the aforementioned studies investigated average F0 representation of simple vowels or monosyllables.

Another important result was that hearing aid amplification partially improved F0 tracking. When hearing-impaired participants were tested in the aided condition, their 5% accuracy scores significantly improved compared to their unaided responses, and similarly RMSE values showed a decreasing (improvement) trend. Previous studies have reported mixed effects of hearing aids on F0 encoding, including enhanced, unchanged, or even reduced neural representation [9,10,13,14]. Our findings from this chapter also contributed to this area by showing that, at least in the context of emotional speech and at the subcortical level, hearing aids can partially restore prosodic cues. However, because our study used a hearing aid simulator instead of real devices, some real hearing aid characteristics like processing delay effects were not reflected. These can interfere with precise neural timing and alter the speech envelope, which plays a key role in neural representation of emotional F0 contours.

In line with our findings presented in Chapter 3, F0 contours were tracked more accurately for sad stimulus than for happy one, regardless of hearing status or amplification condition. All participants exhibited higher F0 tracking accuracy in response to the sad version of "balloon" compared to the happy version. As previously suggested, this advantage may stem from the acoustic characteristics of the sad stimuli, which are more aligned with the temporal characteristics that support robust FFR phase-locking. Since this study focused on happy and sad emotions, exploring a wider range of emotions in future work could further clarify this effect. Overall, this chapter indicated that hearing loss disrupts the neural encoding of dynamic F0 contours essential for emotional prosody, but hearing aid use could restore this ability to near-normal levels.

Chapter 5 broadened the focus of this thesis by moving beyond neural encoding to perceptual processing, examining how hearing loss and hearing aid use affect the perception of various emotions through behavioral responses. Insights from Chapter 4 guided this study into how normal-hearing and hearing-impaired individuals differ in their performance on an emotion identification task under more real-world and cognitively demanding listening conditions. The investigation also assessed whether QuickSIN performance predicts emotion recognition ability in individuals with hearing loss.

One of the main results was that hearing loss adversely affected performance in emotion identification tasks, even when amplification was provided. Hearing-impaired individuals exhibited significantly poorer performance on total correct responses relative to normal-hearing participants. They also required greater cognitive effort to process prosodic information, leaving fewer resources available for the recall task, as reflected by a lower number of recalled words. In the aided condition, although they showed some improvements in response accuracy, reaction time, and word recall, their performance remained consistently lower than that of normal-hearing peers. These findings suggest that while hearing aids improve access to acoustic cues, they do not fully restore perceptual performance to normal levels or reduce the cognitive effort involved in processing emotional speech.

This insufficient compensation from hearing aid contrasted with the electrophysiological findings of Chapter 4, where F0 tracking accuracy in the aided condition was comparable to that of normal-hearing participants. This discrepancy may be attributed to the fact that near-normal neural encoding does not necessarily guarantee improved perceptual performance. Successful emotional speech perception relies not only on the fidelity of acoustic encoding, but also on higher-order cognitive resources. At higher levels of auditory processing, hearing impairment may alter how people attend to or respond to emotional information [15]. Additionally, hearing loss may place greater demands on cognitive resources for processing prosodic information, possibly leading to increased effort, a higher risk of fatigue, and fewer resources available for other aspects of speech perception in HI individuals compared to their NH peers during conversation. It suggests that even when subcortical neural mechanisms are capable of synchronizing to emotion-related pitch variations in HI individuals with amplification, extracting emotional meaning from speech may requires broader cognitive engagement and imposes a heavier cognitive burden compared to individuals with normal hearing. The lower scores observed in our free recall task among hearing-impaired participants, in both aided and unaided conditions, also supported this. The integration of a free recall task was a key contribution of this chapter, offering a new perspective on cognitive load during emotional speech processing. However, this finding should be viewed as indicative, not conclusive, and additional evaluation of cognitive effort is needed.

Consistent with the findings from Chapters 3 and 4, emotion type influenced performance across all conditions. Sadness was the most accurately identified emotion among normal-hearing individuals and hearing-impaired participants tested in unaided and aided conditions. In contrast, and in line with previous studies, emotions with more dynamic spectral or temporal properties (e.g., anger, fear, and happiness) were more vulnerable to misidentification, particularly under unaided conditions.

This chapter also showed a relationship between QuickSIN performance and emotion identification accuracy in individuals with hearing loss. This finding suggests that difficulty in understanding speech in noise, a hallmark of real-world listening challenges, might be associated with poorer prosody perception. Since both tasks rely on shared auditory and cognitive mechanisms, the observed correlation was expected. However, individuals who performed well on the QuickSIN were not always associated with good prosody perception, suggesting some differences on the roles of auditory and cognitive processing for speech perception in noise versus emotion identification.

All in all, Chapter 5 complemented the electrophysiological findings by showing that challenges remain for hearing-impaired individuals in perceiving emotional content in real-world listening situations. It also highlighted the importance of cognitive capacity in speech prosody perception and emphasized the need to consider it in auditory rehabilitation to improve speech perception outcomes in individuals with hearing loss and hearing aid users.

While Chapters 4 (FFR data) and 5 (behavioral test) differed in several methodological aspects, such as the dataset, stimulus type (word vs. sentence), talker gender, and the number of assessed emotions (two vs. seven), the same participants completed both tests. This allows for an exploratory examination of the relationship between electrophysiological and behavioral measurements to gain insight and provide perspective for future investigations. Correlation analyses were conducted using pooled data from all conditions to examine the relationship between 5% Accuracy from the FFR test and behavioral accuracy scores, and between RMSE from the FFR test and behavioral accuracy scores. As shown in Figure 1A, there was a significant moderate positive correlation between FFR 5% Accuracy and behavioral emotion identification scores ($r = 0.52$, $p = 0.0003$). Participants with more accurate neural encoding of F0 tended to perform better in the behavioral task. However, the analysis revealed that the neural measure of F0 tracking ability accounted for 27% of the variance in behavioral scores ($R^2 = 0.27$), suggesting that factors beyond neural encoding also contribute to perceptual performance as we discussed before. Also, as shown in Figure 1B, there was a significant moderate negative correlation between RMSE and behavioral accuracy scores ($r = - 0.44$, $p = 0.003$), indicating that greater error in F0 tracking at the neural level was associated with poorer behavioral performance. However, RMSE explained only 19% of the variability in behavioral accuracy scores, further suggesting that perceptual ability is influenced by multiple factors in addition to the precision of F0 neural encoding.

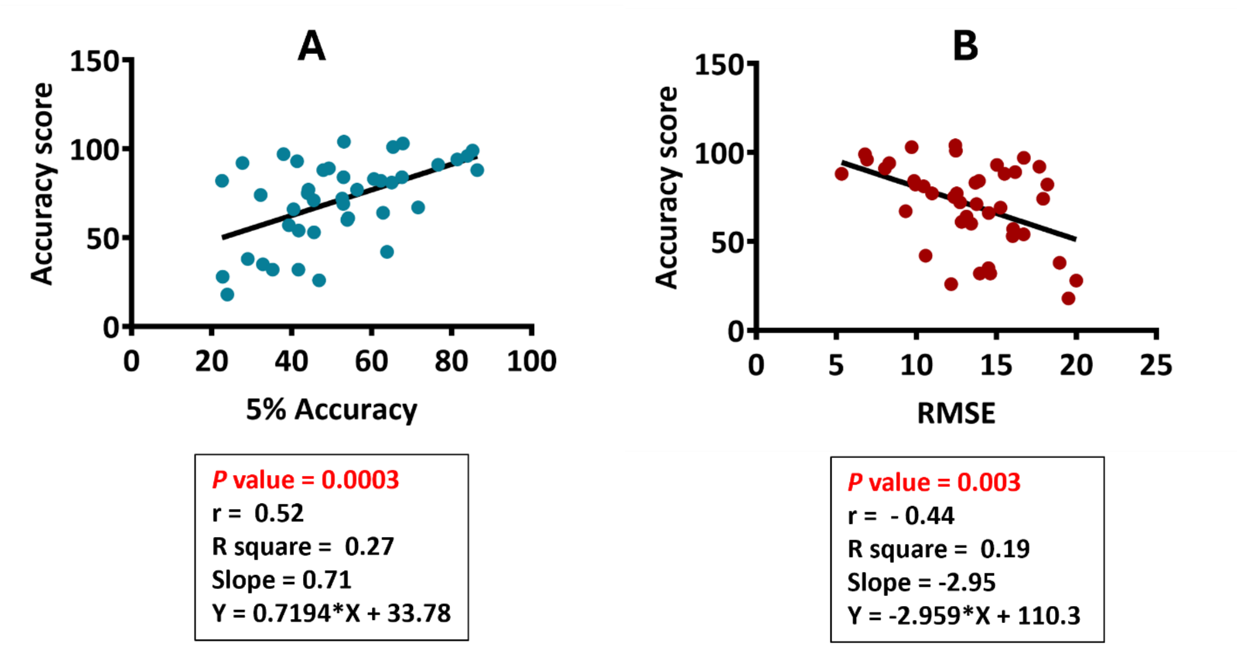


Figure 1. Association between FFR measures and behavioral emotion identification test using pooled data from normal-hearing and HI participants (aided and unaided). (A) Scatterplot showing the Pearson correlation and fitted regression line between 5% Accuracy from the FFR and behavioral accuracy scores. (B) Scatterplot showing the Pearson correlation and fitted regression line between RMSE from the FFR and behavioral accuracy scores.

The reason that our databases for the behavioral experiment and the FFR recording were not compatible was that each of the TESS and ESD datasets was best matched to different parts of our study. For the behavioral test, we selected the Toronto Emotional Speech Set (TESS) because it has been used in a few previous behavioral studies, providing a solid baseline for comparison especially when incorporating a cognitive component. On the other hand, TESS contains only female voices, and all keywords in its sentences are monosyllabic, making it unsuitable for the FFR recordings, where we aimed to include multisyllabic words to better assess prosody. For the

FFR task, we therefore used the ESD, which contains multisyllabic words and encompasses both male and female talkers. Further investigations between electrophysiological and behavioral measurements using matched stimuli and emotions are warranted to clarify the neural mechanisms underlying emotional responses and their behavioral manifestations.

Clinical Implications

Our findings indicate the importance of speech prosody evaluation in audiological assessments for individuals with hearing impairment to identify deficits that are not typically detected by standard speech audiometry. Poor QuickSIN performance may reflect broader communication challenges beyond speech-in-noise perception, including difficulties in perceiving prosodic cues. Clinically, incorporating self-reported questionnaire on emotion perception ability could help identify such deficits. The fact that amplification only partially improved perceptual performance suggests that hearing aids alone may not provide optimal rehabilitation. During hearing aid counseling, this information can guide realistic goal setting. Adding targeted auditory training focused on prosodic and other suprasegmental cues alongside amplification could help behavioral outcomes. Moreover, because hearing-impaired listeners experience greater cognitive load when processing emotional speech, rehabilitation may benefit from considering listening effort through optimized hearing aid fittings and individualized support, ultimately improving communication effectiveness and enriching quality of life.

Study limitations and Future Directions

While our four studies provided multifaceted insights, they also present some limitations that highlight the need for further investigation and guide future research directions.

From an electrophysiological perspective, since only happy and sad emotions were included, using a broader range of emotions in future works would allow for a more comprehensive investigation of how various F0 contours affect the FFR. Also, we only used one male and one female speaker, making it unclear whether the observed effects of male/female acoustic features on the FFR, such as lower F0 and harmonics, can be generalized to other speakers. The use of a simulated hearing aid with a fixed amplification strategy is another limitation, since the effects of real hearing aid features, such as processing delays or comb filtering in vented or open fittings, are not reflected in our results and could distort neural phase locking and desynchronize timing cues. Future studies utilizing real devices, varied amplification parameters, and different fittings could provide a more comprehensive understanding of hearing aid effects and improve generalizability. Moreover, while this thesis focused on F0 contour, future research could explore the neural encoding of other acoustic cues important for emotional prosody such as intensity contour. Additionally, expanding the use of sFFR to cortical responses may help map the pathway from subcortical tracking to perceptual awareness.

The behavioral study also presents several limitations that suggest directions for future research. Our test was limited to a single presentation level, which may have influenced performance differences between groups. Studying more level could help shed more light on the relative roles of audibility and suprathreshold processing on emotion identification. Moreover, the study investigated only one type of hearing aid processing strategy. Like the electrophysiological part, future studies could explore additional hearing aid parameters and amplification strategies to optimize emotion perception. Incorporating real hearing aids, rather than simulations, may better reflect real-world fitting outcomes and enhance ecological validity. Future studies could investigate other hearing aid parameters and amplification strategies to help optimize emotion perception. Given the observed partial restoration of perceptual performance with hearing aids and the critical role of cognitive function in prosody perception, future research could also explore whether auditory training and prosody-focused rehabilitation can further enhance emotion perception outcomes in individuals with hearing loss. Although secondary free recall task is consistent with cognitive effort, further complementary measures are also needed to more directly assess cognitive and listening effort during emotion perception.

In terms of clinical population, this research was limited to cognitively healthy adults aged 18–65. To enhance findings, future studies could include older adults and individuals with hearing loss across a wider range of cognitive abilities to better reflect the diversity seen in clinical settings. Our hearing-impaired participants varied in audiometric configuration, degree of hearing loss, and hearing aid experience, which may have influenced how hearing loss and amplification affected prosody perception. Using a more controlled and homogeneous sample with larger sample sizes in future research could help clarify how different types and severities of hearing loss disrupt prosody perception, and how effectively hearing aids can restore emotion perception across these profiles.

Conclusion

Together, the four studies presented in this thesis provide a comprehensive investigation into how hearing loss and hearing aid amplification influence both the neural encoding and perceptual processing of emotional prosody. Covering the full spectrum from theoretical framework to neural and behavioral findings, the results revealed that hearing loss disrupts access to F0 contours critical for emotional speech. Hearing aid amplification could partially restore the corresponding neural encoding; however, this improvement did not extend to behavioral performance in emotion identification. This suggests that individuals with hearing loss may attend to or process emotional cues differently, as difficulties in emotion identification persist even when audibility is improved through amplification. This thesis also highlights the important role of cognitive function in prosody perception and shows that individuals with hearing loss, whether aided or unaided, experience greater cognitive load when processing emotional speech. These results point to the value of including prosodic features in both assessment and rehabilitative approaches for hearing-impaired individuals. This could not only improve speech communication for people with hearing loss but also enhance their quality of life by facilitating access to the emotional and social dimensions of speech.

References

1. D. B. Pisoni, R. E. Remez, *The handbook of speech perception* (Wiley Online Library, 2005).

2. L. J. Raphael, G. J. Borden, K. S. Harris, *Speech science primer: Physiology, acoustics, and perception of speech* (Lippincott Williams & Wilkins, 2007).

3. M. Karimi-Boroujeni, H. R. Dajani, C. Giguère, Perception of prosody in hearing-impaired individuals and users of hearing assistive devices: An overview of recent advances. *J. Speech. Lang. Hear. Res.* **66**, 775-789 (2023).

4. E. Skoe, N. Kraus, Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* **31**, 302-324 (2010).

5. A. Krishnan, C.H. Suresh & J. T. Gandour, Changes in pitch height elicit both language-universal and language-dependent changes in neural representation of pitch in the brainstem and auditory cortex. *Neuroscience* **346**, 52-63 (2017).

6. J. Van Canneyt, J. Wouters & T. Francart, From modulated noise to natural speech: The effect of stimulus parameters on the envelope following response. *Hear. Res.* **393**, 107993 (2020).

7. M. Ahadi, A. Pourbakht, A. H. Jafari, , Z. Shirjian & A. S. Jafarpisheh, Gender disparity in subcortical encoding of binaurally presented speech stimuli: an auditory evoked potentials study. *Auris Nasus Larynx* **41**, 239-243 (2014).

8. W. Hao *et al.*, Effects of phase-locking deficits on speech recognition in older adults with presbycusis. *Front. Aging Neurosci.* **10**, 397 (2018).

9. S. Ananthakrishnan, A. Krishnan, E. Bartlett, Human frequency following response: neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss. *Ear Hear.* **37**, e91-e103 (2016).

10. S. Anderson, A. Parbery-Clark, T. White-Schwoch, S. Drehobl, N. Kraus, Effects of hearing loss on the subcortical representation of speech cues. J. Acoust. Soc. Am. 133, 3030-3038 (2013).

11. A. Koravand, R. Al Osman, V. Rivest, C. Poulin, Speech-evoked auditory brainstem responses in children with hearing loss. Int. J. Pediatr. Otorhinolaryngol. 99, 24-29 (2017).

12. A. Presacco, J. Z. Simon, S. Anderson, Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PLoS One* **14**, e0213899 (2019).

13. V. Easwar, D. W. Purcell, S. J. Aiken, V. Parsa, S. D. Scollie, Evaluation of speech-evoked envelope following responses as an objective aided outcome measure: Effect of stimulus level, bandwidth, and amplification in adults with hearing loss. *Ear Hear.* **36**, 635-652 (2015).

14. G. BinKhamis, A. Elia Forte, T. Reichenbach, M. O'Driscoll, K. Kluk, Speech auditory brainstem responses in adult hearing aid users: Effects of aiding and background noise, and prediction of behavioral measures. *Trends Hear.* **23**, 2331216519848297 (2019).

15. E. M. Picou *et al.*, Hearing, emotion, amplification, research, and training workshop: Current understanding of hearing loss and emotion perception and priorities for future research. *Trends Hear.* **22**, 2331216518803215 (2018).

Appendices

A. Ethics approval for electrophysiology experiments

27/06/2024

Université d'Ottawa — Bureau d'éthique et d'intégrité de la recherche

University of Ottawa — Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE | CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number	H-03-22-7915
Titre du projet / Project Title	Measurement and Applications of the Speech-evoked Frequency Following Response (sFFR)
Type de projet / Project Type	Recherche de professeur / Professor's research project
Statut du projet / Project Status	Renouvelé / Renewed
Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)	22/07/2022
Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)	21/07/2025

Équipe de recherche / Research Team

Chercheur / Researcher	Affiliation	Role
Hilmi DAJANI	École de science informatique et de génie électrique / School of Electrical Engineering and Computer Science	Chercheur Principal / Principal Investigator
Christian GIGUÈRE	École des sciences de la réadaptation / School of Rehabilitation Sciences	Co-chercheur principal / Co-principal investigator
Maryam KARIMI BOROUJENI	École des sciences de la réadaptation / School of Rehabilitation Sciences	Étudiant-chercheur / Student-researcher
Sajad SADEGHKHANI	École de science informatique et de génie électrique / School of Electrical Engineering and Computer Science	Étudiant-chercheur / Student-researcher

Conditions spéciales ou commentaires / Special conditions or comments

550, rue Cumberland, pièce 154 550 Cumberland Street, Room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

27/06/2024

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Le Comité d'éthique de la recherche (CÉR) de l'Université d'Ottawa, opérant conformément à l'*Énoncé de politique des Trois conseils* (2014) et toutes autres lois et tous règlements applicables, a examiné et approuvé la demande d'éthique du projet de recherche ci-nommé.

L'approbation est valide pour la durée indiquée plus haut et est sujette aux conditions énumérées dans la section intitulée "Conditions Spéciales ou Commentaires". Le formulaire « Renouvellement ou Fermeture de Projet » doit être complété quatre semaines avant la date d'échéance indiquée ci-haut afin de demander un renouvellement de cette approbation éthique ou afin de fermer le dossier.

Toutes modifications apportées au projet doivent être approuvées par le CÉR avant leur mise en place, sauf si le participant doit être retiré en raison d'un danger immédiat ou s'il s'agit d'un changement ayant trait à des éléments administratifs ou logistiques du projet. Les chercheurs doivent aviser le CÉR dans les plus brefs délais de tout changement pouvant augmenter le niveau de risque aux participants ou pouvant affecter considérablement le déroulement du projet, rapporter tout évènement imprévu ou indésirable et soumettre toute nouvelle information pouvant nuire à la conduite du projet ou à la sécurité des participants.

The University of Ottawa Research Ethics Board, which operates in accordance with the *Tri-Council Policy Statement* (2014) and other applicable laws and regulations, has examined and approved the ethics application for the above-named research project.

Ethics approval is valid for the period indicated above and is subject to the conditions listed in the section entitled "Special Conditions or Comments". The "Renewal/Project Closure" form must be completedfour weeks before the above-referenced expiry date to request a renewal of this ethics approval or closure of the file.

Any changes made to the project must be approved by the REB before being implemented, except when necessary to remove participants from immediate endangerment or when the modification(s) only pertain to administrative or logistical components of the project. Investigators must also promptly alert the REB of any changes that increase the risk to participant(s), any changes that considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project or the safety of the participant(s).

Germain ZONGO
Responsable d'éthique en recherche / Protocol Officer
Pour/For **Daniel LAGAREC** Président(e) du/ Chair of the **Comité d'éthique de la recherche en sciences de la santé et sciences / Health Sciences and Sciences Research Ethics Board**

550, rue Cumberland, pièce 154 550 Cumberland Street, Room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

B. Ethics approval for behavioral experiment

07/03/2025

Université d'Ottawa | University of Ottawa

Bureau d'éthique et d'intégrité de la recherche | Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE | CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number	H-01-23-8637
Titre du projet / Project Title	Prosody perception in hearing impaired listeners: effect of hearing aid processing and noise
Type de projet / Project Type	Recherche de professeur / Professor's research project
Statut du projet / Project Status	Renouvelé / Renewed
Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)	30/03/2023
Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)	29/03/2026

Équipe de recherche / Research Team

Chercheur / Researcher	Affiliation	Role
Christian GIGUÈRE	École des sciences de la réadaptation / School of Rehabilitation Sciences	Chercheur Principal / Principal Investigator
Maryam KARIMI BOROUJENI	École des sciences de la réadaptation / School of Rehabilitation Sciences	Étudiant-chercheur / Student-researcher
Hilmi DAJANI	École de science informatique et de génie électrique / School of Electrical Engineering and Computer Science	Co-chercheur / Co-investigator

Conditions spéciales ou commentaires / Special conditions or comments

550, rue Cumberland, pièce 154 Ottawa (Ontario) K1N 6N5 Canada
550 Cumberland Street, Room 154 Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Le Comité d'éthique de la recherche (CÉR) de l'Université d'Ottawa, opérant conformément à l'*Énoncé de politique des Trois conseils* (2014) et toutes autres lois et tous règlements applicables, a examiné et approuvé la demande d'éthique du projet de recherche ci-nommé.

L'approbation est valide pour la durée indiquée plus haut et est sujette aux conditions énumérées dans la section intitulée "Conditions Spéciales ou Commentaires". Le formulaire « Renouvellement ou Fermeture de Projet » doit être complété quatre semaines avant la date d'échéance indiquée ci-haut afin de demander un renouvellement de cette approbation éthique ou afin de fermer le dossier.

Toutes modifications apportées au projet doivent être approuvées par le CÉR avant leur mise en place, sauf si le participant doit être retiré en raison d'un danger immédiat ou s'il s'agit d'un changement ayant trait à des éléments administratifs ou logistiques du projet. Les chercheurs doivent aviser le CÉR dans les plus brefs délais de tout changement pouvant augmenter le niveau de risque aux participants ou pouvant affecter considérablement le déroulement du projet, rapporter tout évènement imprévu ou indésirable et soumettre toute nouvelle information pouvant nuire à la conduite du projet ou à la sécurité des participants.

The University of Ottawa Research Ethics Board, which operates in accordance with the *Tri-Council Policy Statement* (2014) and other applicable laws and regulations, has examined and approved the ethics application for the above-named research project.

Ethics approval is valid for the period indicated above and is subject to the conditions listed in the section entitled "Special Conditions or Comments". The "Renewal/Project Closure" form must be completedfour weeks before the above-referenced expiry date to request a renewal of this ethics approval or closure of the file.

Any changes made to the project must be approved by the REB before being implemented, except when necessary to remove participants from immediate endangerment or when the modification(s) only pertain to administrative or logistical components of the project. Investigators must also promptly alert the REB of any changes that increase the risk to participant(s), any changes that considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project or the safety of the participant(s).

Coordonateur / COORDINATOR
Coordonnateur de l'éthique / Ethics Coordinator
Pour/For **Daniel LAGAREC** Président(e) du/ Chair of the **Comité d'éthique de la recherche en sciences de la santé et sciences / Health Sciences and Sciences Research Ethics Board**

550, rue Cumberland, pièce 154 Ottawa (Ontario) K1N 6N5 Canada
550 Cumberland Street, Room 154 Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

C. Standardized Mini-Mental State Examination (SMMSE)

BCGuidelines.ca

Guidelines & Protocols Advisory Committee

STANDARDIZED MINI-MENTAL STATE EXAMINATION (SMMSE)

NAME OF PATIENT	DATE

Directions for administration of the SSMSE:

1. Before the questionnaire is administered, try to get the person to sit down facing you. Assess the person's ability to hear and understand very simple conversation, e.g. *What is your name?* If the person uses hearing or visual aids, provide these before starting.
2. Introduce yourself and try to get the person's confidence. Before you begin, get the person's permission to ask questions, e.g. *Would it be alright to ask you the same questions about your memory?* This helps to avoid catastrophic reactions.
3. Ask each question a maximum of three times. If the subject does not respond, score 0.
4. If the person answers incorrectly, score 0. Accept that answer and do not ask the question again, hint, or provide any physical clues such as head shaking, etc.
5. The following equipment is required to administer the instrument: A watch, a pencil, Page 3 of this SMMSE with CLOSE YOUR EYES written in large letters and two five-sided figures intersecting to make a four-sided figure, and Page 4, a blank piece of paper.
6. If the person answers: What did you say?, do not explain or engage in conversation. Merely repeat the same directions a maximum of three times.
7. If the person interrupts (e.g. What is this for?), reply: *I will explain in a few minutes, when we are finished. Now if we could proceed please… we are almost finished.*

I am going to ask you some questions and give you some problems to solve. Please try to answer as best as you can.

1. Time: 10 seconds for each reply:

a)	*What year is this?* (accept exact answer only).	/1
b)	*What season is this?* (accept either: last week of the old season or first week of a new season).	/1
c)	*What month is this?* (accept either: the first day of a new month or the last day of the previous month).	/1
d)	*What is today's date?* (accept previous or next date).	/1
e)	*What day of the week is this?* (accept exact answer only).	/1

2. Time: 10 seconds for each reply:

a)	*What country are we in?* (accept exact answer only).	/1
b)	*What province are we in?* (accept exact answer only).	/1
c)	*What city/town are we in?* (accept exact answer only).	/1
d)	(In home) *What is the street address of this house?* (accept street name and house number or equivalent in rural areas). (In facility) *What is the name of this building?* (accept exact name of institution only).	/1
e)	(In home) *What room are we in?* (accept exact answer only). (In facility) *What floor of the building are we on?* (accept exact answer only).	/1

3. Time: 20 seconds

Say: *I am going to name three objects. When I am finished, I want you to repeat them. Remember what they are because I am going to ask you to name them again in a few minutes.* (Say the following words slowly at approximately one-second intervals): *Ball / Car / Man.*

For repeated use: Bell, jar, fan; Bill, tar, can; Bull, bar, pan.
Please repeat the three items for me. (score one point for each correct reply on the first attempt.)
If the person did not repeat all three, repeat until they are learned or up to a maximum of five times (but only score first attempt). /3

Continued Over …

N.B. You may check WORLD reversal scores at: www.attentionMMSE.com

4. Time: 30 seconds
Spell the word WORLD. (you may help the person to spell the word correctly) **Say:** *Now spell it backwards please.* If the subject cannot spell world even with assistance, score 0. Refer to Page 3 for scoring instructions. /5

5. **Time: 10 seconds**
Say: *Now what were the three objects I asked you to remember?*
(score one point for each correct answer regardless of order) /3

6. Time: 10 seconds
Show wristwatch. Ask: *What is this called?*
(score one point for correct response: accept “wristwatch” or “watch”; do not accept “clock” or “time”, etc.). /1

7. Time: 10 seconds
Show pencil. Ask: *What is this called?*
(score one point for correct response; accept ”pencil” only; score 0 for pen) /1

8. Time: 10 seconds
Say: *I would like you to repeat a phrase after me: No ifs, ands or buts*
Score one point for a correct repetition. Must be exact, e.g. no ifs or buts, score 0). /1

9. Time: 10 seconds
Say: *Read the words on this page and then do what it says.* Then, hand the person the sheet with CLOSE YOUR EYES on it. If the subject just reads and does not close eyes, you may repeat: *Read the words on this page and then do what it says*, (a maximum of three times. Score one point only if the subject closes eyes. The subject does not have to read aloud. /1

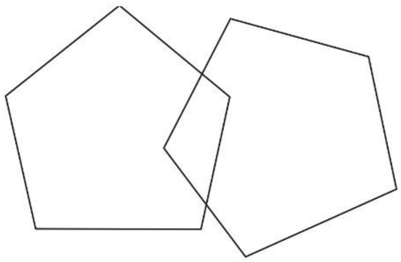
10. Time: 30 seconds
Hand the person a pencil and paper (Page 3). **Say:** *Write any complete sentence on that piece of paper.*
Score one point. The sentence must make sense. Ignore spelling errors. /1

11. Time: 1 minute maximum
Place design, eraser and pencil in front of the person. **Say:** *Copy this design please.* Allow multiple tries. Wait until the person is finished and hands it back. Score one point for a correctly copied diagram. The person must have drawn a four-sided figure between two five-sided figures. /1

12. Time: 30 seconds
Ask the person if he is right or left handed. Take a piece of paper, hold it up in front of the person and **say:** *Take this paper in your right/left hand* (whichever is non-dominant), *fold the paper in half once with both hands and put the paper down on the floor.* Score one point for each instruction executed correctly.

Takes paper in correct hand	/1
Folds it in half	/1
Puts it on the floor	/1
Total Test Score:	**/30**
Adjusted Score	**/**

 Molloy DW, Alemayehu E, Roberts R. Reliability of a standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. American Journal of Psychiatry, 1991; 148(1): 102-105.



FOLD LINE

Scoring WORLD backwards (instructions for item #4)

Write the person's response below the correct response.
Draw lines matching the same letters in the correct response and the response given.
These lines MUST NOT cross each other.
The person's score is the maximum number of lines that can be drawn without crossing any.

Examples:

D	L	R	O	W	
D	L	R	O	W	= Score 5

D	L	R	O	W	
D	R	W	O	D	= Score 3

D	L	R	O	W	
L	O	W	R	O	= Score 3

D	L	R	O	W	
L					= Score 1

D	L	R	O	W	
L	R	R	W	O	= Score 3

D	L	R	O	W	
					= ________

FOLD ALONG THIS LINE AND SHOW INSTRUCTIONS TO PERSON

Close your eyes

Item 10: **Sentence Writing**

Standardized Mini-Mental State Examination (SMMSE) Scoring

Table 1: Stages of Cognitive Impairment as Defined by SMMSE Scores

SCORE	DESCRIPTION	STAGE	DURATION (years)
30-26	could be normal	could be normal	varies
25-20	mild	early	0-23
19-10	moderate	middle	4-7
9-0	severe	late	7-14

Table 2: Areas of Functional Impairment

SCORE	ACTIVITIES OF DAILY LIVING	COMMUNICATION	MEMORY
30-26	could be normal	could be normal	could be normal
25-20	driving, finances, shopping	finding words, repeating, going off topic	three-item recall, orientation to time then place
19-10	dressing, grooming, toileting	sentence fragments, vague terms (e.g., this, that)	spelling WORLD backward, language, and three-step command
9-0	eating, walking	speech disturbances such as stuttering and slurring	obvious deficits in all areas

Adapted from: Vertesi A, Lever JA, Molloy DW, et al. Standardized mini-mental state examination: Use and interpretation. Canadian Family Physician 2001; 47:2018-2023.

The score for WORLD reversal is 17 per cent of SMMSE score (5 of 30 points). Incorrect scoring of WORLD reversal may result in incorrect assumptions of clinical change. One can review the score for WORLD reversal at: www.attentionMMSE.com. Self-learning of this task may also be done at this website.

Reference:
Davey RJ, Jamieson S. The validity of using the mini mental state examination in NICE dementia guidelines. J Neurol Neurosurg Psychiatry. 2004; 75:343-44.

D. Language Experience and Proficiency Questionnaire (LEAP-Q)

Northwestern Bilingualism & Psycholinguistics Research Laboratory
Marian, Blumenfeld, & Kaushanskaya (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research, 50* (4), 940-967.
Adapted to pencil-and-paper version by Marilyn Logan

Language Experience and Proficiency Questionnaire (LEAP-Q), Version for Canada

Last name		First name		Today's Date	
Age		Date of Birth		Male ☐	Female ☐

(1) Please list all the languages you know **in order of dominance**:

1	2	3	4	5

(2) Please list all the languages you know **in order of acquisition** (your native language first):

1	2	3	4	5

(3) Please list what percentage of the time you are *currently* and *on average* exposed to each language.
(*Your percentages should add up to 100%*):

List language here:					
List percentage here:					

(4) When choosing to read a text available in all your languages, in what percentage of cases would you choose to read it in each of your languages? Assume that the original was written in another language, which is unknown to you. (*Your percentages should add up to 100%*):

List language here:					
List percentage here:					

(5) When choosing a language to speak with a person who is equally fluent in all your languages, what percentage of time would you choose to speak each language? Please report percent of total time.
(*Your percentages should add up to 100%*):

List language here					
List percentage here:					

(6) Please name the cultures with which you identify. On a scale from zero to ten, please rate the extent to which you identify with each culture. (Examples of possible cultures include US-American, Chinese, Jewish-Orthodox, etc.):

Culture: ______________

0	1	2	3	4	5	6	7	8	9	10
No identification	Very low identification				Moderate identification					Complete identification

Culture: _______________

0	1	2	3	4	5	6	7	8	9	10
No identification	Very low identification				Moderate identification					Complete identification

Culture: _______________

0	1	2	3	4	5	6	7	8	9	10
No identification	Very low identification				Moderate identification					Complete identification

Culture: _______________

0	1	2	3	4	5	6	7	8	9	10
No identification	Very low identification				Moderate identification					Complete identification

Culture: _______________

0	1	2	3	4	5	6	7	8	9	10
No identification	Very low identification				Moderate identification					Complete identification

(7) How many years of formal education do you have? _________________________________

Please check your highest education level (or the approximate US equivalent to a degree obtained in another country):

- ☐ Less than High School
- ☐ High School
- ☐ Professional Training
- ☐ Some College / CEGEP
- ☐ College / CEGEP
- ☐ Some University
- ☐ University
- ☐ Some Graduate School
- ☐ Masters
- ☐ Ph.D./M.D./J.D.
- ☐ Other:

(8) Date of immigration to Canada, if applicable ____.
If you have ever immigrated to another country, please provide name of country and date of immigration here.

____

(9) Have you ever had a vision problem☐, hearing impairment☐, language disability☐, or learning disability ☐? (Check all applicable).

If yes, please explain (including any corrections):

____

Language:

This is my (**native** **second** **third** **fourth** **fifth**) language.

(1) Age when you…

began acquiring this language:	*became fluent in this language*:	*began reading in this language*:	*became fluent reading in this language*:

(2) Please list the number of years and months you spent in each language environment:

	Years	Months
A country where this language is spoken		
A family where this language is spoken		
A school and/or working environment where this language is spoken		

(3) Please circle your *level of* ***proficiency*** in speaking, understanding, and reading in this language:

Speaking

0	1	2	3	4	5	6	7	8	9	10
None	Very low	Low	Fair	Slightly less than adequate	Adequate	Slightly more than adequate	Good	Very good	Excellent	Perfect

Understanding spoken language

0	1	2	3	4	5	6	7	8	9	10
None	Very low	Low	Fair	Slightly less than adequate	Adequate	Slightly more than adequate	Good	Very good	Excellent	Perfect

Reading

0	1	2	3	4	5	6	7	8	9	10
None	Very low	Low	Fair	Slightly less than adequate	Adequate	Slightly more than adequate	Good	Very good	Excellent	Perfect

(4) Please circle how much the following factors contributed to you learning this language:

Interacting with friends

0	1	2	3	4	5	6	7	8	9	10
Not a contributor	Minimal contributor				Moderate contributor					Most important contributor

Interacting with family

0	1	2	3	4	5	6	7	8	9	10
Not a contributor	Minimal contributor				Moderate contributor					Most important contributor

Reading

0	1	2	3	4	5	6	7	8	9	10
Not a contributor	Minimal contributor				Moderate contributor					Most important contributor

Language tapes/self-instruction

0	1	2	3	4	5	6	7	8	9	10
Not a contributor	Minimal contributor				Moderate contributor					Most important contributor

TV

0	1	2	3	4	5	6	7	8	9	10
Not a contributor	Minimal contributor				Moderate contributor					Most important contributor

Listening to the radio

0	1	2	3	4	5	6	7	8	9	10
Not a contributor	Minimal contributor				Moderate contributor					Most important contributor

(5) Please circle to what extent you are currently exposed to this language in the following contexts:

Interacting with friends

0	1	2	3	4	5	6	7	8	9	10
Never	Almost Never				Half of the time					Always

Interacting with family

0	1	2	3	4	5	6	7	8	9	10
Never	Almost Never				Half of the time					Always

Watching TV

0	1	2	3	4	5	6	7	8	9	10
Never	Almost Never				Half of the time					Always

Listening to radio/music

0	1	2	3	4	5	6	7	8	9	10
Never	Almost Never				Half of the time					Always

Reading

0	1	2	3	4	5	6	7	8	9	10
Never	Almost Never				Half of the time					Always

Language-lab/self-instruction

0	1	2	3	4	5	6	7	8	9	10
Never	Almost Never				Half of the time					Always

(6) In your perception, how much of a foreign accent do you have in this language?

0	1	2	3	4	5	6	7	8	9	10
None	Almost none	Very light	Light	Some	Moderate	Considerable	Heavy	Very heavy	Extremely heavy	Pervasive

(7) Please circle how frequently others identify you as a non-native speaker based on your <u>accent</u> in this language:

0	1	2	3	4	5	6	7	8	9	10
Never	Almost Never				Half of the time					Always

E. Demographic and Music Experience Questionnaire

Demographic and Music Experience

Questionnaire

Hearing Sensitivity Among Young Adult Non-Musicians

Piano Pedagogy Research Laboratory
Audiology and Speech-Language Pathology Program

University of Ottawa

Demographic & Music Experience - CON
Questionnaire

General Information

Participant ID *(to be filled in by the Clinic)*	
Date	

NOTE

Rest assured that this information will remain strictly confidential. Only the members of the research team will have access to this information. Only group data (e.g. group averages) will be made public when the results of this study are presented in scientific conferences or similar contexts.

PLEASE CIRCLE OR ENTER YOUR RESPONSES TO THE QUESTIONS BELOW

SECTION 1: DEMOGRAPHIC AND HEARING-RELATED INFORMATION

QUESTIONS		YOUR RESPONSE
1	Gender	
2	Age	
3	If you are a university student, what is your current program?	
4	Do you have any concerns about your hearing?	• Yes: Specify: ________________ • No
5	Do you ever have difficulty understanding speech?	• Yes • No
6	Do you have a history of ear infections?	• Yes • No
7	Do you have tinnitus (ringing, roaring, hissing, buzzing in your ears?)	• Yes • No
8	Do you have otalgia (ear pain)?	• Yes • No
9	Do you have vertigo/disequilibrium?	• Yes • No
10	Are you regularly exposed to loud sound? (Working in a bar, construction, etc.)	• Yes: Specify: ________________ • No
11	Do you have a history of any noise exposure?	• Yes: Specify: ________________ • No
12	If you answered "yes" to the previous question, did you wear hearing protection when exposed?	• Yes • No
13	Have you ever had your hearing tested?	• Yes • No
14	If you answered "yes" to the previous question, what were the results?	
15	Have you ever seen a doctor who specializes in ears?	• Yes • No
16	Are you currently taking any medications?	• Yes • No
17	If you answered "yes" to the previous question, please list:	

SECTION 2: LANGUAGE INFORMATION

	Language	Age that you first learned it	Place you learned it (at home, at school, in an office, from neighbours, etc)	Your current degree of mastery of the language					
					Highly fluent	Fluent	Some mastery	A little mastery	Not at all
First language				Understand	☐	☐	☐	☐	☐
				Speak	☐	☐	☐	☐	☐
				Read	☐	☐	☐	☐	☐
				Write	☐	☐	☐	☐	☐
					Highly fluent	Fluent	Some mastery	A little mastery	Not at all
Second language				Understand	☐	☐	☐	☐	☐
				Speak	☐	☐	☐	☐	☐
				Read	☐	☐	☐	☐	☐
				Write	☐	☐	☐	☐	☐
					Highly fluent	Fluent	Some mastery	A little mastery	Not at all
Third language				Understand	☐	☐	☐	☐	☐
				Speak	☐	☐	☐	☐	☐
				Read	☐	☐	☐	☐	☐
				Write	☐	☐	☐	☐	☐
					Highly fluent	Fluent	Some mastery	A little mastery	Not at all
Fourth language				Understand	☐	☐	☐	☐	☐
				Speak	☐	☐	☐	☐	☐
				Read	☐	☐	☐	☐	☐
				Write	☐	☐	☐	☐	☐

SECTION 3: MUSIC EXPERIENCE

1	Have you ever played an instrument? (If not, go to question #15 of this section).	• Yes • No
2	If you answered yes, what instrument do you play?	
3	For how many years have you played this instrument?	Years:____________ Age of start: ___________________ Age of end: ____________________
4	If you are still playing, how many hours do you practice per week?	
5	If you play more than one instrument, what is the secondary instrument you play?	
6	For how many years have you played this secondary instrument?	Years:____________ Age of start: ___________________ Age of end: ____________________
7	If you are still playing this instrument, how many hours do you practice per week?	

8	How many years of formal training on a musical instrument (including voice) have you had during your lifetime? (circle one option)	• 0 • 0.5 • 1 • 2 • 3-5 • 6-9 • 10 or more
9	For how many years have you engaged in regular, daily practice of a musical instrument (including voice)? (circle one option)	• 0 • 1 • 2 • 3 • 4-5 • 6-9 • 10 or more
10	At the peak of your interest, how many hours a day did you practice on your primary instrument? (circle one option)	• 0 • 0.5 • 1 • 1.5 • 2 • 3-4 • 5 or more
11	*For how many years have you played or sung in a group, band, choir or orchestra? (circle one option)	• 0 • 1 • 2 • 3 • 4-5 • 6-9 • 10 or more
12	If you are still playing an instrument, where do you usually practice?	• Practice room with sound absorption/dampening • Practice room without sound absorption/dampening • At home • Other: ______________________
13	Do you use any auditory protection when you practice?	• Yes • No
14	If so, what kind?	

15	How many live music events have you attended as an audience member in the past twelve months? (circle one option)	• 0 • 1 • 2 • 3 • 4-6 • 7-10 • 11 or more
16	How much time do you spend per day attentively listening to music? (circle one option)	• 0-15 min • 15-30 min • 30-60 min • 60-90 min • 2 hrs • 2-3 hrs • 4 hrs or more

SECTION 4: PERSONAL MUSIC SYSTEMS

1	Do you regularly listen to a personal music system (PMS) with headphones? (iPod, smartphone, MP3 player, computer etc).	• Yes • No
2	How many minutes or hours per day do you listen to your PMS per day?	
3	At what volume level do you usually listen to your PMS?	• Soft • Medium • Loud • Very loud
4	What kind of headphones do you use when listening to music on your PMS? (You can circle more than one)	• Earbuds (fit in outer ear) • In-ear headphones (extend into the ear canal) • Circumaural (full-size headphones that encompass the entire ear) • Supra-aural headphones (pads that press against the ear)

THANK YOU FOR TAKING THE TIME TO FILL OUT THIS FORM

F. General development questionnaire

GENERAL DEVELOPMENT QUESTIONNAIRE

PARTICIPANT CODE: ______________ DATE: ________________________

Please select the most appropriate answer.

1. *Do you have difficulty hearing?* ☐ yes ☐ no
2. *Have you had many ear infections in the past?* ☐ yes ☐ no

   If so, have you had ventilation "tubes"? ☐ yes ☐ no
3. *Did you experience learning problems in school?* ☐ yes ☐ no
4. *Did you have to repeat a school year?* ☐ yes ☐ no
5. *Do you have difficulty concentrating?* ☐ yes ☐ no
6. *Where are you diagnosed with an attention deficit disorder?* ☐ yes ☐ no
7. *Have you taken (or do you take) medication to help you stay calm?* ☐ yes ☐ no
8. *Did you have any delay in language development?* ☐ yes ☐ no
9. *Have you ever received speech-language pathology services?* ☐ yes ☐ no
10. *Do you play, or have you ever learned to play a musical instrument?* ☐ yes ☐ no

    If so, since when (for how many years)? __________________________________

11. Are you in a choir, or do you sign on a regular basis? ☐ yes ☐ no

If so, since when (for how many years)? ______________________________

12. What is your "preferred language" for communicating? That is, in which language do you express yourself the best?

☐ English ☐ French ☐ Other: __________________

13. At what age did you start speaking English?

☐ 0-5 years old ☐ 6-10 years old ☐ 11-20 years old
☐ 21 years old or later

14. Do any of your parents speak to you in English, or have ever spoken to you in English?

☐ yes ☐ no

15. From kindergarten to the end of high school, how many years did you study full-time in English?
____________ *out of* ___________ *school years*

In which of the following school years did your education take place in English (full-time)?

☐ Kindergarten ☐ Grade 1 ☐ Grade 2 ☐ Grade 3
☐ Grade 4 ☐ Grade 5 ☐ Grade 6

☐ Grade 7 ☐ Grade 8 ☐ Grade 9 ☐ Grade 10
☐ Grade 11 ☐ Grade 12

☐ Secondary 1 ☐ Secondary 2 ☐ Secondary 3 ☐ Secondary 4
☐ Secondary 5

Notes: ____

____

____

G. Consent form for Chapter 3 (Study 1)

uOttawa

L'Université canadienne
Canada's university

Informed Consent Form

Measurement of Speech-Evoked Potentials in Normal Hearing Participants

You are invited to participate in this study. This informed consent form is to make sure that you understand the purpose of this study and the nature of your involvement. The research project is being conducted by Dr. Hilmi Dajani, Full Professor at the School of Electrical Engineering and Computer Science (613-562-5800,). Data from this study may be used by Ms. Maryam Karimi Boroujeni ( ), for her thesis projects, or by her thesis co-supervisor Dr. Christian Giguère, Full Professor at the School of Rehabilitation Sciences (613 562-5800 ext.).

Purpose of the study: The purpose of this study is to investigate the brain's electrical response to speech stimuli. This electrical signal, called the auditory-evoked potential, allows clinicians to diagnose hearing problems and find their origin in the brain. The sounds that clinicians most commonly use for this test are brief clicks or other artificial sounds. These types of sounds do not provide sufficient information to assess the ability to hear speech. As a result, we are interested in developing tests that record the brain's electrical response to speech, or the Speech-Evoked Potential, with a diverse set of speech samples and in different acoustic environments. The knowledge gained is expected to help in the development of better diagnostic tests of hearing loss and tools for evaluating hearing aids.

Procedure: The study is performed in the Audiology Laboratory, 1st floor, Roger-Guindon Hall, 451 Smyth Rd., Ottawa, ON K1H 8M5. Data is collected on a single day, and the entire session lasts about two hours.

The following is the step-by-step procedure that will be followed:

1. The participant's ears will first be examined using an otoscope, which is a device that shines a beam of light to help visualize and examine the condition of the ear canal and eardrum.

2. The hearing of the participant will be tested with pure tones at 500, 1000, 2000 and 4000 cycles per second. We will provide the participant with the result of this test. Those who are not found eligible following the hearing test will be excluded.

3. Participants will be asked to sit in a comfortable chair inside a

Université d'Ottawa
Faculté de génie

École de science informatique et de génie électrique

University of Ottawa
Faculty of Engineering

School of Electrical Engineering and Computer Science

☎ 613 562 5800
800 King-Edward
Ottawa ON K1N 6N5 Canada

www.uOttawa.ca

sound-proof room and will be asked to minimize movement during the recordings.

4. The vertex (at centre of the top of the scalp) or the centre of the forehead, and earlobes will be cleaned with a mild abrasive cream that prepares the skin for attaching electrodes with minimal discomfort and irritation. Then, an electrode filled with conductive gel will be placed on the vertex or centre of forehead, and two other electrodes will be placed on the earlobes to measure the brain's response to sounds. Foam earphones will be inserted into the participant's ear canals. The earphones will deliver the sound at a level calibrated at less than or equal to 85 decibels (SPL or Sound Pressure Level). This level is well within the acceptable safe range for sounds, and well below Canadian occupational exposure limits.

Participant description: Adults with normal hearing and without any external or middle ear disease will be included.

Risks to participating: There is no danger or risk to health associated with this study. All procedures have been pre-tested, and they have been used routinely for many years in hospitals, clinics, and laboratories.

Withdrawing from the study: Your participation in this study is voluntary. You may withdraw from the study at any time, by verbally informing any of the investigators, even after signing this form, or even after data collection has ended. There will be no consequences following this action, and withdrawing from the study will have no bearing on the participants' relationship with the researchers. Moreover, the collected data will be deleted.

Compensation: You will receive $20 for participating in the study. Participants will still receive the compensation should they choose to withdraw from the study.

Confidentiality: Any information about you collected during this study will be kept strictly confidential. Your name will not be associated with the collected data in any way. While the results will appear in students' dissertations and may also be published, you will not be identified.

The data collection is conducted by Dr. Dajani, Dr. Giguère, or their graduate student. The records will be kept on computer files. The files will be password-protected. Only Dr. Dajani, Dr. Giguère, and their graduate student will have access to the data.

The data will be kept in Dr. Dajani's office when not in use and will be conserved for 10 years, after which the computer files will be deleted, and the consent forms will be shredded.

Benefits of the Study: Hearing impairment is an important public health problem. This study will help develop better technologies to measure hearing loss.

Funding Source: Funding for this study is provided by the Natural Sciences and Engineering Research Council of Canada (NSERC)

In closing: With your participation, you will be given a copy of this consent form. You may ask questions at any time, even after signing this consent form.
If you have any questions about this study, you may contact Dr. Dajani, Dr. Giguère, or Ms. Karimi Boroujeni (contact information provided in first paragraph of this form).

If you have any concerns with regards to the ethical conduct of the study, you may contact the Protocol Officer for Ethics in Research, University of Ottawa, Tabaret Hall, 550 Cumberland Street, Room 159, Ottawa, ON K1N 6N5, tel.: 613-562-5841, email: ethics@uottawa.ca.

Signatures: I have read the above description of the study and understand the conditions of participation. My signature indicates that I agree to participate in the study.

Name of the participant (please print):

____

Participant's signature:

________________________________ **Date:** ____________

Researcher's signature:

________________________________ **Date:** ____________

H. Consent form for Chapter 4 (Study 2)

uOttawa

L'Université canadienne
Canada's university

Informed Consent Form

Measurement of Speech-Evoked Potentials in Normal Hearing and Hearing Impaired Participants

You are invited to participate in this study. This informed consent form is to make sure that you understand the purpose of this study and the nature of your involvement. The research project is being conducted by Dr. Hilmi Dajani, Full Professor at the School of Electrical Engineering and Computer Science (613-562-5800,). Data from this study may be used by Ms. Maryam Karimi Boroujeni ( ), for her thesis projects, or by her thesis co-supervisor Dr. Christian Giguère, Full Professor at the School of Rehabilitation Sciences (613 562-5800 ext.).

Purpose of the study: The purpose of this study is to investigate the brain's electrical response to speech stimuli. This electrical signal, called the auditory-evoked potential, allows clinicians to diagnose hearing problems and find their origin in the brain. The sounds that clinicians most commonly use for this test are brief clicks or other artificial sounds. These types of sounds do not provide sufficient information to assess the ability to hear speech. As a result, we are interested in developing tests that record the brain's electrical response to speech, or the Speech-Evoked Potential, with a diverse set of speech samples and in different acoustic environments. The knowledge gained is expected to help in the development of better diagnostic tests of hearing loss and tools for evaluating hearing aids.

Procedure: The study is performed in the Audiology Laboratory, 2nd floor, FHS Building, 200 Lees, Ottawa, ON K1S 5S9. Data is collected in a single session for the normal hearing population and in two sessions for hearing impaired participants. Each session lasts approximately two hours.

The following is the step-by-step procedure that will be followed:

1. The participant's ears will first be examined using an otoscope, which is a device that shines a beam of light to help visualize and examine the condition of the ear canal and eardrum.

2. The hearing of the participant will be tested with pure tones at 500, 1000, 2000 and 4000 cycles per second. We will provide the participant with the result of this test. Those who are not found eligible following the hearing test will be excluded.

Université d'Ottawa
Faculté de génie

École de science informatique et de génie électrique

University of Ottawa
Faculty of Engineering

School of Electrical Engineering and Computer Science

☎ 613 562 5800
800 King-Edward
Ottawa ON K1N 6N5 Canada

www.uOttawa.ca

3. Participants will be asked to sit in a comfortable chair inside a sound-proof room and will be asked to minimize movement during the recordings.

4. The vertex (at centre of the top of the scalp) or the centre of the forehead, and earlobes will be cleaned with a mild abrasive cream that prepares the skin for attaching electrodes with minimal discomfort and irritation. Then, an electrode filled with conductive gel will be placed on the vertex or centre of forehead, and two other electrodes will be placed on the earlobes to measure the brain's response to sounds. Foam earphones will be inserted into the participant's ear canals. The earphones will deliver the sound at a level calibrated at less than or equal to 85 decibels (SPL or Sound Pressure Level). This level is well within the acceptable safe range for sounds, and well below Canadian occupational exposure limits.

Participant description: Adults with normal hearing as well as hearing loss and without any external or middle ear disease will be included.

Risks to participating: There is no danger or risk to health associated with this study. All procedures have been pre-tested, and they have been used routinely for many years in hospitals, clinics, and laboratories.

Withdrawing from the study: Your participation in this study is voluntary. You may withdraw from the study at any time, by verbally informing any of the investigators, even after signing this form, or even after data collection has ended. There will be no consequences following this action, and withdrawing from the study will have no bearing on the participants' relationship with the researchers. Moreover, the collected data will be deleted.

Compensation: You will receive $20 for participating in the study. Participants will still receive the compensation should they choose to withdraw from the study.

Confidentiality: Any information about you collected during this study will be kept strictly confidential. Your name will not be associated with the collected data in any way. While the results will appear in students' dissertations and may also be published, you will not be identified.

The data collection is conducted by Dr. Dajani, Dr. Giguère, or their graduate student. The records will be kept on computer files. The files will be password-protected. Only Dr. Dajani, Dr. Giguère, and their graduate student will have access to the data.

The data will be kept in Dr. Dajani's office when not in use and will be conserved for 10 years, after which the computer files will be deleted, and the consent forms will be shredded.

Benefits of the Study: Hearing impairment is an important public health problem. This study will help develop better technologies to measure hearing loss.

Funding Source: Funding for this study is provided by the Natural Sciences and Engineering Research Council of Canada (NSERC)

In closing: With your participation, you will be given a copy of this consent form. You may ask questions at any time, even after signing this consent form.
If you have any questions about this study, you may contact Dr. Dajani, Dr. Giguère, or Ms. Karimi Boroujeni (contact information provided in first paragraph of this form).

If you have any concerns with regards to the ethical conduct of the study, you may contact the Protocol Officer for Ethics in Research, University of Ottawa, Tabaret Hall, 550 Cumberland Street, Room 159, Ottawa, ON K1N 6N5, tel.: 613-562-5841, email: ethics@uottawa.ca.

Signatures: I have read the above description of the study and understand the conditions of participation. My signature indicates that I agree to participate in the study.

Name of the participant (please print):

____

Participant's signature:

________________________________ **Date:** ____________

Researcher's signature:

________________________________ **Date:** ____________

I. Consent form for Chapter 5 (Study 3)

uOttawa

Université d'Ottawa
Faculté des sciences
de la santé

École des sciences de la
réadaptation

University of Ottawa
Faculty of Health
Sciences

School of Rehabilitation
Sciences

200 Lees
Ottawa, ON, K1S 5S9 Canada

www.uOttawa.ca

Consent Form

---

1. **Title of the research project**

Prosody perception in hearing impaired listeners: effect of hearing aid processing and noise

2. **Name of the researchers**

Lead investigator:
Christian Giguère, Ph.D.
Full Professor, Audiology and Speech-Language Pathology Program
Guindon Hall, room 260, 200 Lees Ave, Ottawa, Ontario, K1S 5S9
Email: Phone:

Co-investigator:
Hilmi Dajani, Ph.D.
Full Professor, School of Electrical Engineering and Computer Science
STE, room 5102, 800 King Edward Avenue, Ottawa, Ontario K1N 6N5
Email: Phone:

Ph.D. student:
Maryam Karimi Boroujeni Email:

3. **Invitation to take part in research study**

Many people experience difficulties in understanding speech such as perceiving rhythm and intonation of speech, also known as prosody, specifically in adverse listening situations such as noisy environments. This research study focuses on how hearing loss, hearing aid processing, and background noise can intensify this problem. Also, the relationship between prosody perception and speech in noise understanding will be assessed in both people with normal hearing and hearing-impaired individuals.

4. Your involvement in the research study

Should you accept to take part in the research study, which is also a requirement for the Ph.D.'s research project of the student mentioned on the first page, you will be asked to take part in a 90-110 minute test session at the Hearing Research Laboratory of the University of Ottawa. This session will include the following measures:

1) Audiometric assessment

This measure is similar to what is done in audiology clinics and allows us to verify your hearing status. Using an otoscope, your auditory canals will be examined to ensure that nothing could prevent you from taking part in the study (for example earwax blockage or active ear infection). Your hearing will be checked by having you listen to soft sounds presented through headphones.

Finally, to assess your ability to understand speech in noise, you will be asked to repeat sentences presented in background noise.

2) Measure of emotion identification

You will be asked to identify series of emotions embedded in sentences presented through headphones, in quiet and in background noise. You will be encouraged to guess the given emotion when unsure of what was heard. These tests will be presented simultaneously with background noise, one ear at a time or in both ears. The competing background noise consists of babble noise (many people talking at once).

Results of the hearing assessment will be explained. However, results on the measures of prosody perception will not be communicated to you since the data analysis will occur at a later time and will performed on pooled data from all participants.

5. Benefits

You will be receiving a free hearing test. You will also be contributing to general knowledge in the field of Audiology, particularly a better understanding of prosody perception and difficulties understanding speech in background noise. This test will allow determining if prosodic deficit and difficulties understanding speech in background noise can be explained by hearing loss or hearing aid processing. Your participation will also result in paving the path for providing or developing better diagnostic or rehabilitation tools for people with hearing loss in the future.

6. Risks and inconvenience

There are no risks involved with your participation in this research study. However, you could experience slight stress of fatigue. To minimize fatigue, short breaks are planned throughout the testing session, and additional breaks can be taken if required.

7. **Confidentiality**

All data collected during this research study will be kept confidential. Your name will be replaced by an identification number on all forms. All information and data will be kept in a locked file at the Hearing Research Laboratory at Guindon Hall and will be destroyed five years after the publication of results. However, to verify the proper conduct of this research and the management of data, a representative from the Ethics Research Board may ask to review the data. You will also receive a copy of the signed consent form for your personal records. In the event that data from this research study is published or presented at a professional conference, none of your personal information will be revealed.

8. **Voluntary participation**

Your participation is fully voluntary. You can, at any time and without consequence, refuse to take part in the study or cancel your appointment. Should you choose to withdraw from the study, the data collected will be destroyed immediately.

9. **Declaring responsibility**

By signing this form, you are not releasing the researchers, or the organization of their legal and professional responsibilities should an event occur that causes you any harm or prejudice.

10. **Further questions or in the case of difficulties**

For any additional information regarding this research study, you can contact the lead researcher, Christian Giguère, at (613) 562-5800 For questions or additional information about the ethical aspects of this research, please contact the Protocol Officer for Research Ethics, University of Ottawa, Tabaret Hall, 550 Cumberland Street, room 154, (613) 562-5387 or ethics@uottawa.ca.

11. **Consent**

The nature and purpose of this study, as well as what will be required from me, have been explained. I've reviewed the consent form. I've also had ample opportunity to ask questions, in which case adequate answers were provided. By signing this form, I agree to take part in the research study.

Name of participant: ____

Signature of participant: ____

Age (at time of participation): ____________________Date: ______________________

J. Table S1. Features of the assessment tools evaluating prosody perception

Assessment tools	Aim	Targeted population	Types of prosody	Subsets	Scoring	Estimated time to complete test
Minnesota Tests of Affective Processing (MNTAP) (Lai et al., 1991)	To evaluates emotion recognition and discrimination in faces and speech samples for children with attention-deficit hyperactivity disorder (ADHD)	Children (6-11)	Affective	4 auditory subsets: Prosody/content preference, Lexical comprehension, Prosody/content congruence, Cross-modal matching 11 visual subsets: Training task, Identity match 1, Identity match 2, Faces teaching, Affect match, Affect naming, Affect choice, Gesture recognition, Localization memory, Face and object recognition memory, Sequential Face Pairs Memory	Speech stimuli are delivered in the form of recorded samples and visual stimuli displayed on the screen. Then, the accuracy of participants' responses is calculated.	2-3 hours
Aprosodia battery (Ross et al., 1997)	1- To distinguish different patterns of affective-prosodic deficits in patients with neurological disorders. 2- To assess the expression of affective prosody	Adults (> 17)	Affective	1- Affective–prosodic comprehension and discrimination: (Word comprehension, monosyllabic comprehension, asyllabic comprehension) 2-Affective–prosodic production: (Word repetition, monosyllabic	For calculating receptive prosody: The number of correct responses for word comprehension,monosyllabic comprehension, asyllabic comprehension, with each score ranging from 0 to 24 For calculating expressive prosody: 1-The computer-assisted techniques measure various	One hour

				repetition, asyllabic repetition, spontaneous affective prosody)	aspects of intonation, timing, and intensity 2- rating the accuracy of prosody production by expert	
Perception of Prosody Assessment Tool (PPAT) (Klieve et al., 1998)	To evaluate the prosody perception abilities of children with CI	Children (7-12)	Linguistic & Affective	Apart from linguistic context: 1-Pitch, intensity, duration Within a linguistic context 2- Question, statement, command forms 3- Tone and affect 4- Grammatical class 5- Stress 6- Compound and abutting words	Calculating the correct percentage of responses	One hour
Profiling Elements of Prosody in Speech-Communication (PEPS-C) (Peppé & McCann, 2003)	To assess understanding and expression of four communication areas with focus on prosody and intonation skills, including interaction, affect, boundary (chunking) and focus	Children (5-4) with autism; later it was applied for adults (18-52)	Linguistic & Affective	6 subsets for prosody perception: Short item discrimination, Turn-End reception, Affect reception, Long item discrimination, Chunking reception, Contrastive Stress Reception 6 subsets for prosody production: Short Item imitation, Turn-End expression, Affect expression, Long item imitation, Chunking expression, Contrastive stress expression	Calculating accurate responses via automatic scoring	One hour
Advanced Clinical Solution (ACS) (Pearson, 2009)	To evaluate social functioning deficits	Adult (16-70)	Affective	1- Facial affect recognition, 2- Affect naming, 3- Prosody-face matching, 4- Prosody-pair matching	Calculating the correct percentage of responses via manual scoring	45 minutes

Emotional Communication in Hearing Questionnaire (EMO-CHeQ) (singh et al., 2019)	A 16-item self-reported scale to assess the experiences of hearing and handicap when listening to signals with vocal emotion information	Adults	Affective	1- characteristics of encountered talkers, 2- communication in challenging listening situations, 3- speech production (e.g., the ability to convey emotion in a subtle manner using one's own voice), and 4- the associated impact of such deficits on socioemotional well-being	Participants should indicate the level of agreement with each statement using the following rating scale: 1 = Strongly disagree; 2 = Slightly disagree; 3 = Neither agree nor disagree; 4 = Slightly agree; 5 = Strongly agree. Then the mean score for total items and each subdomain are computed. The higher scores indicate more difficulty.	Less than 15 minutes
Test of Prosody via Syllable Emphasis (TOPsy) (Nayak et al., 2022)	To measure lexical stress sensitivity in adults	Adults (18-88)	Linguistic	28 items that have been loaded onto two distinct factors corresponding to initially stressed vs. non-initially stressed word	The recordings of a multi-syllabic spoken word are administered over the internet or in lab settings. The participants are asked to identify the syllable that holds the main stress or emphasis in the word. Then, their scores are automatically calculated.	10 minutes

K. Table S2. Summary of studies focused on prosody perception in hearing aid and cochlear implant users

Authors, Year	Sample size	Age range	Study language	Type of auditory prosthesis	Method	Types of prosody	Result
Luo et al, 2007	8 adult NH subjects 8 adult CI users	22_40 years 41_73 years	English	CI	Behavioral measures: The House Ear Institute emotional speech database (HEI-ESD)	Affective	Although CI performance remained lower than that of NH individuals, CI users were able to use some prosodic cues (amplitude and duration) to obtain moderate levels of vocal emotion perception. Removing the amplitude variations worsened the performance of CI users.
Meister et al, 2009	12 CI recipients 12 NH listener	38_75 years 34_68 years	German	CI	Behavioral measures: question/statement test, sentence stress test, and a speaker gender test	linguistic	The identification of sentence stress, questions/statements, and speaker gender in CI users were qualitatively similar to the NH group. In contrast, most CI recipients had poor performance quantitatively, particularly in sentence stress test.
Straatman et al, 2010	14 NH children 19 children with bimodal fitting	6.8_16.7 years 6.0_19.8 years	Dutch	CI alone and bimodal hearing	Behavioral measures: _ Discrimination of pitch movements in bisyllabic non-words _ Discrimination between questions and affirmations	linguistic	Compared to using CI-only, bimodal fitting improved the performance of children in discriminating F0 excursions and identifying questions vs. affirmations. However, even with a hearing aid, children need exaggerated F0 contour.
Meister et al, 2011	13 CI recipients 13 NH listeners	19_81 years 18_73 years	German	CI	Behavioral measures: discrimination and identification task for stress-based sentences (Natural utterances and	linguistic	Discrimination and stress identification based on F0, and intensity had been degraded in CI users compared to NH group. This

					artificial manipulations of stimuli)		effect was not shown for duration cues.
Most et al, 2011	23 adults	17_65 years	English	CI alone and bimodal hearing	Behavioral measures: _ Intonation test _ Syllable stress test _ Word emphasis test	linguistic	Bimodal fitting showed a better score in all three tests compared to CI-only condition, suggesting HA in non-implanted ear provides better low-frequency acoustic hearing which is required for prosody perception.
Cullington & Zeng, 2011	13 adults with CI 13 children with bimodal fitting	18 years and above	English	Bilateral CI and bimodal hearing	Behavioral measures: _ Hearing in Noise Test (HINT) _ Montreal Battery of Evaluation of Amusia _ Aprosodia Battery _ Talker identification	Affective	The mean score of affective prosody discrimination was not significantly different between the bimodal and bilateral groups; however, the bimodal group had a better performance. Such finding was also observed for all the other tests.
Most & Michaelis, 2012	26 children with HL 14 children with NH who mostly wear HA	4.0_6.6 years	Nonsense sentence "bado mino gana"	HA	Behavioral measure: The Emotion Identification Test (EIT) under three presentation conditions: auditory, visual, and auditory–visual	Affective	The corrected mean percent scores were higher for children with NH compared to children with HL in each of the three conditions. However, children with HA performed better in auditory–visual condition than auditory alone condition, suggesting benefiting from auditory cues provided by HA.
Agrawal et al, 2012	28 NH listeners	25–55 years	German	Simulating CI processing	Behavioral measures: reaction time and accuracy rates for neutral sentences and vocoded stimuli which had been simulated two CI speech-coding strategies (PACE and ACE) Electrophysiological	Affective	Faster responses, higher identification rates, and larger P200 amplitude were detected for original stimuli. For vocoded stimuli, better recognition of happy and angry sentences was demonstrated in comparison to natural ones especially for PACE strategy.

					test: An N100-P200 complex		
Agrawal et al, 2013	20 CI users (2 subgroups based on coding strategy: MP3000 and ACE) 20 normal-hearing participants	25–60 years	German	CI	Behavioral measures: Reaction time and accuracy rates for sentences with neutral, happy and angry prosody Electrophysiological tests: Recording N100-P200 complex, Time-Frequency measure	Affective	CI users had difficulty in prosody recognition, as demonstrated by lower accuracy, longer reaction time, prolonged ERP latencies, and reduced amplitude. However, the MP3000 strategy was better than ACE one on happy emotion. In addition, time-frequency results supported the behavioral and ERP findings by showing power increase of Theta and Gamma bands for MP3000 strategy than ACE on emotional stimuli.
Hegarty & Faulkner, 2013	9 children with HL	7.4–14.6 years	English	CI alone and bimodal condition	Behavioral measures: _Discrimination test for bisyllabic word with manipulated F0 and amplitude _ Focus sentence test	linguistic	In both conditions (CI alone and bimodal fitting), better perception of prosodic patterns when children were presented with natural speech stimuli compared to sentences with pitch manipulations. Also, there was no bimodal advantage for stress and intonation perception compared to CI alone.
Van Zyl & Hanekom, 2013	10 adult CI user 10 adults with NH	21–70 years	Afrikaans	CI	Behavioral measures: _ question/statement distinction _ The discrimination of attitude as sarcastic or sincere _ Vowel recognition task Under quiet and noise	linguistic	No significant difference was observed for certain/hesitant task and vowel recognition task between adults with NH and CI users, suggesting CI gives access to durational cues even in noise condition. But CI recipients performed worse in question/ statement discrimination as this task was highly dependent on F0 variation, which were degraded in CI technology especially in the presence of noise.

Kalathottukaren et al, 2015	12 adult CI users 15 NH adults (normative data)	25–78 years 23–57 years	English	CI	Behavioral measures: 1- Profiling Elements of Prosody in Speech-Communication (PEPS-C) 2- DANVA 2 (diagnostic analysis of non-verbal accuracy 2) 3- MBEA (Montreal battery of evaluation of amusia)	Affective and linguistic	Lower score of CI users in turn-end, affect, music, and contrastive stress reception subtests were demonstrated, suggesting degraded pitch perception in electrical hearing. However, they showed no difference in chunking reception subtest.
Chatterjee et al, 2015	31 children with NH 36 children with CI 10 adults with NH 9 adults with CI	6.3 –18.7 years 6.8 –18.4 years mean age= 23.9 years 27.3–69.8 years	English	CI	Behavioral measures: _ 12 sentences from the HINT database spoken with five emotions _ CI simulated of the same sentences	Affective	The mean score of emotion identification for CI children and CI adults were as good as adults with NH listening to a CI simulated speech. However, NH children showed significantly poorer performance in response to CI-simulated speech. Taken together, this finding showed that not only CI technology but also sufficiently developed cognitive system in CI adults and learned adaptive strategies in CI children can help them to extract emotion from degraded speech.
Schmidt et al, 2016	23 older adults with mild to moderate HL 22 NH adults	65–82 years	German	Bilateral HA for at least 2 years	Behavioral measure: Rating level of emotion dimensions arousal (calm vs. aroused) and valence (positive vs. negative attitude) by using utterances from an authentic and affectively colored German conversational speech corpus	Affective	No significant difference was detected between hearing aid users and their NH peers in the perception of both valence and arousal emotions. However, for arousal dimension, HA users were more sensitive to intensity differences than NH adults.

Goy et al, 2016	11 older adults with mild to severe HL 27 NH younger adults	mean age = 77.2 years mean age = 18.3 years	English	HA with a median of 5 years' experience	Behavioral measure: word-recognition measure and emotion-identification measure for Toronto Emotional Speech Set	Affective	Older adults with HL performed worse on both word recognition and emotion identification when they were unaided. In the aided condition, their performance improved significantly for word recognition without much benefit for emotion identification.
Goy et al, 2017	Using simulated materials instead of recruiting people with HL and HA users	–	English	Simulating HA processing	Behavioral measure: Analyzing and comparing the acoustic cues used in emotional sentences of Toronto Emotional Speech Set after being manipulated to simulate sloping bilateral HL and HA outputs in different compressions (linear, slow, and fast)	Affective	Acoustic analysis of received sounds in the ear canal of a mannequin showed that haring aid processing made emotions less distinguishable in aspect of intensity variation and spectral cues. The hearing aid effect varied according to the type of emotion.
Arefi et al, 2017	40 NH young adult subjects	20–35 years	Persian	Simulating HA processing	Behavioral measure: Emotion recognition task before and after hearing aid compression simulation using sentences from Persian emotional speech database (ESD)	Affective	Depending on the type of emotion, recognition of Fear, sad, angry, and happy was reduced following simulating HA amplitude compression.
Laugen et al. (2017)	130 children with NH 35 preschool children with mild to severe HL	4–5 years	Norwegian	HA	Behavioral measure: The test of emotion comprehension	Affective	Children with HA performed similarly to normal hearing children, despite having lower vocabulary scores. Also, parents of hearing-impaired children were more accurate in the estimations of their children's emotional and cognitive development

Christensen et al, 2019	22 adults with mild to severe HL 30 NH adults	22–74 years 21–75 years	English	Bilateral HA	Behavioral measure: Evaluating accuracy rate and reaction time for emotion recognition of 12 sentences from the HINT database recorded in different emotions	Affective	In addition to age-related declines in emotion recognition, hearing loss further impacted the ability to perceive emotions, which was shown by reduced accuracy and increased reaction time.
Cannon and Chatterjee, 2019	19 school-aged children with mild to moderate HL 20 children with NH 11 NH young adults	8–14 years 6–17 years 9_24 years	English	HA	Behavioral measure: Emotion recognition using 12 sentences from the HINT database which had been produced in a child- and adult-directed manner	Affective	Children with mild to moderate hearing loss had similar accuracy and reaction time on emotion recognition tasks using either child- or adult- directed manner, although all participants performed better with child-directed materials. However, in children with HL, linguistic ability had a strong role in emotion recognition. Also, they showed that NH adults outperformed both groups of children, attributed to significant effect of developmental age on performance in the emotion recognition task.
Singh et al, 2019	12 NH adults 10 adults with HL but without (unaided condition) 10 hearing-impaired adults with	mean age = 66.8 years mean age = 73.5 years mean age = 73.5 years	English	Bilateral HA	Behavioral measures: _ 16-item version of the self-report questionnaire (the Emotional Communication in Hearing Questionnaire or EMO-CHeQ) _ Emotion-identifica- tion task under audio and audiovisual conditions	Affective	Both the HI and HA groups showed greater vocal emotion communication handicap on the EMO-CHeQ compared to NH group. However, the groups of aided and unaided listeners reported similar degrees of handicap. On the emotion-identification task, the HI and HA group performed worse than to the NH group in audio alone condition, with no group

	HA (aided condition)						differences being observed for audiovisual condition.
Legris et al, 2021	29 subjects with mild to moderate presbycusis 29 NH subjects	50–90 years 50–75 years	French	Bilateral HA	Behavioral measures: - Montreal Affective Voice Test (MAV) - The Profile of Emotional Competence questionnaire * Comparing patients' scores before and after hearing aid fitting	Affective	Despite the use of HA, the presbycusis individuals' score remained poorer compared to normal subjects. Based on the MAV results, the impaired auditory emotion perception of presbycusis patients was correlated with pure-tone average and age.

Abbreviations: HL=hearing loss, NH=normal hearing, HA= hearing aid, CI=cochlear implant.

L. Table S3. Pure tone hearing thresholds (dB HL) across test frequencies for normal-hearing (NH) and hearing-impaired (HI) participants.

Participants	Tested Ear	250 Hz	500 Hz	1000 Hz	2000 Hz	3000 Hz	4000 Hz	6000 Hz	8000 Hz
NH-P1	Right	15	15	15	10	20	10	5	15
NH-P2	Right	15	15	15	10	10	10	15	20
NH-P3	Right	25	25	15	15	15	15	20	15
NH-P4	Right	10	10	15	15	10	20	20	5
NH-P5	Right	25	15	10	15	10	15	5	15
NH-P6	Right	5	5	5	15	15	20	10	15
NH-P7	Right	15	15	15	15	20	25	25	20
NH-P8	Right	5	5	10	10	5	10	5	10
NH-P9	Right	5	0	5	-5	5	0	5	5
NH-P10	Right	10	10	15	10	15	10	-5	5
NH-P11	Right	20	15	15	15	15	10	0	5
NH-P12	Right	5	5	0	5	0	0	10	10
NH-P13	Right	15	20	20	15	10	15	10	20
NH-P14	Right	20	10	10	10	15	20	20	10
HI-P1	Right	55	45	30	10	10	10	10	5
HI-P2	Right	45	50	65	50	30	30	15	20
HI-P3	Right	60	65	65	70	55	50	35	40
HI-P4	Left	20	20	20	20	35	35	40	55
HI-P5	Left	70	70	70	65	60	55	55	70
HI-P6	Left	55	55	60	55	55	60	65	75
HI-P7	Right	30	45	40	30	25	20	15	25
HI-P8	Left	25	25	35	45	45	50	50	75
HI-P9	Left	30	35	35	45	45	45	50	50
HI-P10	Left	15	25	35	40	50	55	60	65
HI-P11	Left	15	10	15	30	30	35	35	50
HI-P12	Right	10	10	20	35	35	55	70	75
HI-P13	Left	15	15	15	30	60	60	60	65
HI-P14	Right	40	40	40	40	15	10	25	35

M. Table S4. Mean SNR Loss (dB) at 70 and 55 dB presentation levels for normal-hearing (NH) and hearing-impaired (HI) participants.

Participants	SNR Loss (70 dB)	SNR Loss (55 dB)
NH-P1	1	0.5
NH-P2	1	0.5
NH-P3	0.5	-1.5
NH-P4	0.5	1.5
NH-P5	0.5	0.5
NH-P6	2.5	2
NH-P7	1.5	4.5
NH-P8	3.5	2
NH-P9	2.5	1.5
NH-P10	1.5	0
NH-P11	2.5	2
NH-P12	2.5	2.5
NH-P13	0.5	-0.5
NH-P14	2.5	0.5
HI-P1	11	7.5
HI-P2	1.5	8
HI-P3	14.5	24.5
HI-P4	2.5	1.5
HI-P5	25	25.5
HI-P6	20.5	25.5
HI-P7	0.5	0.5
HI-P8	2.5	4.5
HI-P9	4.5	7
HI-P10	3.5	7.5
HI-P11	5.5	5
HI-P12	6	1.5
HI-P13	3.5	3
HI-P14	4.5	9