

Spring 2014

The neural representation and behavioral detection of frequency modulation

Daniel Elliott Shearer
James Madison University

Follow this and additional works at: <https://commons.lib.jmu.edu/diss201019>

 Part of the [Communication Sciences and Disorders Commons](#)

Recommended Citation

Shearer, Daniel Elliott, "The neural representation and behavioral detection of frequency modulation" (2014). *Dissertations*. 59.
<https://commons.lib.jmu.edu/diss201019/59>

This Dissertation is brought to you for free and open access by the The Graduate School at JMU Scholarly Commons. It has been accepted for inclusion in Dissertations by an authorized administrator of JMU Scholarly Commons. For more information, please contact dc_admin@jmu.edu.

The Neural Representation and Behavioral Detection of Frequency Modulation

Daniel Elliott Shearer

A dissertation submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Doctor of Audiology

Communication Sciences and Disorders

May 2014

Dedication

This dissertation is dedicated my parents, Lee and Mary Shearer, my siblings, Justin and Lauren, and my girlfriend, Courtney. To my parents, I owe so many of my accomplishments to you for your unending encouragement and support. To my siblings, after whom I have modeled my own character and personal aspirations, thank you for your countless pearls of wisdom, whether genuine or playfully flighty. To Courtney, who cools my nerves and delivers me from panic, I might not have survived this without you.

Acknowledgments

The completion and defense of this dissertation would not have been possible without the contributions of several individuals who offered invaluable assistance to me. First, I am tremendously appreciative of my advisor, Dr. Christopher Clinard for his willingness to allow me to work in his laboratory. His guidance, collaborative work, and resources were instrumental in this accomplishment. Additionally, his reliable mentorship, persistent instruction, and professionalism kept me on track toward completion. I must thank my committee members, Drs. Brenda Ryals and Lincoln Gray for their constructive feedback and support throughout the dissertation process. I am additionally thankful for my fellow students, Caitlin Cotter and Lacey Heckler, who graciously assisted in data collection and laboratory work.

In addition to thanking my loving family, who could always be relied on at times when their words of encouragement were most needed, I must thank my girlfriend, Courtney. Her comforting nature and unwavering support kept me afloat at times when the stresses of graduate schoolwork seemed too immense.

Finally, I must recognize the individuals within the Harrisonburg community who offered their time to participate in this study. Each of these individuals underwent hours of careful testing, and for this I am grateful.

Table of Contents

Dedication.....	ii
Acknowledgments.....	iii
Chapter I.....	1
Review of Literature	1
Introduction.....	1
Temporal envelope vs. temporal fine structure.....	2
Temporal fine structure processing is related to speech-in-noise understanding	3
Behavioral frequency modulation detection	5
Neural representation of temporal encoding within the auditory system	7
Neural encoding and the behavioral perception of sound.....	8
Purpose.....	10
Chapter II	12
Methods.....	12
Subjects	12
Stimuli.....	12
Behavioral procedure	13
Electrophysiology procedure	14
Speech-in-noise procedure.....	18
Statistical approach	18
Chapter III.....	20
Results.....	20

FMDLs	20
FFR stimulus-to-response cross correlations	21
QuickSIN	22
Relationships between behavioral and physiological measures	23
Chapter IV	27
Discussion	27
Introduction	27
Clinical relevance: behavioral assessment	31
Clinical relevance: Cochlear implant processing strategies	32
Methodological issues	33
Future directions	35
Conclusions	36
Chapter V	37
References	37

List of Figures

- Figure 1. Schematic spectrograms of Frequency-Following Response stimuli
- Figure 2. Time-domain FFR waveforms from an individual, age 23
- Figure 3. Boxplots showing mean listener FMDLs
- Figure 4. Scatterplots showing individual FMDLs
- Figure 5. Boxplots showing mean FFR stimulus-to-response correlations
- Figure 6. Boxplot showing mean listener QuickSIN performance
- Figure 7. Scatterplots between FFR stimulus-to-response correlations and FMDLs
- Figure 8. Scatterplots between FMDLs and SNR Loss
- Figure 9. Scatterplots between FFR stimulus-to-response correlations and SNR Loss

Abstract

Understanding a speech signal is reliant on the ability of the auditory system to accurately encode rapidly changing spectral and temporal cues over time. Evidence from behavioral studies in humans suggests that relatively poor temporal fine structure (TFS) encoding ability is correlated with poorer performance on speech understanding tasks in quiet and in noise. Electroencephalography, including measurement of the frequency-following response, has been used to assess the human central auditory nervous system's ability to encode temporal patterns in steady-state and dynamic tonal stimuli and short syllables. To date, the FFR has been used to investigate the accuracy of phase-locked auditory encoding of various stimuli, however, no study has demonstrated an FFR evoked by dynamic TFS contained in the modulating frequency content of a carrier tone. Furthermore, the relationship between a physiological representation of TFS encoding and either behavioral perception or speech-in-noise understanding has not been studied. The present study investigated the feasibility of eliciting FFRs in young, normal-hearing listeners using frequency-modulated (FM) tones, which contain TFS. Brainstem responses were compared to the behavioral detection of frequency modulation as well as speech-in-noise understanding. FFRs in response to FM tones were obtained from all listeners, indicating a reliable measurement of TFS encoding within the brainstem. FFRs were more accurate at lower carrier frequencies and at shallower FM depths. FM detection ability was consistent with previously reported findings in normal-hearing listeners. In the present study, however, FFR accuracy was not predictive of behavioral performance. Additionally, FFR accuracy

was not predictive of speech-in-noise understanding. Further investigation of brainstem encoding of TFS may reveal a stronger brain-behavior relationship across an age continuum.

Chapter I

Review of Literature

Introduction

Auditory evoked potentials have long been used in scientific and clinical settings to study and to assess the encoding of various properties of sound within the mammalian inner ear and central auditory pathway. The recorded periodic stimulation of the outer hair cells of the cochlea (Withnell, 2001) and the phase-locked neural firings of both single-units (Kiang, Watanabe, Thomas, & Clark, 1965) and populations of auditory nerves (Young & Sachs, 1979) indicate that the auditory system is capable of processing signals containing high-rate fluctuations over time.

The frequency-following response (FFR) is a measurable auditory evoked potential in humans that can illustrate the neural firings within brainstem nuclei in response to the temporal quality of the auditory stimulus (Worden & Marsh, 1968). In other words, the FFR can reflect a phase-locked neural representation of an acoustic stimulus. Though the FFR has been elicited with a variety of static (e.g., tones) and dynamic acoustic stimuli (e.g., consonant-vowel stimuli), fewer papers have focused on the ability of the FFR to represent stimuli with dynamic frequency content, or fine structure that changes in frequency. Here we assess the feasibility of evoking FFRs using FM tones, which contain dynamic frequency content. In addition, we investigated the relationship between the accuracy of the FFR to such stimuli and the perception of frequency modulation. Finally, the relationship between the physiological and behavioral responses to this fine temporal information and the ability to accurately understand SIN has yet to be explored.

If feasible, the FFR in response to a sound whose TFS varies in frequency (i.e., an FM tone) could extend the utility of this auditory evoked potential in the experimental setting, offering additional insight into the effects of numerous variables (e.g., age and hearing loss) on TFS encoding. Furthermore, establishing the presence or absence of a relationship between physiologic encoding of TFS and SIN understanding could support the use of TFS processing in cochlear implants.

Temporal envelope vs. temporal fine structure

The auditory system, from the cochlea to the auditory cortex, processes sound, in part, along the time domain. Complex signals such as speech typically contain multiple temporal characteristics, which include temporal envelope, periodicity, and TFS. Within the frequency channels of the cochlea, temporal cues are present in either the patterns of neural firings synchronized to a certain period of the stimulus (TFS encoding via neural phase-locking), or in the lower-rate fluctuations of these firing patterns (envelope fluctuations). Envelope amplitude fluctuations in the range of 2-50 Hz primarily provide the cues to the manner of articulation, which aids in determining the presence of voicing. The high-rate frequency content of TFS (600-10,000 Hz) may be observed within the envelope of a speech signal (Rosen, 1992). For example, if an amplitude-modulated tone had a carrier frequency of 500 Hz and a modulation rate of 10 Hz, the fine structure would be the oscillations within the envelope – 500 Hz. Within a speech signal, TFS often contributes to the recognizable formant patterns that distinguish vowels. The formants provide segmental cues to the identification of the place of articulation of consonants and the dynamic characteristics of neighboring vowels; the interaction between the two helps to identify the spoken word.

Accurate processing of these cues allows for phonetic discrimination based on formant transitions.

Temporal fine structure processing is related to speech-in-noise understanding

It is well established that the normal auditory system utilizes, in part, temporal cues within a speech signal (i.e., envelope and TFS) to encode the signal in both quiet and in noise (see Moore, 2008 for review). Several different approaches have been taken to investigate the role of TFS processing in the understanding of speech in quiet and in noise. Previous studies have compared listener performance on psychoacoustical tasks in which thresholds are based on TFS processing to SIN test scores, while others have more directly assessed speech-in-noise (SIN) understanding in conditions in which the TFS of the signal has been isolated or altered.

Many psychoacoustical approaches to measuring TFS processing ability in normal-hearing and hearing-impaired listeners have used stimuli designed to target specific auditory mechanisms important for TFS encoding (i.e., phase-locking). Buss, Hall, & Grose (2004) investigated the relationship between TFS processing and speech understanding by comparing FM detection limens (FMDLs) and speech recognition in quiet using filtered (i.e., low-pass filtered at 1800 Hz) and unfiltered speech. FM stimuli with carrier frequencies of 500 and 1000 Hz were frequency modulated at a slow rate (2 Hz), with the expectation that these stimulus parameters were best represented by phase-locked neural activity. When hearing thresholds were controlled for statistically, significant correlations were found between FM detection limens for a 1000 Hz carrier tone and for filtered and unfiltered

speech understanding. The results of this study suggest that TFS processing, as reflected by frequency modulation detection limens, are related to speech understanding.

Strelcyk and Dau (2009) assessed binaural unmasking and lateralization in response to phase differences in dichotically presented tones in order to investigate TFS processing in normal hearing and hearing impaired listeners. FM and speech detection tasks similar to those in Buss et al., (2004) were also used, though various maskers were presented to interfere with speech while FM stimuli were sometimes presented in noise or in both ears. The authors found that better TFS processing ability was correlated with improved SIN understanding. These results suggest that TFS processing plays an important role in SIN understanding for listeners with and without normal hearing thresholds.

More direct methods of establishing the importance of TFS processing in speech perception include alterations of a speech signal that isolate or otherwise manipulate the temporal qualities of the stimulus. Using Hilbert transforms several authors have separated the envelope and fine structure components of speech stimuli, allowing for the two to be altered or presented separately. In quiet and with adequate training, no difference was observed in speech understanding ability for normal-hearing listeners who were provided with only either the envelope or TFS components of the speech signal (Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006). This finding suggests that TFS information alone provides sufficient information for speech understanding in quiet. In noise, TFS cues contribute to understanding speech, specifically when the background interferer is modulated in amplitude (Hopkins & Moore, 2008). It is thought that TFS cues present in the “dips” (i.e., periods of low amplitude) in noise aid in improving the short-term signal-to-noise ratios (SNR) of the

signal. This model was supported by Drennan, Won, Dasika, and Rubinstein (2007), who used the Hilbert transform technique to randomize the TFS within speech stimuli. As the fine structure moved along a continuum from being unaltered to being completely randomized, understanding of speech in the presence of background noise was diminished as the degree of randomized fine structure increased. Together, these studies demonstrate that the TFS components of speech play an important role in the understanding of speech in both quiet and in noise.

Behavioral frequency modulation detection

In order to ensure the validity of using particular stimuli (i.e., FM tones) in studies of TFS processing, the parameters of these stimuli must be designed to target the temporal mechanisms of the auditory system. Multiple studies have addressed whether temporal (e.g., phase locking) or rate-place coding more greatly influence FMDLs. These studies have examined the effects of various stimulus parameters (e.g., carrier frequency and frequency modulation rate) to see if performance agrees with temporal or rate-place models. In a series of studies investigating FM detection, Moore and Sek (1995, 1996) and Sek and Moore (1995) reported worsened FMDLs compared to amplitude modulation (AM) detection when FM rate was increased for carrier frequencies below 4 kHz. At higher carrier frequencies (i.e., 6 kHz), FM and AM detection were similarly affected by FM rate. These findings suggest that FM detection can be achieved through either place and/or temporal coding in the auditory system, and that the influence of one mechanism instead of the other was determined by modulation rate and carrier frequency. Behavioral FMDLs were most like

those predicted by temporal coding when the carrier frequency was low (e.g., 500 and 1000 Hz) and the frequency modulation rate was also low (i.e., 2 Hz).

Moore and Sek (1996) further investigated the roles for temporal and place coding in FM detection by testing the effect of added amplitude modulation (AM) on FMDLs. Added AM disrupted excitation-pattern cues and resulted in increased (poorer) FMDLs in conditions in which place coding was thought to be dominant. At carrier frequencies below 4 kHz and at FM rates below 10 Hz, the effect of AM was diminished, suggesting that temporal cues were more dominant. These findings were supported by Ernst and Moore (2010), who investigated similar interactions between FM rate, carrier frequency, and AM. The authors reported a greater effect of AM at higher FM rates (e.g., 10 Hz) when the carrier frequency was either 1 or 4 kHz, suggesting that rate-place cues were more responsible for processing stimuli with these parameters. Additionally, the disruptive effects of AM were diminished when stimuli were presented at 20 dB SL compared to 60 dB SL. It is possible that sharpened tuning occurring at the relatively low presentation level improved place coding and weakened phase-locking, resulting in similar thresholds across a range of modulation rates.

Although multiple studies have indicated that behavioral FM detection at low FM rates (< 5 Hz) and low carrier frequencies (≤ 1000 Hz) is likely facilitated by a temporal, phase-locking mechanism, no study has compared the behavioral detection of FM to its physiological representation in the same individual. Further, no data have been published regarding the neural representation of a FM tone in humans.

Neural representation of temporal encoding within the auditory system

Recordings of single-unit and population auditory nerve fibers have shown that the auditory nerve is capable of encoding TFS information via neural phase-locking (Kiang et al., 1965; Young & Sachs, 1979). In humans, the FFR is an AEP that may be recorded from the scalp and reflects phase-locked neural activity from neural populations within the upper brainstem (Worden & Marsh, 1968, Smith, Marsh, & Brown 1975). The inferior colliculus is believed to be the primary neural generator of the FFR recorded in humans, although other brainstem sub nuclei also generate FFRs (Smith et al., 1975). Since its discovery, studies have investigated the accuracy of the FFR in recording the temporal encoding of static or time-variant tonal or speech-like stimuli.

Early investigations of the FFR found that the response could be elicited by tone bursts and long-duration tones (Worden & Marsh, 1968; Glaser et al., 1976), with response limits of about 70-1500 Hz and thresholds of 40 dB sensation level (Glaser, Suter, Dasheiff, & Golderberg, 1976). More recent evidence has shown that the FFR is capable of reflecting the underlying neural activity occurring during the encoding of steady-state tonal complexes and speech-like stimuli (Bidelman & Krishnan, 2011), time-variant tonal and speech stimuli (Krishnan & Parkinson, 2000), synthetic consonant-vowels (Plyler & Ananthanarayan, 2001; Johnson, Nicol, & Kraus, 2005), synthetic speech with Mandarin tone contours (Krishnan, Xu, Gandour, & Cariani, 2004; Krishnan & Gandour, 2009) , and iterated ripple noise with Mandarin pitch contours (Swaminathan, Krishnan, & Gandour, 2008; Krishnan, Gandour, Bidelman, & Swaminathan, 2009).

Krishnan and Parkinson (2000) demonstrated that the FFR is capable of accurately following tones that rise or fall steadily in frequency. The response, however, was not equivalent for sweeps in both directions, and stimuli that had frequency content that was increasing, or rising, had more robust responses, presumably because of better neural synchrony. Plyler and Ananthanarayan (2001) used several synthetic consonant-vowel speech stimuli with varying degrees of upward or downward frequency sweeps as a second formant. In normal-hearing listeners, the FFR faithfully encoded the frequency content of the formant transitions, suggesting that the response was capable of exhibiting the neural representation of dynamic speech stimuli within the brainstem. Using four similar Mandarin Chinese lexical tones containing individually distinct time-varying fundamental frequencies, Krishnan and Gandour (2009) found that the FFR was capable of following the fundamental-frequency tone contours inherent in some Mandarin speech syllables.

Neural encoding and the behavioral perception of sound

Several studies have investigated the relationships between the neural encoding and behavioral perception of numerous stimuli. Many of these studies reported significant correlations between the FFR and perception. It was therefore a purpose of the present study to investigate the possible relationship between the FFR elicited by FM tones and the behavioral detection of frequency modulation as well as SIN understanding.

Aiken, LeClair, & Kieft (2011) measured FFRs in response to dual tone multi-frequency (DTMF) signals, which contained two tonal partials that created a perceivable pitch and compared these responses to subjects' ability to match the pitch of DTMF signals to a pure tone. FFR spectra were analyzed to determine if the frequency of the matched pure

tone was emphasized. Across subjects, the identified pitch typically fell close to one of the two partials of the DTMF and the frequencies of the two partials were present in the FFR spectra. While the neural representation and behavioral identification of the two partials were not correlated, there was clearly demonstrable association between the FFRs and pitch perception.

Krishnan and colleagues (Bidelman & Krishnan, 2011; Krishnan, Bidelman, Smalt, Ananthkrishnan, & Gandour, 2012) investigated the pre-attentive, sub-cortical encoding of sounds with pitch qualities that could be behaviorally discriminated. Using three-tone complexes resembling musical triads, Bidelman and Krishnan (2011) demonstrated a positive correlation between the reported musical consonance of a triad and the amplitude of the FFR to the corresponding complex. Krishnan et al. (2012) used iterated rippled noises (IRNs) to compare pitch discrimination ability with sub-cortical and cortical neural encoding. IRNs are comprised of multiple Gaussian noises added periodically, which contribute to a salient pitch. The study revealed that the behavioral discrimination of IRN pitch could be predicted by analysis of the neural encoding of IRNs within the brainstem and auditory cortex. Additionally, both pitch discrimination performance and FFR magnitude improved with increased IRN pitch saliency.

To investigate the role of subcortical neural encoding of sound on the perception of speech, Kraus and colleagues have used synthetic /da/ speech stimuli to compare the physiological and behavioral responses of children with learning problems to those of normal children (see Skoe & Kraus, 2010 and Johnson et al., 2005 for review). These studies have used a /da/ syllable as a stimulus means of measuring two separate brainstem responses. The

onset burst of the consonant of the stimulus elicits an onset neural response with morphological landmarks similar to click and tone-evoked auditory brainstem responses. The long-duration vowel component contains a formant structure that elicits an FFR with spectral emphasis at the fundamental frequency and its harmonics. The /da/-evoked stimulus was determined to be reliably obtained from normal hearing children and appeared to faithfully reflect the acoustic properties of the stimulus (Russo, Nicol, Musacchia, & Kraus, 2004). Cunningham, Nicol, Zecker, Bradlow, & Kraus (2001) reported that the FFR to the /da/ in silence was similar for normal and learning impaired listeners. When the /da/ was presented in noise, the stimulus-to-response correlations of the FFR were significantly reduced in learning impaired children but remained stable in normal children. This degradation of the FFR due to noise was well correlated with SIN performances of the two groups of listeners. Using a similar test paradigm, King, Warrier, Hayes, & Kraus (2002) later reported increased latencies of the FFR response to /da/ in children with learning impairment. In normal hearing adults, the FFR in response to /da/ was compared to SIN understanding using the QuickSIN test. Listeners who performed more poorly on the QuickSIN demonstrated greater degradation of the fundamental frequency component of the FFR when the /da/ was presented in noise (Song, Skoe, Banai, & Kraus, 2011).

Purpose

The majority of FFR studies that have used dynamic frequency content have focused on the encoding of fundamental-frequencies; few have examined the neural representation of dynamic TFS that has changing frequency content (i.e., frequency modulation). The present study investigated the feasibility of measuring the FFR in response to frequency-modulated

tones, thereby measuring the phase-locked neural encoding of TFS within the brainstem. A brain-behavior relationship was investigated by comparing the physiologic response to FM tones with behavioral detection and SIN understanding. A low-rate FM detection task was chosen to assess TFS processing ability, which would be analyzed beside results from a clinical SIN test. These measures were then analyzed to determine if correlations existed between the neural response and either behavioral thresholds of TFS processing or SIN understanding. Hypotheses for the current study included:

- 1) Measurable frequency-following responses will be elicited by FM tones and will reflect the neural representation of the TFS of the stimuli.
- 2) The degree of neural representation of frequency modulation reflected by the frequency-following response will be significantly correlated with behavioral frequency modulation detection limens. Individuals whose frequency-following responses are more robust and have higher (better) stimulus-to-response correlation will demonstrate better (smaller) detection limens.
- 3) Performance on the QuickSIN test will be significantly correlated with both behavioral frequency modulation detection limens and the neural representation of FM tones. Individuals who demonstrate better SIN understanding will have better (smaller) FMDLs and produce more robust FFRs with greater correlation to the eliciting stimuli

Chapter II

Methods

Subjects

21 subjects (ages 21-40[*mean* = 30.33, *SD* = 6.82]) were recruited to participate in both behavioral and physiology-based conditions. All subjects had clinically normal behavioral pure tone thresholds (thresholds \leq 25 dB HL at octave frequencies 250 – 8000 Hz) with normal tympanometric findings and no history of chronic otologic pathology. Subjects were recruited if they met the age criteria of being 21 to 40 years old. Additional inclusion criteria included each subject being a native, monolingual English speaker with no extended duration of musical training. Current use of prescription medications for uses such as sleep, seizures, memory, attention, or mood was an exclusion criterion. Testing for all procedures was performed in a double-walled, sound-treated booth. All procedures had prior approval by the institutional review board at James Madison University. Subjects participated in one session of approximately four hours.

Stimuli

Similar stimuli were used for both the behavioral FMDL and physiological FFR conditions. Tone bursts were presented at 80 dB SPL to the right ear via an Etymotic Research ER-3A insert earphone and had a carrier frequency of either 500 Hz or 1000 Hz. For the FMDL procedure, tones had duration of 1000 ms, including 15 ms rise and fall times with a Hanning envelope, and a frequency modulation rate of 2 Hz. These 1000 ms stimuli contained two complete cycles of frequency modulation. For the electrophysiologic procedure, stimuli of 500 ms duration including 15 ms rise and fall times were used; these 500 ms tone bursts contained one complete cycle of frequency modulation.

Calibration of stimulus level was performed using a Larson Davis Model 824 sound level meter with a one-half inch microphone and a 2 cc coupler. For FMDL and FFR conditions, all stimuli were presented at 80 dB SPL.

Behavioral procedure

Stimuli used during FMDL task were routed through a Tucker Davis Technologies attenuator (TDT PA4), mixer (TDT SM3) and headphone buffer (TDT HB6) to an ER3A earphone. Stimuli presented during FMDL testing were created in a custom Matlab program.

An FM detection task based on the methods used by Moore and Sek (1996) was used to assess FMDLs across all listeners. Carrier frequencies of 500 and 1000 Hz were presented in two separate conditions in random order. FMDLs were measured using an adaptive two-interval, two-alternative forced-choice procedure with a two-down and one-up rule. Each trial consisted of one unmodulated and one frequency-modulated pure tone, presented in random, sequential order. Subjects were seated in front of a computer monitor, which featured a display of two boxes that corresponded to the presentation of the first and second intervals of the task. For each trial, subjects were instructed to listen carefully to each stimulus and determine which one appeared to change over time (i.e., the pitch was perceived as modulating). Subjects then mouse-clicked on the box that corresponded to the changing stimulus. Feedback was provided in the form of the selected box illuminating green (correct selection) or red (incorrect selection). An adaptive procedure regulated FM depth, which changed by a factor of 1.5 for the first four reversals, followed by changes by a factor of 1.26 for the remaining reversals. Each trial continued for twelve reversals. Prior to the start of behavioral testing, an abbreviated training trial was completed to assess each subject's

understanding of the behavioral task. Subjects completed at least three separate trials under each stimulus condition (i.e., 500 Hz or 1000 Hz carrier frequency). The order of the carrier frequency conditions was chosen randomly, and additional trials were completed if subject performance was variable. Participants were encouraged to take breaks if their FMDLs or tracking functions became variable or they appeared fatigued.

Electrophysiology procedure

To examine a neural representation of the stimuli used in the behavioral task, FFRs were elicited by stimuli with similar parameters to those used for the FMDLs. Four FFR conditions were recorded: 1) 500 Hz with 0.4% FM depth, 2) 500 Hz with 2% FM depth 3) 1000 Hz with 0.4% FM depth and 4) 1000 Hz with 2% FM depth. Figure 1 illustrates how FFR conditions corresponded to the behavioral FM detection task. The 0.4% FM depth would have been difficult to detect behaviorally while the 2.0% FM depth would have been relatively easier to detect. Stimuli used in the physiology conditions had duration of 500 ms and contained one period of frequency modulation.

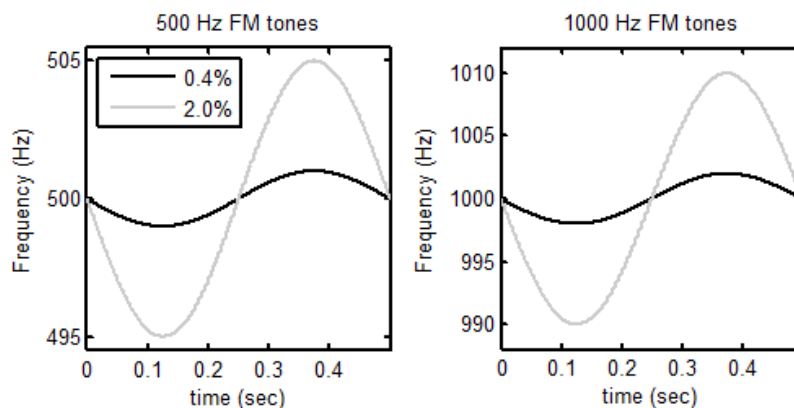


Figure 1 Schematic spectrograms illustrating the peak-to-peak frequency deviation from the carrier frequency. **Left Panel:** For stimuli of 500 Hz carrier frequency and depths of 0.4 and 2.0%, peak-to-peak deviations were 2 and 10 Hz respectively. **Right Panel:** For stimuli of 1000 Hz carrier frequency and depths of 0.4 and 2.0%, peak-to-peak deviations were 4 and 20 Hz, respectively.

FFRs to FM tones were recorded using a three-channel recording. Electrodes were placed at five points along the scalp according to the 10-20 system (Jasper, 1958): the left and right ear lobes (A1 and A2, respectively); the vertex of the head (Cz); the forehead (Fpz); and the nape of the neck. The Fpz electrode acted as the ground, with Cz being used as the inverting electrode. The signal from the Cz electrode was jumped to three separate inverting channels (A1, A2, and the nape of the neck). Stimuli were presented using a Neuroscan Stim Audio System P/N 1105 and AEPs were recorded through Neuroscan SynAmps RT hardware and SCAN software. To reduce the potential for transducer artifact in physiological recordings, the ER-3A transducer was shielded, and a double-length sound tube was used for calibration and data collection.

Evoked potentials were processed within an epoch time window of -100 to 520 ms relative to stimulus onset, at a sampling rate of 20 kHz, and through an online band-pass filter at 30-3000 Hz. An artifact rejection criterion of $\pm 30 \mu\text{V}$ was used to prevent the recording of large myogenic activity. If specific electroencephalogram (EEG) activity (i.e., cardiac artifact) regularly peaked at or above this artifact-rejection criterion, the criterion was raised to $\pm 35 \mu\text{V}$. Impedance levels at all electrode sites did not exceed 5 k Ω , and relative impedance levels between electrodes were limited to no more than 2 k Ω . Response recording was concluded after 1000 accepted sweeps were collected. Figure 2 shows FFRs recorded from a 23-year-old subject for each stimulus condition. Prior to initial data collection for the present study, control conditions were recorded within the laboratory to verify that the FFR recordings were not contaminated by stimulus artifact. These conditions included recordings where the earphone was not placed in the ear, as well as recordings from a watermelon. Stimulus artifact was not observed during these control recordings.

During recording, subjects were instructed to lie quietly in a reclining chair and relax. Breaks were provided upon request or if subjects became restless.

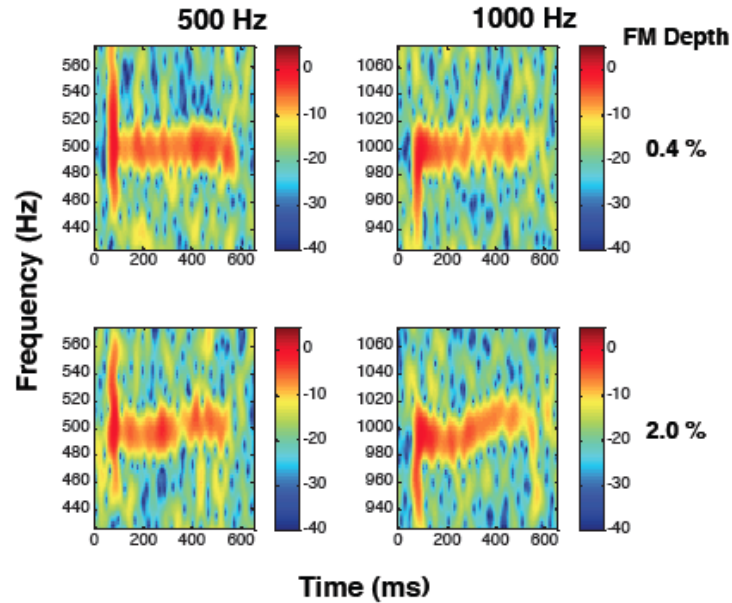


Figure 2. The time-domain FFR waveforms from an individual, age 23, for the four FFR conditions are shown.

Electrophysiology analysis

FFRs were analyzed in respect to their stimulus-to-response correlations, which were calculated to determine how well the FFR waveform resembled that of the stimulus waveform for the same condition (i.e., 500 Hz carrier frequency with 0.4% frequency modulation depth). A 0 to 15 ms lag time was applied to stimulus waveforms and cross-correlation coefficients were calculated at different points along this time window. The highest correlation coefficient and its corresponding lag time were recorded. Higher correlation coefficients would indicate that the neural response represented a more accurate encoding of the dynamic frequency content contained in a stimulus.

Speech-in-noise procedure

SIN understanding was tested using the Etymotic Research QuickSIN Speech in Noise Test, specifically lists 1, 2, 10, 11 were used; McArdle and Wilson (2006) reported that these four lists have the most homogenous scores in normal-hearing and hearing impaired listeners. A QuickSIN practice list was presented prior to the presentation of these four lists in order to assess the subject's understanding of the task. The QuickSIN compact disc was played by a Sony CD player (model CDP-CE375) with an output to a TDT PA4 set to 33 dB of attenuation. This achieved an output level of 70 dB SPL (A-weighted) through an ER-3A insert earphone in the right ear.

In accordance with QuickSIN procedure, subjects listened to each list and repeated sentences spoken by a target female speaker. Each of the six sentences within a list contained five target words and the number of target words correctly repeated determined the score for each sentence. The signal-to-noise ratio (SNR) of each list decreased from +25 dB to 0 dB, decreasing by 5 dB steps with each successive sentence. Sentence scores were aggregated into a list score, which was subtracted from 25.5 to determine a total SNR Loss (Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004). QuickSIN scores reflect the speech-to-babble ratio at which 50% of the target words are understood. Scores from the first four lists were averaged together to determine a mean SNR loss, in dB, for each subject.

Statistical approach

Analyses of variance (ANOVA) were used to analyze dependent variables of behavioral FMDLs and FFR stimulus-to-response correlations. A one-way ANOVA was conducted for FMDLs with a factor of carrier frequency (within-subject on 2 levels, 500 and

1000 Hz). The FFR stimulus-to-response correlations were analyzed with factors of frequency (within-subjects on two levels, 500 and 1000 Hz), and FM depth (within-subjects on two levels, 0.4% and 2.0%). Partial η^2 was used for a measure of effect size; small, medium, and large effect sizes were defined as partial η^2 values of 0.0099, 0.0588, and 0.1379, respectively (Cohen, 1988). Pearson-product moment correlations were used to examine relationships between behavioral and physiological measures.

Chapter III

Results

FMDLs

A one-way repeated-measures ANOVA was conducted on FMDLs with a factor of frequency (within-subjects on two levels, 500 and 1000 Hz). The main effect for frequency was not significant, $F(1,21) = 3.82$, $p = .064$, partial $\eta^2 = .154$, indicating that FMDLs were not significantly different between the two frequencies. Figure 3 summarizes the FMDL data and compares listener performance across the two FMDL conditions. Figure 4 represents individual listener performance between the two FM detection conditions.

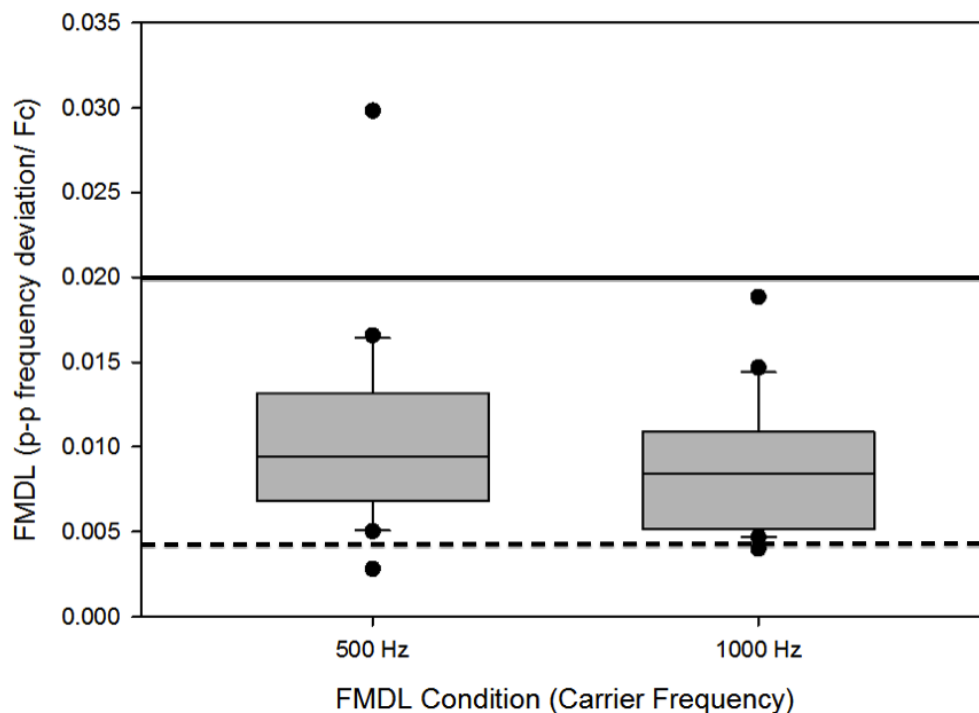


Figure 3. Boxplot illustrating mean listener performance on FMDL tasks in which carrier frequencies were 500 and 1000 Hz. The FM depths selected for the physiological task are represented by a solid (2.0%) and dashed line (0.4%). FMDLs for carrier frequencies of 500 and 1000 Hz were not significantly different ($p > .05$).

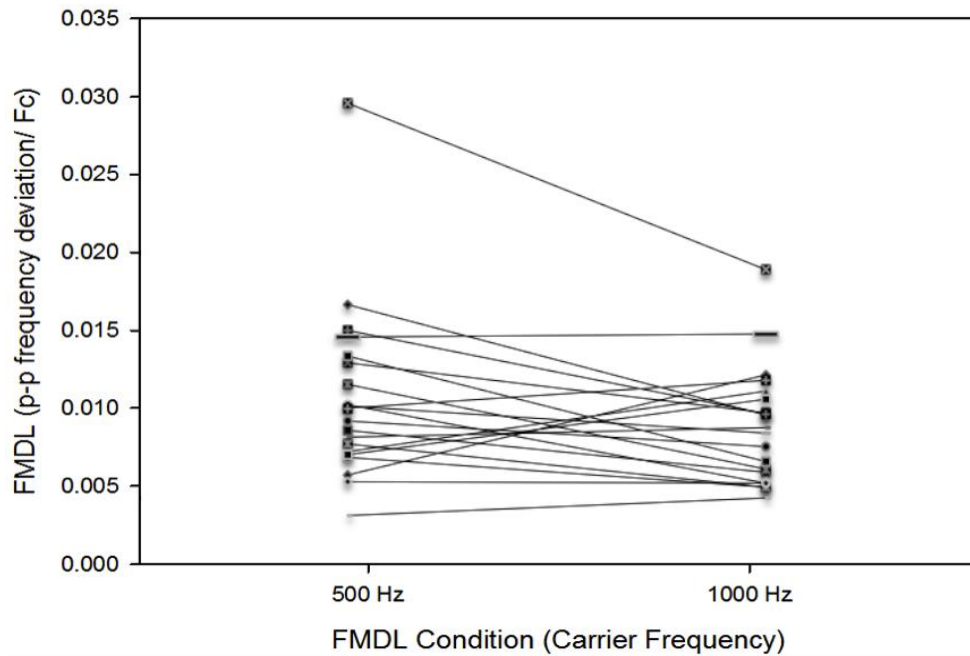


Figure 4. Individual subject trends of FMDL tasks. Each subject's performance on the 500 Hz and 1000 Hz FMDL tasks is represented by a different line.

FFR stimulus-to-response cross correlations

A two-factor repeated measures ANOVA was conducted on FFR stimulus-to-response correlations. Factors were frequency (within-subjects on two levels, 500 and 1000 Hz) and FM depth (within-subjects on two levels, 0.4% and 2.0%). There was a significant main effect of frequency, $F(1,19) = 42.16, p < .001$, partial $\eta^2 = .689$ as well as a significant main effect of FM depth, $F(1,19) = 49.21, p < .001$, partial $\eta^2 = .721$. There was a significant interaction between frequency and FM depth, $F(1,19) = 15.09, p < .001$, partial $\eta^2 = .443$. Figure 5 shows stimulus-to-response correlations across each FFR condition. FFR stimulus-to-response correlations were lower (poorer) for 1000 Hz than for 500 Hz, and were more robust for the conditions with the least amount of frequency modulation.

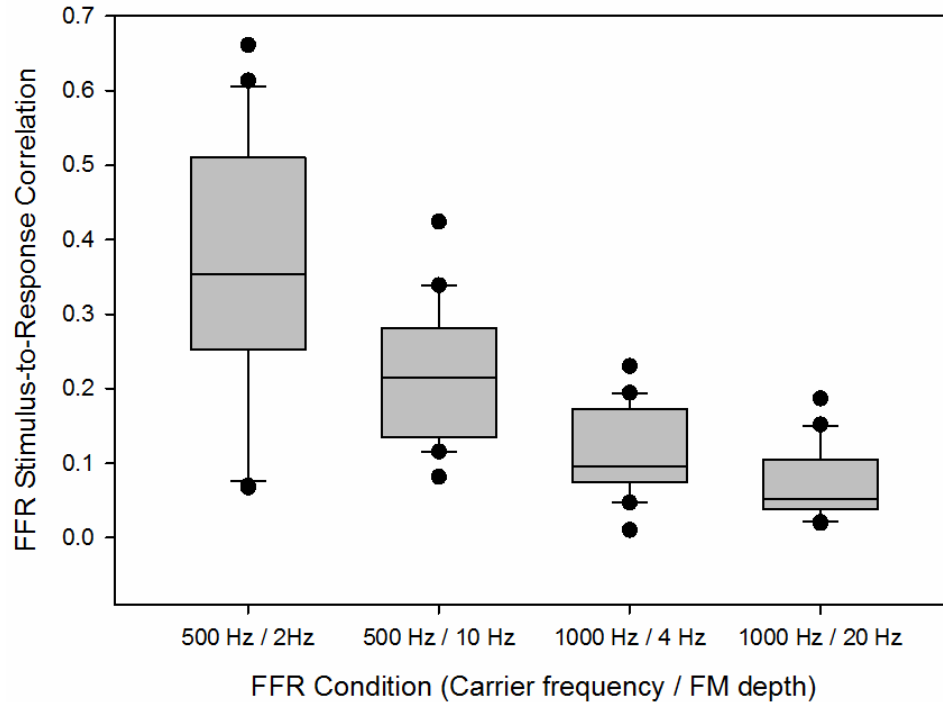


Figure 5. Boxplot illustrating FFR stimulus-to-response correlations in the four different FFR conditions. There was a significant effect of frequency ($p < .001$) and FM depth ($p < .05$); FFR stimulus-to-response cross-correlation coefficients became poorer as carrier frequency and FM depth increased.

QuickSIN

QuickSIN data for the four sentence lists were averaged together for each listener and a mean SNR loss was calculated across all listeners (mean = 0.24 dB, SD = 1.15 dB). Figure 6 illustrates overall mean listener QuickSIN performance. These data were used in statistical analyses of the relationship between SIN understanding and the detection and neural encoding of FM tones.

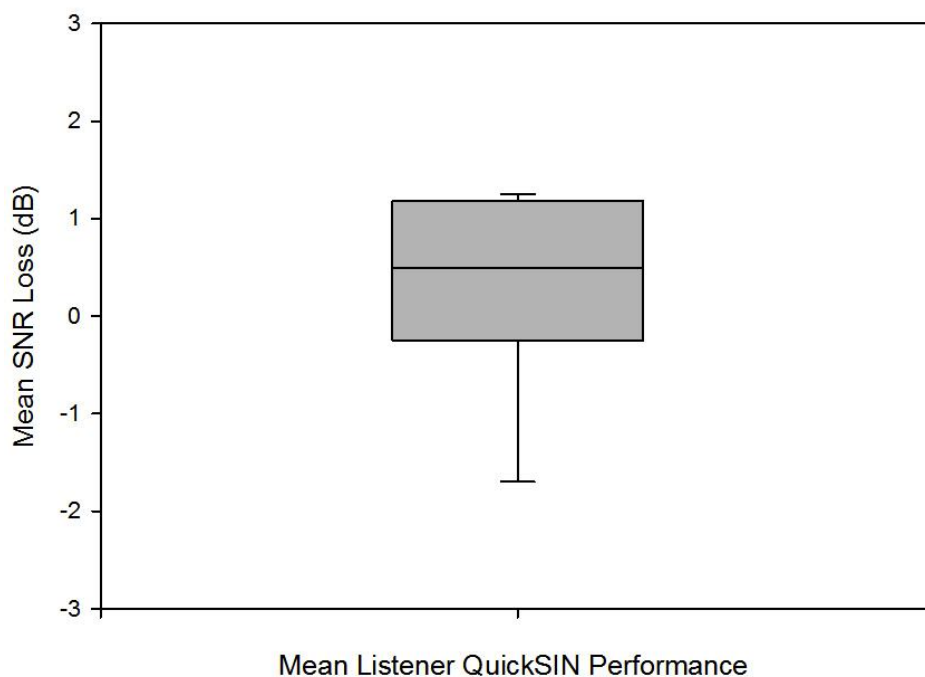


Figure 6. Boxplot illustrating mean QuickSIN performance (averaged across four sentence lists) across all listeners.

Relationships between behavioral and physiological measures

Pearson-product moment correlation coefficients were calculated to determine the relationships between measures of behavioral perception (FM detection and QuickSIN SNR loss) and a physiological coding (FFR stimulus-to-response correlation). At 500 Hz, there were no significant correlations between FMDLs and FFR stimulus-to-response correlation coefficients, 0.4% FM depth ($r = -.117$, $p = .623$) and 2.0% FM depth ($r = -.259$, $p = .271$). At 1000 Hz, there were also no significant correlations between FMDLs and FFR stimulus-to-response correlation coefficients, 0.4% ($r = -.411$, $p = .720$) or 2.0% FM depth ($r = -.251$, $p = .285$) conditions. Figure 7 illustrates how FFR stimulus-to-response cross-correlation coefficients are related to FMDLs across all subjects.

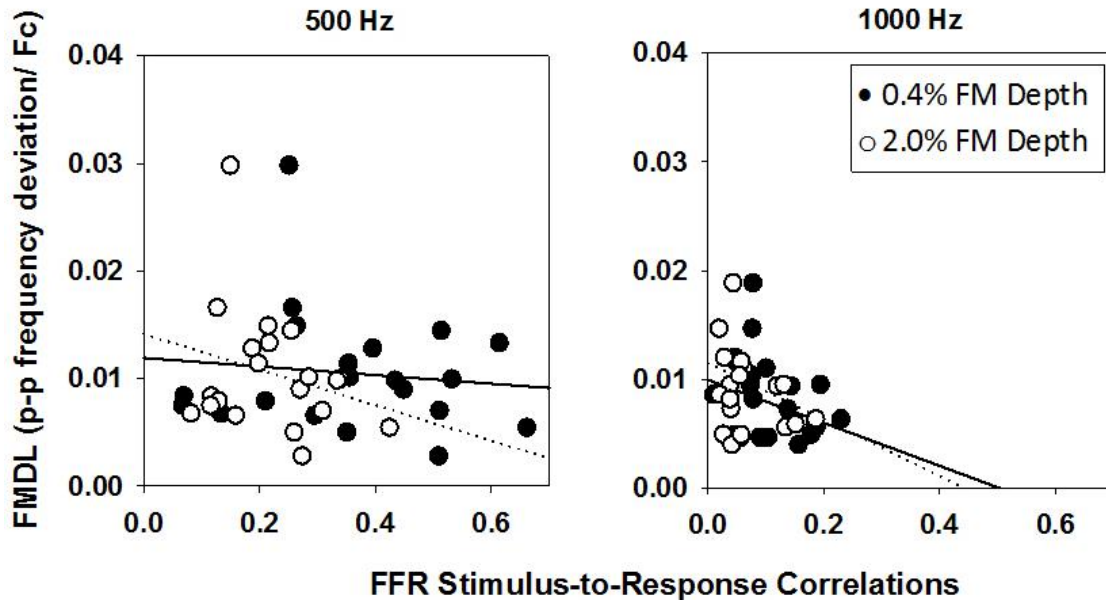


Figure 7. Scatterplots illustrating the relationship between FFR stimulus-to-response correlations and FMDLs across all listeners for both the 500 Hz (left panel) and 1000 Hz (right panel) conditions. Correlations between the perception of frequency modulation, FMDLs, and FFR stimulus-to-response correlations were not significant ($p > .05$). In each panel, lines of best fit are shown for the 0.4% FM depth (solid line) and 2.0% FM depth (dashed line).

To determine how FMDLs and FFR stimulus-to-response correlations were related to SIN understanding reflected by QuickSIN SNR loss, Pearson-product moment correlations were calculated. There was no significant correlation between QuickSIN SNR loss and FMDLs for either the 500 Hz ($r = .124$, $p = .594$) or 1000 Hz ($r = -.069$, $p = .766$) conditions (Figure 8). QuickSIN SNR loss was not significantly correlated with 500 Hz FFR stimulus-to-response cross correlations in the 0.4% ($r = -.115$, $p = .630$) and 2.0% FM depth ($r = -.245$, $p = .298$) conditions, or with 1000 Hz stimulus-to-response cross correlations in the 0.4% ($r = -.036$, $p = .880$) and 2.0% ($r = .082$, $p = .731$) conditions (Figure 9).

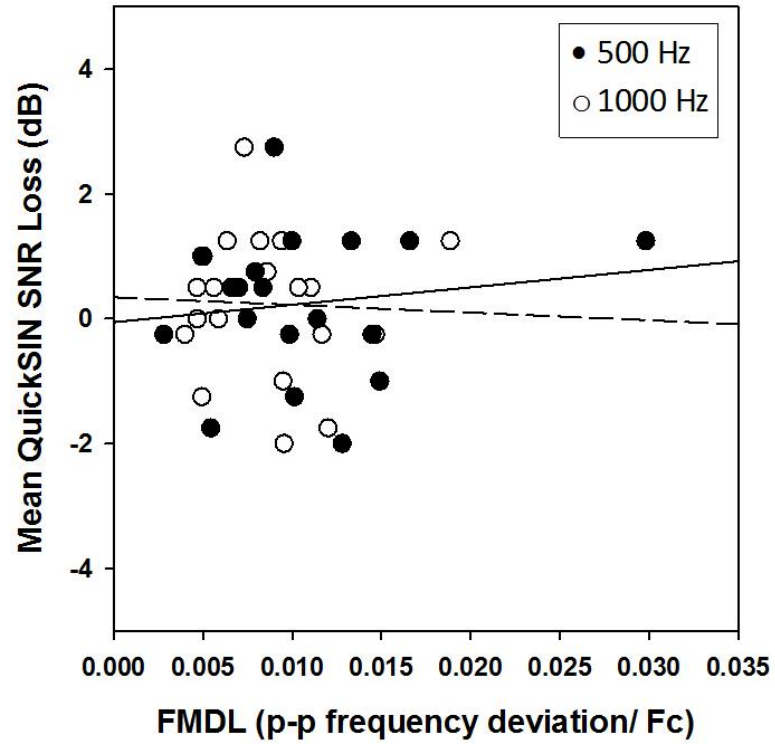


Figure 8. Scatterplots illustrating the relationship between 500 Hz and 1000 Hz FMDLs and mean QuickSIN SNR losses across all listeners. There was no significant correlation between QuickSIN performance and FMDLs ($p > .05$). Lines of best fit are shown for the 500 Hz (solid line) and 1000 Hz (dashed line) data.

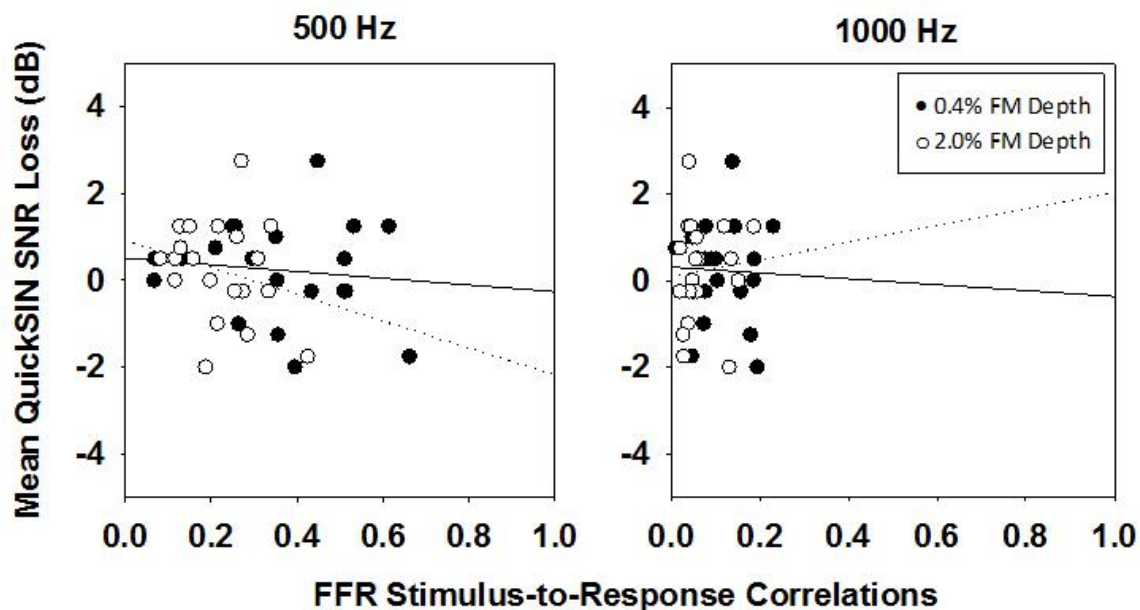


Figure 9. Scatterplots illustrating the relationship between FFR stimulus-to-response correlations for the 500 Hz (left panel) and 1000 Hz (right panel) conditions and mean QuickSIN SNR losses across all listeners. Mean QuickSIN performance and FFR stimulus-to-response correlations were not significantly related ($p > .05$). In each panel, lines of best fit are shown for the 0.4% FM depth (solid line) and 2.0% FM depth (dashed line).

Chapter IV

Discussion

Introduction

The main purpose of the present study was to investigate the feasibility of using the FFR to measure the neural encoding within the auditory brainstem of TFS information contained in FM tones. A secondary objective was to use this measure to determine to what degree neural encoding of TFS correlated with the behavioral detection of FM and SIN understanding. Several psychoacoustic studies have demonstrated that TFS processing contributes to speech understanding in quiet and in noise as well as to the detection of FM when parameters are selected that reduce the influence of place coding. A strong correlation between these behavioral measures and the neural encoding of TFS may indicate a relationship between physiological processing of TFS and SIN understanding. Additionally, such a correlation may support the use of TFS in speech processing algorithms for amplification devices.

Given the objectives of the present study, it was hypothesized that:

- 1) FFRs would be measurable in response to dynamic FM tones and would accurately reflect the frequency content of the stimulus.
- 2) The FFR stimulus-to-response correlation would be significantly correlated with an FMDL of the same carrier frequency. For example, listeners with better (smaller) FMDLs would be expected to have FFRs with higher (better) stimulus-to-response correlations, and those individuals with poorer (larger) FMDLs would be expected to have poorer (smaller) FFR stimulus-to-response correlations.

- 3) Behavioral and physiological measures of TFS processing, as defined by FFR metrics and FMDLs, were expected to be predictive of SIN performance. For example, better FMDLs and higher FFR stimulus-to-response correlations would each be significantly correlated with a smaller (better) SNR loss.

The results of the present study suggest that the encoding of FM tones containing dynamic TFS within the human auditory system may be measured using the FFR elicited by FM tones. This brainstem response is reliably present and is capable of faithfully representing the dynamic frequency content of FM tones, as reflected by stimulus-to-response correlations. In addition to this finding, a repeated-measures ANOVA of FFR data revealed a significant interaction between carrier frequency and FM depth. Significant main effects of carrier frequency and FM depth also suggest that phase-locking was most robust in the 500 Hz conditions and when the FM depth was lowest (i.e., 0.4% compared to 2.0%).

For the physiological procedure, a significant interaction was observed between carrier frequency and FM depth; FFR accuracy was better overall for the 500 Hz condition and FM depth had a greater effect within the 500 Hz condition. The effect of frequency is likely due to an inherent upper frequency limit of the FFR and the degradation of the response with increasing frequency (Krishnan, 2007). Due to an overall weaker response obtained in the 1000 Hz condition, the effect of FM depth at this carrier frequency was less pronounced than for the more robust 500 Hz FFR. The significant effect of FM depth is possibly a result of differences in the rate of frequency change between the two FM depths for each carrier. For example, within a constant time frame (i.e., 500 ms), one cycle of 0.4%

frequency modulation results in a more gradual frequency change than one cycle of 2.0% modulation. Improved FFR accuracy at lower FM depths may reflect easier encoding of more gradual frequency change

Though a significant effect of frequency was not hypothesized, the finding is not surprising. A more robust neural response would presumably be elicited by a higher carrier frequency due to synchronized firings of a greater neural population. However, the poorer stimulus-to-response correlations in the 1000 Hz conditions observed in the present study are consistent with the upper frequency limits of phase-locking within the rostral brainstem (1500-2000 Hz) as well as the frequency response of the FFR, reviewed by Krishnan (2007). FFR responses for a given stimulus level are typically most robust at 500 Hz, with response amplitude and accuracy decreasing as stimulus frequency increases. Many previous studies of the FFR in response to dynamic stimuli have primarily investigated a frequency range below that of the stimuli used in the present study. Those studies also evaluated FFRs related to fundamental frequency processing, as opposed to fine structure processing. However, a study performed by Clinard, Tremblay, and Krishnan (2010) recorded FFRs in response to six different pure tones (500 Hz, 1000 Hz, and four neighboring frequencies) in a sample of subjects whose ages ranged from 22-77 years old. FFR metrics of phase coherence and amplitude indicated a more robust response to tones in the range of 500 Hz compared to those in the range of 1000 Hz. This finding is consistent with the significant effect of frequency reported in the present study.

The effect of FM depth suggests that the phase-locking mechanism functions more accurately when frequency deviation within a certain time window is minimized. It can be

argued that this finding is consistent with a “sluggish” mechanism for decoding of TFS proposed by Moore and Sek (1995). The authors suggested that temporal decoding within the auditory brainstem has a limited sampling rate of neural firings produced by the auditory nerve. Due to this, high-rates of frequency change may not be sufficiently processed within the brainstem. Though the modulation rate was constant across each FFR stimulus and a full period of frequency modulation occurred for each stimulus, FM depth varied within each frequency condition. As an example, for the 1000 Hz condition, frequency modulated between either 998-1002 Hz (0.4% FM depth) or 990-1010 Hz (2.0% FM depth). A greater rate of frequency change was necessary to complete one full cycle of modulation for the 2.0% condition.

The behavioral portion of the present study assessed listener performance on a subjective task of FM detection in order to determine behavioral FMDLs and compare these data to physiological responses. Performance on a norm-referenced test of SIN understanding was also assessed. Data obtained from these behavioral tasks were analyzed using one-way repeated measures ANOVAs. For the behavioral FM detection task, the main effect of frequency was not significant; FMDLs, expressed as peak-to-peak frequency deviation of the determined threshold divided by the carrier frequency, were not significantly different across the 500 Hz and 1000 Hz conditions. The similarity in listener FMDLs across the two frequencies is consistent with previous findings of FM detection (Sek & Moore, 1995), which reported a comparatively smaller effect of frequency for FMDLs compared to frequency difference limens. Buss et al., (2004) reported lower FMDLs when carrier frequency was 500 Hz compared to performance on a 1000 Hz condition. These

investigators presented FM tones with a duration of 400 ms, whereas the present study used 1000 ms tones for the behavioral task. It is possible that this difference in stimulus presentation contributed to the inconsistent findings across studies.

An aim of the present study was to determine if the physiologic response to FM was correlated with the behavioral FM detection limen and the ability to understand speech in the presence of noise. It was hypothesized that more robust and more accurate neural representations of FM would be correlated with smaller (better) FMDLs and better SIN understanding. An analysis used to determine Pearson product moment correlation coefficients revealed that no significant correlations existed between the stimulus-to-response correlation of the FFR response and behavioral FM detection in either frequency condition. Additionally, overall QuickSIN performance was not significantly correlated with the FFR to FM stimuli or with behavioral FM detection.

Clinical relevance: behavioral assessment

The FFR is an AEP that is not commonly utilized in the clinical setting. The FFR is typically not evoked until stimulus levels exceed 40-50 dBnHL (for a review, see Krishnan, 2007) and is therefore considered a poor estimator of hearing sensitivity. Analysis of the spectral characteristics of the FFR waveform in response to a /da/ syllable has shown strong correlations between response morphology and learning impairment (Cunningham et al., 2001; King et al., 2002; Johnson et al., 2005), however it is not known to what extent this measure is applied in audiology practice. In addition to the studies that have demonstrated FFR elicitation using pure tones, frequency sweeps, synthetic speech, and speech syllables, the present study adds FM tones and the underlying TFS to the current repertoire of FFR

stimuli. The FFR to FM tones was not predictive of FM detection or SIN understanding, however further study is needed to determine if correlations exist between the FFR and other domains of auditory processing. It is possible that the FFR, with its expanding catalog of eliciting stimuli, may serve as a reliable predictor of auditory processing. Additionally, for patients who cannot provide reliable, subjective judgment of auditory stimuli, a battery of AEPs including brainstem responses (i.e., auditory brainstem response and FFR) and cortical potentials (i.e., middle and late latency responses) may provide some insight into the function of the structures responsible for auditory processing.

Analysis of performances on the QuickSIN test across listeners did not indicate a high variance in total SNR loss. This finding is possibly due to the considerably homogenous subject pool that was recruited, which had a relatively narrow age range and included only listeners with normal hearing, no formal musical training, and unremarkable otologic and medical histories. Therefore, a statement regarding the test's ability to separate listeners by performance and further identify a cause of the separation cannot be made.

Clinical relevance: Cochlear implant processing strategies

The role of temporal envelope and TFS information in speech understanding in quiet and noise is still not clearly understood. Particular studies have combined the envelope of one speech signal with the TFS of another into what is described as an acoustic "chimera" (Smith, Delgutte, & Oxenham, 2002; Zeng, Nie, Liu, Stickney, Del Rio, Kong & Chen, 2004). Listener perception of these sounds suggests that TFS information more greatly influences pitch perception while envelope information provides greater cues to speech understanding. Consistent with these findings, common methods of signal processing in

cochlear implants used within the past ten years have placed emphasis on the encoding of the temporal envelope information of the speech signal (Rubinstein, 2004). Recent evidence, however, has found that combination processing strategies that include both the envelope processing of high frequencies and TFS processing of low frequencies provide comparable SIN understanding (Schatzer, Krenmayr, Au, Kals, & Zierhofer, 2010; Riss, Hamzavi, Selberherr, Kaider, Blineder, Starlinger, ... & Amoldner, 2011b) if not significant improvement compared to envelope-only strategies (Riss, Hamzavi, Katzinger, Baumgartner, Kaider, Gstoettner, & Amoldner, 2011a; Vermeire, Kleine Punte, & Vande Heyning, 2010). Additionally, FM encoding of speech has been shown to provide significant improvement in SIN understanding by implant wearers compared to traditional amplitude modulation-based encoding strategies (Nie, Stickney, & Zeng, 2005).

It was hypothesized that the findings of the present study would be consistent with the evidence of improved SIN understanding by implant wearers utilizing TFS processing strategies; the accuracy of the neural representation of TFS and FM detection would be correlated with QuickSIN performance. However, no such relationship existed. It is possible that the absence of a significant correlation was reflective of the parameters of the experiment rather than the role of TFS encoding and QuickSIN performance.

Methodological issues

The present study demonstrated that the physiological response to frequency modulation was not significantly correlated with behavioral detection of a similar stimulus. To examine the relationship between the FFR and FMDL data, the differences between the stimuli used in the FFR and FM detection tasks must be considered. Stimulus duration

differed between the two tasks with 500 ms presentations used in the physiological condition compared to 1000 ms tones in the behavioral condition. The 500 ms duration of the FFR stimuli allowed for the necessary amount of sweeps to be recorded more quickly, which shortened the overall test time. The 1000 ms FM stimulus presented during the behavioral task was thought to be of an adequate length to allow the listener to subjectively determine if frequency was modulated over time. Reducing the stimulus duration for the behavioral task to 500 ms, which is comparable to the 400 ms duration used by Buss et al., (2004), may affect FMDLs and possibly reveal a stronger brain behavior relationship.

Analysis of data collected from the three experiments within the present study suggests that neither the physiological response nor the behavioral detection of frequency modulation were significantly correlated with SIN understanding. This may be due to the fact that QuickSIN scores did not vary across subjects. It is possible that diversifying subject demographics (i.e., age, musical training, or monolingualism) may result in greater variance in the data, which could reveal a clearer brain-behavior relationship.

Previous studies have demonstrated an effect of age on SIN understanding (Frisina & Frisina, 1997) as well as behavioral frequency discrimination and the neural representation of frequency in FFR analyses (Clinard et al., 2010), however, each of these studies had subject pools with age ranges of at least five decades. The participant pool of the present study ranged from 21 – 40 years of age, less than two decades. The primary focus of the present study was toward the feasibility of using the FFR to investigate the neural representation of TFS and to determine if a behavioral and physiological relationship existed. Investigating an age effect was therefore not considered during the subject selection.

Future directions

The present study demonstrates that the FFR can be used to reflect the neurological representation of TFS when presented in the form of an FM tone. The hypothesized correlation between the physiological response and the behavioral detection of FM tones was not supported. Given these findings, future FFR studies may be designed with the goal of investigating the neural encoding of more complex stimuli. Additionally, a more diverse subject pool may be utilized to further investigate brain-behavior relationships and the effects of some extrinsic factors such as age and hearing sensitivity.

The FFR is capable of reliably reflecting the neural encoding of the TFS contained within a dynamic stimulus, as demonstrated in the present study. Future studies may utilize more complex stimuli in order to test the limits of the FFR. The FFR has already been elicited by steady-state and dynamic approximations of speech sounds (Krishnan, 1999, 2002; Krishnan et al., 2004), as well as the single-syllable speech stimulus, /da/ (Skoe & Kraus, 2010). Future studies utilizing novel speech and speech-like sounds may expand the variety of usable FFR stimuli, and possibly allow for more detailed brain-behavior relationships to be assessed.

Clinard et al., (2010) demonstrated that the FFR reflected a decline in neural phase-locking of 1000 Hz tones with increasing age, though this finding was not correlated with the perceptual threshold of frequency discrimination. The present study recorded FFRs from listeners with ages ranging 21 – 40 years old, though an age effect was not investigated. Future expansion upon the present study would include a greater age range across all

listeners. It is expected that findings would be in agreement with Clinard et al., (2010), demonstrating that FFRs at frequencies around 1000 Hz are weaker in older populations.

In order to rule out the possible mitigating factors of aging and hearing loss, the present study investigated the FFR of a relatively young group of individuals who all had normal audiometric thresholds. Future studies may include one or more subject groups that include listeners with some degree of hearing loss or auditory processing difficulty. While countless studies have investigated perceptual and physiological differences between normal hearing and hearing impaired individuals, it is notoriously difficult to determine the precise mechanism responsible for the hearing deficits. For this reason, careful study design would be needed to control for deficits in auditory mechanisms other than neural phase-locking.

Conclusions

Based on analyses of data measured from subjects who were evaluated in accordance with the methods described above, the following conclusions can be made:

1. - Stimulus-to-response correlations suggest that the FFR is capable of faithfully reflecting the dynamic nature of FM stimuli, particularly at 500 Hz.
2. - Larger depths of frequency modulation, which had larger peak-to-peak frequency deviations, had negative effects on the accuracy of the FFR.
3. - No significant correlations existed between the FFR elicited by FM tones and the behavioral detection of FM. Additionally, FFR measures were not predictive of SIN understanding.
4. - It is feasible to elicit FFRs using FM tones, allowing this approach to be applied to populations that have impaired perception of frequency modulation (e.g., older adults).

Chapter V

References

- Aiken, S. J., LeClair, K., & Kiefte, M. (2011). Pitch perception and frequency- following response in inharmonic signals. *The Journal of the Acoustical Society of America*, 129(4), 2591-2591.
- Bidelman, G. M., & Krishnan, A. (2011). Brainstem correlates of behavioral and compositional preferences of musical harmony. *Neuroreport*, 22(5), 212.
- Buss, E., Hall III, J. W., & Grose, J. H. (2004). Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear and hearing*, 25(3), 242-250.
- Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. *Hearing research*, 264(1), 48-55.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Psychology Press.
- Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clinical Neurophysiology*, 112(5), 758-767.
- Drennan, W. R., Won, J. H., Dasika, V. K., & Rubinstein, J. T. (2007). Effects of temporal fine structure on the lateralization of speech and on speech understanding in noise. *Journal of the Association for Research in Otolaryngology*, 8(3), 373-383.
- Ernst, S. M., & Moore, B. C. (2010). Mechanisms underlying the detection of frequency modulation. *The Journal of the Acoustical Society of America*, 128(6), 3642-3648.

- Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hearing research, 106*(1-2), 95-104.
- Glaser, E. M., Suter, C. M., Dasheiff, R., & Goldberg, A. (1976). The human frequency-following response: its behavior during continuous tone and tone burst stimulation. *Electroencephalography and Clinical Neurophysiology, 40*(1), 25-32.
- Hopkins, K., & Moore, B. C. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *The Journal of the Acoustical Society of America, 125*(1), 442-446.
- Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and clinical neurophysiology, 10*, 371-375.
- Johnson, K. L., Nicol, T. G., & Kraus, N. (2005). Brain stem response to speech: a biological marker of auditory processing. *Ear and hearing, 26*(5), 424-434.
- Kiang, N. Y. S., Watanabe, T., Thomas, E. C., & Clark, L. F. (1965). Discharge patterns of single fibres in the cat's auditory nerve. *Cambridge, Mass.: MIT Press.*
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). 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, 116*(4), 2395-2405.
- King, C., Warrier, C. M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neuroscience letters, 319*(2), 111-115.

- Krishnan, A. (1999). Human frequency-following responses to two-tone approximations of steady-state vowels. *Audiology and Neurotology*, 4(2), 95-103.
- Krishnan, A. (2002). Human frequency-following responses: representation of steady-state synthetic vowels. *Hearing research*, 166(1), 192-201.
- Krishnan, A. (2007). Frequency-following response. *Auditory evoked potentials: Basic principles and clinical application*, 313-335.
- Krishnan, A., Bidelman, G. M., Smalt, C. J., Ananthakrishnan, S., & Gandour, J. T. (2012). Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia*, 50(12), 2849-2859.
- Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and language*, 110(3), 135-148.
- Krishnan, A., Gandour, J. T., Bidelman, G. M., & Swaminathan, J. (2009). Experience dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*, 20(4), 408.
- Krishnan, A., & Parkinson, J. (2000). Human frequency-following response: representation of tonal sweeps. *Audiology and Neurotology*, 5(6), 312-321.
- 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), 1-12.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences*, 103(49), 18866-18869.

- McArdle, R. A., & Wilson, R. H. (2006). Homogeneity of the 18 QuickSIN lists. *Journal of the American Academy of Audiology, 17*(3).
- Moore, B. C. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology, 9*(4), 399-406.
- Moore, B. C., & Sek, A. (1995). Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation). *The Journal of the Acoustical Society of America, 97*(4), 2468-2478.
- Moore, B. C., & Sek, A. (1996). Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *The Journal of the Acoustical Society of America, 100*(4), 2320-2331.
- Nie, K., Stickney, G., & Zeng, F. G. (2005). Encoding frequency modulation to improve cochlear implant performance in noise. *Biomedical Engineering, IEEE Transactions on, 52*(1), 64-73.
- Palmer, A. R., & Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hearing research, 24*(1), 1-15.
- Plyler, P. N., & Ananthanarayan, A. K. (2001). Human frequency-following responses: representation of second formant transitions in normal-hearing and hearing-impaired listeners. *Journal of the American Academy of Audiology, 12*(10).

- Riss, D., Hamzavi, J. S., Katzinger, M., Baumgartner, W. D., Kaider, A., Gstoettner, W., & Arnoldner, C. (2011). Effects of fine structure and extended low frequencies in pediatric cochlear implant recipients. *International journal of pediatric otorhinolaryngology*, 75(4), 573-578.
- Riss, D., Hamzavi, J. S., Selberherr, A., Kaider, A., Blineder, M., Starlinger, V., ... & Arnoldner, C. (2011). Envelope versus fine structure speech coding strategy: a crossover study. *Otology & Neurotology*, 32(7), 1094-1101.
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 336(1278), 367-373.
- Rubinstein, J. T. (2004). How cochlear implants encode speech. *Current opinion in otolaryngology & head and neck surgery*, 12(5), 444-448.
- Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, 115(9), 2021-2030.
- Schatzer, R., Krenmayr, A., Au, D. K., Kals, M., & Zierhofer, C. (2010). Temporal fine structure in cochlear implants: preliminary speech perception results in Cantonese-speaking implant users. *Acta oto-laryngologica*, 130(9), 1031-1039.
- Sek, A., & Moore, B. C. (1995). Frequency discrimination as a function of frequency, measured in several ways. *The Journal of the Acoustical Society of America*, 97(4), 2479-2486.
- Skoe, E., & Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear and hearing*, 31(3), 302.

- Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*(6876), 87-90.
- Smith, J. C., Marsh, J. T., & Brown, W. S. (1975). Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalography and clinical neurophysiology*, *39*(5), 465-472.
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Perception of speech in noise: neural correlates. *Journal of Cognitive Neuroscience*, *23*(9), 2268-2279.
- Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing). *The Journal of the Acoustical Society of America*, *125*(5), 3328-3345.
- Swaminathan, J., Krishnan, A., & Gandour, J. T. (2008). Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport*, *19*(11), 1163-1167.
- Vermeire, K., Kleine Punte, A., & Van de Heyning, P. (2010). Better speech recognition in noise with the fine structure processing coding strategy. *ORL*, *72*(6), 305-311.
- Withnell, R. H. (2001). Brief report: the cochlear microphonic as an indication of outer hair cell function. *Ear and hearing*, *22*(1), 75-77.
- Worden, F. G., & Marsh, J. T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography and clinical neurophysiology*, *25*(1), 42-52.

Young, E. D., & Sachs, M. B. (1979). Representation of steady- state vowels in the temporal aspects of the discharge patterns of populations of auditory- nerve fibers. *The Journal of the Acoustical Society of America*, 66(5), 1381-1403.

Zeng, F. G., Nie, K. B., Liu, S., Stickney, G., Del Rio, E., Kong, Y. Y., & Chen, H. B. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues (L). *The Journal of the Acoustical Society of America*, 116(3), 1351-1354.