Neural Coding of Sound Envelope in Reverberant Environments

by

Michaël C. C. Slama

Diplôme d'Etudes Musicales, Conservatoire National de Région de Metz, 2004
Ingénieur, Ecole Supérieure d'Electricité, 2004
M.S., Georgia Institute of Technology, 2005
S.M., Massachusetts Institute of Technology, 2008

Submitted to the Harvard-MIT Division of Health Sciences and Technology
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

Signature of Author: .....
Harvard-MIT Division of Health Sciences and Technology
August 5, 2011

Certified by: .....
Bertrand Delgutte, Ph.D.
Professor of Otology and Laryngology, Harvard Medical School
Thesis Supervisor

Accepted by: .....
Ram Sasisekharan, Ph.D.
Director, Harvard-MIT Division of Health Sciences and Technology
Edward Hood Taplin Professor of Health Sciences and Technology and Biological Engineering

Neural Coding of Sound Envelope in Reverberant Environments

by

Michaël C. C. Slama

Submitted to the Harvard-MIT Division of Health Sciences and Technology
on August 5, 2011 in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Abstract

Speech reception depends critically on temporal modulations in the amplitude envelope of the speech signal. Reverberation encountered in everyday environments can substantially attenuate these modulations. To assess the effect of reverberation on the neural coding of amplitude envelope, we recorded from single units in the inferior colliculus (IC) of unanesthetized rabbit using sinusoidally amplitude modulated broadband noise stimuli presented in simulated anechoic and reverberant environments.

Consistent with the attenuation of amplitude modulation (AM) in the stimulus, both rate and temporal coding of AM were degraded in IC neurons. However, in most neurons, the degradation in temporal coding was smaller than the degradation in the stimulus. In many neurons, this compensation could be accounted for by the modulation input-output function (MIOF), which describes the nonlinear transformation of modulation depth from the sound stimulus into the neural response. However, in a subset of neurons, the MIOF underestimated the strength of temporal coding, suggesting that reverberant stimuli may have a coding advantage over anechoic stimuli with the same modulation depth. Additional experiments suggest that interaural envelope disparities and interaural decorrelation introduced by reverberation may partly explain this coding advantage.

In another set of experiments, we tested the hypothesis that temporal coding of AM is not static, but depends dynamically on the modulation depth statistics of preceding stimulation. In a subset of neurons, preceding stimulation history significantly altered the MIOF. On average, temporal coding of modulation frequency was more robust in conditions when low modulation depths predominate, as in reverberant environments.

Overall, our results suggest that the auditory system may possess mechanisms for reverberation compensation, and point to an important role of binaural and dynamic neural processes for robust coding of AM in reverberant environments.

Thesis Supervisor: Bertrand Delgutte, Ph.D.
Title: Professor of Otology and Laryngology, Harvard Medical School

Acknowledgments

First and foremost, I would like to thank my advisor, Bertrand Delgutte, for all of his help and guidance throughout my thesis project. The past years under his supervision have been incredibly intellectually stimulating, and lots of fun. I am grateful for having learned from his knowledge, intuitions, and advice, usually communicated with his unique sense of humor.

I would also like to express my gratitude to my thesis committee: Michale Fee and Barbara Shinn-Cunningham. I have always felt energized after a committee meeting thanks to their enthusiasm and ideas.

This thesis would not have been possible without the tremendous help from two individuals: Sasha Devore and Ken Hancock. I am grateful to Sasha for patiently teaching me the unanesthetized rabbit preparation and providing me with invaluable insights in the early days of my research project. I am grateful to Ken for our numerous discussions (sometimes related to science) and for his 24/7 help with the experimental software.

Many other people provided me with technical and administrative support over the years. I would like to thank in particular Jessica Cunha, Shig Kuwada, Connie Miller, Mike Ravicz, Dianna Sands, Chris Scarpino, Ishmael Stefanov-Wagner and Melissa Wood.

When I started in the SHBT program, I first worked with John Rosowski, who I would like to thank for his mentoring and for introducing me to research in auditory physiology. I am also extremely grateful to the entire SHBT faculty for their dedication to students and their passion for teaching.

I would also like to thank the many SHBT students I interacted with during my time at MIT. I feel fortunate to have shared this experience with such talented and passionate individuals.

Day-to-day work in the lab would not have been as pleasant had not it been for the cheerful interaction with EPL students and postdocs. In particular, I would like to thank my labmates: Yoojin Chung, Mitch Day, Luke Shaheen, Grace Wang, Bo Wen and Nathaniel Zuk.

Finally, I am extremely grateful to my parents for their unconditional support over the years, and to my wife Rachel and daughter Aviva with whom even the Boston winters are enjoyable.

Table of contents

List of figures

Chapter 1

General introduction

Sound waveforms can be decomposed into the product of a slowly-varying amplitude "envelope" and a fast-varying carrier signal. Slow temporal modulations in the amplitude envelope, or amplitude modulations (AM), are ubiquitous in natural sounds such as speech, animal vocalizations and environmental sounds. Acoustic analysis of natural sounds reveals that their average AM power spectra as a function of modulation frequency f_m are low-pass, with a $1/f_m$ decay over a few decades (Attias and Schreiner, 1997). Moreover, separation of natural sounds into the product of a temporal envelope and a wideband carrier is possible in many cases, as coherent fluctuation of temporal envelopes across different spectral frequency regions is common (Nelken *et al.*, 1999).

AM in the envelope of speech signals are crucial to speech understanding by human listeners. For example, speech intelligibility can reliably be predicted from an index based on AM in various frequency bands (the Speech Transmission Index or STI: Houtgast *et al.*, 1980). Another indication of the importance of AM for understanding speech comes from Shannon *et al.* (1995; 1998) who showed that speech reception is robust to the degradation of spectral information if AM in the envelope are preserved. Most communication situations, however, occur in the presence of reverberation, which degrades the sound envelope: While in an ideal anechoic environment, the speech signal is transmitted from a speaker to a listener with no degradation, a realistic acoustic environment contains boundary surfaces that reflect a portion of the direct sound, thereby adding reverberant energy that combines with the original signal. As the propagation paths of echoes are longer than the direct path between the speaker and the

listener, reverberant energy tends to attenuate AM by filling in the amplitude gaps in the signal envelope. This degradation of the envelope is not specific to enclosed spaces such as rooms (in which reflections mostly come from the floor, walls, and ceiling), but also occurs in natural environments such as forests, where sounds are reflected onto trees and foliage (Richards and Wiley, 1980).

In spite of the degradation of temporal modulations important for speech understanding, performance of normal-hearing listeners in speech intelligibility tasks is usually robust in moderate reverberation (e.g. Poissant *et al.*, 2006; Sato *et al.*, 2007; Yang and Bradley, 2009), suggesting that the normal auditory system may possess compensation mechanisms that counteract the degradation of envelopes in reverberation. *However, to date, no neurophysiological study has investigated the effects of reverberation on the neural coding of sound envelopes. The goals of this thesis are to quantify the degradation of envelope coding due to reverberation in the auditory midbrain, an important stage of the auditory system for envelope coding, and to test whether there exist neural compensation mechanisms that help envelope coding in realistic reverberant environments.*

Envelopes can be decomposed into Sinusoidally Amplitude Modulated (SAM) components of different modulation frequencies by Fourier analysis. Therefore, for simplicity, neurophysiological and psychophysical studies of AM have mostly used SAM stimuli such as SAM tones, SAM narrowband noise, or SAM broadband noise (Joris *et al.*, 2004). Perception of AM has been investigated at least since the 19th century, with von Helmholtz reporting how the sensation produced by a beating stimulus changes with beating rate (von Helmoltz, 1863). For SAM broadband noise, AM detection thresholds vs. f_m functions are typically low-pass (Viemeister, 1979), i.e. AM detection is more sensitive at low f_m. However, the dependence of

detection thresholds on f_m is very different for narrowband noise, and depends on bandwidth (Ewert and Dau, 2004). The hearing-impaired have higher modulation detection thresholds (Bacon and Viemeister, 1985). For SAM broadband noise, discrimination thresholds for modulation depth m (quantity indicating by how much the modulated amplitude varies around its mean) increase with m and are roughly independent of overall level (in the range 0 – 40 dB SPL), and f_m (in the range 25 – 100 Hz) (Wakefield and Viemeister, 1990).

Neurophysiological studies of AM coding have been carried out in a variety of species and throughout the auditory system. Tuning to AM frequency has been traditionally investigated by measuring average firing rate and the modulation depth of the period histogram response (response modulation depth or RMD) in response to SAM stimuli. A rate Modulation Transfer Function (rMTF) is obtained by plotting average firing rate as a function of f_m, while a temporal Modulation Transfer Function (tMTF) is obtained by plotting RMD as a function of f_m (e.g., Eggermont, 1991). The larger the RMD, the more phase-locked the neurons is to f_m. The modulation gain, expressed in dB, can be computed from the tMTF by taking the ratio of RMD to stimulus modulation depth m.

rMTF and tMTF measurements reveal a progressive shift along the ascending auditory pathway from a temporal code in the auditory nerve, to a dominant rate code in primary auditory cortex for mid- and high-frequencies (Joris *et al.*, 2004). In the AN, the tMTFs are essentially low-pass: For example in anesthetized cat (Joris and Yin, 1992), the tMTFs measured in response to SAM tones have a positive gain up to about 1 kHz, before envelope synchrony dramatically decreases. rMTFs are essentially flat. Temporal coding remains dominant in the cochlear nucleus (CN), where modulation gains are generally larger than in the AN, and band-pass tMTFs arise (e.g. Møller, 1974, in anesthetized rat). The superior olivary complex (SOC) is

the first place in the auditory pathway where a conversion between temporal and rate code has been reported for modulation coding (e.g. Kuwada and Batra, 1999, in unanesthetized rabbit). In the Inferior Colliculus (IC), the modulation gains of the tMTFs are even larger than for AN or CN, and the shapes of the rMTFs are much more varied (e.g. Rees and Palmer, 1989, in guinea pig), suggesting that the auditory midbrain is the stage of a major transformation in the neural processing of AM. The upper limit of phase-locking to AM gets progressively lower along the ascending auditory pathway: In the thalamus and the auditory cortex, this cutoff is generally below 100 Hz (e.g. Miller *et al.*, 2001).

We chose to record from the IC because it plays a primordial role in the coding of AM, and is a major processing center of the auditory system. The main ascending inputs to the IC include projections from the ventral and dorsal CN, from the medial, lateral, and periolivary nuclei of the SOC, and from the nuclei of the lateral lemniscus (LL) (e.g. Adams, 1979). Descending inputs to the IC include the medial geniculate body (MGB) of the thalamus (e.g. Kuwabara and Zook, 2000) and most areas of the auditory cortex (e.g. Winer *et al.*, 1998). Commissural connections from the other IC and intracollicular connections are also important (e.g. Saldaña and Merchàn, 2005). The main target of IC neurons is the MGB (e.g. Wenstrup and Grose, 1995), but there are also numerous descending projections from the IC, mainly to the ipsilateral dorsal nucleus of the LL, the ipsilateral periolivary nuclei of the SOC (e.g. Caicedo and Herbert, 1993) and the ipsilateral and contralateral dorsal CN (e.g. Vater and Feng, 1990).

Several subdivisions have been characterized within the IC based on anatomical and physiological properties of their neurons. The main subdivisions of the IC are the central nucleus (ICC), the dorsal cortex (DC), and the external cortex (ICX) located laterally, anteriorly,

and posteriorly. The ICC is composed of parallel tonotopic laminae with characteristic frequency increasing dorsolaterally to ventromedially (e.g. Merzenich and Reid, 1974).

Many neurons in the IC are modulation sensitive. Most tMTFs are low-pass or band-pass, with modulation gains of the order of 15 dB at the temporal best modulation frequency (tBMF: frequency for which the tMTF is maximal); rMTF types are more varied but the most common seems to be band-pass (e.g. Rees and Palmer, 1989). The best modulation frequencies (BMF) found in the ICC are typically lower than 200 Hz: For example in unanesthetized rabbits (Nelson and Carney, 2007), most temporal and rate BMFs were found between 16 and 128 Hz. It is not clear whether these BMFs are topographically organized in the ICC: Evidence for a gross map of rBMF orthogonal to the tonotopic axis was reported in cat (Schreiner and Langner, 1988) and awake macaque (Baumann *et al.*, 2011). Interpretation as evidence for a modulation filter bank is difficult as MTF shapes and BMFs are highly dependent on stimulus level (Krishna and Semple, 2000) and modulation waveform (Sinex *et al.*, 2002; Krebs *et al.*, 2008).

In chapter 2, we assess the effect of reverberation on the neural coding of AM by measuring neural MTFs in single units of unanesthetized rabbit IC using SAM broadband noise stimuli presented in simulated anechoic and reverberant environments. Consistent with the attenuation of AM in the stimulus, we find that both rate and temporal coding of AM are degraded in IC neurons. However, in most neurons, the degradation in temporal coding is smaller than the degradation in the stimulus, suggesting that the auditory system partially compensates for the acoustic degradation. In many neurons, this compensation could be accounted for by the compressive shape of the modulation input-output function (MIOF), which describes the nonlinear transformation of modulation depth from the sound stimulus into the neural response.

In chapter 3, we further investigate the temporal coding of AM in reverberation by comparing the RMD to reverberant and anechoic stimuli whose modulation depths are matched. We find that in a subset of neurons, phase-locking to the modulation is stronger for reverberant stimuli than for anechoic stimuli with the same modulation depth. To explain this coding advantage, we explore the influence of various monaural and binaural features of reverberant stimuli on temporal coding of AM. We find that monaural features such as envelope distortion and spectral coloration introduced by reverberation do not play an important role in the reverberant response, while binaural features such as interaural envelope disparities and interaural cross-correlation may partly explain the reverberant advantage.

Finally, in chapter 4, we test the hypothesis that temporal coding of AM is not static, but depends dynamically on the modulation depth statistics of preceding stimulation. In a subset of neurons, preceding stimulation history significantly altered the MIOF. On average, temporal coding of modulation frequency was more robust in conditions when low modulation depths predominate, as in reverberant environments.

This thesis expands the knowledge of AM coding with stimuli more relevant to everyday listening situations. Overall, our results suggest that the auditory system may possess mechanisms for reverberation compensation, and point to an important role of binaural and dynamic neural processes for robust coding of AM in reverberant environments.

Chapter 2

Effects of reverberation on rate and temporal coding of amplitude modulation: Steady state analysis

Abstract

Speech reception depends critically on temporal modulations in the amplitude envelope of the speech signal. These modulations are substantially attenuated by reverberation encountered in everyday environments. To assess the effect of reverberation on the neural coding of amplitude envelope, we recorded from single units in the inferior colliculus of unanesthetized rabbit using sinusoidally amplitude modulated broadband noise stimuli presented in simulated anechoic and reverberant environments.

Both the average firing rate and the modulation depth of the period histogram were studied as a function of modulation frequency to obtain rate and temporal modulation transfer functions (rMTF and tMTF), respectively. The maximum response modulation depth of the tMTF as well as the signal-to-noise ratio of the rMTF were consistently lower in the reverberant condition than in the anechoic condition, indicating that reverberation degrades both temporal and rate coding, as expected. However, in most neurons, the degradation in temporal coding was smaller than the acoustic degradation, suggesting that a neural modulation gain compensates for the degradation in the stimulus. The compressive shape of the modulation input/output function describing the transformation of anechoic modulation depths into neural modulations was successful in explaining this compensation gain across the population. However, in a subset of neurons, the prediction from the modulation input/output function underestimated the strength of temporal coding, suggesting that, in these neurons, reverberant stimuli have a coding advantage over anechoic stimuli with the same modulation depth.

Overall, our results indicate that amplitude modulation coding in the inferior colliculus is degraded in reverberation, but that there are mechanisms in the auditory system that counteract the degradation of the temporal envelope.

Introduction

Reverberation is present in most communication situations as sounds are reflected from the walls, floor, and ceiling of rooms. Performance of normal-hearing subjects in speech intelligibility tasks can be degraded in the presence of extreme reverberation and/or noise (e.g. Payton *et al.*, 1994; Neuman *et al.*, 2010) but is usually robust in moderate reverberation (Poissant *et al.*, 2006; Sato *et al.*, 2007; Yang and Bradley, 2009). The decrease in performance from a quiet to a reverberant environment is typically larger for the hearing-impaired than for the normal-hearing listeners (Plomp and Duquesnoy, 1980; Irwin and McAuley, 1987; Payton *et al.*, 1994; Sato *et al.*, 2007). The more severe the hearing loss, the larger the drop in intelligibility (Plomp and Duquesnoy, 1980; Harris and Swenson, 1990).

The temporal envelope of speech signals provides important cues for understanding speech. Reducing the amplitude of slow temporal modulations in the envelope of speech decreases intelligibility (Drullman *et al.*, 1984). Conversely, increasing the modulation depths in the speech envelopes can improve speech intelligibility in noise (Lorenzi *et al.*, 1999). Performance in sentence and consonant recognition tasks in quiet remains good after the spectral information is reduced to a few frequency bands, as long as amplitude modulations are preserved (Shannon *et al.*, 1995; 1998; Apoux and Bacon, 2004; 2008).

Reverberation degrades the amplitude modulations (AM) in the temporal envelope of speech signals important for speech intelligibility, essentially acting as a low-pass filter in the modulation domain. The effect of reverberation on speech intelligibility can approximately be predicted by the Speech Transmission Index, a physical measurement based on temporal modulations (Houtgast *et al.*, 1980; Steeneken and Houtgast, 1980). Moreover, speech intelligibility in reverberation for the normal-hearing and the hearing-impaired is correlated to

performance in temporal resolution tasks such as gap detection (Irwin and McAuley, 1987; Dreschler and Leeuw, 1990; Lacher-Fougère and Demany, 2005). Therefore temporal envelope cues are crucial to understand speech in reverberation. The underlying motivation of this study is to test whether there are neural mechanisms that help the normal auditory system compensate for the degradation in temporal envelope due to reverberation.

Tuning to AM frequency is a widespread feature of neurons in the auditory system. Traditionally, rate and temporal tuning are characterized by neural modulation transfer functions (e.g., Eggermont, 1991). The rate modulation transfer function (rMTF) is the average firing rate as a function of modulation frequency (f_m), while the temporal modulation transfer function (tMTF) and phase modulation transfer function (pMTF) are respectively the magnitude and phase of the modulation depth of the neural response, as a function of f_m. rMTF and tMTF measurements reveal a progressive shift along the ascending auditory pathway from a temporal code in the auditory nerve, to a dominant rate code in primary auditory cortex for mid- and high-frequencies (Joris *et al.*, 2004). The Inferior Colliculus (IC), a major processing center of the auditory system, is a key stage for the coding of AM, as IC neurons, overall, exhibit larger tMTF gains as well as sharper rMTF tuning than subcollicular neurons (Joris *et al.*, 2004).

AM coding has been investigated in the IC of animal models with a variety of stimuli, but to our knowledge, never in the presence of reverberation. To investigate the effects of reverberation on AM coding, we recorded from single-units in the IC of unanesthetized rabbits in response to sinusoidally amplitude modulated (SAM) broadband noise in simulated anechoic and reverberant environments. Consistent with the degradation of AM in the stimulus, we find that both rate and temporal coding of modulation frequency are degraded in IC neurons. However, in most neurons, the degradation in temporal coding was smaller than the degradation in the

stimulus, suggesting that a neural gain compensates for the degradation in the stimulus. The compressive shape of the nonlinear input-output function describing the transformation of anechoic modulation depths into neural modulations accounted for much of this compensation gain across the population. However, in a subset of neurons, the prediction from the modulation input/output function underestimated the strength of temporal coding in reverberation, suggesting that, in these neurons, reverberant stimuli have a coding advantage over anechoic stimuli with the same modulation depth.

Methods

Surgical Preparation

Surgical procedures to prepare dutch-belted rabbits (*Oryctolagus cuniculus*) for chronic unanesthetized recordings of IC single units were based on the techniques of Kuwada *et al.* (1987), Nelson and Carney (2007), and Devore and Delgutte (2010). The use of an unanesthetized preparation guarantees that the neurophysiological results are not biased by an effect of anesthesia, which was shown to be important in the IC (Tollin *et al.*, 2004; Ter-Mikaelian *et al.*, 2007). We chose the rabbit as an animal model because it is a docile animal that can be easily trained to stay still during an unanesthetized experiment.

Animals underwent two separated aseptic surgeries before being used for experiments. In the first aseptic surgery, animals were anesthetized with an intramuscular injection of acepromazine (1 mg/kg), ketamine (44 mg/kg), and xylazine (6 mg/kg). Anesthesia was monitored by periodically checking pedal withdrawal and corneal reflexes, and supplemental doses of ketamine (15 mg/kg) and xylazine (2 mg/kg) were administered as necessary. An incision was made along the midline, and connective tissue was removed to expose the skull. After implanting stainless steel screws into the skull, a brass head bar (to keep the animal head immobilized during an experiment) and a stainless steel cylinder were attached to the screws with dental acrylic. The anterior part of the cylinder was aligned with bregma and the head bar was positioned at a 20-30° angle with the tooth-orbit line. At the end of the procedure, ear molds were made with vinyl polysiloxane impression material (Reprosil®). After the procedure, the surgical site was treated twice a day for a week with a topical antibiotic (Bacitracin) and animals were monitored for signs of pain or discomfort. Buprenorphine (Buprenex®, 0.015 mg/kg) was administered subcutaneously as an analgesic up to 48 hours after surgery.

The animals were given a week to recover, before starting a daily training protocol designed to habituate them to the experimental setup. Animals were restrained in a spandex sleeve and sat in a cradle. Head movements were prevented by locking the implanted brass head bar in place. Sound stimuli were delivered through speakers connected to the ear molds. Training was conducted for two weeks until the animals were able to sit comfortably for two hours with their body and head restrained while listening to sound.

After completion of training, a second aseptic surgical procedure was performed. For this procedure, animals were anesthetized either by IM injection of a mixture of acepromazine, ketamine and xylazine as described for the first procedure, or by inhalation of isoflurane. Isoflurane anesthesia was induced by placing the animals in a hermetic Plexiglas box ventilated with a 1 L/min flow of isoflurane (5% mixture in oxygen), and maintained throughout the procedure with mask delivery of a 0.5-1 L/min flow of isoflurane (1-2.5 % mixture in oxygen). Isoflurane concentration was adjusted to maintain a suppressed pedal withdrawal reflex and high oxygen blood saturation. A small (≈1-2 mm diameter) craniotomy was performed about 10 mm posterior from bregma and 3 mm lateral from the midline, inside the implanted stainless steel cylinder. Bacitracin was applied to the exposed dura, and the cylinder filled with a sterile elastopolymer (Sammons-Preston). During the course of several months of chronic recordings, additional surgeries were periodically done to clean the exposed dura off of scar tissue and/or extend the size of the existing craniotomy.

All surgical procedures were approved by the Institutional Animal Care and Use Committees of the Massachusetts Eye and Ear Infirmary and the Massachusetts Institute of Technology.

Recording Procedures

Recording sessions took place in an electrically-shielded sound attenuating chamber. At the beginning of a recording session, animals were securely restrained as described above, and the elastopolymer cap filling the implanted stainless steel cylinder was removed. The inside of the cylinder and exposed skull and dura were flushed with sterile saline, and any scar tissue that had grown since the last recording session was removed. A few drops of a topical anesthetic (Marcaine) were applied to the surface of the dura for a few minutes to reduce pain upon insertion of a guide tube and/or electrode through the meninges.

The two ear molds made at the end of the first aseptic surgery described above were then inserted in the animals' ears, and two Beyer-Dynamic (DT-48) sound speakers were coupled to ≈5 mm diameter sound delivery plastic tubes encased in the ear molds. A probe-tube microphone (Etymotic ER-7C) passing through the sound delivery tubes was used to measure sound pressure near the tympanic membrane in response to broadband chirp stimuli. The transfer function of the acoustic system was computed from these measurements, and inverse filters compensating for this transfer function were digitally created. All sound stimuli were generated by a 24-bit D/A converter (National Instruments NIDAC 4461) at a sampling rate of 50 kHz, and filtered by the inverse filters resulting from the acoustic calibration of that experimental session.

In earlier experiments, we inserted a 25-gauge stainless steel guide tube 1-2 mm below the dura through which an electrode was descended, in order to precisely position the electrode and protect its tip. In later experiments, we stopped using a guide tube so as to limit dural scar tissue growth, and encased the part of the electrode that would not penetrate the brain in a glass pipette to better control position in the horizontal plane. We recorded from single units in the IC

using either epoxy-coated (AM Systems) or glass-coated (custom made) tungsten electrodes. Electrodes were descended vertically into the IC (centered around 7-8 mm below the surface of the brain) using a remote-controlled hydraulic micropositioner (Kopf 650). Neural activity from the electrode was amplified and band-pass filtered between 0.3 and 3 kHz. Custom software was used to measure spike times by threshold crossing and save them to disk.

Experimental sessions typically occurred 6 days a week for up to 3 months in each IC. Session duration varied but did not exceed 2.5 hours. Animals were monitored with a video system, and sessions were interrupted if they showed any sign of discomfort. At the end of a session, the electrode was pulled out, the exposed dura was flushed with sterile saline, and covered with Bacitracin ointment to prevent infection. A new elastopolymer cap was then made to fill the stainless steel cylinder and protect the exposed skull.

Virtual Auditory Space Stimuli

We simulated binaural room impulse responses (BRIRs) using the room-image method (Allen and Berkley, 1979; Shinn-Cunningham *et al.*, 2001) with room dimensions and simulation parameters similar to those of Devore *et al.* (2009). The virtual room measured 11x13x3 m and two receivers, separated by 12 cm, were located near its center (Fig. 2.1 A). We included a sphere as a model for the rabbit's head so that the resulting BRIRs would contain both interaural time and level difference cues. BRIRs were simulated for a source positioned at 0º azimuth and at distances of 1.5 and 3 m from the center of the simulated head. The direct-to-reverberant energy ratio was 0 dB for the 1.5 m condition (moderate reverberation) and -6 dB for the 3 m condition (strong reverberation). Anechoic impulse responses were obtained by isolating the first peak (direct sound) from the reverberant BRIRs (see Fig. 2.1 B for an example of a

reverberant impulse response with a zoom onto the first 50 ms where direct sound and individual reflections can be resolved). For a given source-receiver distance, the energy in a reverberant BRIR was larger than the energy of the corresponding anechoic BRIR, due to the addition of reverberant energy. In order to control for sound pressure level, we scaled both channels of a reverberant BRIR with a common factor chosen so that the energy in the reverberant BRIR contralateral to the recording site matched that of the anechoic BRIR. Virtual auditory space stimuli were created by convolving Sinusoidally Amplitude Modulated (SAM) broadband noise tokens with the left and right BRIRs (Fig. 2.1 C). The standard stimulus was 100% modulated, but lower modulation depths were used occasionally.

Reverberation is a dynamic process: The earlier portion of the reverberant stimuli is more modulated than the later portion, as reflections degrading amplitude modulations gradually build up over time following stimulus onset. The decay of modulation depth as a function of time in our reverberant stimuli depended on both f_m and reverberant condition, but had a similar profile in all cases: A sharp decrease in the first 20-100 ms, followed by a plateau (Fig. 2.1 D). At all frequencies, modulation depth was nearly constant (less than 5% variation) after 250 ms. In this chapter, we investigate the effect of reverberation on AM coding in the steady state part of the reverberant stimuli, by excluding from our analyses an integer number of cycles equal to or greater than 250 ms. The time course of the neural responses in the early part of the reverberant stimuli is studied in Chapter 3.

The degradation in AM between the source and the receivers due to reverberation was quantified by room Modulation Transfer Functions (MTFs, Fig. 2.1 E). Room MTFs were obtained by computing the change in modulation depth magnitude and phase of the envelope of 100% SAM broadband noise stimuli as a function of f_m, in the steady state region. To compute

the modulation depth for modulation frequency f_m, we created 50 simulated reverberant stimuli by convolving our BRIRs with 50 different tokens of 100% SAM broadband noise (2 sec duration, frequency f_m), then averaged the full-wave rectified reverberant stimuli across tokens to obtain the amplitude envelope, and took the discrete Fourier transform of the steady state part of this envelope (FT_{env}). The complex modulation depth at a receiver is given by (Houtgast *et al.*, 1980; Schroeder, 1981):

$$m(f_m) = 2 \times \frac{FT_{env}(f_m)}{FT_{env}(0)} \qquad \text{(Eq. 2.1)}$$

Since the modulation depth at the source is 100%, the dB magnitude of the MTF is simply:

$$MTF_{dB}(f_m) = 20 \times \log_{10} |m(f_m)| \qquad \text{(Eq. 2.2)}$$

For reverberant conditions, we refer to the modulation depth at the source as the "distal modulation depth", and the modulation depth at the ear drum as the "proximal modulation depth". For anechoic conditions, proximal and distal modulation depths are essentially equal, and referred to simply as the "stimulus modulation depth".

The simulated BRIRs did not include the initial delay corresponding to the propagation time of the direct sound from the source to the receivers. The MTF phases plotted in Fig. 2.1 E are the phases of the modulation waveforms relative to the direct sound at the receivers.

Experimental Procedures

A search stimulus (40 Hz SAM broadband noise bursts presented binaurally at 60 dB SPL) was played while descending the microelectrode through the brain towards the IC. Well isolated single units were subsequently studied. First, a rate level function was measured using 200 ms diotic broadband noise bursts presented in random order with levels between 0 and 70 dB

SPL (5-10 repetitions), from which an acoustic threshold was determined. Characteristic frequency (CF) was determined using either an iso-rate tracking algorithm (Kiang and Moxon, 1974) or by presenting tone pips of various frequencies near threshold. All subsequent stimuli were presented at the same sound level, about 15-20 dB above the broadband noise threshold.

Neural Modulation Transfer Functions (MTFs) were measured in response to anechoic and reverberant virtual auditory space stimuli. The sound source was 100% modulated SAM broadband noise (different tokens for every trial) with f_m between 4-256 Hz (octave spacing, plus 45, 90, and 180 Hz). Either the moderate or the strong reverberant condition was used first, and the other reverberant condition was studied subsequently, time permitting. Stimuli were 2 second long, followed by a 1 second silent interval, and presented 3-5 times each. Presentation order was randomized across f_m and between anechoic and reverberant conditions.

Modulation Input/Output Functions (MIOFs) were measured in response to anechoic SAM broadband noise of different modulation depths presented in random order. We typically used 5-12 modulation depths between 0 and 1. MIOFs were usually measured at one f_m, chosen to elicit both a large firing rate and strong phase-locking to the modulation. In a few experiments, we measured MIOFs for a range of frequencies similar to the range used for the MTF measurements. Stimuli were 2 second long followed by a 1 second silence, repeated 4-5 times, and presented randomly across modulation depths and f_m (in the experiments where the MIOF was measured for a range of f_m).

Data Analysis

Neural Modulation Transfer Functions

Only well isolated single units were included in our data analysis. To limit the influence of turning on the sound stimulus, which often elicited a strong onset response, and to quantify AM coding in the steady state part of the reverberant stimuli, where proximal modulation depth was constant (Fig. 2.1 D), we discarded the action potentials occurring in a time interval corresponding to the smallest integer number of modulation cycles equal to or greater than 250 ms.

Rate Modulation Transfer Functions (rMTFs) were computed for each room condition as the average firing rate across trials as a function of f_m,. The strength of envelope frequency representation in the rMTF was estimated by a signal-to-noise ratio (SNR) metric based on ANOVA (Hancock *et al.*, 2010). The rMTF SNR was computed as the variance in firing rate due to changes in f_m, divided by the intrinsic variance in firing rate across multiple repetitions of the same stimulus:

$$SNR = 10\log_{10}\left(\frac{SS_{fm}}{SS_{total} - SS_{fm}}\right) \qquad \text{(Eq. 2.3)}$$

with $SS_{fm} = N\sum_{fm}\left[rMTF(f_m) - \overline{rMTF}\right]^2$, $SS_{total} = \sum_{fm}\sum_{k=1}^{N}\left[r(k, f_m) - \overline{rMTF}\right]^2$, N the number of repetitions of each f_m, $r(k, f_m)$ the firing rate for repetition k of frequency f_m, $rMTF(f_m) = \frac{1}{N}\sum_{k=1}^{N} r(k, f_m)$, and $\overline{rMTF}$ the mean $rMTF$ across frequencies.

Temporal Modulation Transfer Functions (tMTFs) were defined as the modulation depth (magnitude) of the response period histogram as a function of f_m. To avoid any numerical errors

resulting from binning spike times to construct period histograms, we computed the response modulation depth (RMD) from the spike times themselves, as twice the vector strength:

$$tMTF(f_m) = \frac{2}{n}\sqrt{\left[\sum_{i=1}^{n}\cos(2\pi f_m t_i)\right]^2 + \left[\sum_{i=1}^{n}\sin(2\pi f_m t_i)\right]^2} \qquad \text{(Eq. 2.4)}$$

where n is the total spike count and t_i are the spike times.

A Rayleigh test of uniformity was used ($\alpha<0.05$) to assess the significance of the computed RMD (Mardia, 1972). The standard deviation of the RMD for each f_m was estimated using an approximation derived for von Mises distributions and described in Eq. (4.8.18) of Mardia and Jupp (1999). Both the Rayleigh statistic and the standard deviation estimate only depend on total spike count and vector strength.

Phase Modulation Transfer Functions (pMTFs) were defined as the phase of the complex modulation depth of the period histogram response relative to the phase of the sound source envelope. Similarly to the tMTF, we computed the phase directly from the spike times. Since our f_m spacing was rather wide, it was not possible to determine with certainty how to unwrap the phase as a function of f_m. Instead, we assumed the anechoic case to have a linear phase and a maximum latency of 15 ms, consistent with studies in rat IC (Rees and Møller, 1983) and gerbil IC (Krishna and Semple, 2000). We unwrapped the phase by adding multiples of ± 1 cycles so that the phase difference between consecutive data points was less than 0.5 cycles. We then fitted a regression line to the low frequency phase data points ($f_m \leq 32$ Hz[1]), and arranged the high frequency data points so as to minimize the distance from the regression line. We estimated the mean anechoic response group delay as the slope of the regression line. Since changes in

[1] Under the assumption of linear phase with maximum latency of L_{max}=15 ms, phases are unambiguous up to frequency f_{max} such that $L_{max} \times f_{max} < 0.5$ cycles, i.e. $f_{max} = 32$ Hz

phase introduced by reverberation were small (typically < 0.2 cycles) compared to the accumulation of unwrapped phase with f_m (typically 3-5 cycles between 4 and 256 Hz) due to neural delays, we removed the mean anechoic delay before plotting anechoic and reverberant pMTFs. We did not use phases from tMTF points that were not Rayleigh significant. Standard deviations of the phases at each f_m were estimated using an approximation derived for von Mises distributions and described in Eq. (4.8.19) of Mardia and Jupp (1999). Similarly to the estimate for standard deviation of the RMD, this estimate only depends on total spike count and vector strength.

From the tMTF data in the anechoic and reverberant cases, and from the room MTF, we defined the neural compensation gain as the difference (in dB) between the neural degradation and the acoustic degradation due to reverberation:

$$G_{dB}(f_m) = 20\log_{10}\left(\frac{tMTF_{reverberant}(f_m)}{tMTF_{anechoic}(f_m)}\right) - MTF_{dB}(f_m) \qquad \text{(Eq. 2.5)}$$

Similarly, we defined the phase shift between neural and acoustic degradations as:

$$\Phi(f_m) = pMTF_{reverberant}(f_m) - pMTF_{anechoic}(f_m) - MTF_{phase}(f_m) \qquad \text{(Eq. 2.6)}$$

Since the receivers were positioned near the center of the room, with the source at 0º azimuth relative to the simulated head, left and right room MTFs for the same source-receiver distance were very similar. For the purpose of comparisons with and predictions of neural data, we used the room MTF for the ear contralateral to the IC recorded from.

Modulation Input/Output Functions (MIOFs)

Modulation Input/Output Functions (MIOFs) were constructed similarly to the tMTFs, by computing the RMD of the period histogram as a function of stimulus modulation depth, after

removing the spikes occurring before 250 ms. Rayleigh tests of uniformity were performed at level $\alpha<0.05$. Standard deviations of the RMD were estimated as described for the tMTFs. A scaled incomplete beta function was fitted to the data points as a function of stimulus modulation depth using a weighted least square procedure, with the reciprocals of the RMD variances used as weights. The fitted curve has 3 parameters A, B, and C, and has the form:

$$MIOF(m) = A\int_0^m u^{B-1}(1-u)^{C-1}\,du \qquad \text{(Eq. 2.7)}$$

We chose an incomplete beta function because it provides both compressive and expansive shapes that encompass the diversity of MIOFs encountered, and is constrained to have a value of 0 for $m = 0$ (no response modulation for an unmodulated input).

Goodness of fit was assessed with the weighted coefficient of determination r^2. We analyzed and used MIOFs only when r^2 was greater than 0.5. This criterion excluded 3 neurons (out of 95), where synchronization was poor and variability was large. For the 92 neurons that passed this criterion, the fit was usually very good (median r^2 was 0.98, range was 0.55 to 0.99).

From the MIOF, we predicted the neural compensation gain by computing the change in RMD between fully modulated ($m = 1$, corresponding to the anechoic case) and partially modulated anechoic stimuli ($m < 1$, corresponding to the proximal modulation depth of the reverberant case), relative to the change in stimulus modulation depth:

$$G_{dB,predicted}(m) = 20\log_{10}\left(\frac{MIOF(m)}{MIOF(1)}\right) - 20\log_{10}\left(\frac{m}{1}\right) \qquad \text{(Eq. 2.8)}$$

By rearranging the terms, $G_{dB,predicted}$ is simply the neural modulation gain of the MIOF for a stimulus modulation depth of m relative to the neural modulation gain for a stimulus modulation depth of 1:

$$G_{dB,predicted}(m) = G_{dB,MIOF}(m) - G_{dB,MIOF}(1) \qquad \text{(Eq. 2.9)}$$

$$\text{with } G_{dB,MIOF} = 20\log_{10}\left(\frac{MIOF(m)}{m}\right)$$

The standard deviation of $G_{dB,predicted}$ was inferred from the standard deviation of RMD by using a Taylor expansion:

$$std_G(m) \approx \left|\frac{\partial G_{dB,predicted}(m)}{\partial MIOF}\right| \times std_y(m) \approx \frac{20}{\ln(10) \times MIOF(m)} \times std_y(m) \qquad \text{(Eq. 2.10)}$$

k-means clustering of the MIOFs

To investigate whether different populations of IC neurons had specific MIOF characteristics, we partitioned the set of fitted MIOFs using k-means clustering. Considering our set of MIOFs, $\{MIOFj\}$, and $\mathbf{S} = \{C_1, C_2, ..., C_k\}$ a set of k clusters partitioning $\{MIOFj\}$, the k-means clustering procedure finds a partition **S** that minimizes the within-cluster sum of squares $WCSS = \sum_{i=1}^{k} \sum_{IOMF_j \in C_i} \left\| MIOF_j - \mu_i \right\|^2$. To choose the number of clusters, we plotted the minimized *WCSS* as a function of *k*. In our case, increasing the number of clusters beyond 3 did not substantially reduce *WCSS* (Fig. 2.7 C, inset). We therefore partitioned the population into 3 clusters.

Results

Reverberation degraded rate and temporal coding of amplitude modulation

We measured neural modulation transfer functions with virtual auditory space stimuli (Fig. 2.1, Methods) to study the effects of reverberation on the coding of amplitude modulation in 110 well isolated single units from the IC of unanesthetized rabbits. The virtual auditory space stimuli were 100% modulated SAM broadband noise produced by a sound source located 1.5 or 3 m away from a simulated spherical head, in a classroom-size virtual room (Fig. 2.1 A). The Direct-to-Reverberant (D/R) energy ratio was 0 dB for the 1.5 m source-to-receiver distance ("moderate" reverberation) and -6 dB for the 3 m source-to-receiver distance ("strong reverberation"). Simulated reverberation degraded the modulation depth of the stimuli in a modulation frequency-dependent fashion, as illustrated by the room Modulation Transfer Functions (Fig. 2.1 D-E). Intuitively, reflections from the walls, ceiling and floor are superimposed on the source stimulus (Fig. 2.1 B) and partially fill the gaps in the envelope of the SAM stimulus, reducing its modulation depth.

Figure 2.2 A-C shows the responses of an example neuron to simulated anechoic and reverberant stimuli. The spike rasters (Fig. 2.2 A) display spike times for each repetition of each modulation frequency condition in the anechoic (left panel, in blue) and moderate reverberant (right panel, in red) conditions. Firing rate varied substantially with f_m in the anechoic case, with few spikes for modulation frequencies above 64 Hz, except near the onset. The rMTF for the anechoic condition (Fig. 2.2 C, top panel, in blue) was constructed by averaging the firing rates across repetitions for each f_m, and had a low-pass shape. The high frequency region of lesser activity in the anechoic case is much more responsive in the reverberant case (Fig. 2.2 A, right panel), resulting in an almost flat reverberant rMTF (Fig. 2.2 C, top panel, in red).

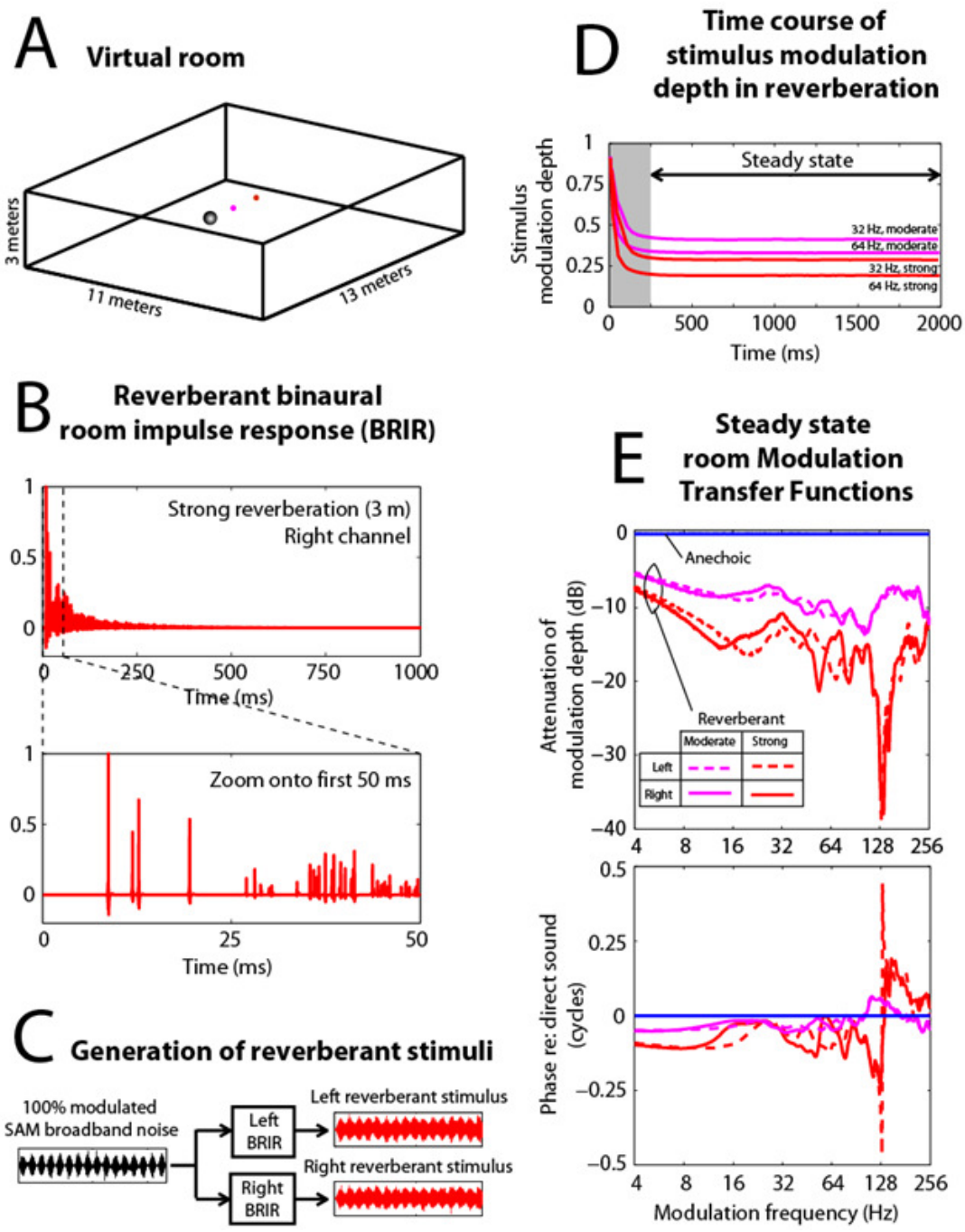

Figure 2.1 Virtual Auditory Space Stimuli

A. Geometry of the virtual room. Reverberant Binaural Room Impulse Responses (BRIR) were simulated using the image method for a source positioned at 0º azimuth either 1.5 m ("moderate" reverberation, magenta color) or 3 m ("strong" reverberation, red color) away from the simulated spherical head (grey color). **B.** Example BRIR (Right channel, strong reverberation). The direct sound is rapidly followed by reflections with approximately exponential energy decay. **C.** Reverberant stimuli were obtained by convolution of 100% modulated SAM broadband noise with the simulated BRIR. **D.** Stimulus modulation depth in reverberation sharply decreases from a high value before reaching a plateau by 250 ms. **E.** Room Modulation Transfer Functions (MTF) describe the degradation of amplitude modulation between a source and a receiver due to reverberation. MTFs were computed in the steady state part of the reverberant stimuli (Methods).

Phase-locking to f_m was a clear feature of firing patterns in this neuron: Repeated firing around a preferred phase of the stimulus is visible in the anechoic spike raster at low frequencies (Fig. 2.2 A, left panel). Period histogram (40 bins/period) were constructed for each f_m (Fig. 2.2 B) to better visualize temporal coding of AM. For frequencies below 180 Hz, the anechoic period histograms have a modulation depth of 1 or higher. Plotting the RMD as a function of f_m to form a tMTF (Fig. 2.2 C, middle panel, in blue) reveals a band-pass shape, with a best temporal modulation frequency (tBMF) near 90 Hz. The reverberant period histogram was less modulated than the anechoic one (Fig. 2.2 B, right panel) with modulation depths of 1 or greater only between 16 and 64 Hz. The reverberant tMTF (Fig. 2.2 C, middle panel, in red) also has a band-pass shape, but with a decreased RMD at all frequencies relative to the anechoic condition.

The mean response modulation phases were computed relative to the phase of the direct sound at the receivers (after removing the mean response latency: see Methods) and plotted as a function of f_m to form a pMTF (Fig. 2.2 C, bottom panel). In this neuron, the mean neural delay was rather large (≈18 ms), but reverberation had little effect on phase, even at frequencies for which the RMD and the firing rate were substantially altered (e.g. 90 Hz). The nearly flat phase profile for $f_m \leq 64$ Hz is consistent with an 18 ms neural delay at these frequencies.

Figure 2.3 A-C shows data from three additional example neurons illustrating the diversity of neural MTFs. The shape of the anechoic rMTF differs between the three cases: In A, the rMTF is low-pass; In B it is band-pass with high firing rates and a best rate modulation frequency – rBMF – at 90 Hz; In C, the rMTF is band-pass with lower firing rate and an rBMF near 45 Hz. However, for all three cases, the effect of reverberation is to flatten the rMTF.

Anechoic tMTFs also had a variety of shapes (middle panels): Sharp band-pass with very poor phase-locking at high frequencies in A, essentially all-pass in B, broad band-pass in C. In

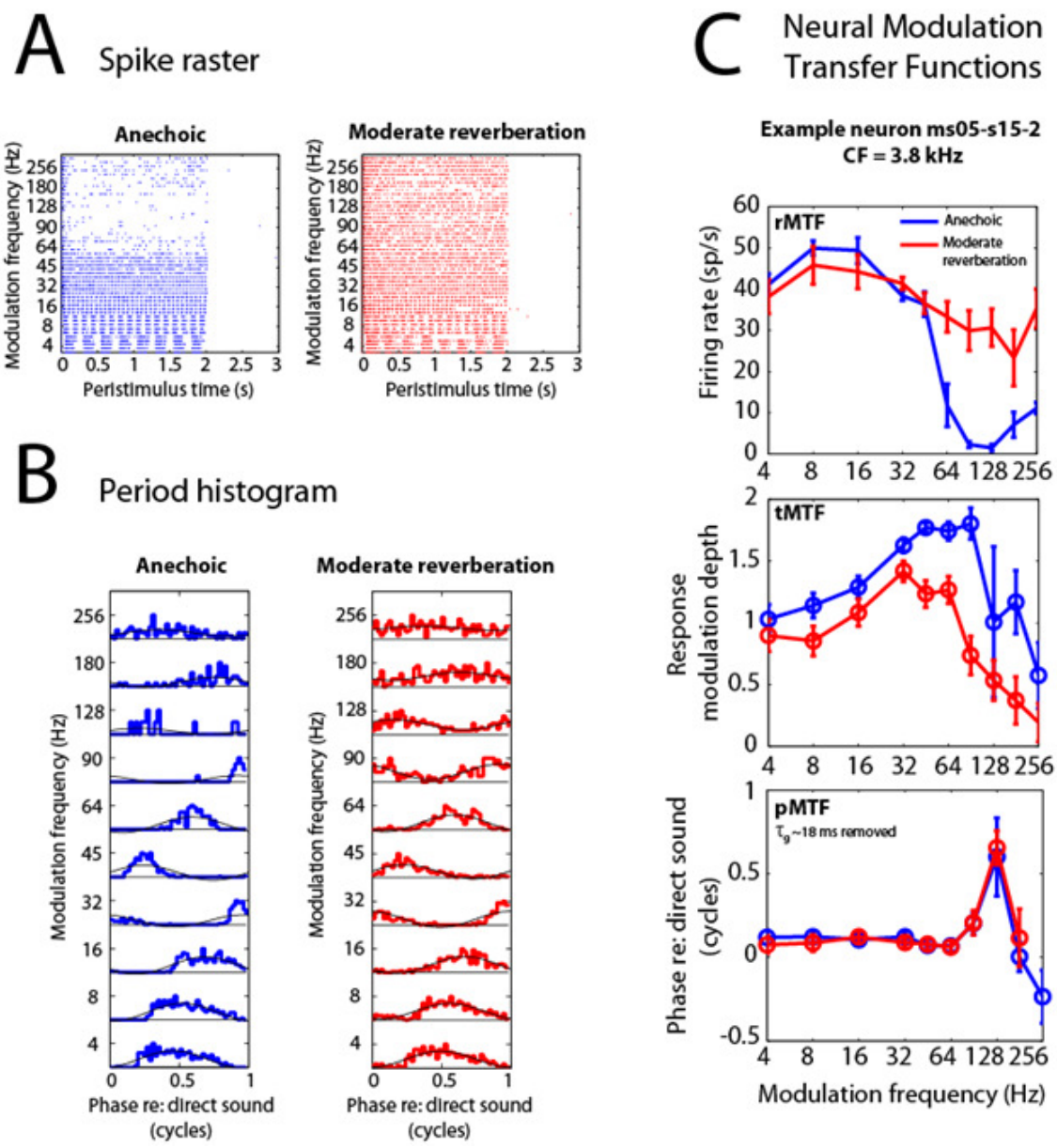

Figure 2.2 Example neural responses to virtual auditory space stimuli
A. Spike raster for neuron ms05-s15-2. Each modulation frequency was presented 5 times. Presentations were random across f_m, reverberation condition, and repetition. *Left panel (blue):* Anechoic case. *Right panel (red):* Moderate reverberation. **B.** Period histogram constructed for the same neuron (normalized units), with 40 bins/period. Fourier components were superimposed as sinusoids to represent response modulation depths (RMD). **C.** Neural Modulation Transfer Functions (MTF) constructed from the same data set. Anechoic MTFs are drawn in blue, reverberant MTFs in red. *Top panel:* Rate MTFs (mean firing rate ± 2 standard errors). *Middle panel:* Temporal MTFs (RMD ± 2 standard deviations). *Bottom panel:* Phase MTFs (response modulation phase ± 2 standard deviations). The standard deviations of the tMTFs and pMTFs were computed assuming a von Mises distribution (see Methods). RMDs that were significant (Rayleigh test of uniformity, $\alpha<0.05$) are marked with an open circle in the tMTFs and pMTFs. Phase data points for which RMD was not significant were omitted. Phases were unwrapped assuming linear phase, and a mean neural delay τ_g = 18 ms was removed to facilitate visualization (see Methods).

all three cases, reverberation decreased phase-locking to f_m, but the neural degradation was frequency dependent, resulting in very different reverberant tMTFs: All-pass in A with very low RMD, high-pass in B, and sharp band-pass tMTF in C with a tBMF near 32 Hz.

The anechoic pMTFs measured in these three neurons showed relatively large group delays (17 ms in A, 14 ms in B, and 12 ms in C), but effects of reverberation differed (Fig. 2.3 A-C, bottom panels): In A, the reverberant responses had a small phase advance at some frequencies, while neuron C showed a consistent phase delay in reverberation between 16 and 90 Hz; the phase in neuron B was not very affected by reverberation.

The best modulation frequencies of these three neurons in the anechoic condition were representative of the range we encountered (rBMFs were 8 Hz in A, 90 Hz in B, and 45 Hz in C, while tBMFs were 45 Hz in A and B, and 90 Hz in C), although among these examples, only example A had an anechoic tMTF with a clear maximum. The distribution of BMFs in Figure 2.3 D shows our broad sampling of IC neurons: Anechoic rBMFs ranged from 4 to 256 Hz with a median at 32 Hz, and anechoic tBMFs ranged from 4 to 180 Hz, with a median at 45 Hz. There was little or no correlation between rBMF and tBMF ($r=0.13$, $p=0.19$).

We used simple coding strength metrics to summarize the effects of reverberation across our data set. For rate coding, we computed a signal-to-noise ratio (SNR) metric of the rMTF, based on ANOVA (Methods), and expressed in dB. The larger the variance in firing rates across frequency (signal), the larger the SNR; The larger the variance across repetitions (noise), the lower the SNR. Reverberation significantly degraded the SNR across the population (Fig. 2.4 A-B). A paired Student's t-test performed on the dB values between the anechoic and reverberant conditions was highly significant ($p<0.001$). The mean degradation in SNR across the population was ≈7.5 dB when data from both strong and moderate reverberation were combined.

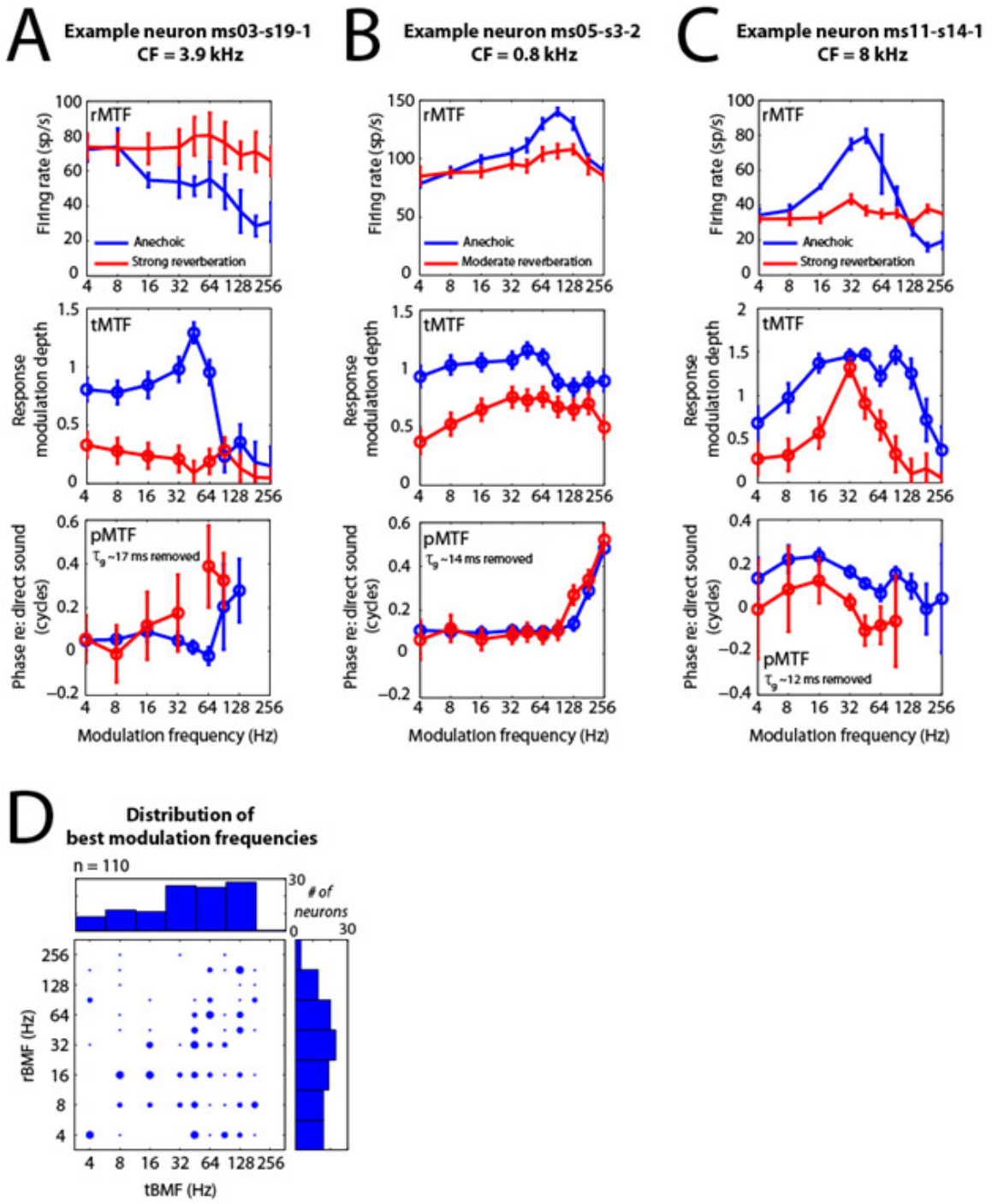

Figure 2.3 Diversity of neural MTFs

A-C. Neural MTFs (anechoic in blue, reverberant in red) for 3 example neurons. *Top panels:* Rate MTFs (mean firing rate ± 2 standard errors). *Middle panels:* Temporal MTFs (RMD ± 2 standard deviations). *Bottom panels:* Phase MTFs (response modulation phase ± 2 standard deviations). Open circles for tMTF and pMTF correspond to significant RMDs (Rayleigh test of uniformity, $\alpha<0.05$). Phases were not plotted when RMDs were not significant. The mean neural delay τ_g was removed. **D.** Scatter plot and histograms of the best rate and temporal modulation frequencies (rBMFs and tBMFs) of the population of anechoic neural MTFs.

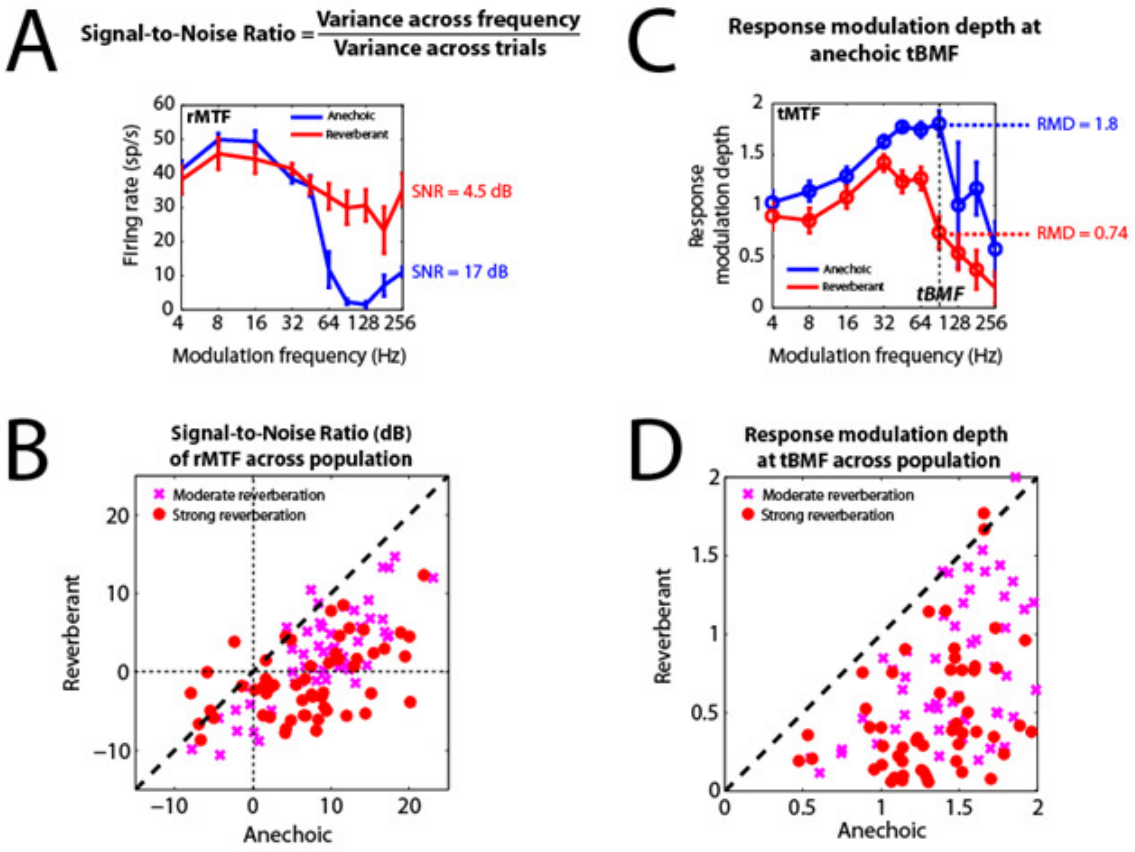

Figure 2.4 Effects of reverberation on the strength of envelope frequency representation
A. Strength of envelope frequency representation in the rMTF was quantified by a signal-to-noise ratio (SNR) metric based on ANOVA (Methods). **B.** SNR was clearly degraded in reverberation across the population (bottom panel). **C.** Strength of envelope frequency representation in the tMTF was quantified by the RMD at the best modulation frequency of the anechoic tMTF (tBMF). **D.** Temporal coding was clearly degraded across the population (bottom panel).

The mean degradation in strong reverberation (≈8 dB) was not significantly larger than that in moderate reverberation (≈6.5 dB) ($p=0.1$, two-sample t-test).

For temporal coding, we compared the anechoic and reverberant RMDs at the anechoic tBMF (Fig. 2.4 C). Degradation of temporal coding due to reverberation was highly significant ($p<0.001$, paired t-test) across the population (Fig. 2.4 D). Expressed in dB, the mean degradation at tBMF was ≈9 dB. The mean degradation in strong reverberation (≈11.5 dB) was significantly larger than that in moderate reverberation (≈7.5 dB) ($p<0.001$, two-sample t-test).

Degradation of temporal coding was smaller than acoustic degradation

We compared the degradation of temporal coding due to reverberation to the degradation in modulation depth in the acoustic stimuli (Fig. 2.5). An example of this comparison is shown for one neuron in Figure 2.5 A. The left panel represents the tMTF and pMTF in anechoic (blue) and reverberant (red) conditions for this neuron. In this neuron, RMD decreased in reverberation at every tested frequency. Response phases were very similar between anechoic and reverberant conditions. The differences (in dB) between reverberant and anechoic tMTFs and pMTFs are shown in red in the center panels. The differences between tMTFs quantify the degradation in temporal coding due to reverberation. In this neuron, the degradation was largest at low frequencies, and ranged between ≈2 and ≈8 dB. In contrast, the degradation in modulation of the acoustic signals, i.e. the magnitude of the room MTF (Methods, Fig. 2.1 E) was more pronounced, with values ranging between ≈5 and ≈12 dB (green line in Fig. 2.5 A, center panel) and had a different frequency dependence than the neural degradation. The difference between neural and acoustic degradation is plotted on the right panel, and reveals that at most frequencies, there was a net gain in modulation depth in the neural response, of the order of 7 dB at the tBMF (Methods: Eq. 2.5). This "neural compensation gain" partly counteracted the acoustic degradation due to reverberation.

The difference between reverberant and anechoic pMTFs (Fig. 2.5 A) is plotted in red in the middle panel, and represents the change in response phase introduced by reverberation. The phase shift was small at most frequencies (<0.05 cycles), except for a 0.14 cycle phase lead of the reverberant response at 128 Hz. The dependence of this phase shift on f_m was similar to that of the room MTF phase (in green). The difference between neural and acoustic phase shifts provided the phase of the neural compensation (Fig. 2.5 A, right panel and Methods, Eq.

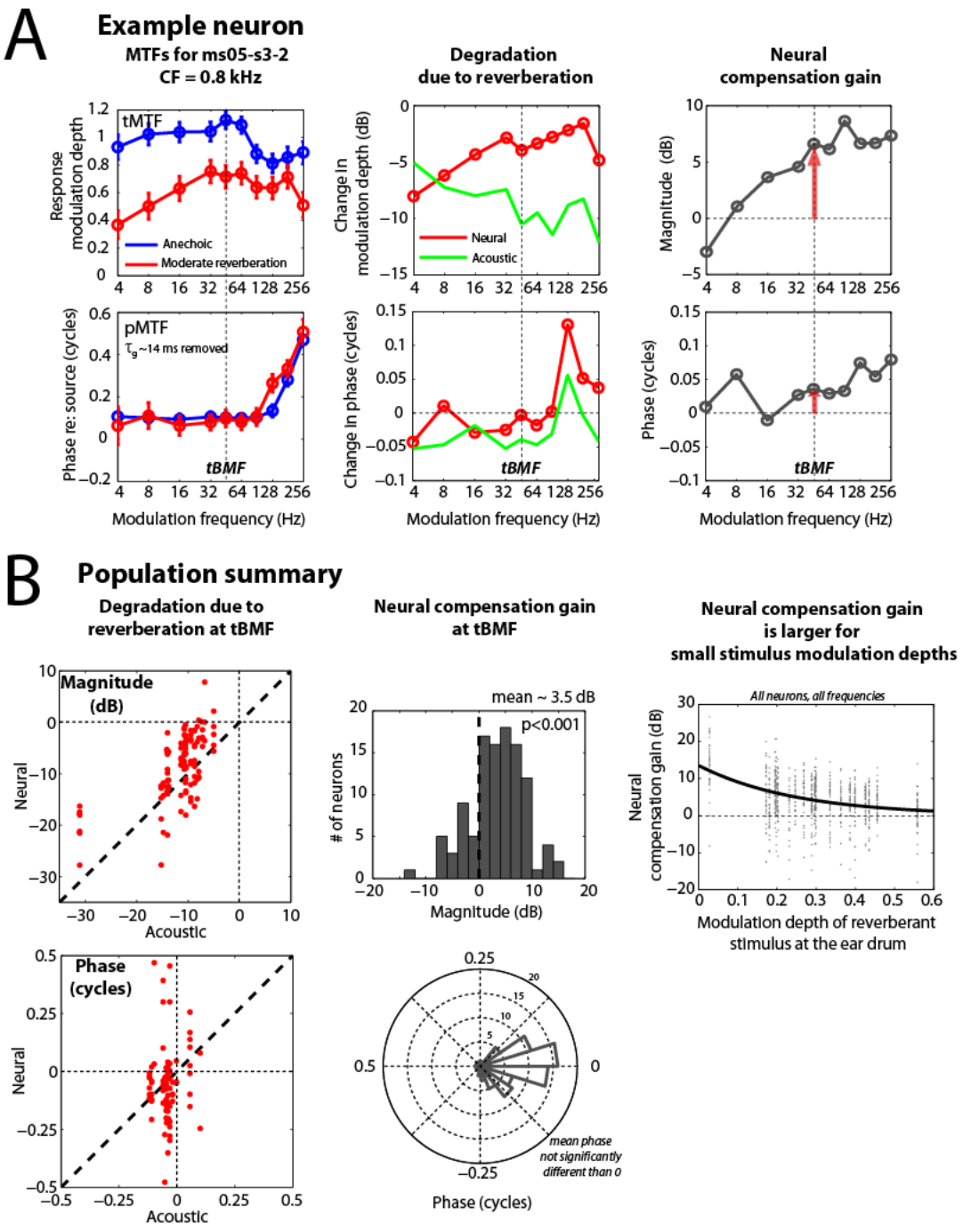

Figure 2.5 Partial compensation for the acoustic degradation in temporal coding
A. Example neuron (ms05-s3-2). *Left panel:* tMTF and pMTF (anechoic in blue, reverberant in red). *Center panel:* Degradation due to reverberation (neural in red, i.e. change in tMTF and pMTF between anechoic and reverberant responses; acoustic in green, i.e. room MTF). *Right panel:* Neural compensation gain (difference between neural degradation and acoustic degradation from center panel). **B.** Population summary. *Left panel:* Comparison between acoustic and neural degradation due to reverberation at tBMF. *Center panel:* Neural compensation gain at tBMF (magnitude and phase). *Right panel:* Dependence of neural compensation gain on stimulus modulation depth at the ear drum. Thick black line: Decaying exponential fit.

2.6). It represents the phase shift between anechoic and reverberant neural responses that is not accounted for by the phase shift between anechoic and reverberant stimuli.

Comparisons between neural and acoustic degradations are summarized across the population in Fig. 2.5 B. The neural compensation gain at each neuron's anechoic tBMF was on average significantly larger than 0 dB ($p<0.001$, paired t-test) with a mean of roughly 3.5 dB, suggesting that the degradation in stimulus modulation depth due to reverberation was partially compensated at the level of the IC. The phase of the neural compensation did not significantly differ from 0 ($p>0.05$, single-sample test assuming a von Mises distribution; p. 123 in Mardia and Jupp, 1999), suggesting that on average, the room MTF accounted for the phase shifts observed between anechoic and reverberant neural responses.

We plotted the neural compensation gain from all neurons and modulation frequencies as a function of the modulation depth of the reverberant stimulus at the ear drum – or proximal modulation depth (Fig. 2.5 B, rightmost panel). Although there were large variations in gain for each proximal modulation depth, the gain clearly decreased with increasing proximal modulation depth: Gains had a moderate ($r=0.37$) but highly significant ($p<0.001$) correlation coefficient with the fit to a decaying exponential.

Neural compensation gain was largely explained by the gain of the Modulation Input/Output Function (MIOF)

The Modulation Input/Output Function (MIOF) describes the nonlinear transformation of anechoic modulation depths into neural modulations. The decreasing trend of the neural compensation gain with increasing proximal modulation depth is reminiscent of the dependence of the modulation gain of the Modulation Input/Output Function (MIOF) on stimulus modulation

depth (see e.g. Joris *et al.*, 2004). To test the hypothesis that the neural compensation gain is explained by the modulation gain of the MIOF, we measured MIOFs in response to anechoic SAM broadband noise with modulation depths between 0 and 1 (Methods) in 92 well-isolated IC single-units. Neural MTFs for both anechoic and reverberant stimuli were also obtained in 44 of these units. MIOFs were usually measured at one f_m, chosen to elicit both a large firing rate and strong phase-locking to the modulation. Figure 2.6 A shows an example MIOF, measured for a modulation frequency of 64 Hz. In this example, RMD increased steeply at low stimulus modulation depths, and plateaued at higher modulation depths. Assuming that RMD for both anechoic and reverberant stimuli only depends on the proximal modulation depth (i.e. the modulation depth of the stimulus at the ear drum), we can predict the reverberant RMD from the MIOF and the acoustic MTF. In this example, the proximal modulation depth of the reverberant stimulus was ≈0.3, which, based on the MIOF in this neuron should elicit an RMD of ≈0.87 (green dot in Fig. 2.6 A). The measured reverberant RMD (≈0.93) was not significantly different from the prediction (red dot) ($p<0.05$; test of equality of concentration parameters assuming a von Mises distribution, p. 133 in Mardia and Jupp, 1999).

RMDs in this example MIOF were always larger than stimulus modulation depths, revealing a net gain of amplitude modulation at the level of the IC. This gain is defined by $G_{dB,MIOF}(m) = 20\log_{10}\left(\frac{MIOF(m)}{m}\right)$, with m the proximal modulation depth. To test whether the neural compensation gain observed in the responses to reverberant stimuli could be accounted for by the gain of the MIOF, we defined the predicted neural compensation gain as the difference between $G_{dB,MIOF}(m_{rev})$, with m_{rev} the proximal modulation depth of the reverberant stimulus ($m_{rev} < 1$), and $G_{dB,MIOF}(1)$, since the modulation depth of the anechoic stimulus is

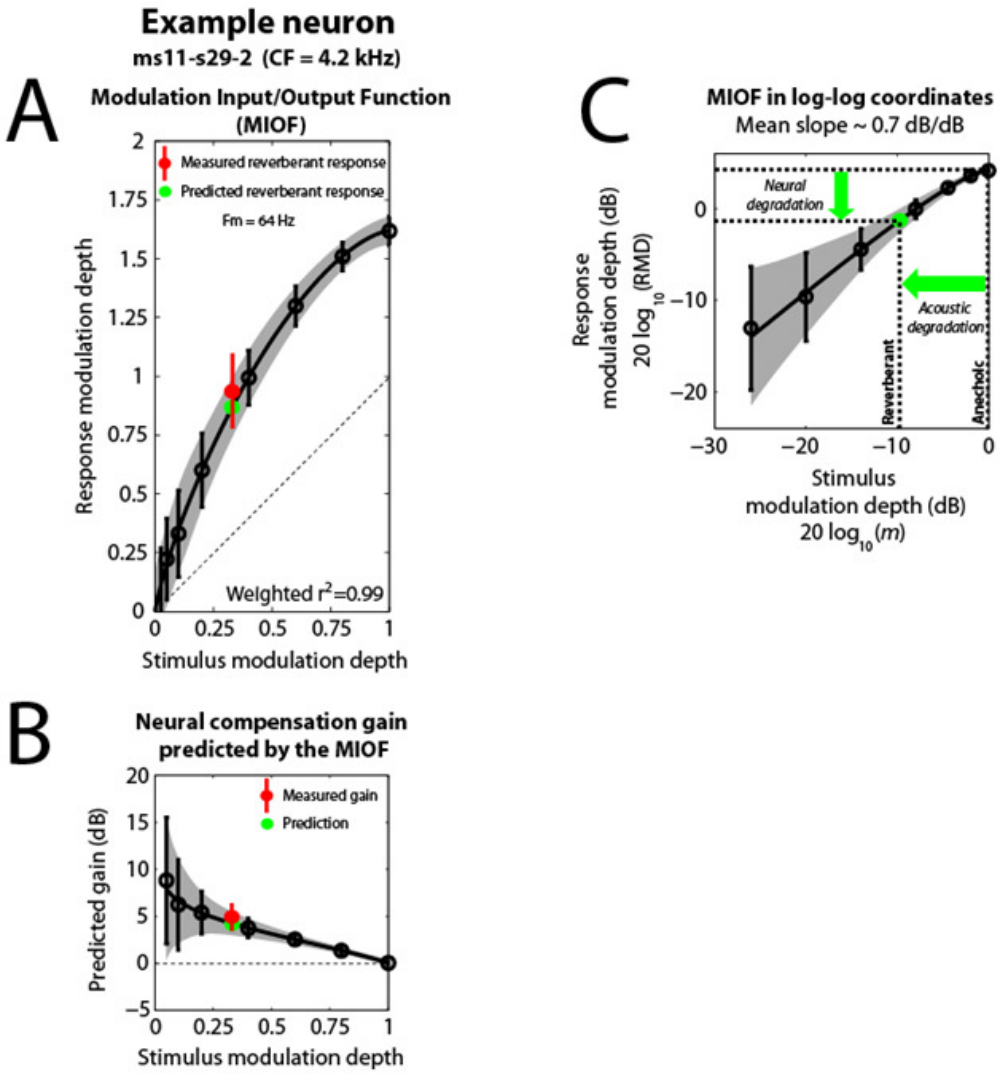

Figure 2.6 Prediction of the neural compensation gain from the Modulation Input/Output Function in an example

Example neuron (ms11-s29-2). **A.** Modulation Input/Output Function (MIOF) showing RMD (mean ± 2 standard deviations) as a function of stimulus modulation depth. MIOFs were measured in response to SAM broadband noises with modulation depths between 0 and 1 (Methods). Open circles: Significant RMDs (Rayleigh test of uniformity, $\alpha<0.05$). Thick black line: Weighted least-square fit with an incomplete beta function (Methods, Eq. 2.7). Green dot: Prediction of reverberant modulation depth from the MIOF. Red dot and error bars: Measured reverberant modulation depth. Dashed line: Identity line. **B.** Predicted neural compensation gain from the MIOF as a function of stimulus modulation depth (mean ± 2 standard deviations; Methods, Eq. 2.8, 2.9 & 2.10). Only gains computed from a significant RMD were displayed. Green dot: Prediction of neural compensation gain obtained at the modulation depth of the reverberant stimulus. Red dot and error bars: Measured neural compensation gain. **C.** MIOF in log-log coordinates. Neural compensation gain arises from the compressive shape of the MIOF.

always 1 (Eq. 2.9, Methods and Fig. 2.6 B). In this example, the predicted compensation gain decreased with increasing proximal modulation depth, and accounted for the measured compensation gain in reverberation.

In this neuron, the positive value of the predicted compensation gain can be explained by the compressive shape of the MIOF. In log-log coordinates (Fig. 2.6 C), the MIOF was roughly linear, with slope ≈ 0.7 dB/dB. Because the slope is less than 1 dB/dB, the neural degradation (green vertical arrow) is smaller than the acoustic degradation (green horizontal arrow), resulting in partial neural compensation.

Across the population, the MIOF accounted for much of the compensation gain, in that predictions from the MIOF were highly correlated with the measured RMD in reverberation (Fig. 2.8 B, $r = 0.9$, $p<0.001$). Most MIOFs had a compressive shape, with mean slopes in log-log coordinates typically < 1 dB/dB (Fig. 2.7 A). The mean slope across the population (≈ 0.85 dB/dB) was significantly smaller than 1 ($p=0.001$; paired t-test). The decreasing dependence of predicted gain on proximal stimulus modulation depth (Fig. 2.7 B) was similar to that of the measured compensation gain (Fig. 2.5 B, rightmost panel). This dependence also arises from the compressive shape of the MIOFs.

We partitioned the population of MIOFs into 3 clusters using a k-means clustering algorithm (Fig. 2.7 C). We chose the number of clusters by computing the Within Cluster Sum of Squares (WCSS) as a function of the number of clusters (Fig. 2.7 C: inset; Methods). WCSS first decreased as a function of number of clusters, and then plateaued. We chose to partition our data into 3 clusters because the incremental decay of WCSS fell below 5% of the total WCSS with more than 3 clusters.

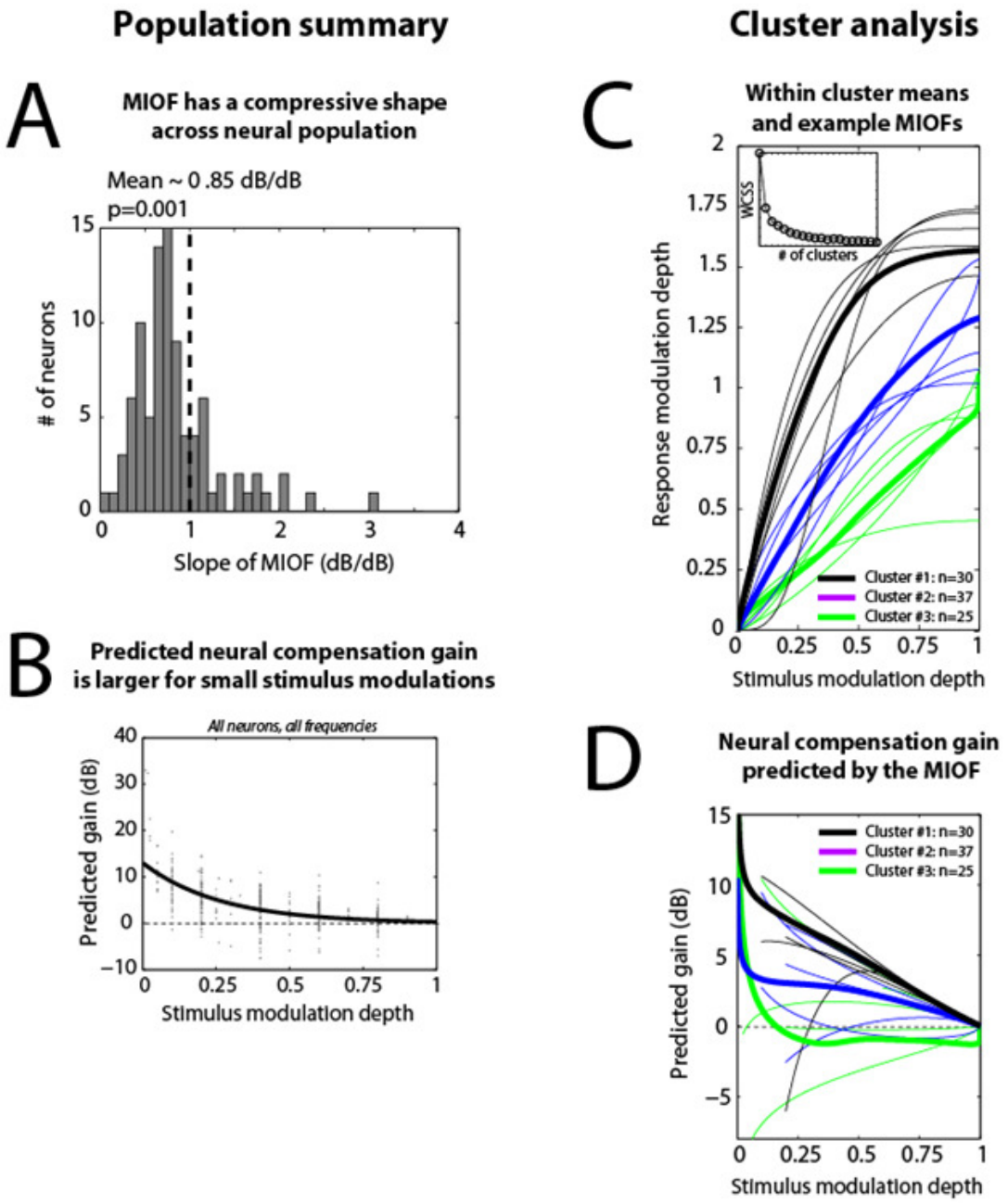

Figure 2.7 Modulation Input/Output Functions and predictions of the neural compensation gain across the population

A. Histogram of the mean slopes of the MIOF in log-log coordinates across the population. Mean slope was significantly smaller than 1, consistent with the MIOF being compressive on average. **B.** Dependence of the predicted compensation gain on stimulus modulation depth. Only gains computed from a significant RMD were displayed. Thick black line: Decaying exponential fit. **C.** k-means clustering of the population of MIOFs (3 clusters). Thick lines: Within cluster means. Thin lines: Randomly selected examples (5 examples per cluster). Inset: Within Cluster Sum of Squares as a function of the number of clusters (Methods). **D.** Predicted compensation gains from the means of each MIOF cluster (thick lines) and from the example MIOFs (thin lines) in C. Example gains were truncated where RMD was not significant.

The first cluster (black lines) included 30 MIOFs, and was characterized by high modulation gains, steep slopes for low stimulus modulation depths, and a plateau at high stimulus modulation depths. The second cluster (blue lines) included the largest number of MIOFs (37). The response modulation gains were smaller than for the first cluster, and the dependence on stimulus modulation depth was less compressive. Finally, the third cluster (green lines, 25 MIOFs) was characterized by an almost linear growth, with the lowest gain among the three clusters.

We computed the predicted compensation gain of the mean MIOF for each cluster (Fig. 2.7 D). The first cluster (black lines) had the largest gains (≈0-10 dB at most stimulus modulation depths), while the second cluster (blue lines) had moderate gains (≈0-3 dB). The third cluster (green lines) had the highest variability in predicted gains among the three clusters, but on average had negative gains (of the order of -1 to -2 dB) for stimulus modulation depths above 0.2. For most neurons in clusters 1 and 2, the predicted gains decreased as a function of stimulus modulation depth, although at different rates. This was not the case for cluster 3.

In some neurons, temporal coding in reverberation was more robust than predicted by the MIOF

Although reverberant RMD and predictions from the MIOF were highly correlated[1] as expected (Fig. 2.8 B), reverberant responses were, on average, slightly more robust than the predictions (paired Student's t-test, $p=0.05$). In some neurons (open circles in Fig. 2.8 B), the reverberant RMD differed significantly from the prediction ($p<0.05$; test of equality of

[1] We compared reverberant RMD and predictions from the MIOF only in the cases when reverberant RMD was significant with a Rayleigh test of uniformity ($\alpha<0.05$).

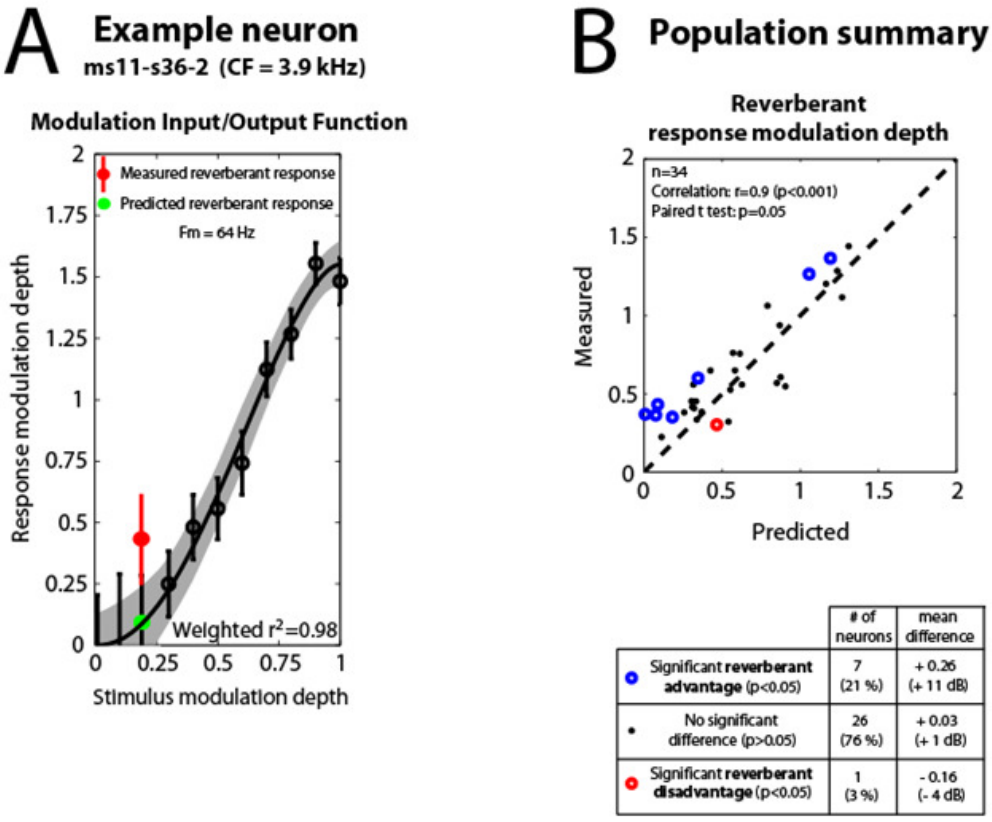

Figure 2.8 In a subset of neurons, temporal coding in reverberation was more robust than predicted by the MIOF

A. Example neuron (ms11-s36-2). Input/Output Modulation Function (MIOF) showing RMD (mean ± 2 standard deviations) as a function of stimulus modulation depth. In this example, the measured reverberant response (red dot) was significantly greater than the prediction from the MIOF ($p<0.05$; test of equality of concentration parameters, p. 133 in Mardia and Jupp, 1999). **B.** Population summary. *Top panel:* Comparison between measured reverberant RMD and prediction from the MIOF. Points for which reverberant RMD was not significant (Rayleigh test of uniformity, $\alpha<0.05$) were omitted. Measurements and predictions were strongly correlated ($r=0.9$, $p<0.001$). However, 8 responses (out of 34) were significantly different than the prediction. *Bottom table:* Breakdown of the differences found between reverberant RMD and predictions. Reverberant stimuli had a significant coding advantage in 7 neurons, a significant coding disadvantage in 1 neuron, and did not elicit a significantly different RMD than the prediction in 26 neurons.

concentration parameters, p. 133 in Mardia and Jupp, 1999). Specifically, 7 of 34 neurons (21%) showed a significant coding advantage for reverberant stimuli, compared to anechoic stimuli with the same modulation depth (see example in Fig. 2.8 A). In this subset of neurons, the mean difference (ratio) between the reverberant RMD and the prediction was 0.26 (11 dB). In one

case, the reverberant stimulus had a significant coding disadvantage over the prediction (difference was roughly -0.16 or -4 dB).

Discussion

Our neurophysiological experiments show that reverberation degrades rate and temporal coding of modulation frequency in IC neurons of unanesthetized rabbits. In most neurons, the degradation in temporal coding is smaller than the degradation in the stimulus, and this compensation gain is largely explained by the compressive shape of the modulation input-output function (MIOF) describing the transformation of stimulus modulation depths into neural modulations. However, in a subset of neurons, the prediction from the MIOF underestimated the modulation depth of the reverberant response, suggesting that, in these neurons, reverberant stimuli had a coding advantage over anechoic stimuli with the same modulation depth.

Are Virtual Auditory Space Stimuli sufficiently realistic?

We used Virtual Auditory Space techniques to study the effects of reverberation on the coding of sound envelope in unanesthetized rabbits. Our stimuli were the convolution products of SAM broadband noise (source) and BRIRs simulated using the image method (Allen and Berkley, 1979; Shinn-Cunningham *et al.*, 2001). These techniques are convenient because they allow for a complete control over the reverberation parameters, and tremendously simplify the study of neurophysiological effects of reverberation.

The image method is widely used in a variety of fields (from psychophysics to architectural acoustics) to simulate room acoustics, and several studies have specifically assessed the resemblance of simulations to actual measurements in rooms or theoretical predictions. Allen and Berkley (1979) showed that simulations based on the image method accurately predicted theoretical rms deviations of sound pressure levels from mean pressure for different room dimensions and a variety of D/R ratios. Shinn-Cunningham *et al.* (2001) compared

acoustic features of simulated BRIRs to measurements made with a manikin in a small room for a variety of source and receiver positions. The long term spectra were similar between measurements and predictions. Comb-filtering was visible in both cases, probably due to the early, well resolved reflections[1]. Distortion of interaural time differences (ITD) were also qualitatively similar. A limitation of the model was a small error in predicting reverberant energy for small (0.15 – 1 m) source-receiver distances (likely due to not including reflections from the head itself), but this limitation is not relevant to our simulations as the source-receiver distance was larger (1.5 and 3 m).

Zahorik (2009) investigated the perceptual resemblance of speech samples convolved with simulated and measured BRIRs. Simulation parameters for the virtual room were matched to the characteristics of the rooms in which the BRIR measurements were carried out. Only small perceptual differences were found between modeled and measured rooms.

Overall, these studies support the use of the room-image method as a reasonable approximation for medium-size room acoustics. However, a possible caveat of our study is the use of a rigid sphere as an approximation for a rabbit head. Kim *et al.* (2010) showed that there are important differences at high frequency between the Head Related Transfer Functions (HRTFs) measured in rabbits and HRTFs simulated with a spherical head model. Across source and receiver positions, the range of interaural time differences predicted by the spherical model was similar to that experienced by rabbits, but the range of interaural level differences (ILDs) was underestimated by the model. However, for a source positioned directly in front of the receivers, similar to our simulations, both measured and predicted ILDs were close to 0.

---

[1] In our simulations, room MTFs for the strong reverberation (Fig. 2.1 E) also showed evidence of comb-filtering from the early, discrete reflections, with, for example, a large notch at 128 Hz, which disappeared after we removed the first two reflections in the BRIR.

Comparison with other studies of AM coding in IC

Our measurements of anechoic neural MTFs and MIOFs are in line with previous studies of AM coding in the IC. In particular, Nelson and Carney (2007) measured rMTFs, tMTFs, and MIOFs in response to SAM tones in unanesthetized rabbits. Although we used SAM broadband noise rather than the SAM tones at the neuron's CF used in Nelson and Carney (2007), the range of best modulation frequencies (both rBMF and tBMF) were similar. In both studies, rBMFs and tBMFs were distributed mostly between 16 and 128 Hz. Moreover, the MTF shapes we encountered roughly matched their description: rMTFs were mostly band-pass or low-pass, while tMTFs were mostly band-pass.

Our MIOFs are also consistent with previous studies in the IC. These studies reported monotonic increases in RMD with stimulus modulation depth, significant phase-locking to f_m for stimulus modulation depths smaller than 0.1, and saturation of the MIOF in most neurons at high modulation depths (e.g. Rees and Møller, 1983; Krishna and Semple, 2000; Nelson and Carney, 2007). Our study identifies three clusters of MIOFs (Fig. 2.7 C): One cluster (33% of the population) is characterized by large modulation gains and saturation of RMD, especially at low stimulus modulation depth; The second cluster (40%) does not saturate as much and has lower modulation gains; The third cluster (27%) does not saturate and usually has modulation gains less than 0 dB. We did not find any correlation between cluster and characteristic frequency, nor between cluster and modulation frequency. However, the absence of relationship between cluster and modulation frequency may be due to our method for choosing modulation frequencies: We usually measured the MIOF at a modulation frequency that elicited strong phase locking to the fully modulated stimuli, therefore favoring clusters 1 and 2.

It is interesting to note that the slopes of the mean MIOF for clusters 1 and 2 are largest in the lower range of stimulus modulation depths: The largest slope was 0.84 dB/dB at $m \approx 0.08$ for cluster 1 and 0.97 dB/dB at $m \approx 0.22$ for cluster 2. This may help listening in noisy and/or reverberant situations by improving sensitivity to changes in modulation depth in the range where signal envelopes are degraded and modulation depths are typically lower than in quiet, anechoic environments. We can further interpret the role of the 3 clusters as optimally encoding different ranges of modulation depth: Neurons in cluster 1 are mostly sensitive to changes in the lowest stimulus modulation depths, whereas neurons in cluster 2 are mostly sensitive to a midrange near 0.25, and neurons in cluster 3 provide better differential sensitivity at higher modulation depths. Although MIOFs from each clusters were grouped around their means (see the 5 randomly picked examples for each cluster in Fig. 2.7 C), cluster 3 displayed more variability, and the point of maximum slope of the mean MIOF in cluster 3 near 0.45 is not representative of all individual neurons, although maximum sensitivity in individual neurons of cluster 3 tended to occur near or above 0.45.

Shape of the MIOF at subcollicular levels

The observation that the MIOFs are usually compressive in the IC begs the question of the origin of this compression. Is it created at the level of the IC, or inherited, at least partially, from subcollicular stages? A limited number of studies have investigated temporal coding of AM as a function of *m* at subcollicular levels. In the auditory nerve (AN) of anesthetized cat, the MIOFs increase monotonically with *m* (Joris and Yin, 1992), and are compressive with positive modulation gains in at least a subset of neurons. The maximum RMD tends to be larger for high spontaneous rate fibers than for low spontaneous rate fibers, although it is not clear if the shape

of the MIOF is correlated to spontaneous rate as well. In the ventral cochlear nucleus (VCN) of anesthetized cat, a large proportion of MIOFs seems to be compressive (Rhode, 1994). The modulation gain and shape of the MIOF depends on cell type: The MIOF shape for primary like and chopper units displays a lot of variability (from linear with negative gains to highly compressive with large gains), while MIOFs from onset units are more consistently and more strongly compressive. In the superior olivary complex (SOC) of unanesthetized rabbit, MIOF shapes are correlated to properties of pure tone responses (Kuwada and Batra, 1999): Neurons with offset responses to pure tones have highly compressive MIOFs with small dynamic ranges, while neurons with sustained responses to pure tones tend to have less compressive MIOFs, more similar to the AN data of Joris and Yin (1992). Overall, the picture emerging from the literature supports the view that the compressive shape of the MIOF originates at subcollicular stages of the auditory system, and that the MIOF tends to be less compressive in the more peripheral stages.

Importance of the MIOF in compensating for the degradation of AM in reverberation

The degradation in RMD due to reverberation was on average smaller than the degradation in the envelope of the stimuli (Fig. 2.5). We defined the difference between neural and acoustic degradations as a "neural compensation gain", and showed that this neural compensation gain was, in general, well predicted by the MIOF (Fig. 2.6; Fig. 2.8 B). The phase of the compensation was on average not significantly different from 0, which is also consistent with a lack of systematic phase changes with stimulus modulation depth that we observed (not shown; see also Nelson and Carney, 2007).

Our argument that the gain of the MIOF, $G_{dB,MIOF}(m) = 20\log_{10}\left(\frac{MIOF(m)}{m}\right)$, can act as a compensation mechanism could be criticized in that the neural compensation gain was only measured for a source that was 100% modulated. The emergence of a positive compensation gain could be exaggerated from using this high source modulation depth. Since the compensation gain is largely dependent on the shape of the MIOF, the compensation gain might simply result from the saturation of the MIOF at high modulation depths. In fact, $G_{dB,MIOF}(m)$ can be large for small m, but is bounded for m = 1 by a theoretical maximum of 6 dB, as RMD cannot exceed 2. As a result, the predicted compensation gain $G_{dB,predicted}(m) = G_{dB,MIOF}(m) - G_{dB,MIOF}(1)$, with m the proximal modulation depth of the reverberant stimulus, is almost guaranteed to be positive.

However, the compressive shape of the MIOF in most neurons, especially those in clusters 1 and 2, ensures that this compensation will also occur for partially modulated stimuli at the source. As detailed in the Appendix, the compressive shape of the MIOF implies that the MIOF gain $G_{dB,MIOF}(m)$ is a decreasing function of stimulus modulation depth. Therefore, if we call m_d the modulation depth at the source (or distal modulation depth), and m_p the modulation depth at the ear drum (or proximal modulation depth, with $m_p < m_d$ due to reverberation), a compressive MIOF necessarily leads to $G_{dB,MIOF}(m_p) > G_{dB,MIOF}(m_d)$. As a result, there should still be a positive compensation gain $G_{dB,MIOF}(m_p) - G_{dB,MIOF}(m_d)$ even in the case of a partially modulated stimulus at the source.

Significance of a reverberant advantage in the steady state response

Since the modulation depth of our reverberant stimuli sharply decreases over time to reach a plateau by 250 ms (Fig. 2.1 D), the comparison between the reverberant response in the steady state portion and the prediction from the MIOF (Fig. 2.6 A; Fig. 2.8 A-B) addresses the following question: Do reverberant stimuli have a coding advantage over anechoic stimuli when we control for modulation depth at the ear drum? The MIOF predicts the response to an anechoic stimulus whose modulation depth matches that of the plateau in the reverberant stimulus. If the MIOF were a static nonlinearity invariant to other stimulus characteristics, steady state reverberant responses should never be significantly different than predictions from the MIOF.

Contrary to this prediction, we found a subset of neurons (Fig. 2.8 B) for which steady state reverberant responses were significantly more modulated than predicted by the MIOF. One possibility is that these differences are simply artifacts of our experimental procedures: The MIOF was usually measured after the MTFs, and although we are only analyzing data for well isolated single-units, it is possible that the operating range of the modulation response changed due to a variety of uncontrolled factors (e.g. small change in spike quality, effect of auditory attention). However, in experiments described in next chapter, we changed our experimental procedure to interleave the reverberant stimuli with an anechoic stimulus whose modulation depth matched that of the reverberant steady state, and we still found a significant reverberant advantage in a subset of neurons.

Another possible explanation for the observed differences between reverberant and depth-matched anechoic responses is an effect of cochlear band-pass filtering on the acoustic signals. To predict the reverberant RMD from the MIOF, we used stimulus modulation depths

computed from the broadband acoustic waveforms. However, the modulation depth for cochlea filtered reverberant stimuli can differ from the broadband modulation depth, especially for narrow bandwidths (e.g. Steeneken and Houtgast, 1980). To address this issue, we modeled the filtering effect of the cochlea by band-pass filtering our reverberant stimuli with gammatone filters of equivalent rectangular bandwidths similar to those of rabbit auditory nerve fiber tuning curves (Borg *et al.*, 1988). Acoustic MTFs were computed from the filtered stimuli, exactly as in the broadband case (Methods). At the characteristic frequencies we encountered (usually > 1 kHz), narrowband and broadband stimulus modulation depths were usually similar. Further, we used the narrowband stimulus modulation depths to predict the reverberant RMD from the MIOF. Across the population, the differences between reverberant RMD and broadband predictions were significantly larger than ($p=0.001$, paired t-test) and uncorrelated to ($r=-0.08$, $p=0.55$) the differences between narrowband and broadband predictions. This suggests that the coding advantage of reverberant stimuli over depth-matched anechoic stimuli is not due to the small differences we found between narrowband and broadband stimulus modulation depths.

Several other factors could potentially, alone or in combination, explain the differences we observed between reverberant and anechoic responses. First, the interaural cross-correlation (IACC) of the stimulus is different in the two cases (IACC = 1 in the anechoic condition, < 1 in the reverberant condition). Secondly, small interaural differences in modulation depth and envelope phase are present in the reverberant stimuli. Other features could also play a role, such as a coloration of the long-term spectrum magnitude of the reverberant sounds, or a small distortion in the reverberant envelopes. We investigate the influence of these various factors in the next chapter.

Overall, our results indicate that AM coding in the IC is degraded in reverberation, but that the degradation is less severe than that in the acoustic stimulus, suggesting that the early auditory system may possess mechanisms for reverberation compensation. In particular, the compressive nature of the nonlinearity describing the transformation of stimulus modulations into neural modulations contributes to this robustness.

Appendix

Reverberation degrades amplitude modulations between a sound source and the ear drum. The modulation depth at the ear drum (or proximal modulation depth m_p) is therefore smaller than the modulation depth at the sound source (the distal modulation depth m_d):

$m_p < m_d$, with their ratio being characterized by the room Modulation Transfer Function.

The Modulation Input/Output Function (MIOF) describes the nonlinear transformation of proximal modulation depth into neural modulations. Because MIOFs are monotonically increasing, the degradation in stimulus modulations results in a degradation of neural modulation:

$$MIOF(m_p) < MIOF(m_d)$$

Here, we demonstrate that a sufficient condition for the degradation in the neural modulations to be smaller than the degradation in the acoustic stimulus is that the MIOF has a compressive shape.

If the MIOF is compressive over the interval [0,1], we can write:

$$\forall (m_1, m_2) \in [0,1]^2, \forall t \in [0,1], \; MIOF(t \times m_1 + (1-t) \times m_2) \geq t \times MIOF(m_1) + (1-t) \times MIOF(m_2)$$

This means that the curve is always above any secant (Fig. 2.9 A). Since $m_p < m_d$, we have $\frac{m_p}{m_d} < 1$. Therefore, using $t = \frac{m_p}{m_d}$, $m_1 = m_d$, and $m_2 = 0$, we obtain:

$$MIOF\left(\frac{m_p}{m_d} \times m_d + \left(1 - \frac{m_p}{m_d}\right) \times 0\right) \geq \frac{m_p}{m_d} \times MIOF(m_d) + \left(1 - \frac{m_p}{m_d}\right) \times MIOF(0)$$

Since the MIOF is constrained to have a value of 0 for $m_p = 0$ (no response modulation for an unmodulated input), we can simplify the previous inequality as:

$$MIOF(m_p) \geq \frac{m_p}{m_d} \times MIOF(m_d)$$

Dividing left and right terms by m_p, we obtain:

$$\frac{MIOF(m_p)}{m_p} \geq \frac{MIOF(m_d)}{m_d}$$

Taking the logarithm, we get:

$$20\log_{10}\left(\frac{MIOF(m_p)}{m_p}\right) \geq 20\log_{10}\left(\frac{MIOF(m_d)}{m_d}\right)$$

$$\text{i.e. } G_{dB,MIOF}(m_p) \geq G_{dB,MIOF}(m_d)$$

$$\text{and } G_{dB,predicted}(m_p) = G_{dB,MIOF}(m_p) - G_{dB,MIOF}(m_d) \geq 0$$

Moreover, by rearranging the terms in the logarithm, we obtain:

$$20\log_{10}\left(\frac{MIOF(m_p)}{MIOF(m_d)}\right) \geq 20\log_{10}\left(\frac{m_p}{m_d}\right)$$

This means that the neural degradation $20\log_{10}\left(\frac{MIOF(m_p)}{MIOF(m_d)}\right)$ is less negative than the acoustic degradation $20\log_{10}\left(\frac{m_p}{m_d}\right)$ (Fig. 2.9 B).

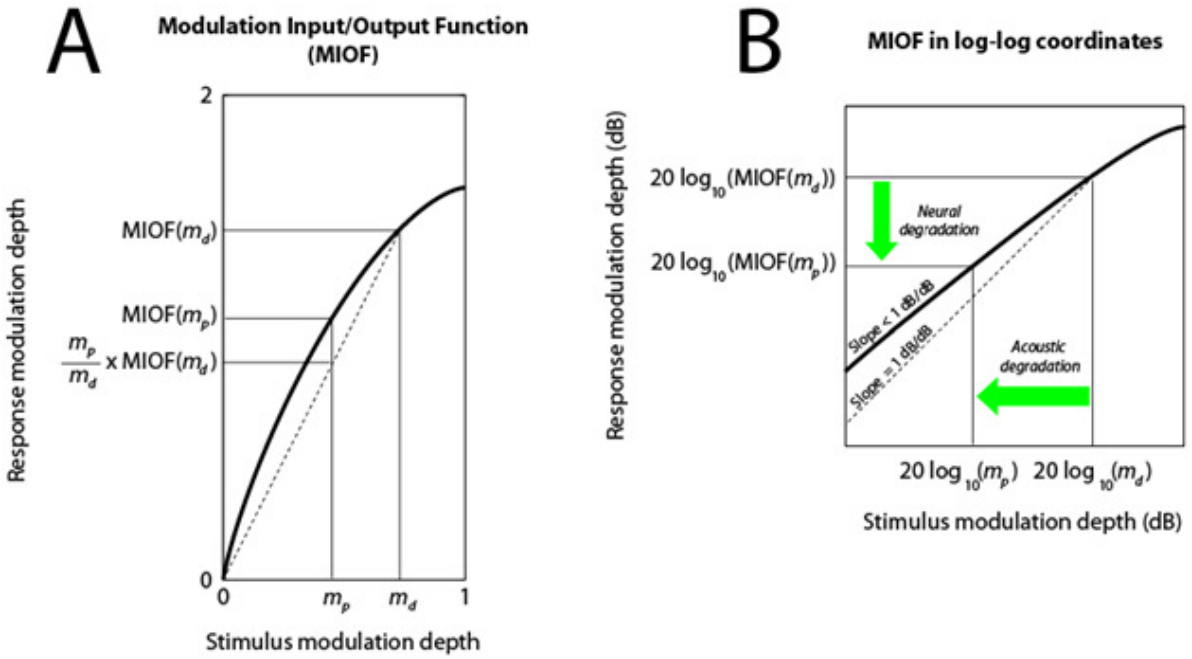

Figure 2.9 The compressive shape of the MIOF ensures partial compensation for the acoustic degradation, regardless of the modulation depth at the source

A. Schematic MIOF with a compressive shape: The curve is always above any secant (dotted line). The modulation depth at the ear drum (proximal modulation depth m_p) is smaller than the modulation depth at the sound source (distal modulation depth m_d) due to the acoustic degradation from reverberation. **B.** Schematic MIOF in log-log coordinates. The slope is < 1 dB/dB due to the compressive shape. As a result, neural degradation (vertical arrow) is smaller than acoustic degradation (horizontal arrow), even for distal modulations depths < 1.

Chapter 3

Effects of binaural and temporal features of reverberation on temporal coding of amplitude modulation

Abstract

In chapter 2, we showed that reverberation degrades the temporal coding of amplitude modulation (AM), and that the Modulation Input/Output Function (MIOF), which describes the nonlinear transformation of acoustic modulations into neural modulations, was successful in predicting this degradation in a majority of cases. However, in a subset of neurons, predictions from the MIOF underestimated the strength of temporal coding in the steady state reverberant response, suggesting that, in these neurons, reverberant stimuli had a coding advantage over anechoic stimuli with the same modulation depth.

In this chapter, we compared the time course of the reverberant responses with a prediction from the time course of stimulus modulations transformed by the MIOF. Consistent with Chapter 2, we found a subset of neurons in which the MIOF underestimated the strength of temporal coding in the later portion of the reverberant response. To explain this reverberant advantage, we carried out additional experiments with modified anechoic stimuli that included some features of the reverberant stimuli. We found that envelope distortion and spectral coloration introduced by reverberation had a negligible effect on neural responses, whereas binaural features such as interaural envelope disparities and changes in interaural cross-correlation partly explained the observed differences between anechoic and reverberant responses. Moreover, we found that diotic reverberant stimuli also had a coding advantage over anechoic stimuli with the same modulation depth, despite having the same binaural characteristics, and that altering the temporal properties of the reverberant room impulse responses had an effect on the neural responses. This suggests that IC neurons may exploit the temporal properties of reverberant room impulse responses to partially compensate for the AM degradation in the stimulus.

Overall, our results indicate that both binaural and temporal features mediate the robustness of temporal coding of AM in reverberation.

Introduction

Reverberation degrades amplitude modulations (AM) in the envelope of acoustic signals. The room Modulation Transfer Function (Houtgast *et al.*, 1980; Schroeder, 1981) quantifies the average degradation in AM between a sound source and a receiver, assuming a static degradation. However, reverberation is a dynamic process: Near stimulus onset, only the direct sound is present, and degradation is minimal, whereas later on in the stimulus, more and more reflections are superimposed to the direct sound, and signal degradation increases before reaching a plateau. In a study of sound localization in reverberant environments, Devore *et al.* (2009) showed in a virtual room that sound localization cues, although degraded in the later portion of reverberant stimuli, are mostly intact near stimulus onset. They also showed that directional sensitivity of neurons of the cat Inferior Colliculus (IC) in response to simulated reverberant stimuli, follows a similar time course. We therefore ask if for the temporal coding of AM, a similar parallel exists between the time course of stimulus degradation and time course of the neural responses.

In Chapter 2, we effectively assumed that the degradation due to reverberation was a static decrease in modulation depth, as modeled by the room MTF. We showed that reverberation degrades temporal coding of amplitude modulation (AM), but that this neural degradation is smaller than the degradation in the stimulus, suggesting that a neural gain compensates for the acoustic degradation. We also showed that the Modulation Input/Output Function (MIOF), which describes the nonlinear transformation of acoustic modulations into neural modulations, largely predicted this neural compensation in a majority of cases. However, in a subset of neurons, predictions from the MIOF underestimated the strength of temporal coding in the steady state, fully degraded reverberant response, suggesting that, in these neurons,

reverberant stimuli had a coding advantage over anechoic stimuli with the same modulation depth.

The observation that coding of reverberant stimuli is more robust than expected in the later portion of the stimulus cannot be entirely explained by the idea that the time course of the reverberant response parallels the time course of the reverberant stimulus, and suggests that the hypothesis that neural modulations only depend on the instantaneous modulation depth of the stimulus (as modeled by the MIOF) is not entirely valid. There are many differences between reverberant and anechoic stimuli, even when their modulation depths are matched. Specifically, for sinusoidal amplitude modulated (SAM) sound sources, reverberation results in a small distortion in the envelope, making the reverberant envelope slightly non-sinusoidal. In IC of anesthetized chinchilla, Sinex *et al.* (2002) demonstrated that phase-locking to modulation frequency was dependent upon the shape of the modulation waveform. However, they were comparing sinusoidal and trapezoidal waveforms, whereas the envelope distortion from reverberation is smaller. Other studies have looked at the effect of changing the pulse duration and interpulse interval of pulse trains (e.g. Krebs *et al.*, 2008) but focused on rate coding. Therefore the effect of envelope distortions introduced by reverberation on temporal coding is not clear from the literature.

Other potential differences between reverberant and anechoic stimuli are related to their binaural properties. The Interaural Cross-Correlation (IACC) of the stimulus differs between anechoic signals (high IACC) and reverberant signals (lower IACC). The effect of IACC has been studied extensively in psychophysical and neurophysiological studies. Human psychophysical experiments show a high sensitivity to changes in IACC, especially for correlations close to 1 (e.g. Culling *et al.*, 2001). Neurophysiological studies have found a linear

or nonlinear monotonic dependence of firing rate on IACC in the IC (e.g. Albeck and Konishi, 1995; Coffey *et al.*, 2006). In unanesthetized rabbit IC (Coffey *et al.*, 2006), firing rate usually increased with IACC at the cell's best Interaural Time Difference (ITD), and decreased with IACC at the cells worst ITD. These findings are consistent with a cross-correlation model for binaural interaction (Yin *et al.*, 1987). However, a description of the dependence of phase-locking to modulation frequency on the IACC of a pair of modulated noises is lacking.

Other binaural features potentially differing between reverberant and anechoic conditions are envelope delays between the two ears, and small differences in modulation depth present at the two ears, which we group under the term "Interaural Envelope Disparities" (IEDs). Again, effects of IEDs on firing rate have been reported in IC neurons (e.g. Yin *et al.*, 1984 in cat; Griffin *et. al*, 2005 in guinea pig), but possible effects on temporal features of the response to SAM noise are unclear.

In this chapter, we measured responses of IC neurons in unanesthetized rabbit to simulated anechoic and reverberant SAM broadband noise, and compared the time course of the reverberant responses to the time course of the modulations in the reverberant stimuli. Consistent with Chapter 2, we found that in a subset of neurons, reverberant stimuli had a coding advantage over anechoic stimuli whose modulation depth matched that of the reverberant stimuli. To identify which specific acoustic features of the reverberant stimuli were responsible for this advantage, we performed additional experiments with modified stimuli that included some features, alone or in combination, of the reverberant stimuli. Such features included envelope distortion, IACC, IEDs, and spectral coloration. We found that envelope distortion and spectral coloration introduced by reverberation had a negligible effect on neural responses, whereas binaural features such as IEDs and IACC partly explained the observed differences

between anechoic and reverberant responses. Moreover, we found that altering the temporal properties of the reverberant room impulse responses had an effect on temporal coding of AM. This suggests that IC neurons may exploit the temporal and binaural properties of reverberant room impulse responses to partially compensate for the AM degradation.

Methods

Surgical Preparation and Recording Procedures

Methods for chronic recordings from single units in IC of unanesthetized dutch-belted rabbits (*Oryctolagus cuniculus*) were based on the techniques of Kuwada *et al.* (1987), Nelson and Carney (2007), and Devore and Delgutte (2010), and are described in Chapter 2 (Methods). All procedures were approved by the Institutional Animal Care and Use Committees of the Massachusetts Eye and Ear Infirmary and the Massachusetts Institute of Technology.

Acoustic Stimuli

Virtual Auditory Space

We used the same simulated binaural room impulse responses (BRIRs) as in Chapter 2: BRIRs were simulated for a source positioned at 0º azimuth and at distances of 1.5 and 3 m from the center of a spherical head. The direct-to-reverberant (D/R) energy ratio was 0 dB for the 1.5 m condition (moderate reverberation) and -6 dB for the 3 m condition (strong reverberation). Anechoic impulse responses were obtained by isolating the first peak (direct sound) from the reverberant BRIRs. Virtual auditory space stimuli were created by convolving Sinusoidally Amplitude Modulated (SAM) broadband noise tokens with the left and right BRIRs.

Diotic anechoic stimuli with matched modulation depth (depth-matched anechoic stimuli)

The dynamic effect of reverberation on AM was visualized by computing the modulation depth of the reverberant stimuli as a function of time. The modulation depth was computed according to Eq. 2.1, in time windows with a length of one modulation cycle shifted every 1 ms over the 2-second total duration. The resulting stimulus modulation time course consisted of

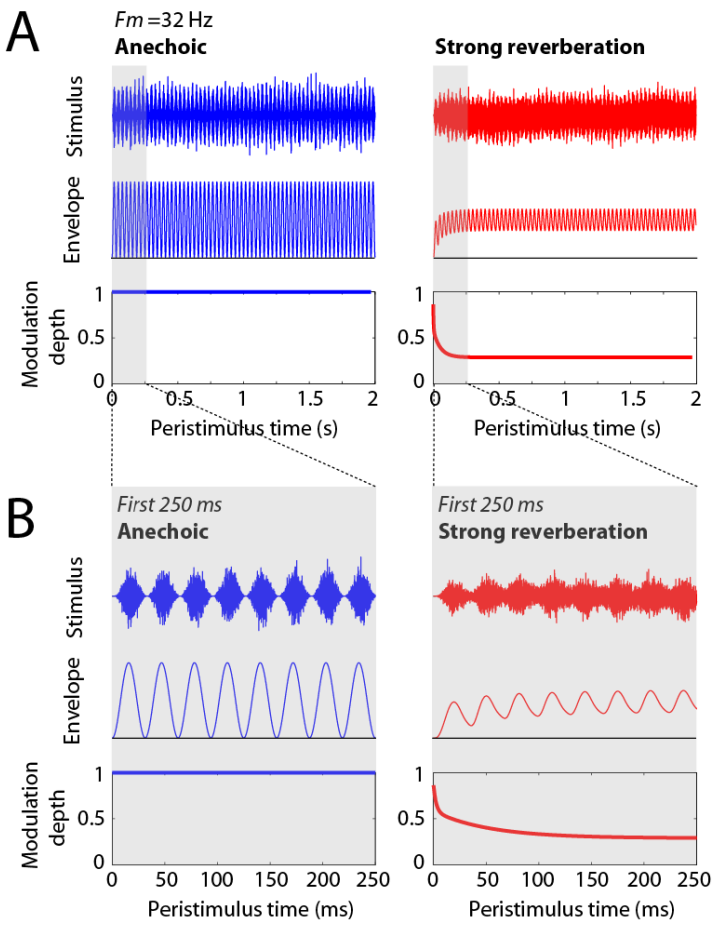

Figure 3.1 Time course of amplitude modulations in simulated anechoic and reverberant stimuli
A. Stimulus waveform, envelope, and modulation depth as a function of time over the 2-second duration of example stimuli with f_m = 32 Hz. *Left panel* (in blue): Anechoic case. Stimulus was Sinusoidally Amplitude Modulated (SAM) broadband noise with a static average modulation depth of 1. *Right panel* (in red): Reverberant case. A fully modulated broadband noise was filtered by the right channel of the binaural room impulse response for the strong reverberation condition. Envelope quickly reached steady state within the first 250 ms. Time course of modulation depth was fitted with the sum of two decaying exponentials. Modulation depth started close to 1, before sharply decreasing as reverberant energy built up. A plateau was reached by 250 ms. **B.** Zoom onto first 250 ms.

a sharp decay followed by a plateau (steady state portion, reached within 250 ms). This time course was fitted with the sum of two decaying exponentials (Fig. 3.1).

SAM noise was generated with the same modulation depth as the reverberant stimuli in the steady state portion (in the ear contralateral to the IC recorded from) (Fig. 3.2). The resulting stimuli, once filtered by the anechoic BRIRs, were anechoic SAM broadband

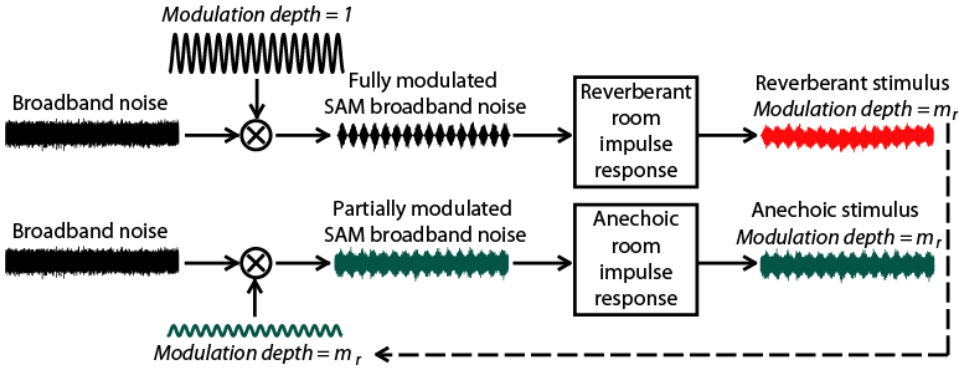

Figure 3.2 Generation of depth-matched anechoic stimuli
Schematic illustrating the procedure for the generation of a partially-modulated anechoic stimulus with modulation depth matching that of a reverberant stimulus. Modulation depth of the reverberant stimulus in the contralateral ear was used to sinusoidally modulate broadband noise. The resulting stimulus was then filtered with the anechoic room impulse response. The depth-matched anechoic stimuli were diotic, while the standard reverberant stimuli were dichotic.

noises with modulation depths matching those of the reverberant stimuli in their steady state portion. Contrary to the reverberant stimuli, the depth-matched anechoic stimuli were presented diotically.

Diotic anechoic stimuli with matched envelope waveform (envelope-matched anechoic stimuli)

Reverberation introduces small distortions in the sinusoidal envelopes of the anechoic stimuli (Fig. 3.3). To investigate the effect of these envelope distortions, we synthesized anechoic stimuli with the same average envelope shape as the reverberant stimuli. Average envelope shape was extracted for each reverberant condition and each modulation frequency f_m by taking 50 tokens of reverberant SAM noise, full-wave rectifying, low-pass filtering (order 3 Butterworth filter with cutoff frequency $5 \times f_m$), and averaging across tokens. The resulting envelope was divided into 1-period time bins, and the bins corresponding to the steady state portion of the stimulus (≥ 250 ms) were averaged together. One cycle of the average envelope

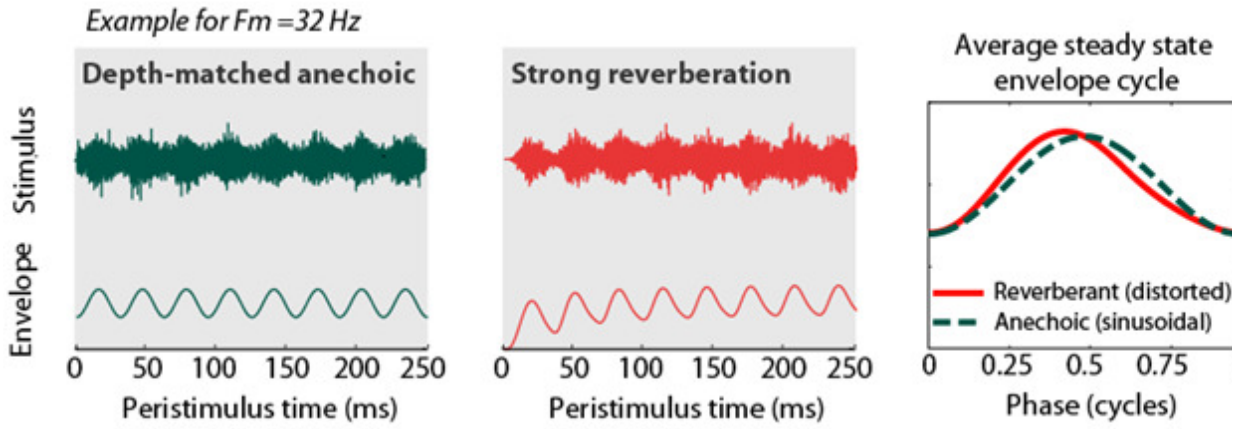

Figure 3.3 Reverberation creates envelope distortions
Left and middle panels: Example depth-matched anechoic and reverberant stimuli (f_m = 32 Hz, strong reverberation) and envelopes. *Right panel*: Average envelope cycle after 250 ms (steady state portion of the reverberant stimulus). Anechoic envelopes (green, dashed line) were sinusoidal, whereas reverberant envelopes (red, solid line) were slightly distorted.

waveform of the channel contralateral to the IC was used to modulate 2-sec tokens of white noise, which were subsequently filtered with the anechoic BRIRs.

Anechoic stimuli with matched static interaural cross-correlation (IACC-matched anechoic stimuli)

In our study, the modulated sound source was positioned at 0˚ azimuth from the spherical head. As a result, the anechoic stimuli at the left and right ear drums were nearly identical. The peak interaural cross-correlation (IACC) was close to 1 at all times. In contrast, reverberation decorrelates the stimulus waveforms in the left and right stimulus channels (Fig. 3.4). Near stimulus onset, the peak IACC was 1, as the direct sound dominates, after which the peak IACC decreased and increased in a cyclic pattern at the modulation frequency: When the amplitude of the source stimulus is large, the energy of the direct sound dominates over the reverberant energy, and the interaurally correlated direct sound boost the peak IACC. Conversely, as the

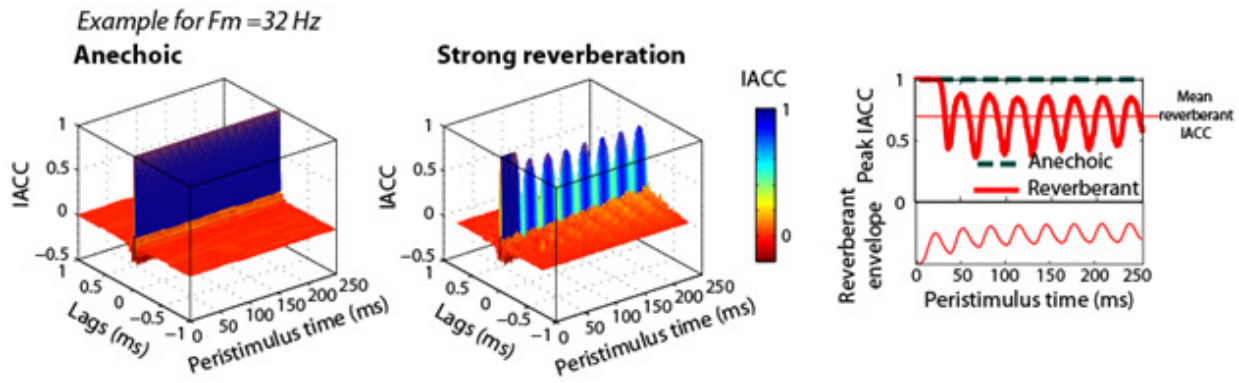

Figure 3.4 Reverberation decreases mean Interaural Cross-Correlation (IACC)
Interaural Cross-Correlation (IACC) as a function of time for anechoic (*left panel*) and reverberant (*middle panel*) stimuli (f_m = 32 Hz, strong reverberation) computed in 780 µs windows (40 windows/cycle). Only lags < 1 ms and the first 250 ms of the stimulus are represented. *Right panel*: Peak IACC was constant with a value of 1 for the anechoic stimulus (green, dashed line), whereas peak IACC sharply decreased before oscillating in the reverberant case (red, thick line). Mean reverberant IACC: red, thin line.

amplitude of the source stimulus is small, the energy of the direct sound becomes small compared to the reverberant energy, and peak IACC decreases.

Acoustic analysis showed that the mean peak IACC in the later portion (≥ 250 ms) of our stimuli depended on reverberant condition (the D/R energy ratio) as expected but not on f_m. We estimated the mean IACC by averaging IACCs of 50 tokens of reverberant stimuli computed over the stimulus duration excluding the early portion. Mean IACC was ≈0.85 in moderate reverberation, and ≈0.74 in strong reverberation. To investigate the effect of IACC on the neural response, we synthesized anechoic stimuli with the same mean IACC as the reverberant stimuli. The procedure to create dichotic stimuli with a specified IACC was based on the techniques of Culling (2001) and Devore and Delgutte (2010): A pair of uncorrelated noise tokens x_1[k] and x_2[k] were created with a Gram–Schmidt procedure; A third noise x_3[k] was synthesized as a mixture of x_1 and x_2 using the following formula (with ρ the desired IACC):

$$x_3[k] = \rho \cdot x_1[\mathrm{k}] + \sqrt{1-\rho^2} \cdot x_2[\mathrm{k}] \qquad \text{(Eq. 3.1)}$$

With this procedure, x_1 and x_3 have a correlation of ρ. As the anechoic BRIRs are essentially pure delays with identical waveforms in the two ears, the anechoic stimuli resulting from filtering x_1 and x_3 with the anechoic BRIRs also had a mean correlation of ρ. However, these IACC-matched stimuli differ from reverberant stimuli in that their IACC is static.

Anechoic stimuli with matched Interaural Envelope Disparities (IED-matched anechoic stimuli)

Reverberation introduced a small delay between the envelopes at the two ears (envelope Interaural Time Difference or ITD_{env}), a small difference in sound level (Interaural Level Difference or ILD) and a small difference in stimulus modulation depth (Interaural Depth Difference or IDD) (Fig. 3.5). Together, we refer to these properties as Interaural Envelope Disparities (IEDs). These differences arise because the simulated head was positioned slightly away from the center of the virtual room. The anechoic stimuli had no IEDs, because the source was positioned at a 0° azimuth with respect to the receivers. To investigate the effect of IEDs, we created anechoic stimuli with IEDs matching those of the reverberant stimuli. To do so, we measured the average magnitude and phase of the modulation waveform in the steady state portion of both channels of the reverberant stimulus, as described in Chapter 2 (Methods, Eq. 2.1). The mean ITD_{env} across modulation frequencies was ≈ 174 μs in moderate reverberation, and ≈ 678 μs in strong reverberation. Although these differences appear large when expressed in μs, the corresponding Interaural Phase Differences (IPD_{env}) are small. Across modulation frequencies, IPD_{env} ranged from ≈0 to 0.02 cycles in moderate reverberation, and from ≈0 to 0.07 cycles in strong reverberation. The ILDs were also small at all modulation frequencies (<0.1 dB in moderate reverberation and < 0.2 dB in strong reverberation). Across modulation frequencies,

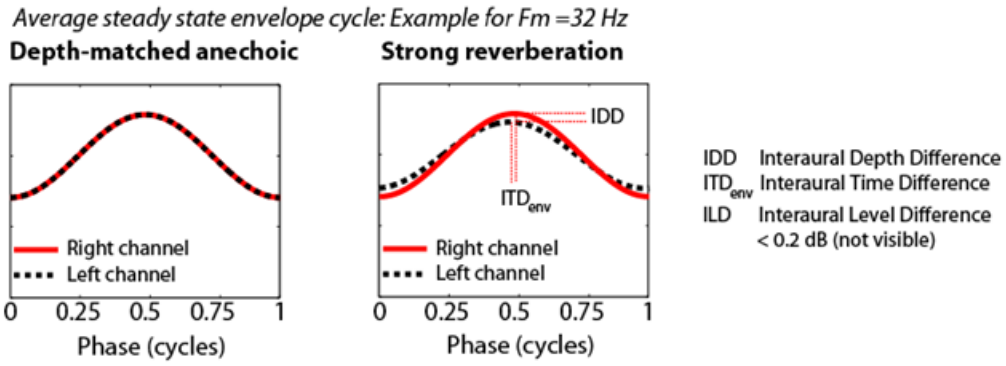

Figure 3.5 Reverberation creates Interaural Envelope Disparities (IEDs)
Comparison of the left and right average envelope cycles of the depth-matched anechoic (left) and reverberant (right) stimuli (f_m = 32 Hz, strong reverberation). Left and right channels were identical in the anechoic case, whereas reverberation introduced small Interaural Envelope Disparities (IEDs), consisting of a small time difference (ITD_{env}), a small level difference (ILD, not visible because <0.2 dB), and a small modulation depth difference (IDD).

IDD ranged from ≈0 to 0.07 (≈0 to 2 dB) in moderate reverberation, and from ≈0.01 to 0.07 (≈0 to 4 dB) in strong reverberation.

Anechoic stimuli with matched spectral coloration (spectrum-matched anechoic stimuli)

Reverberation alters the long term power spectrum of the stimuli (Fig. 3.6 A). To study the effect of this spectral coloration, we created anechoic stimuli with similar long term spectra as the reverberant stimuli. The procedure to synthesize these stimuli is schematized in Figure 3.6 B. We created a coloration filter h_c from the ratio of the Fourier transform magnitude of the contralateral reverberant impulse response (h_r) to that of the anechoic impulse response (h_a), with a random phase ϕ:

$$h_c(t) = iFFT\left[\left|\frac{FFT(h_r)}{FFT(h_a)}\right| e^{j\phi}\right] \quad \text{(Eq. 3.2)}$$

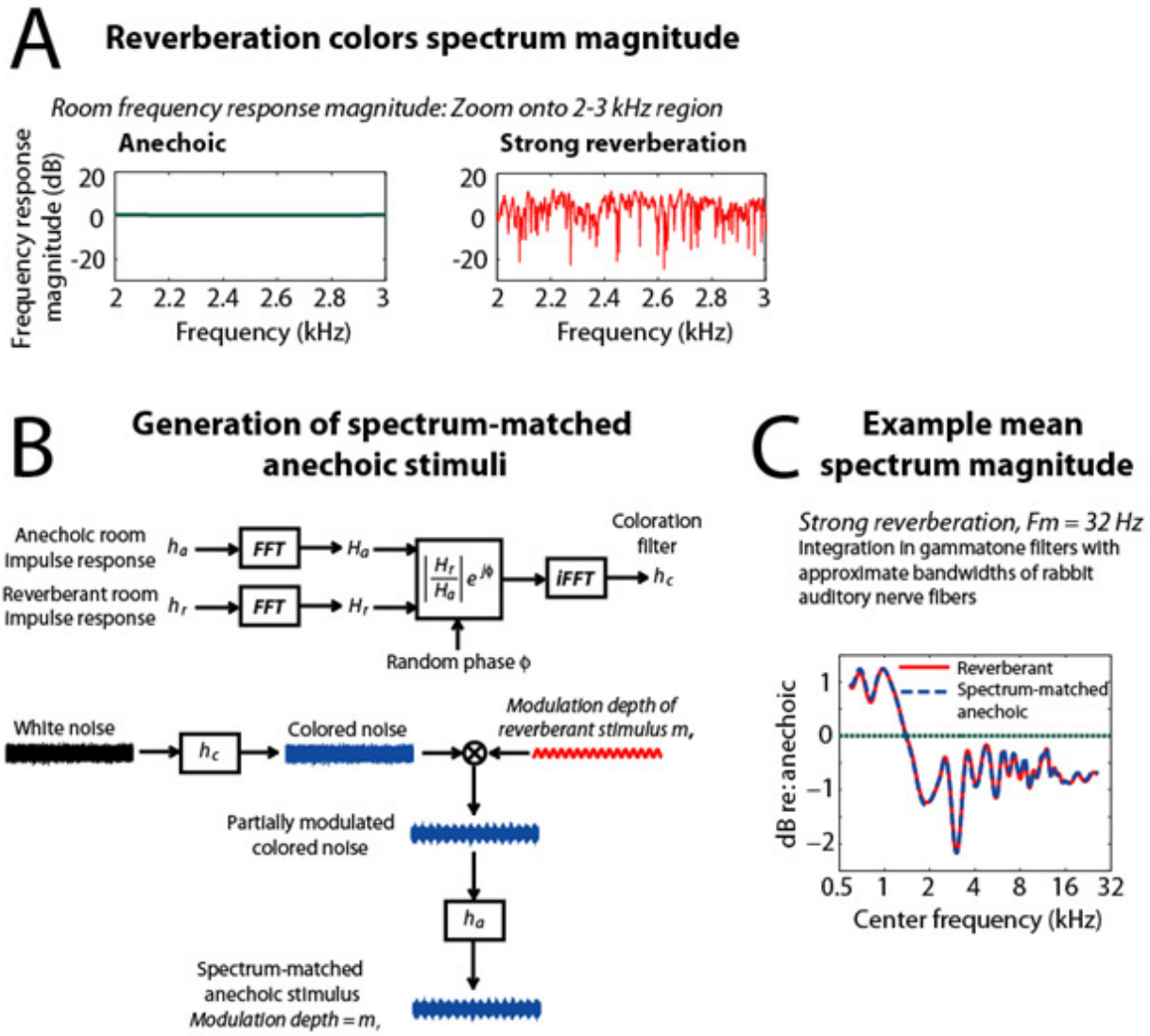

Figure 3.6 Generation of spectrum-matched anechoic stimuli
A. Frequency response magnitude of the anechoic (left) and reverberant (right) room impulse response (strong reverberation, right channel), zoomed in the 2-3 kHz region. The anechoic response magnitude was roughly flat whereas the reverberant response magnitude consisted of tightly superimposed notches corresponding to individual reflections, therefore coloring the spectrum magnitude of the stimuli. **B.** Schematic illustrating the procedure we used to create colored anechoic stimuli whose mean spectrum magnitude was similar to that of the reverberant stimuli (see text) **C.** Comparison between the mean spectrum magnitude of a reverberant stimulus (f_m = 32 Hz, strong reverberation) and that of a spectrum-matched anechoic stimulus. Spectrum magnitudes were integrated in gammatone filters with equivalent rectangular bandwidths similar to that of peripheral rabbit filters reported in the literature (Borg *et al.*, 1988).

Applying this filter to a modulated sound source would result in complete demodulation due to the random phase. To overcome this difficulty, we applied this coloration filter to the unmodulated noise source, then applied a sinusoidal modulation to the colored noise (with modulation depth matched to the reverberant stimulus), and filtered the resulting partially modulated colored noise with the anechoic BRIRs.

As h_c had a duration of ≈1.5-2 seconds (depending on the reverberant condition), and did not have its energy concentrated near onset due to the random phase, the convolution product of noise with h_c started with a ≈1.5-2 second gradual rise in amplitude before reaching a steady state. To obtain a static anechoic, colored stimulus, we used 4 seconds of noise instead of 2, and removed the first 2 seconds of the filtered stimulus.

Although this procedure cannot produce the exact same long term power spectrum as the reverberant stimulus because the order of modulation and filtering are reversed, we verified that the average spectrum magnitudes, integrated over gammatone filters with equivalent rectangular bandwidths of rabbit auditory nerve filters (Borg *et al.*, 1988) were very similar for colored anechoic and reverberant conditions (Fig. 3.6 C). Differences did not exceed 0.1 dB.

We synthesized both diotic and dichotic spectrum-matched anechoic stimuli. In both conditions, we used the same coloration filter in the two ears, derived from the contralateral reverberant impulse response. In the dichotic condition, we also matched the IACC to that of the reverberant stimuli using the procedure described in a previous section.

Diotic reverberant stimuli

We created diotic reverberant stimuli by convolving fully modulated SAM broadband noise with the same reverberant room impulse response for both ears. We used the reverberant

BRIR contralateral to the IC. The resulting stimuli have the envelope and fine structure of dynamic reverberant stimuli, but are identical in the two ears.

Diotic reverberant stimuli with truncated reflections

To study the effect of the temporal structure of the reverberant BRIRs, we manipulated the BRIR contralateral to the IC by truncating a specific time window (Fig. 3.7): In one case, we kept the direct sound and the early reflections (t < 18 ms), but removed the later reflections ("No late reflections" impulse response, in orange); In the other case, we kept the direct sound and the late reflections, but removed the early reflections ("No early reflections" impulse response, in blue). We chose the cutoff between early and late reflections at 18 ms because, for both reverberant conditions, it marked a natural separation in the impulse response between large, well resolved reflections, and smaller, overlapping reflections.

Manipulating the reverberant impulse responses affected the amount of AM degradation they caused. The D/R ratio of the "No late reflections" impulse responses was +4.5 dB for the moderate reverberation condition, and +0.5 dB for the strong reverberation condition. The D/R ratio of the "No early reflections" impulse responses was +1.5 dB for the moderate reverberation condition, and -4 dB for the strong reverberation condition. This "No late reflections" impulse responses created substantially less degradation (≈ 4-10 dB differences in modulation depth across reverberation condition and f_m) than the full impulse responses, whereas the degradation from the "No early reflections" impulse responses was more similar to that created by the unmodified reverberation (≈0-4 dB differences at most frequencies and conditions). To obtain stimuli with the same modulation depths as the reverberant (full impulse response) stimuli, we used partially- rather than fully-modulated sound sources, and chose their modulation depths to

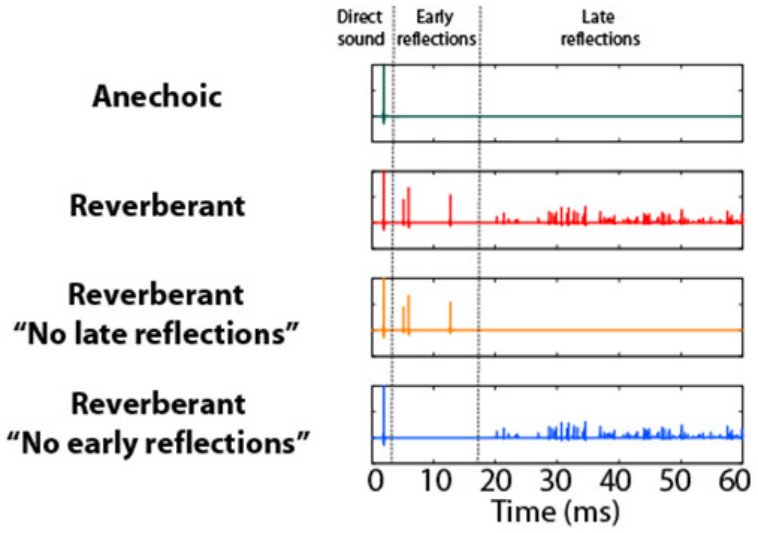

Figure 3.7 Diotic reverberant stimuli with truncated reflections
The temporal fine-structure of diotic reverberant stimuli was manipulated by truncating the reverberant impulse responses. "No late reflections" impulse responses were created by removing reflections occurring after 18 ms, whereas "No early reflections" impulse responses were created by removing reflections occurring before 18 ms. Stimuli synthesized from these different conditions were matched for modulation depth (see text).

counterbalance the differences in the amount of degradation caused by the truncated impulse responses.

Recording Procedures

Experimental procedures for isolating single unit, measuring rate level functions, determining Characteristic Frequency (CF), and choosing a sound level were as described in Chapter 2. Neural Modulation Transfer Functions (MTFs) and Modulation Input/Output Functions (MIOFs) for anechoic stimuli were also obtained as described in Chapter 2.

Responses to virtual auditory space stimuli were measured at one modulation frequency, chosen between 4-256 Hz based on the anechoic MTFs to elicit both a large firing rate and strong phase-locking to the modulation. If time allowed, recordings were performed at another

modulation frequency, or in another reverberant condition. We most frequently used modulation frequencies of 16 – 90 Hz (median = 64 Hz across the population). All stimuli were 2 second long followed by a 1 second silence, and repeated 4-71 times (the median number of trials was 10). A large number of presentations was required to characterize the time course of response modulation depth with fine resolution.

Results come from protocols with various stimulus conditions interleaved with the standard reverberant stimulus. For each protocol, presentations were randomized across stimulus conditions at one modulation frequency. In earlier experiments, standard reverberant stimuli were interleaved with fully-modulated anechoic stimuli, but MIOFs were measured separately in the same neurons. Later, standard reverberant stimuli were interleaved with depth-matched anechoic, envelope-matched anechoic, and diotic reverberant stimuli. In the final protocol, designed to simultaneously test multiple hypotheses, standard reverberant stimuli were interleaved with 9 other stimuli, all matched in modulation depth: (1) Diotic anechoic; (2) IED-matched anechoic; (3) IACC-matched anechoic; (4) Diotic spectrum-matched anechoic; (5) Spectrum- and IACC-matched anechoic; (6) IED- and IACC-matched anechoic; (7) Diotic reverberant; (8) Diotic reverberant with "No early reflections"; (9) Diotic reverberant with "No late reflections".

Data Analysis

Time course of response modulation depths

Only well isolated single-units were included in our data analysis. The time course of the neural modulation response was obtained by computing response modulation depths (RMD) separately in time bins whose duration was an integer number of modulation cycles. The number

of cycles was chosen so that RMD was significant (Rayleigh test of uniformity, $\alpha<0.05$) in at least 95% of all time bins for a given condition. In the examples shown in this Chapter, bin width ranged from 44 to 500 ms, depending on f_m and spike count. RMDs were computed from the spike times as described in Eq. 2.4. To avoid aliasing, the analysis windows were sometimes overlapping. The time course of the responses was smoothed with a rectangular moving average filter (usually with a 3-point span).

Steady state response modulation depths

To compute the steady state RMDs to all of our stimulus conditions, we only included the spikes that occurred in the later portion of the response, corresponding to the steady state portion of the reverberant stimulus. To do so, we removed the spikes that occurred in an onset window with duration the smallest integer number of modulation cycles equal to or greater than 250 ms. Only steady state RMDs that were significant (Rayleigh test of uniformity, $\alpha<0.05$) were included in our population statistics.

Modulation Input/Output Functions

Modulation Input/Output Functions (MIOFs) were constructed and fitted with smoothed curves as described in Chapter 2 (Methods).

Orthogonal linear regressions using Principal Component Analysis

To compare the RMDs between 2 stimulus conditions across the population of neurons, we modeled the dependence of one condition on the other with a linear regression. Instead of using ordinary least-squares linear regression analysis, which introduces an artificial asymmetry

between the coordinates, we Principal Component Analysis (PCA) to find the direction of the data's principal component: RMDs for all neurons were arranged in two vectors, one for each stimulus condition. Vector means were subtracted, and the covariance matrix of the two vectors was computed and diagonalized. The direction of the eigenvector corresponding to the largest eigenvalue was used as the regression line. We tested the null hypothesis that the slope of the regression line was equal to 1 with a statistical test specifically derived for orthogonal regressions (Wong, 1989). We tested the linear dependence of the 2 conditions by computing the Pearson correlation coefficient.

Results

Time course of reverberant response

We measured the time course of responses to simulated reverberant stimuli in 91 well isolated single units from the IC of unanesthetized rabbits. The virtual auditory space reverberant stimuli were synthesized using 100% modulated SAM broadband noise played by a sound source located 1.5 or 3 m away from a simulated spherical head, in a medium size virtual room (See chapter 2). Reverberation degraded the modulation depth of the stimuli in a time-dependent fashion, as illustrated for 32 Hz modulation in the strong reverberant condition in Figure 3.1. Unlike the anechoic stimulus for which modulation depth was constant, the modulation depth of the reverberant stimulus started close to 1, but sharply decreased within the first 250 ms before reaching a plateau near 0.3. The large modulation depth near stimulus onset reflects the dynamic nature of reverberation: Near stimulus onset, the direct sound dominates at the receivers, whereas as time goes by, more and more reflections superimpose and degrade the stimulus envelope.

Modulation Input/Output Functions (MIOF) were measured in 51 of these 91 single units, in response to modulated broadband noise whose modulation depth was varied between 0 and 1. As the MIOF describes the transformation of stimulus modulations into neural modulations, we tried to predict the time course of the reverberant response modulation depth (RMD) from the time course of stimulus modulation depth and the MIOF. Figure 3.4 shows results from an example neuron (f_m = 45 Hz) where reverberant RMD and prediction are compared. The time course of stimulus modulation depth (Fig. 3.8 A) is characterized, as described above, by a large modulation depth near stimulus onset, followed by a sharp decay leading to a plateau by 250 ms. The MIOF for this neuron (Fig. 3.8 B) increases steeply for stimulus modulation

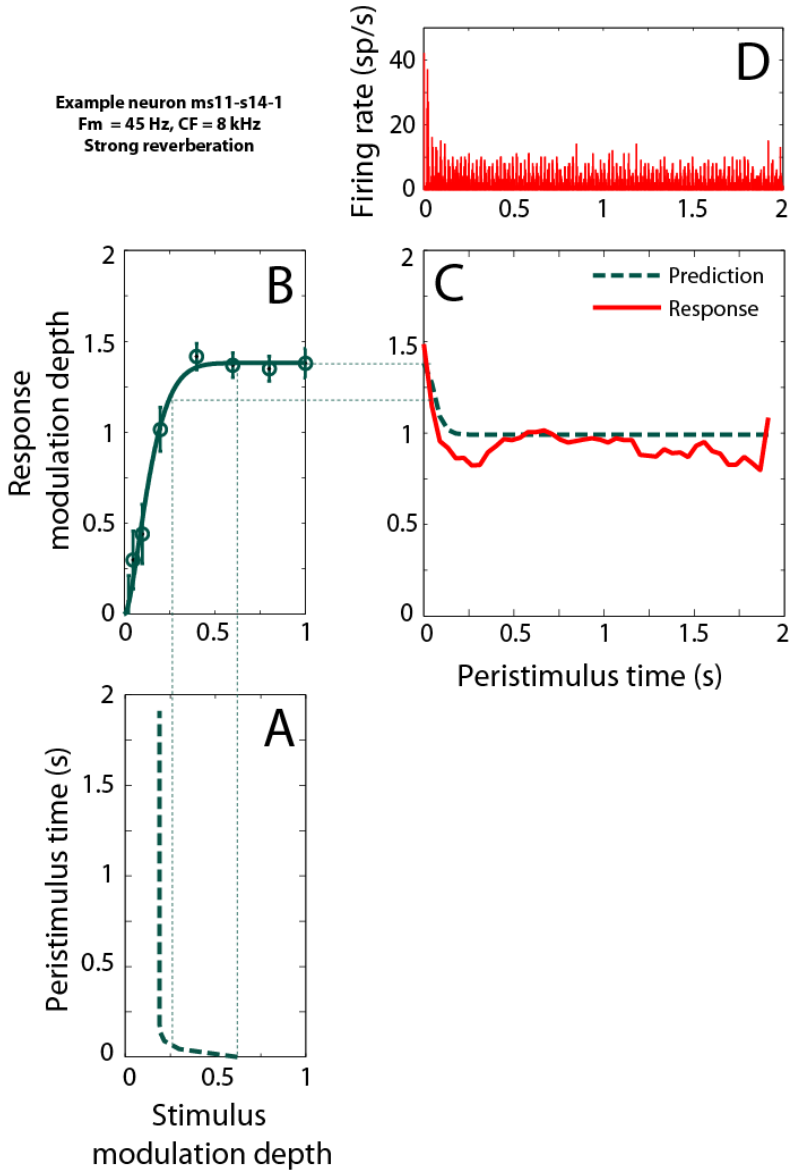

Figure 3.8 Time course of reverberant response modulation depth and prediction from the Modulation Input/Output Function in an example neuron

A. Time course of stimulus modulation depth (f_m = 45 Hz, strong reverberation). Modulation depths were computed in 2-period bins (≈ 44 ms), and fitted with the sum of two decaying exponentials. **B.** Modulation Input/Output Function (MIOF) measured in response to SAM noises with modulation depths varied between 0 and 1 in an example neuron (ms11-s14-1). Open circles: Significant response modulation depths (Rayleigh test of uniformity, $\alpha<0.05$). Thick line: Weighted least-square fit with an incomplete beta function. **C.** Reverberant response modulation depths (red, solid line) and prediction (green, dashed line) from the time course of stimulus modulations (A) and the MIOF (B) in the same neuron. Response modulation depths were computed in 2-period bins (≈ 44 ms) and smoothed with a 7-point moving average filter. **D.** Peristimulus Time Histogram of the reverberant response (10 bins/period)

depths below 0.3, before saturating to an RMD of ≈ 1.4. Using the MIOF and the time course of stimulus modulation depth, we predicted the time course of the reverberant RMD (Fig. 3.8 C, green, dashed line). Qualitatively, the prediction resembles the time course of the stimulus modulation depth, consistent with the monotonic shape of the MIOF. The predicted RMDs are larger than the stimulus modulation depths, because of the neural modulation gain (see Chapter 2). The Peristimulus Time Histogram (PSTH) of the measured reverberant response in this neuron (Fig. 3.8 D) shows a peak in firing rate near stimulus onset, followed by clear phase-locking to the modulation frequency throughout the entire 2-second duration of the stimulus. This phase-locking was quantified by the RMD (Fig. 3.8 C, red, solid line), which shows a sharp decay, followed by a plateau at an RMD of ≈1. In this neuron, the time course of the reverberant RMD was fairly well predicted by the MIOF.

In another neuron (Fig. 3.9), the reverberant RMD and the prediction were radically different. In this neuron (f_m = 16 Hz), the MIOF (Fig. 3.9 B) had a non-saturating profile with smaller modulation gains than in the previous example. The sharp decaying profile of the stimulus modulation depth time course (Fig. 3.9 A) was mirrored in the prediction of RMD (Fig. 3.9 C, green, dashed line), reaching a plateau with a value of ≈0.4. The measured RMD (Fig. 3.9 C, red, solid line), was also characterized by a sharp decay near stimulus onset, but in contrast to the prediction, increased between 200 ms and 1 second to reach a high plateau of ≈1.1. This increase following the sharp decay is clearly visible in the PSTH as well (Fig. 3.9 D). The initial peak in RMD is due to the notch in the PSTH immediately following the onset response. The decay in RMD before 200 ms corresponds to high, weakly-synchronized firing rate, whereas the gradually emerging PSTH peaks correspond to the slow increase in RMD after 200 ms. In this example, the MIOF poorly predicted the reverberant RMD. Further, the actual RMD was

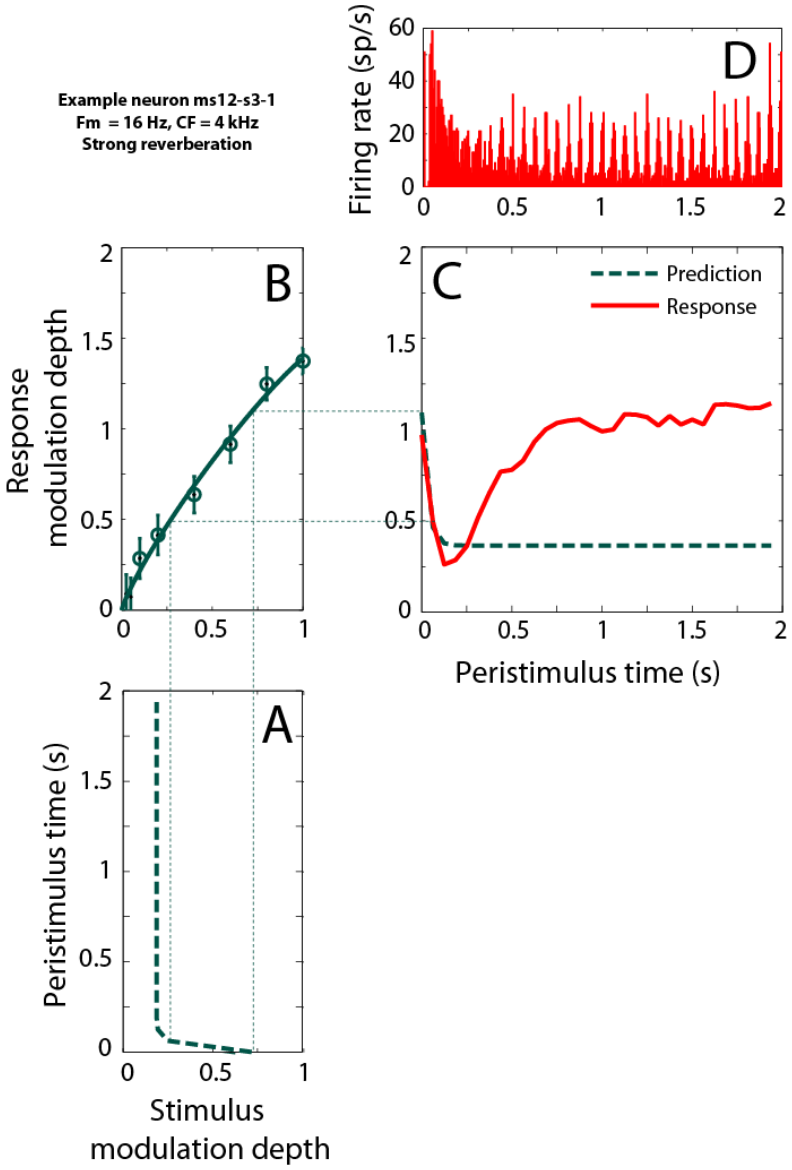

Figure 3.9 Time course of reverberant response modulation depth and prediction from the Modulation Input/Output Function in another example neuron: Reverberant advantage
A. Time course of stimulus modulation depth (f_m = 16 Hz, strong reverberation). Modulation depths were computed in 1-period bins (62.5ms), and fitted with the sum of two decaying exponentials. **B.** Modulation Input/Output Function (MIOF) neuron ms12-s3-1. Open circles: Significant response modulation depths (Rayleigh test of uniformity, $\alpha<0.05$). Thick line: Weighted least-square fit with an incomplete beta function. **C.** Reverberant response modulation depths (red, solid line) and prediction (green, dashed line) from the time course of stimulus modulations (A) and the MIOF (B) in the same neuron. Response modulation depths were computed in 1-period bins (62.5ms) and smoothed with a 3-point moving average filter. **D.** Peristimulus Time Histogram of the reverberant response (10 bins/period).

significantly higher than the prediction almost throughout the stimulus duration.. This suggests that, in this example, the reverberant stimulus has a coding advantage over an anechoic stimulus with the same modulation depth. Surprisingly, this advantage is largest in the later portion of the response, where the reverberant stimulus modulation depth has reached steady state.

Reverberant stimuli had a coding advantage over anechoic stimuli with the same modulation depth in some neurons

We compared the reverberant RMDs after 250 ms (steady state portion of the stimulus) to the predictions from the MIOF across our neural population (Fig. 3.10 A). Consistent with our findings in Chapter 2, the MIOF provided an overall good prediction of the reverberant RMD in that the correlation between reverberant RMD and prediction across the neural population was high ($r=0.82$, $p<0.001$). The mean difference between reverberant response and prediction was not significantly different from 0 (Paired Student's t-test, $p=0.07$). The majority of neurons (27, i.e. 56%) did not show a significant difference between response and prediction (black dots in Fig. 3.10 A). However, 15 of the 48 neurons (31%) that had a significant RMDs had significantly greater reverberant RMD than predicted from the MIOF (blue circles in Fig. 3.10 A). The mean difference between response and prediction for this group showing a reverberant advantage was 0.3 (+ 9 dB). In contrast, 6 neurons (13%) had a significant reverberant disadvantage (red circles in Fig. 3.10 A), with a mean difference in RMD of -0.32 (or – 4 dB).

In most cases, the response to the reverberant stimulus was measured first, followed by the MIOF measurement. It is thus possible that the differences we report above between reverberant RMD and predictions from the MIOF are due to long term changes in the response properties of the neuron between the two measurements. To rule out this possibility, we

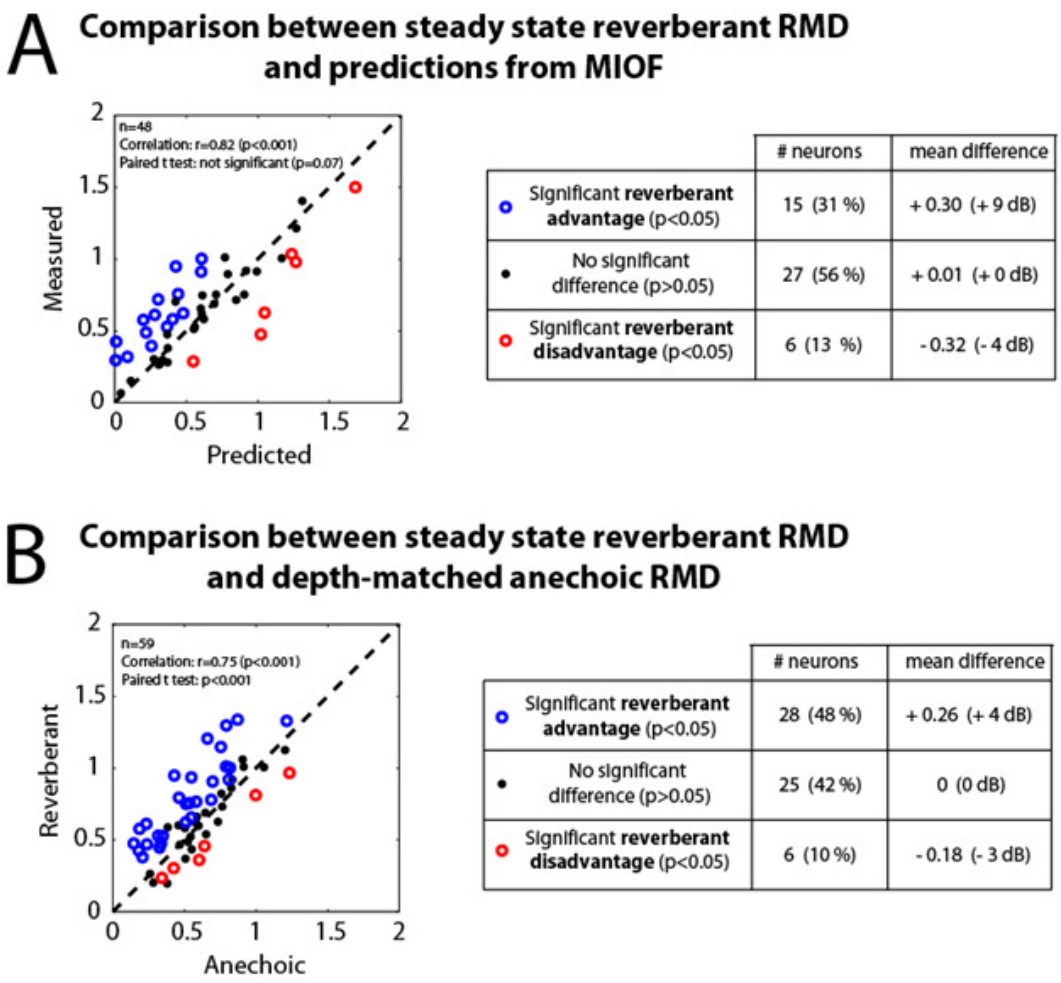

Figure 3.10 Population summary of the coding advantage of reverberant stimuli over anechoic stimuli with the same modulation depth

A. Comparison between measured reverberant response modulation depth after 250 ms (steady state portion of the stimulus) and prediction from Modulation Input/Output Function (MIOF). Point for which reverberant response was not significant (Rayleigh test of uniformity, $\alpha<0.05$) were omitted. Open circles: 21 out of 48 responses were significantly different than prediction ($p<0.05$), of which 15 showed a coding advantage for the reverberant stimulus. **B.** Comparison between reverberant response and depth-matched anechoic response after 250 ms. Point for which reverberant response was not significant (Rayleigh test) were omitted. Reverberant responses were significantly more robust ($p<0.05$) than anechoic responses in 48% of the population.

modified our experimental protocol by interleaving the reverberant stimuli with anechoic stimuli whose modulation depths were designed to match those in the steady state portion of the reverberant stimuli (Fig. 3.2). RMDs were measured for the reverberant and depth-matched anechoic stimuli in another set of neurons, and compared in the steady state part of the

reverberant stimulus (Fig. 3.10 B). Using these methods, the mean difference in RMD between reverberant and depth-matched anechoic conditions was ≈0.1, and significant with a paired t-test ($p<0.001$). The majority of the neurons (28 out of 59, or 48%; blue circles in Fig. 3.10 B) for which both reverberant and depth-matched anechoic RMDs were significant (Rayleigh test) had significantly larger reverberant steady state RMD than the depth-matched anechoic RMD ($p<0.05$; test of equality of concentration parameters). The mean reverberant advantage for these neurons was 0.26 (+ 4 dB). In contrast, 6 neurons (10%) showed a significant reverberant disadvantage (red circles in Fig. 3.10 B), with a mean difference between reverberant and anechoic RMD of -0.18 (- 3 dB). A large proportion of neurons (25 out of 59, or 42%) did not show any significant difference in RMD between reverberant and depth-matched anechoic conditions (black dots in Fig. 3.10 B).

The finding that reverberant and anechoic RMDs are significantly different in a large proportion of neurons when the stimulus modulation depths were matched, suggests that the nonlinear transformation from stimulus modulation depth to neural modulations (MIOF) depends not only on stimulus modulation depth, but also on other factors that differ between reverberant and anechoic stimuli. In the following sections, we explore the influence of several of these factors, such as envelope distortion, interaural cross-correlation, interaural envelope disparities, and spectral coloration.

Effect of envelope distortion

One difference between reverberant and depth-matched anechoic stimuli that might explain the differences in RMD described above is in the shape of their envelopes. Anechoic stimuli had a sinusoidal envelope throughout the stimulus duration. In contrast, reverberation

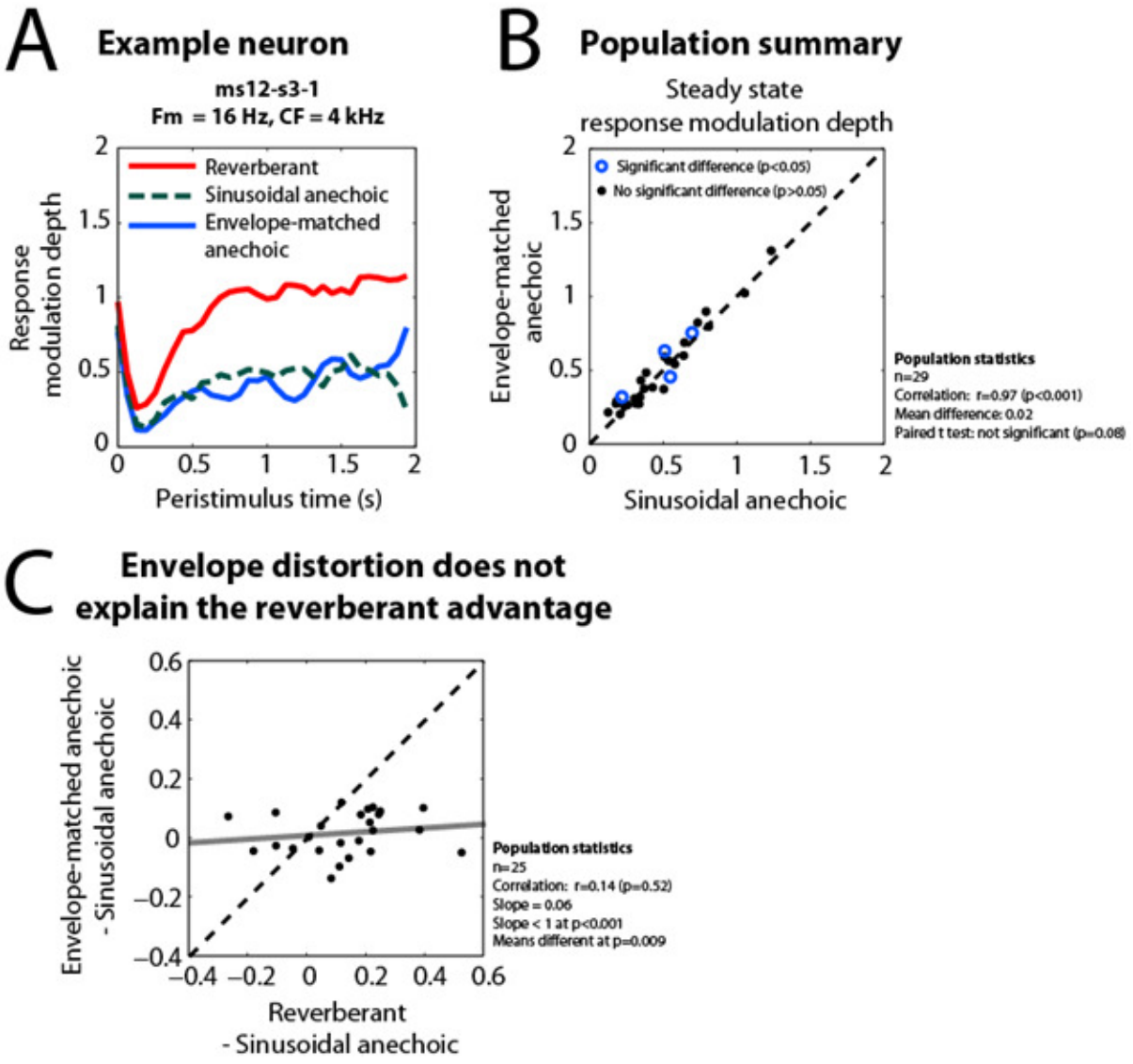

Figure 3.11 Effect of envelope distortions created by reverberation

A. Example neuron ms12-s3-1. Red, solid line: Reverberant. Green, dashed line: Depth-matched anechoic. Blue, solid line: Response to an anechoic stimulus with the mean distorted envelope of the reverberant stimulus. Responses were computed in 1-period bins (62.5 ms) and smoothed with a 3-point moving average filter. **B.** Population summary: Comparison between envelope-matched anechoic and sinusoidal anechoic conditions. Points for which response modulation depths were not significant (Rayleigh test of uniformity, $\alpha<0.05$) were omitted. Pearson correlation was very high ($r=0.97$, $p<0.001$). 4 out of 29 neurons showed significant differences between sinusoidal and distorted stimuli ($p<0.05$), but differences were small. **C.** Comparison between effect of envelope distortion and effect of reverberation. Thick, grey line: Direction of first principal component. Pearson correlation was small (0.14) and not significant ($p=0.52$). Slope was small (≈ 0.06) and effect of reverberation was significantly greater than effect of envelope distortion ($p=0.009$).

slightly distorted the envelope, making the average envelope period skewed with the rising edge somewhat sharper than the falling edge (Fig. 3.3). For all modulation frequencies and D/R energy ratios, the deviation from a sinusoid was small. However, since the shape of the envelope can have notable effects in IC neurons (Sinex *et al.*, 2002; Krebs *et al.*, 2008), we tested the hypothesis that envelope distortions introduced by reverberation are responsible for the differences we observed between reverberant RMD and anechoic RMD.

To test this hypothesis, we extracted the average envelope period in the steady state portion of the reverberant stimulus (Methods), and used this distorted envelope to modulate broadband noise. The resulting modulated noise was then filtered by the anechoic room impulse response, resulting in an anechoic stimulus with the same modulation depth and average envelope waveform as the reverberant stimulus. Figure 3.11 A shows data for an example neuron (f_m = 16 Hz, same neuron as in Fig. 3.9), in which the reverberant RMD (red, solid line) was significantly larger than the sinusoidal anechoic RMD (green, dashed line). In this neuron, the RMD of the envelope-matched anechoic stimulus (blue, solid line) was very similar to that of the sinusoidal anechoic stimulus, inconsistent with the hypothesis that envelope distortions account for the reverberant advantage.

Across the population (Fig. 3.11 B), the envelope-matched anechoic RMDs were highly correlated with the sinusoidal anechoic RMDs ($r=0.97$, $p<0.001$) and the means were not significantly different between the two conditions (Paired t-test). Only 4 out of 29 neurons had significantly different RMDs between the two conditions ($p<0.05$; test of equality of concentration parameters), but even in these cases, differences were small (≈0.1). Overall, this suggests that envelope distortions introduced by reverberation had only a small influence on RMD in our stimuli.

We further tested whether these small differences between sinusoidal and envelope-matched anechoic conditions could explain the reverberant advantage across the population. If envelope distortions were responsible for the differences observed between reverberant and sinusoidal anechoic RMDs, the effect of envelope distortion (defined as the difference between envelope-matched anechoic RMD and sinusoidal anechoic RMD) should be similar to the effect of reverberation (defined as the difference between reverberant RMD and sinusoidal anechoic RMD). Figure 3.11 C shows a scatter plot across the neural population of the effect of reverberation against the effect of envelope distortion. The correlation between the two effects was very small and not significant ($r=0.14$, $p=0.52$), and the slope of the regression line was small (≈ 0.06) and significantly smaller than 1 ($p<0.001$). Moreover, the mean effect of reverberation was significantly larger than the mean effect of envelope distortions ($p=0.009$). Together, these tests indicate that the small envelope distortions created by reverberation do not explain the reverberant advantage.

Effect of Interaural Cross-Correlation (IACC)

Another factor that may explain differences between depth-matched anechoic and reverberant responses is differences in the Interaural Cross-Correlations (IACC) of the stimuli: In the anechoic case, the signals at the two ears are nearly identical, and therefore highly correlated (IACC ≈ 1) whereas in the reverberant case, the reflections decorrelate the input to the two ears. The effect of reverberation on IACC is visualized in Figure 3.4. The short-time IACC of the anechoic stimulus computed over windows of 780 μs (40 windows per cycles at 32 Hz) has a distinct maximum at 0 lag, with a value close to 1 throughout the duration of the stimulus. In contrast, in the reverberant case, the peak IACC starts at a high value (close to 1) near stimulus

onset before reflections could reach the ears, and then oscillates in a cyclic pattern with a frequency equal to the modulation frequency (Fig. 3.4, right panel). The mean peak IACC in the oscillating portion of our stimuli was ≈0.74 in the strong reverberation condition, and ≈0.85 in the moderate reverberation condition for all modulation frequencies.

To test the hypothesis that the mean decrease in IACC introduced by reverberation is responsible for the observed differences between reverberant RMD and depth-matched anechoic RMD, we synthesized anechoic stimuli for which both mean IACC and modulation depth were matched to those of the steady state portion of the reverberant stimuli (Methods). Figure 3.12 A shows results for two example neurons with different effects of mean IACC on the response. In the first example (top panel), the reverberant RMD (red, solid line) was significantly more robust than the IACC-matched anechoic RMD (blue, solid line), while the latter was very similar to the diotic anechoic RMD (green, dashed line). Therefore, in this neuron, mean IACC did not play a major role in the response, and the reverberant advantage could not be explained by differences in IACC. In contrast, in the other example (bottom panel), the RMD for the IACC-matched anechoic stimulus was significantly lower than that of the diotic anechoic stimulus. The reverberant RMD was also less robust than the diotic anechoic RMD, consistent with the hypothesis that, in this neuron, mean IACC partly explains the differences observed between anechoic and reverberant responses.

Across the population (Fig. 3.12 B), the IACC-matched RMDs were well correlated to the diotic anechoic responses ($r=0.74$, $p<0.001$) and the means were not significantly different between the two conditions (Paired t-test, $p=0.15$). However, in 2 neurons (out of 29), decorrelation significantly increased RMD, and in 7 neurons, decorrelation significantly

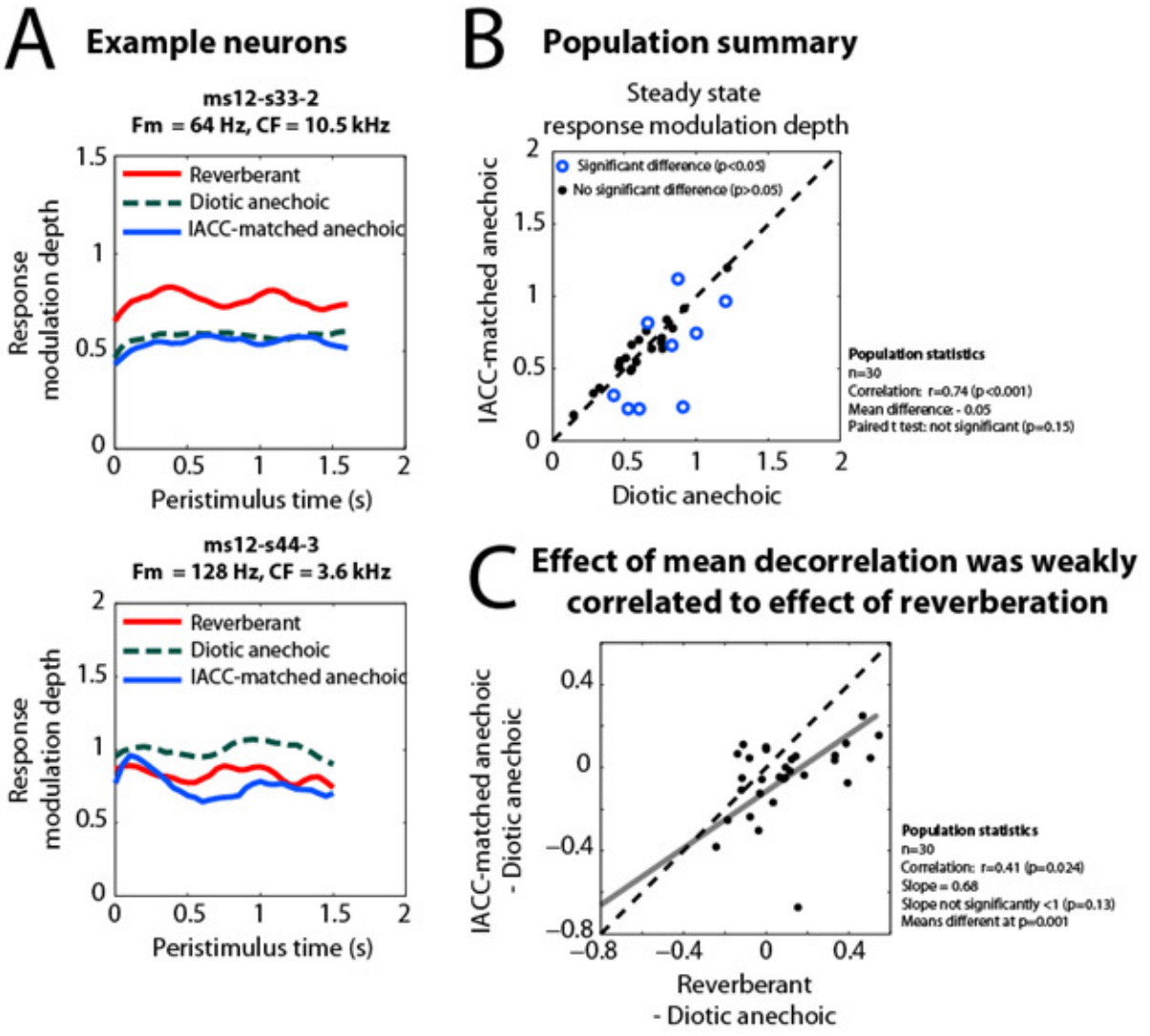

Figure 3.12 Effect of mean interaural decorrelation introduced by reverberation
A. Example neurons showing responses to diotic anechoic stimuli (green, dashed line), reverberant stimuli (red, solid line), and decorrelated anechoic stimuli (blue, solid line) for which IACC was matched to the mean IACC of the reverberant stimulus. *Top panel*: Example where mean IACC did not have an effect. Responses were computed in ≈391 ms time bins with 90% overlap, and smoothed with a 3-point moving average filter. *Bottom panel*: Example where decorrelation significantly decreased response modulation depth and may explain the difference observed between diotic anechoic and reverberant responses. Responses were computed in 500 ms time bins with 90% overlap, and smoothed with a 3-point moving average filter. **B.** Population summary: Comparison between IACC-matched anechoic and diotic anechoic conditions. Points for which response modulation depths were not significant (Rayleigh test of uniformity, $\alpha<0.05$) were omitted. 9 of 30 neurons showed a significant effect of mean IACC. **C.** Comparison between effect of mean IACC and effect of reverberation. Thick, grey line: Direction of first principal component. Pearson correlation was small (0.41) but significant (p=0.024). Slope was not significantly different from 1 (p=0.13), but effect of reverberation was significantly greater than effect of mean IACC (p=0.001).

decreased RMD. Overall, RMD was significantly affected by a static decorrelation in nearly 1/3 of the neurons.

We further tested whether these differences between diotic and IACC-matched anechoic conditions could explain the reverberant advantage across the population. If mean IACC were responsible for the observed differences between reverberant and diotic anechoic RMDs, the effect of mean IACC (defined as the difference between IACC-matched RMD and diotic anechoic RMD) should be similar to the effect of reverberation (defined as the difference between reverberant RMD and diotic anechoic RMD). Figure 3.12 C shows the comparison across the neural population between the effect of mean IACC and the effect of reverberation. The correlation between the two effects was small ($r=0.41$) but significant ($p=0.024$), and the slope of the regression line was not significantly different from 1 ($p=0.13$). However, the mean effect of reverberation was significantly larger than the mean effect of IACC ($p=0.001$). Overall, the effect of reverberation may be partly attributed to the static decorrelation in a subset of neurons.

Effect of Interaural Envelope Disparities (IEDs)

Aside from interaural differences in the temporal fine structure of reverberant stimuli reflected in the lowered IACC, small differences between the envelopes of the left and right ear input signals are also created by reverberation (Fig. 3.5). In the anechoic case, the signals at the two ears are identical, and therefore the envelopes at the two ears are also identical. However, in the reverberant case, a small interaural time difference (ITD_{env}), a small difference in sound level (Interaural Level Difference or ILD) and a small difference in stimulus modulation depth

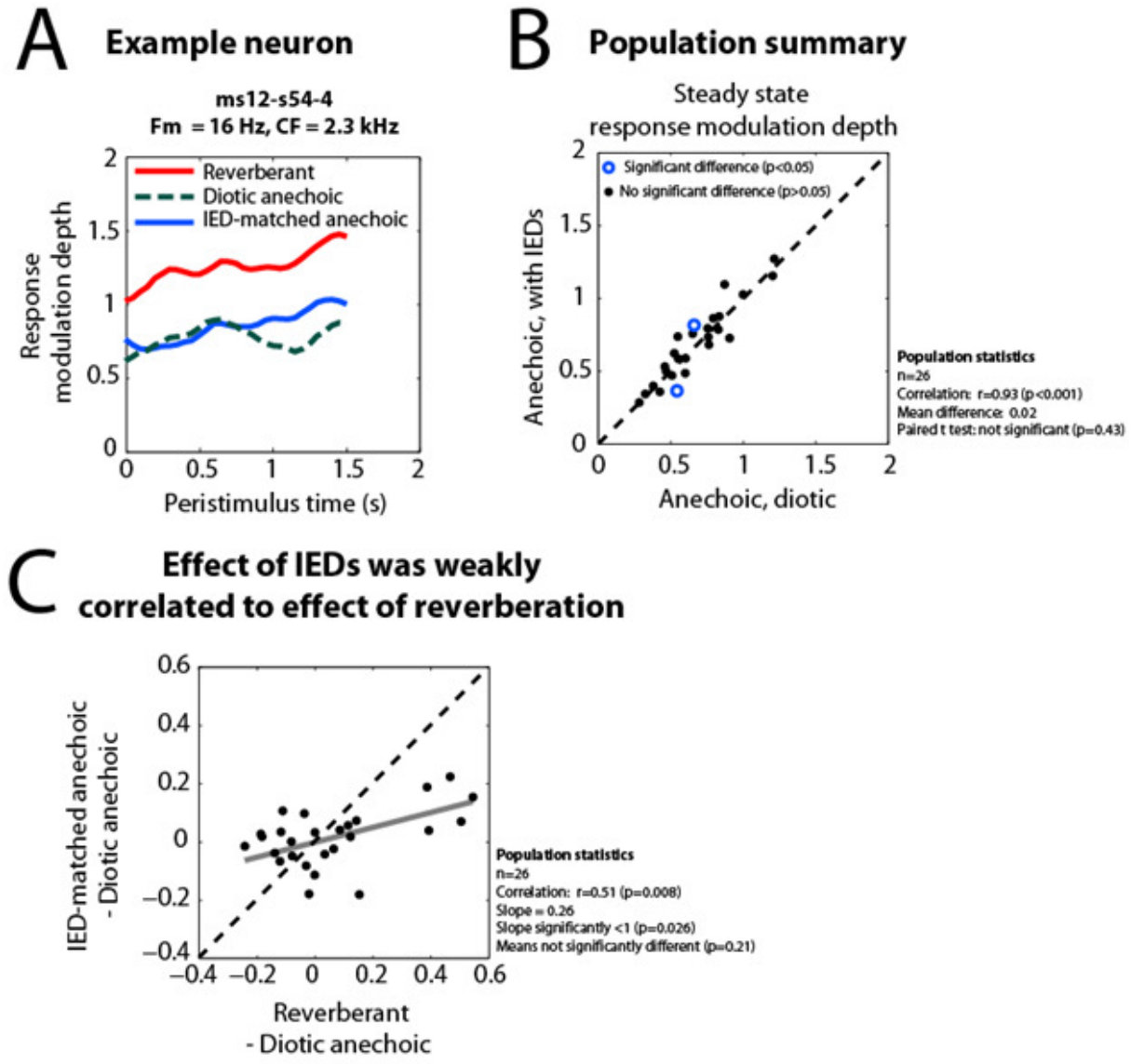

Figure 3.13 Effect of Interaural Envelope Disparities (IEDs) introduced by reverberation
A. Example neuron showing responses to the diotic anechoic stimulus (green, dashed line), the reverberant stimulus (red, solid line), and the IED-matched anechoic stimulus (blue, solid line). Responses were computed in 500 ms time bins with 90% overlap, and smoothed with a 3-point moving average filter. **B.** Population summary: Comparison between IED-matched anechoic and diotic anechoic conditions. Points for which response modulation depths were not significant (Rayleigh test of uniformity, $\alpha<0.05$) were omitted. Only 2 out of 26 neurons showed a significant effect of IEDs. **C.** Comparison between effect of IEDs and effect of reverberation. Thick, grey line: Orthogonal regression line. Pearson correlation was moderate (0.51) but significant (p=0.008). Slope of orthogonal regression was significantly smaller than 1 (p=0.026), but effect of reverberation was not significantly greater than effect of IEDs (p=0.21).

(Interaural Depth Difference or IDD) were present across the two ears (Fig. 3.5, Methods). We regroup these three effects under the term "Interaural Envelope Disparities" (IEDs).

To test the hypothesis that IEDs introduced by reverberation are responsible for the differences observed between reverberant RMDs and depth-matched anechoic RMDs, we synthesized dichotic anechoic stimuli for which the interaural time, level, and modulation depth differences in the envelope matched those in the steady state portion of the reverberant stimuli, at the modulation frequency being used. Figure 3.13 A shows data from an example neuron where the reverberant RMD (red, solid line) was significantly larger than the diotic anechoic RMD (green, dashed line). In this neuron, the RMD profile for the IED-matched anechoic stimulus (blue, solid line) was very similar to that of the diotic anechoic stimulus. Therefore, in this neuron, IEDs introduced by reverberation did not strongly affect the RMD.

Across the population (Fig. 3.13 B), the IED-matched anechoic RMDs were highly correlated to the diotic anechoic RMDs ($r=0.93$, $p<0.001$) and the means were not significantly different across conditions (Paired t-test, $p=0.43$). IEDs had a significant effect ($p<0.05$) in only 2 neurons (out of 26), and these effects were small. Although IEDs did not elicit significant differences in most individual neurons, the overall effect of IEDs (difference between IED-matched anechoic and diotic anechoic RMDs) was moderately correlated to the effect of reverberation ($r=0.51$, $p=0.008$), and their means were not significantly different ($p=0.21$). However, the slope of the orthogonal regression line (≈ 0.26) was significantly smaller ($p=0.026$) than 1, suggesting that the IEDs introduced by reverberation contributed only partially to the differences observed between reverberant and diotic anechoic RMDs.

We also tested the hypothesis that the combined effect of mean IACC and IEDs can explain the differences between reverberant and diotic anechoic RMDs. To do so, we

synthesized dichotic stimuli that were matched to the reverberant stimuli for both mean IACC and IEDs. The RMDs of these IED- and IACC-matched anechoic stimuli were highly correlated with those of the diotic anechoic stimuli ($r=0.71$, $p<0.001$) but 11 out of 26 neurons showed significantly different RMDs between the matched and the diotic conditions. The correlation between the combined effect of IACC and IEDs, and effect of reverberation was moderate ($r=0.53$, $p=0.005$), and not significantly higher ($p=0.59$; test based on the Fisher r to z transformation) than the correlations obtained with IACC or IEDs only (Fig. 3.12 C and 3.13 C). This suggests that the reverberant advantage was not due to an interaction between IACC and IEDs.

Effect of spectral coloration

Filtering by the reverberant BRIRs causes a coloration of the stimulus power spectrum. In contrast to the anechoic impulse responses, which have roughly flat frequency response magnitudes (Fig. 3.6 A, left), the frequency response magnitudes of the reverberant filters consist of tightly superimposed notches corresponding to intervals between individual reflections (Fig. 3.6 A, right). To investigate whether this spectral coloration is responsible for the differences between reverberant and depth-matched anechoic RMDs, we synthesized anechoic stimuli with mean spectrum magnitudes similar to those of the reverberant stimuli (Methods).

Figure 3.14 A shows data from an example neuron with a strong reverberant advantage. In this neuron, the RMD to the spectrum-matched anechoic stimulus (blue, solid line) was very similar to that of the anechoic, white stimulus (green, dashed line), inconsistent with the hypothesis that spectral coloration explains the reverberant advantage. Across the population (Fig. 3.14 B), the spectrum-matched anechoic RMDs were very similar to the white anechoic

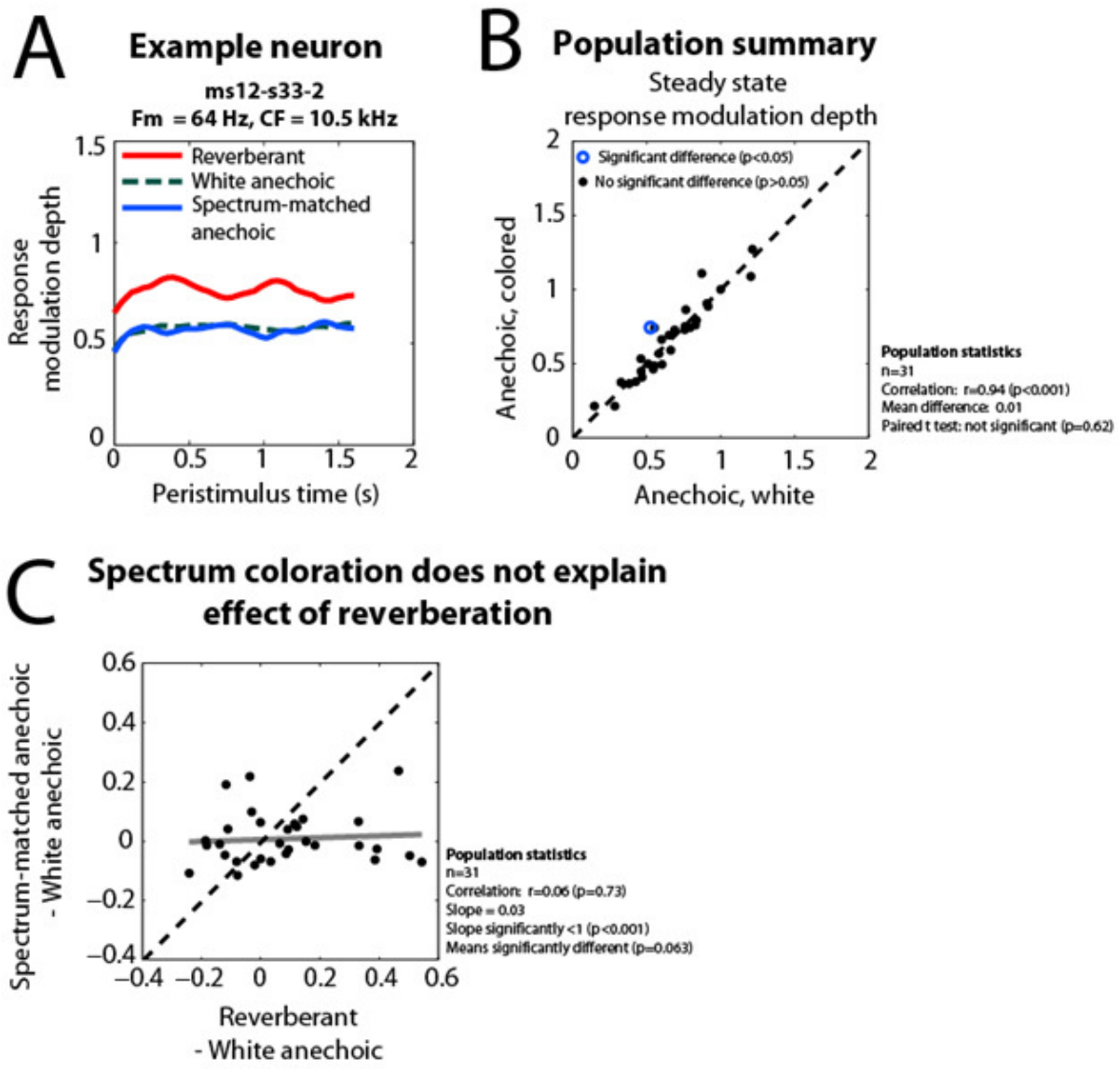

Figure 3.14 Effect of spectral coloration introduced by reverberation
A. Example neuron showing responses to the white anechoic stimulus (green, dashed line), the reverberant stimulus (red, solid line), and the spectrum-matched anechoic stimulus (blue, solid line). Responses were computed in ≈391 ms time bins with 90% overlap, and smoothed with a 3-point moving average filter. **B.** Population summary: Comparison between spectrum-matched anechoic and white anechoic stimuli. **C.** Comparison between effect of coloration and effect of reverberation. Thick, grey line: Direction of first principal component. Pearson correlation was very small (0.06) and insignificant (p=0.73).

RMDs ($r=0.94$, $p<0.001$, only 1 neuron out of 31 with significant differences at $p<0.05$), suggesting that spectral coloration had a negligible effect on RMD. Further, the small differences between spectrum-matched and white anechoic stimuli were not correlated to the differences between reverberant and white anechoic stimuli (Fig. 3.14 C), indicating that spectral coloration alone did not explain the reverberant advantage.

To investigate the combined effect of mean IACC and spectral coloration on RMD, we also synthesized anechoic stimuli that were both spectrum- and IACC-matched to mimic the spectral and static binaural properties of the reverberant stimuli. Across the population, coloration and decorrelation had a significant effect in 11 neurons out of 30, of which 3 showed an increase and 10 showed a decrease in RMD relative to the white, diotic condition. The combined effect of coloration and mean IACC was moderately correlated to the effect of reverberation ($r=0.59$, $p=0.001$). Although this correlation is higher than the correlation between effect of IACC alone and effect of reverberation (Fig. 3.12 D; $r=0.41$), the difference is not significant ($p=0.37$; test based on the Fisher r to z transformation), suggesting that the reverberant advantage is not due to an interaction between IACC and coloration.

Diotic reverberant stimuli also had a coding advantage over depth-matched anechoic stimuli in some neurons

In previous sections, we investigated the influence of binaural factors such as mean IACC and IEDs by incorporating the binaural characteristics of reverberant stimuli into depth-matched anechoic stimuli. Another way of investigating the influence of binaural characteristics on the reverberant RMDs is to make a reverberant stimulus diotic. If the differences in RMD between reverberant (dichotic) stimuli and depth-matched anechoic (diotic) stimuli disappear when the

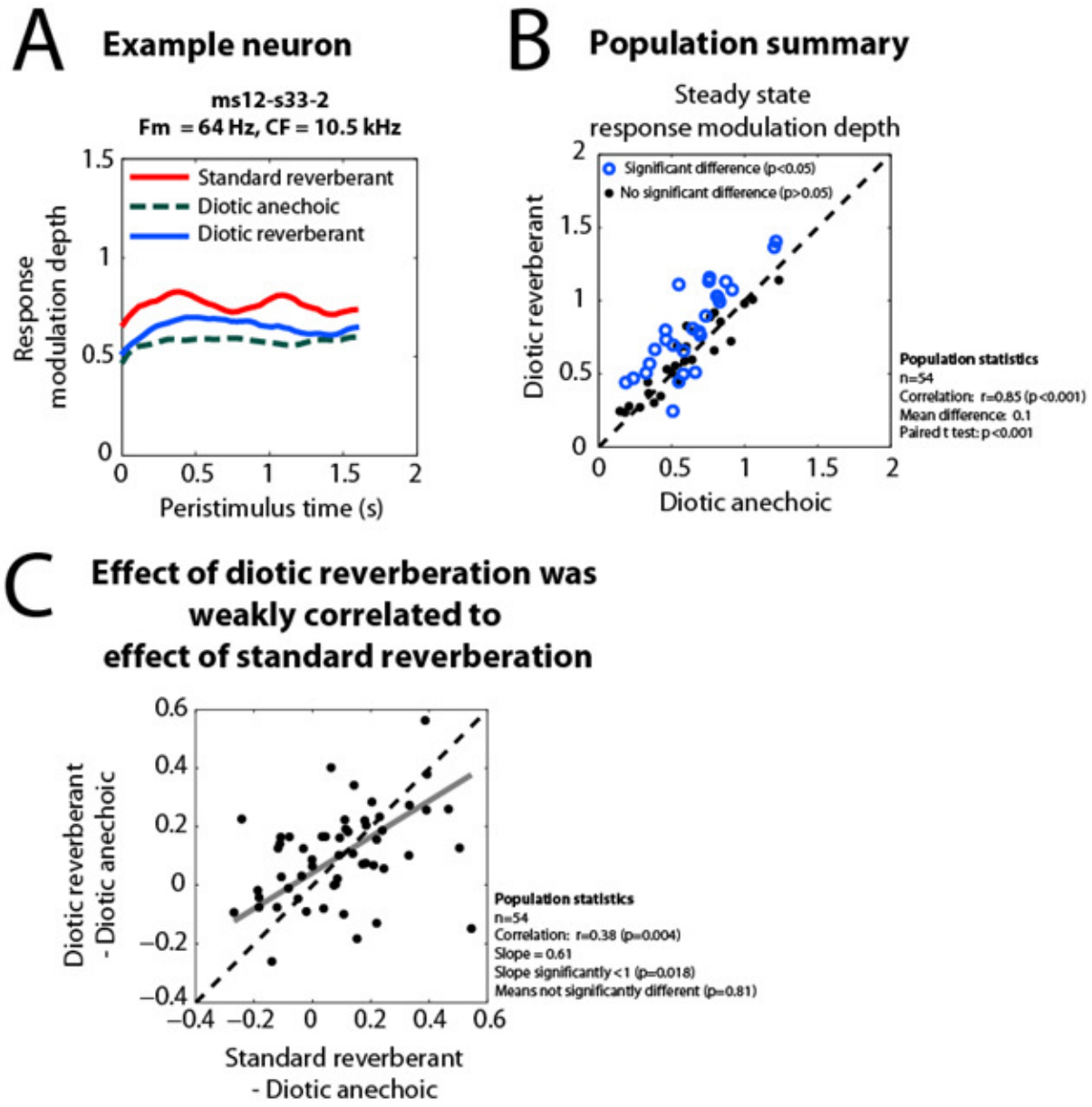

Figure 3.15 Diotic reverberant stimuli had a coding advantage over depth-matched anechoic stimuli

A. Diotic reverberant stimuli were synthesized by using the same reverberant filter for both left and right ears. Example neuron showing responses to the diotic anechoic stimulus (green, dashed line), the standard reverberant stimulus (red, solid line), and the diotic reverberant stimulus (blue, solid line). Responses were computed in ≈391 ms time bins with 90% overlap, and smoothed with a 3-point moving average filter. **B.** Population summary: Comparison between diotic reverberant and anechoic stimuli. In half of the neurons (27 out of 54), the diotic reverberant stimuli had a significant coding advantage over the anechoic stimuli with the same modulation depth. **C.** Comparison between effect of diotic reverberation and effect of standard reverberation. Thick, grey line: Direction of first principal component. Pearson correlation was small (0.38) but significant (p=0.004).

reverberant stimuli are made diotic, these differences could be attributed to binaural properties of the reverberant stimuli. We synthesized diotic reverberant stimuli as described in the Methods. The resulting reverberant stimuli were therefore identical in both ears, although they possessed the temporal and spectral characteristics inherited from realistic reverberant room impulse responses.

Figure 3.15 A shows results from an example neuron with a strong reverberant advantage. In this neuron, the RMD to the diotic reverberant stimulus (blue, solid line) was intermediate between the RMD to the diotic anechoic stimulus (green, dashed line) and the RMD to the standard reverberant stimulus (red, solid line). This result is inconsistent with the hypothesis that binaural effects alone can explain the reverberant advantage, since diotic reverberation also leads to an advantage. Across the population (Fig. 3.15 B), the diotic reverberant stimuli had a significant coding advantage ($p<0.05$) over depth-matched anechoic stimuli in half of the neurons (27 out of 54), while 4 neurons showed a significant coding disadvantage in diotic reverberation, and 23 did not show a significant difference. The mean difference between diotic reverberant and diotic anechoic was 0.1 ($p<0.001$, paired t-test).

We also compared the effect of diotic reverberation to the effect of standard reverberation (Fig. 3.15 C). The two conditions showed a moderate positive correlation ($r=0.39$, $p=0.004$), indicating that monaural factors also contribute to the differences observed between standard reverberant and depth-matched anechoic stimuli. The relatively low correlation is consistent with binaural phenomena such as IACC and IEDs playing an important role in the reverberant RMD.

Effect of early vs. late reflections of the room impulse response

To identify the stimulus factors underlying the diotic reverberant advantage documented in the previous section, we asked whether neurons can exploit the temporal structure of the reverberant filter to partially compensate for the acoustic degradation in modulation. To test this hypothesis, we manipulated the reverberant room impulse response (Fig. 3.7, Methods): In one case, we kept the direct sound and the early reflections, but removed the later reflections ("No late reflections" impulse response, in orange); In another case, we kept the direct sound and the late reflections, but removed the earlier ones ("No early reflections" impulse response, in blue).

Results from an example neuron are shown in Figure 3.16 A. In this example, the diotic reverberant RMD (red, solid line) was significantly larger than the depth-matched anechoic RMD (green, dashed line). The "No early reflections" response (blue, solid line) was also larger than the depth-matched anechoic RMD, but lower than the reverberant RMD. Finally, the "No late reflections" RMD was more similar to the depth-matched anechoic RMD than to the reverberant RMD. This example demonstrates that RMD can be affected by the temporal structure of the impulse response. Across the population (Fig. 3.16 B), both types of truncated stimuli had a small advantage over the depth-matched anechoic case on average (paired Student's t-test; $p=0.05$ for the "No late reflections" condition; $p=0.017$ for the "No early reflections" condition). In the "No late reflections" manipulation, only 4 neurons (out of 31) had a significant advantage over the anechoic condition. In contrast, the "No early reflections" manipulation led to significant differences over anechoic in 11 out of 32 neurons, 8 of which were an advantage. The effect of each manipulation (relative to anechoic) was compared to the effect of diotic reverberation (Fig. 3.16 C). The effect of the "No early reflections" manipulation was significantly more correlated with the effect of diotic reverberation than the effect of the

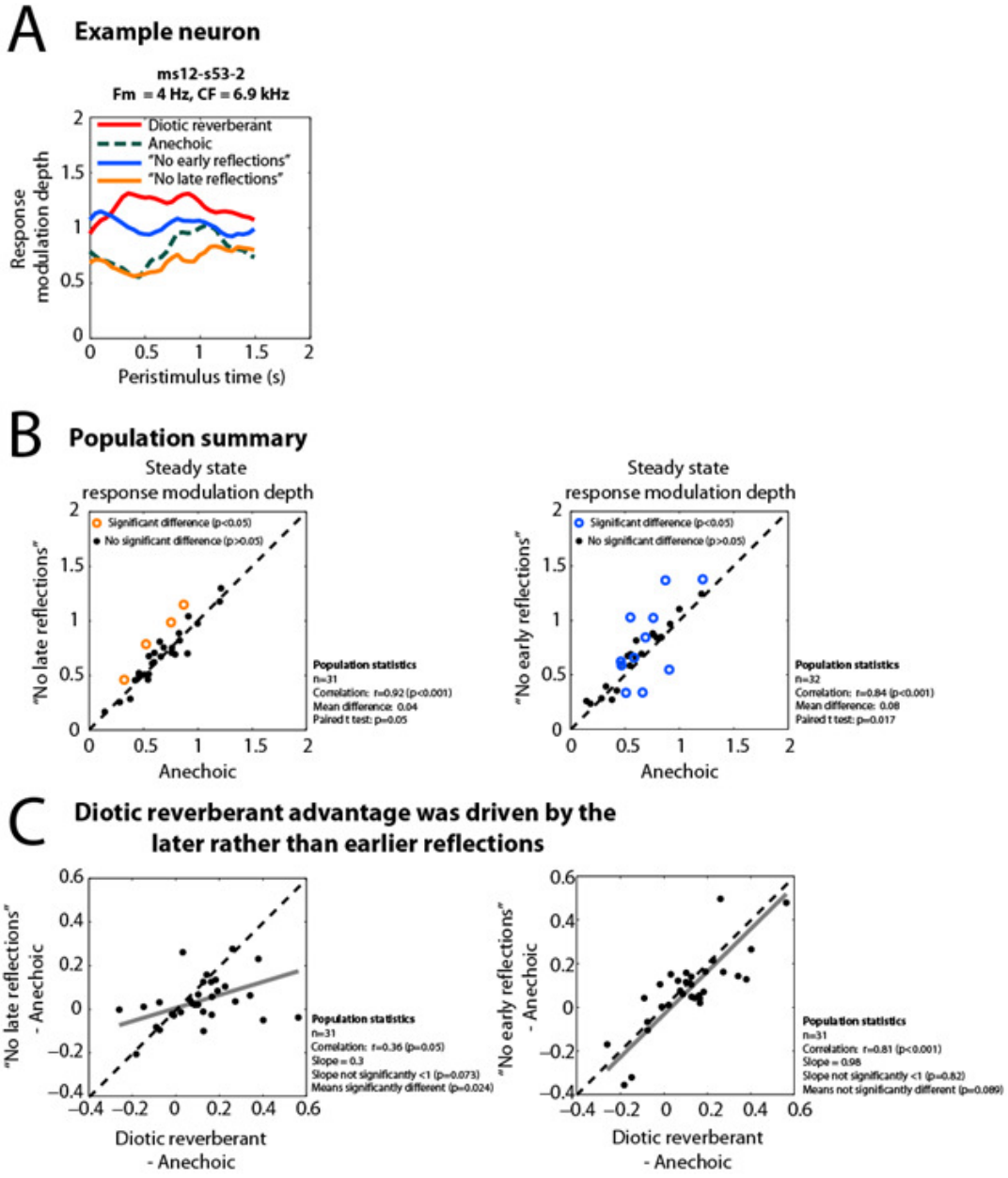

Figure 3.16 Effect of earlier vs. later reflections of the room impulse response
A. Example neuron showing responses to the anechoic stimulus (green, dashed line), the diotic reverberant stimulus (blue, solid line), the "No late reflections" reverberant stimulus (orange, solid line), and the "No early reflections" reverberant stimulus (brown, solid line). Responses were computed in 500 ms time bins with 90% overlap, and smoothed with a 3-point moving average filter. **B.** Population summary. *Left*: Comparison between "No late reflections" and anechoic stimuli. *Right*: Comparison between "No early reflections" and anechoic stimuli. The "No early reflections" stimuli had a significant coding advantage over the anechoic stimulus in more neurons than the "No late reflections" stimuli. **C.** *Left*: Comparison between effect of "No late reflections" reverberation and effect of diotic reverberation. *Right*: Comparison between effect of "No early reflections" reverberation and effect of diotic reverberation. Thick, grey lines: Direction of first principal components. Correlation was significantly stronger (p=0.005) for the comparison with "No early reflections" reverberation, than for the comparison with "No late reflections" reverberation, suggesting that later rather than earlier reflections mediated the diotic reverberant advantage.

"No late reflections" manipulations (r=0.81 for the "No early reflections" stimuli, and r=0.36 for the "No late reflections" stimuli; test based on the Fisher r to z transformation at p=0.005). This suggests that later, rather than earlier reflections mediate the diotic reverberant advantage.

Comparison of predictions of the reverberant advantage across the neural population

In the previous sections, we synthesized different stimuli that mimicked some of the characteristics of reverberant stimuli (envelope waveform, spectral coloration, mean IACC, IEDs). We compared the effect of reverberation to the effect of one or a combination of these features by performing an orthogonal linear regression analysis (Methods). The correlation coefficient and the slope of the regression line should be near 1 if the feature or combination of features explains the effect of reverberation across the neural population. In Figure 3.17, we show the slope of the regression line against the correlation coefficient across stimulus conditions. In the left panel, this comparison was done using the data from the entire neural population. The condition for which the slope was closest to 1 was the IED, IACC-matched anechoic condition (slope=0.84, not significantly different from 1, p=0.37). The correlation coefficient for this condition was 0.53. The spectrum- and IACC-matched condition led to a somewhat larger correlation coefficient (r=0.59), but this correlation was not significantly larger than that for the IED- and IACC-matched condition. The slope for the spectrum- and IACC-matched condition was smaller (0.63) than for the IED- and IACC-matched condition, and significantly different from 1 (p=0.023). This suggests that the combined effect of IEDs and IACC best predicts the effect of reverberation across the neural population. However, the variance explained with this combination of features is still rather small (r^2 =0.28) and the effect of reverberation is significantly larger than the combined effect of IEDs and IACC

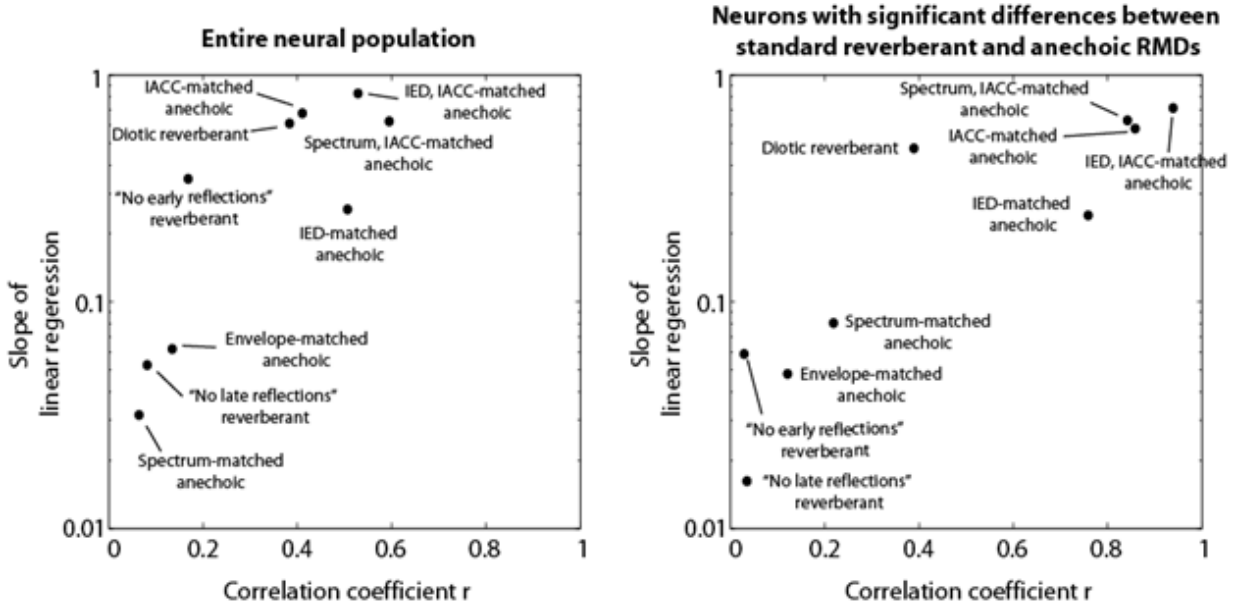

Figure 3.17 Comparison of predictors of the reverberant advantage
In the previous sections, the effect of reverberation (difference between reverberant RMD and anechoic RMD) was compared to the effect of a stimulus imitating some of the features of the reverberant stimulus (difference between featured stimulus RMD and anechoic RMD). For each featured stimulus, we computed the slope of the regression line obtained with PCA analysis (Methods) and the Pearson correlation coefficient. A large correlation coefficient and a slope close to 1 are consistent with the featured stimulus partly explaining the effect of reverberation. *Left panel:* Scatter plot of the regression slopes and correlation coefficients when including the entire neural population. *Right panel:* Same comparison for the subset of neurons for which the effect of reverberation was significant ($p<0.05$).

($p=0.014$). Therefore, IEDs and IACC only partially explain the effect of reverberation across the population.

We performed a similar analysis using only data from neurons for which there was a significant difference ($p<0.05$) between reverberant and depth-matched anechoic RMDs (Fig. 3.17, right panel). The 3 conditions that best explained the significant effects of reverberation on the basis of their correlation coefficients and the slopes of their regression lines were again the 3 conditions with matched IACC. The IED- and IACC-matched condition both had the highest correlation coefficient ($r=0.94$) and the slope closest to 1 (slope=0.72, not significantly different

from 1 with p=0.19). However, the effect of reverberation was significantly larger than the combined effect of IEDs and mean IACC ($p<0.001$). Again, this suggests that our best predictor of the effect of reverberation only partially explains the effect.

Discussion

Our experiments show that in a subset of neurons, reverberant stimuli have a coding advantage over anechoic stimuli with the same modulation depth. We found that envelope distortion and spectral coloration introduced by reverberation had negligible effects on response modulation depth (RMD), whereas binaural features such as IEDs and mean IACC explained part of the observed differences between reverberant and depth-matched anechoic RMDs. Moreover, diotic reverberant stimuli also had a coding advantage over depth-matched anechoic stimuli, and truncating reverberant room impulse responses had an effect on temporal coding in some neurons. This suggests that some IC neurons may exploit the temporal properties of reverberant room impulse responses to partially compensate for the AM degradation.

Time course of reverberant stimuli: Is there an onset advantage?

One of our initial motivations for looking at the time course of reverberant RMD was the finding by Devore *et al.* (2009) that directional sensitivity of IC neurons follows a similar time course as the degradation in sound localization cues in the stimulus. Further, they showed that simple mechanisms such as firing rate adaptation helped emphasize the early, non degraded part of the response, over the ongoing, fully degraded response. We hypothesized that this resemblance in time course between stimulus and response might translate to AM coding in reverberation, and that simple mechanisms of "onset dominance" might similarly help emphasize the early, less degraded portion of the response.

At first sight, several neurons seemed to follow, at least in the first 250 ms, a profile similar to stimulus modulations. For example, neurons presented in Figures 3.8 and 3.9 show an initially high RMD, followed by a rapid decay paralleling the prediction from the MIOF and the

time course of stimulus modulation depth. In other neurons, (Fig. 3.12 A, Fig. 3.13 A), the reverberant RMD did not start at a high value, although the longer integration window, necessary to provide a significant estimate of RMD, may have blurred a robust onset RMD. Across the population, we compared the RMD computed in a short time window at onset (of duration the smallest integer number of modulation cycles greater or equal to 20 ms) to the ongoing RMD computed in the remaining portion of the response. On average, onset RMD was larger than ongoing RMD (Paired Student's t-test, $p<0.001$) consistent with the time course of stimulus modulations. However, the same analysis performed on the static, depth-matched anechoic stimuli, led to a similar result. For example in Figure 3.11 A, the RMD to the depth-matched anechoic stimulus (green, dashed line), for which stimulus modulation depth was constant over time, had a significant onset advantage similar to that of the reverberant RMD. Across the population, reverberant and anechoic onset advantages were highly correlated ($r=0.87$, $p<0.001$), suggesting that they are due to properties of the neuron rather than to the strong modulations in the earliest portion of the reverberant stimuli. For example in the reverberant PSTH of Fig. 3.9 D, a clear notch is present immediately following a burst of activity at the onset of the stimulus. This notch is shorter than a stimulus period, and present in the anechoic response as well. While the notch modulates the response, and therefore elicits a large RMD, it is not due to the stimulus, but to an intrinsic firing pattern of the neuron.

Therefore, the hypothesis that AM coding in reverberation is helped by robust coding near stimulus onset, where the degradation is minimal, is not supported, at least for the low modulation frequencies we investigated.

Negligible effects of envelope distortion and spectral coloration

In this study, envelope distortions and spectral coloration introduced by reverberation had very little effect on RMD compared to the sinusoidal and white anechoic conditions (Fig. 3.11 A-B, Fig. 3.14 A-B). Moreover, the small differences observed relative to the sinusoidal and white anechoic condition were uncorrelated to the differences observed between reverberant and anechoic conditions (Fig. 3.11 C, Fig.3.14 C).

At first sight, our finding of a negligible effect of envelope distortion seems to be in contradiction with a study by Sinex *et al.* (2002) in IC of chinchilla, who showed that the synchronization of IC neurons to the modulation depended on the modulation waveform. However, their study compared rather large envelope differences (sinusoidal vs. trapezoid shape) whereas our reverberant BRIRs only introduced small envelope distortions (Fig. 3.3). On the other hand, the resemblance between anechoic and reverberant envelopes in our study may have been biased by our use of sinusoidal modulations. It is possible that using more asymmetric modulation waveforms, such as the sharp amplitude transitions common in speech would result in greater distortion of the envelope.

The lack of effect of spectral coloration on RMD is unsurprising given the small differences in the average magnitude spectra integrated over frequency bands equivalent to rabbit peripheral filter bandwidths: For example in Figure 3.6 C, the differences between reverberant and anechoic magnitude spectra were less than 1 dB at most frequencies. While sound level can have a strong effect on phase-locking to modulation in some IC neurons, changes in RMD are usually less than 0.2 for 10 dB increments (Krishna and Semple, 2000), which is consistent with our observations that small level differences of less than 1 dB did not affect RMD.

Effects of binaural features on AM coding

An unexpected result of this study was the influence of binaural features on temporal coding of AM in a subset of neurons. We found that the effect of IEDs, the effect of IACC, and the combined effect of IEDs and IACC, were moderately correlated to the effect of reverberation (Fig. 3.8 D and Fig. 3.9 D; Results). Overall, the combined effect of IEDs and IACC was the best predictor of effects of reverberation relative to depth-matched anechoic (Fig. 3.17).

The effect of IACC on RMD was larger than the effect of IEDs. In a majority of neurons for which IACC had an effect (7 out of 9), the decorrelation caused a decrease in RMD, compared to the diotic anechoic condition. It is unclear what mechanism could explain this effect. We tested whether a simple coincidence detection mechanism could explain the effect by computing the short-term correlation function for the IACC-matched and diotic anechoic stimuli with integration windows and delays ranging from 0.01 to 10 ms. The modulation depth of the output of the correlation function was nearly identical for the two stimuli. This suggests that a simple coincidence detection mechanism receiving delayed inputs from the two ears cannot account for the decrease in RMD with a decrease in IACC.

Limitation due to differences between broadband and narrowband binaural features

A potential limitation of our study is that we used broadband stimuli to compute the reverberant IEDs and IACC. In the auditory system, however, acoustic signals are first band-pass filtered in the cochlea before being transduced into neural impulses in the auditory nerve. The IEDs and IACC that an IC neuron "sees" are therefore based on the band-pass rather than broadband waveforms. We modeled the filtering effect of the cochlea by band-pass filtering our reverberant stimuli with gammatone filters of equivalent rectangular bandwidths similar to those

of rabbit auditory nerve tuning curves (Borg *et al.*, 1988). IEDs and IACC were computed from the filtered reverberant stimuli exactly as in the broadband case (Methods). The narrowband IEDs and IACC differed somewhat from the broadband case at the characteristics frequencies we encountered (usually > 1 kHz): The median absolute difference between narrowband and broadband case was ≈ 0.001 cycles for IPD_{env}, ≈ 0.03 for IDD, and ≈0.05 for IACC.

Although these differences are small, it is possible that using anechoic stimuli with the binaural features of the narrowband reverberant stimuli would have resulted in better predictions of the effect of reverberation. However, the choice of which IACC and IEDs to use is not trivial, as these binaural features depend on the exact bandwidth and shape of the band-pass filter, which is unknown and nonlinear.

Can dynamic changes in IACC explain the reverberant advantage?

Our best predictors of the reverberant advantage (combined effect of IEDs and IACC) only partially explained the effect across the neural population, in that the correlation coefficient between the combined effect of IED and IACC and the effect of reverberation was only 0.53. Another feature of the reverberant stimuli is the dynamic nature of IACC. As shown in Figure 3.4, IACC oscillates at the modulation frequency of the stimulus: When the instantaneous amplitude of the sound source is large, the reverberant stimulus at the ear drum is dominated by the direct sound, thereby increasing IACC, whereas when the amplitude of the sound source is small, the direct-to-reverberant energy ratio decreases, thereby decreasing IACC. The depth of this modulation in IACC[1] depended on f_m and reverberation condition (range was 0.08 – 0.33 in

[1] We defined the depth of IACC modulation similarly to the depth of amplitude modulation, as twice the magnitude of the Fourier component at the modulation frequency, divided by the DC component.

moderate reverberation, 0.07 – 0.36 in strong reverberation). There was a small phase difference between the modulation in IACC and the modulation in the envelope (< 0.1 cycles). This phase difference depended on f_m.

It is likely that this modulation in IACC at the envelope frequency has an influence on reverberant RMDs. Joris *et al.* (2006) showed that neurons in the IC of anesthetized cat phase-lock to the frequency of sinusoidal oscillations in IACC of unmodulated broadband noise. At the modulation frequencies we used, they routinely found RMDs as high as 1 – 1.5. However, in their case, the IACC oscillated between 0 and 1 (i.e. the modulation depth of IACC was 1) whereas the modulation depth of IACC was smaller in our case. Moreover, Joris *et al.* used unmodulated noise, whereas IACC and envelope are both modulated as the same frequency in our case. It is possible that the reverberant advantage is due to an interaction between envelope and IACC modulations. We compared the effect of reverberation to the modulation depth of stimulus IACC across our population, and found a small positive correlation (r=0.44, p=0.01). It is not surprising that this correlation is small, as even for large modulations of IACC, the effect on RMD likely depends on the neuron's binaural properties as well as on the stimulus.

Possible explanations for the diotic reverberant advantage

The finding of a significant advantage for diotic reverberation over depth-matched anechoic stimuli (Fig. 3.15) is surprising, because differences in binaural features of the stimuli cannot explain the effect. The fact that truncating the reverberant BRIRs had an impact on RMD in some neurons (Fig. 3.16) suggests that some IC neurons are sensitive to the temporal structure of the BRIRs. Theoretically, an echo cancellation mechanism could be implemented to counteract the effect of reverberation: An array of coincidence detector neurons could perform

an autocorrelation of the input, and the timing of the main reflections of the reverberant filter could be extracted from the maxima of the autocorrelation. A simple neural circuit with inhibited inputs delayed to match the timing of the main reflections, and added to the spike train coming from the auditory periphery, could enhance the RMD by effectively cancelling the responses to sound reflections. However, such a mechanism seems unlikely as the delays to apply to the inhibitory input would need to be computed with high speed and precision.

Another potential explanation for the diotic reverberant advantage is an effect of the dynamic pattern of amplitude modulations: Contrary to the depth-matched anechoic stimuli, which have a constant, typically small, modulation depth throughout the stimulus duration, the reverberant stimuli start off with high modulation depths which rapidly decay toward a plateau at a lower modulation depth. The onset response to the large modulation depth might affect the later response. We tested this hypothesis in a few neurons by presenting an anechoic stimulus whose onset was fully modulated, and followed (typically after 62.5 or 125 ms) by a segment with modulation depth matched to the reverberant plateau. Even in neurons for which the reverberant advantage was large, these experiments did not show any significant differences between the static anechoic stimuli, and the dynamic anechoic stimuli with large onset modulation. Furthermore, in a study on forward masking of AM in IC of unanesthetized rabbit, Wojtczak *et al.* (2010) have shown that effects of a fully modulated precursor on the RMD of a partially-modulated subsequent signal were consistent with a decrease in phase-locking, rather than an increase; these effects were found in only a small subset of neurons.

Another hypothesis to explain the diotic reverberant advantage is related to the observation, detailed in the previous section, that IC neurons phase-lock to the frequency of IACC oscillations. Although the diotic reverberant stimuli have a constant peak IACC of 1,

introducing a small delay between the two ears results in interaural decorrelation. We computed the short-term cross-correlation function for diotic reverberant stimuli and found that for some delays, IACC oscillates at the modulation frequency. Such oscillations in IACC were usually not present in the anechoic case, suggesting that for a binaural cell with an appropriate internal delay, phase-locking to modulation frequency may also be enhanced in diotic reverberation due to oscillations in IACC.

Overall, our results point to the importance of interactions between binaural processes and coding of AM in reverberation. They further suggest that IC neurons may be sensitive to the temporal structure of reverberant impulse responses, and may use this information to partially compensate for the degradation in AM.

Chapter 4

Dynamic effects of stimulus statistics on temporal coding of amplitude modulation: Implications for reverberant environments

Abstract

In chapters 2 and 3, we assumed that the coding of amplitude modulation (AM) can be described by a static nonlinearity, the Modulation Input/Output Function (MIOF), primarily dependent on stimulus modulation depth. In this chapter, we study the effect of stimulation history on the temporal coding of AM. This is relevant to reverberation, as reverberant environments have different modulation depth statistics than anechoic environments. In particular, the mean stimulus modulation depths are lower in reverberant environments than in anechoic environments, due to the degradation of amplitude envelopes.

We measured MIOFs of IC neurons using continuous Sinusoidally Amplitude Modulated (SAM) broadband stimuli, whose modulation depth was dynamically drawn from specific distributions designed to imitate the statistical characteristics of modulation depths in anechoic and reverberant environments. We found that modulation depth statistics had a significant effect on the MIOF in a subset of neurons. Neural sensitivity to changes in stimulus modulation depth was not systematically affected across the population. However, on average, temporal coding of modulation frequency was stronger in conditions when low modulation depths predominate, as in reverberant environments. These effects occurred over a wide range of time scales (up to ≈13 seconds).

Overall, our findings point to the importance of considering stimulation history in investigating the temporal coding of AM.

Introduction

The auditory system adapts its coding strategy to the statistics of the stimulus features of interest. In the Inferior Colliculus (IC) of anesthetized guinea pig, Dean *et al.* (2005) reported that rate-level functions rapidly adjust to optimize the coding of the most probable sound levels in a dynamic noise or tone stimulus. The time course of this dynamic range adaptation is of the order of hundreds of milliseconds (Dean *et al.*, 2008). Wen *et al.* (2009) used similar stimuli in the auditory nerve of anesthetized cat, and showed that dynamic range adaptation to sound level statistics originates in the auditory periphery, although the peripheral effects are not as marked as they are in the midbrain. In the auditory cortex of anesthetized and unanesthetized ferrets, Rabinowitz *et al.* (2011) used dynamic random chord sequences to study the effect of spectrotemporal contrast, which was varied by changing the variance of the distribution of levels. Neural sensitivity to changes in the stimulus increased when contrast was low, on a timescale of hundreds of milliseconds. Similarly to these studies on the coding of sound level, Dahmen *et al.* (2010) demonstrated in IC of anesthetized ferrets that the coding of interaural level differences (ILD), important for sound localization, adapts to the mean and variance of the ILD in the stimulus set. Their neurophysiological findings were complemented by human psychophysical studies, which showed that the perceived midline was biased in the direction of the most frequent range of ILDs, and that sensitivity to changes in ILD decreased with increasing variance of the ILD distribution.

Reverberation degrades the amplitude envelope of acoustic stimuli. Therefore, the statistical properties of AM differ between reverberant and anechoic environments. In particular, since modulation depths decrease in reverberation, the probability of weak to moderate modulation depths in a reverberant environment will be higher than the probability of large

modulation depths. For a given sound source, the smaller the direct-to-reverberant energy ratio, the lower the mean modulation depths.

Adaptation to the statistics of amplitude modulated stimuli was studied in IC of anesthetized cats (Kvale and Schreiner, 2004). Sinusoidal carriers were modulated with a step-wise rectangular function (800 steps/sec) whose amplitudes were drawn from a Gaussian distribution. The effect of the variance of this distribution on firing rate as well as on temporal receptive fields was investigated. They found small changes in the latency and amplitude of the receptive fields, which occurred within hundreds of milliseconds. However, this study focused on the effects of changing the variance of a distribution of amplitudes, rather than investigating the effects of changing the mean of the distribution.

The questions we ask in this study are two-fold: (1) Do IC neurons adapt their temporal coding of AM to the changes in the mean of a distribution of stimulus modulation depth? (2) Does this adaptation help the coding of AM in reverberant environments? To answer these questions, we recorded from single units from unanesthetized rabbit IC in response to Sinusoidally Amplitude Modulated (SAM) broadband noise stimuli. The modulation depth was randomly drawn every 250 ms from distributions designed to imitate the properties of anechoic and reverberant environments. We find that stimulus modulation statistics have a significant effect on the Modulation Input/Output Functions (MIOFs) in a subset of neurons. On average, response modulation depths (RMD) were larger in the reverberant-like condition than in the anechoic-like condition. In contrast, sensitivity to changes in stimulus modulation depth was not systematically affected across the population. The time course of adaptation to modulation depth statistics was slow in some neurons ($\approx 3 - 13$ seconds) and faster in other neurons (< 3 seconds).

Methods

Surgical Preparation and Recording Procedures

Methods for chronic recordings of IC single units from unanesthetized dutch-belted rabbits (*Oryctolagus cuniculus*) were based on the techniques of Kuwada *et al.* (1987), Nelson and Carney (2007), and Devore and Delgutte (2010), and are described in Chapter 2 (Methods). All procedures were approved by the Institutional Animal Care and Use Committees of the Massachusetts Eye and Ear Infirmary and the Massachusetts Institute of Technology.

Dynamic Stimuli

Stimuli were Sinusoidally Amplitude Modulated (SAM) broadband noise with a fixed modulation frequency f_m chosen between 8-128 Hz to elicit both a large firing rate and strong phase-locking to the modulation. Modulation depth was drawn randomly, usually every 250 ms, from one of 3 non-uniform distributions designed to simulate 3 environments with various degrees of reverberation (Fig. 4.1 A-C). For each distribution, modulation depths could take 20 different values (from m = 0.05 to m = 1, in 0.05 increments). In the "anechoic-like" distribution, a region of high probability was defined at large stimulus modulation depths ($m \geq 0.8$) to simulate an environment in which envelopes are strongly modulated. In the two "reverberant-like" distributions, the region of high probability was defined at moderate ($0.4 \leq m \leq 0.6$) and low ($0.05 \leq m \leq 0.25$) stimulus modulation depths, to simulate environments in which envelopes are degraded by reverberation, hence increasing the probability of weak to moderate modulations. In all three distributions, 80% of the probability density was concentrated in the region of high probability. The high-probability regions in the reverberant-like conditions

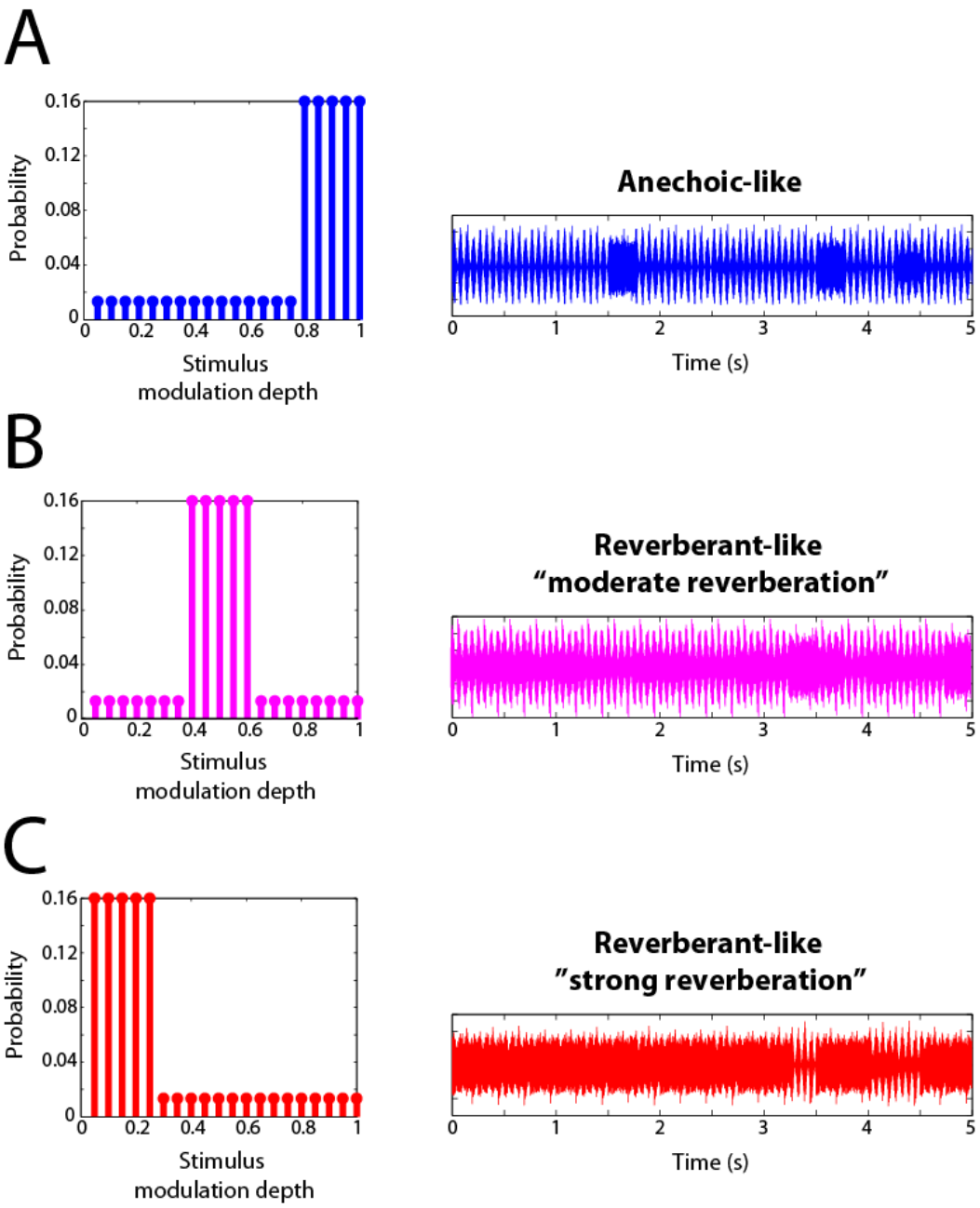

Figure 4.1 Distributions of modulation depth used to generate the dynamic stimuli
A. *Left*: Anechoic-like distribution. Stimulus modulation depths are drawn 80% of the time from a region of large modulations ($0.8 \leq m \leq 1$), simulating an anechoic environment where envelopes are sharp. *Right:* Example SAM noise (f_m=16 Hz, first 5 seconds) generated from the anechoic-like distribution. Stimulus modulation depth changes every 250 ms. **B.** *Left*: Reverberant-like distribution simulating moderate reverberation. Stimulus modulation depths are drawn 80% of the time from a region of moderate modulations ($0.4 \leq m \leq 0.6$), simulating a reverberant environment where envelopes are moderately degraded. *Right:* Example stimulus (f_m=16 Hz, first 5 seconds). **C.** *Left*: Reverberant-like distribution simulating strong reverberation. Stimulus modulation depths are drawn 80% of the time from a region of small modulations ($0.05 \leq m \leq 0.25$), simulating a reverberant environment where envelopes are strongly degraded. *Right:* Example stimulus (f_m=16 Hz, first 5 seconds).

roughly correspond to the attenuated modulation depths of fully modulated sound sources in the moderate and strong reverberant environments used in Chapters 2 and 3.

To avoid any bias in our computation of response modulation depths due to windowing, the time interval in which stimulus modulation depth was constant was always an integer number of modulation cycles. In most cases, this interval was 250 ms. In some cases, it was 200 ms to accommodate frequencies of 45 and 90 Hz.

In early experiments (Fig. 4.2), the anechoic-like and reverberant-like stimuli were presented sequentially (usually strong-reverberant condition first, followed by the anechoic-like condition, and, time permitting, the moderate reverberant-like condition). In later experiments (Fig. 4.3 – 4.7), anechoic-like and strong reverberant-like distributions were alternated at regular intervals in a continuous "switching" stimulus. For this switching paradigm, distributions were usually switched every 37.5 seconds, although in some neurons, switches could be either less frequent (every 75 seconds) or more frequent (every 18.75 seconds). The first distribution to start the stimulus was selected at random for each neuron.

Experimental Procedures

Experimental procedures for isolating single units, measuring rate level functions, determining Characteristic Frequency (CF), and choosing a sound level, were similar to those in Chapter 2. Neural Modulation Transfer Functions (MTFs) were obtained in response to static anechoic stimuli separated by silent intervals as described in Chapter 2.

Responses to dynamic stimuli were usually measured at one modulation frequency. If time allowed, recordings were made at another modulation frequency. Stimuli were 3-5 minute long in the experiments where we presented the anechoic-like and reverberant-like stimuli

separately. In the experiments using the switching stimuli, the total stimulus duration ranged from 90 seconds to 18 minutes, with a median of 8 minutes. Stimulus duration differed between neurons because we continued to record as long as single-unit isolation was good. When recording quality degraded (usually due to a movement of the animal), we interrupted the recording, and excluded the action potentials following the degradation from data analysis.

Data Analysis

Modulation Input/Output Functions (MIOFs) and Rate-Depth Functions (RDFs)

Modulation Input/Output Functions (MIOFs) characterize the nonlinear transformation of stimulus modulation depth *m* into neural response modulation depth (RMD). These functions were measured and fitted as described in Chapter 2. The RMD was directly computed from the spike times as twice the vector strength at f_m.

For stimuli periodically switching between the anechoic-like and reverberant-like distributions, we computed the MIOF in the steady-state portion of the response (after the neuron has adapted to the new distribution) by removing the spikes occurring in a fixed time window following each switch. This time window was specific to each neuron and chosen to exceed the time constant of the effect (see "Time course" section). In most cases, we removed 3 seconds of data following each switch. In some cases, when the adaptation to modulation depth statistics was slow, we removed up to 13.5 seconds of data.

Spike counts for each *m* were averaged across trials, and mean firing rate was fitted as function of *m* using a sigmoidal curve, to construct Rate-Depth Functions (RDFs). For switching stimuli, we computed the steady-state RDF in the same portion of the response as the steady-state MIOF.

Variability of response modulation depth (RMD) across stimulus presentations

A large number of action potentials is needed to obtain significant RMDs (Rayleigh test of uniformity, $\alpha<0.05$). RMDs estimated in individual 250 ms epochs (during which m was constant) were usually very biased and variable due to the small number of spikes in each interval. To estimate the variability across trials while limiting this bias, we used a bootstrap procedure. We first identified the minimum number of epochs N that any value of m was presented. For each value of m, we then formed 200 random sets, with replacement, of N epochs, extracted the spike times from each set to compute 200 RMD values, and computed the standard deviation of these 200 RMDs.

The standard deviations estimated with this procedure were plotted as a function of m, and fitted with an increasing or decreasing sigmoidal function of the form:

$$std_{MIOF}(m) = \frac{A}{B + e^{-m/C}} + D \quad \text{(Eq. 4.1)}$$

Discriminability index (d')

To investigate the effect of stimulus modulation statistics on the response sensitivity to changes in stimulus modulation depth m, taking into account neural variability across repetitions, we computed a discriminability index (*d'*), as the ratio of the separation between mean responses to two adjacent modulation depths m and $m + \delta m$, to the square root of the mean variances across repetitions in these two responses:

$$d'(m) = \frac{MIOF(m+\delta m) - MIOF(m)}{\sqrt{\frac{std_{MiOF}(m+\delta m)^2 + std_{MiOF}(m)^2}{2}}} \quad \text{(Eq. 4.2)}$$

We used the fitted MIOFs and fitted standard deviations to compute *d'*. We plotted *d'* as a function of *m*, and compared the mean *d'* across distributions in various ranges of *m*.

Effect significance

For each stimulus modulation depth *m*, we tested the hypothesis that the RMD obtained in the anechoic-like condition differ from the RMD in the reverberant-like condition, with a test of equality of concentration parameters assuming von Mises distributions ($p<0.05$; p. 133 in Mardia and Jupp, 1999). Since 20 different modulation depths were presented, we performed 20 tests for each neuron. Therefore, the probability that at least one of the 20 tests would come positive by chance was $1-0.95^{20} = 64\%$. To determine whether there was an effect of stimulus statistics in a neuron at the 5% significance level, we computed the probability of obtaining at least *k* positive tests by chance as a function of *k*:

$$P_{false\_positive}(k) = \sum_{i=1}^{k} C_i^n p^i (1-p)^{n-i} \qquad \text{(Eq. 4.3)}$$

with n=20 the number of tests performed and p=0.05 the significance level of each test. We found that $P_{false\_positive}$ fell below 5% for k=3 ($P_{false\_positive}(3) \approx 0.016$). Therefore, we considered the effect of stimulus statistics to be significant in neurons where the RMDs were significantly different ($p<0.05$) between anechoic-like and reverberant-like conditions for at least 3 stimulus modulation depths.

Time course

In the switching stimulus paradigm, the stimulus switches back and forth between the anechoic-like and the strong reverberant-like distributions. In the neurons for which stimulus statistics had a significant effect (as defined in the previous section) on the MIOF, we estimated

the time course of the change in the MIOF by performing likelihood ratio tests as a function of time. We assumed that the likelihood of a spike phase θ_i given a stimulus modulation depth *m* and a particular distribution of modulation depths (anechoic-like or reverberant-like) followed a von Mises distribution:

$$p(\theta_i \mid m, distribution) = \frac{1}{2\pi I_0(\kappa(m))} \exp(\kappa \cos(\theta_i - \mu(m))) \quad \text{(Eq. 4.4)}$$

where $I_0(\kappa(m))$ is the modified Bessel function of the first kind and order 0, $\kappa(m)$ is the concentration parameter of the von Mises distribution, and $\mu(m)$ is its mean direction. We estimated $\kappa(m)$ and $\mu(m)$ from the spike times occurring in the steady-state portion of the response for each stimulus modulation depth *m*. We used the maximum likelihood estimates of κ and μ as described in Mardia and Jupp (1999, p 85). We thus obtained 40 likelihood functions for each neuron (20 values of *m* times 2 modulation distributions –anechoic-like and reverberant-like).

For each spike phase θ_i, a logarithmic likelihood ratio was computed from the estimated likelihood functions as:

$$LR(\theta_i) = \log\left(\frac{p(\theta_i \mid m, anechoic-like)}{p(\theta_i \mid m, reverberant-like)}\right) \quad \text{(Eq. 4.5)}$$

A positive LR indicates that the anechoic-like distribution is a better model for the spike phase θ_i than the reverberant-like distribution. Assuming independent spike phases θ_i, we summed the LRs of all spikes occurring in set time bins relative to switch onset. Time bins were 2.5 second long, with 98% overlap (50 ms delay between the onsets of two time bins). LRs from each bin were smoothed with a moving average filter with a 2.5 second span. By plotting the smoothed LR as a function of time, we determined at what point in time after a switch the spike

phases became more likely to originate from the steady-state distribution from which the stimuli were actually drawn. Time constants were defined as the time needed for the smoothed LR to change sign. The anechoic time constant τ_a is defined as the time elapsed from a switch from reverberant-like to anechoic-like distributions before the smoothed LR becomes positive, while the reverberant time constant τ_r is the time elapsed from anechoic-like to reverberant-like distributions before the smoothed LR becomes negative. Since the smoothed LR was integrated in 2.5-second time bins, there was a 2.5 second resolution limit on the estimated time constants.

Results

We measured Modulation Input/Output Functions (MIOFs) and Rate-Depth Functions (RDFs) in response to dynamic Sinusoidally Amplitude Modulated (SAM) broadband noise stimuli (Fig. 4.1) in 47 well isolated single units of IC of unanesthetized rabbit. 7 of these units were studied with the anechoic-like and reverberant-like stimulus distributions presented separately, while 40 were studied with a switching stimulus for which the distribution of modulation depths alternated, usually every 37.5 seconds, between the anechoic-like and the strong reverberant-like distributions.

Adaptation to modulation depth statistics with the non-switching paradigm

Modulation depth statistics had a significant effect on the MIOF in 4 out of 7 neurons studied with the non-switching paradigm. Figure 4.2 shows the MIOF and RDF measured in 3 neurons. In these 3 examples, the strong reverberant-like stimulus was presented first, followed by the anechoic-like stimulus, and the moderate reverberant-like stimulus. In the example of Fig. 4.2 A, the anechoic-like and moderate reverberant-like conditions produced very similar MIOFs. In contrast, the MIOF measured in the strong reverberant-like condition was significantly different from the other MIOFs. Specifically, response modulation depth (RMD) was significantly larger over almost the entire range of stimulus modulation depths, and the MIOF saturated at a lower *m*. In this example, the operating range of the MIOF seemed to shift toward the region of high probability in the strong reverberant-like condition, as the maximum slope occurred at a lower *m* than in the other conditions. The RDFs for this neuron also showed marked differences between the strong reverberant-like condition and the two other conditions. The RDFs were monotonically increasing in all three conditions, but the slope and range of firing

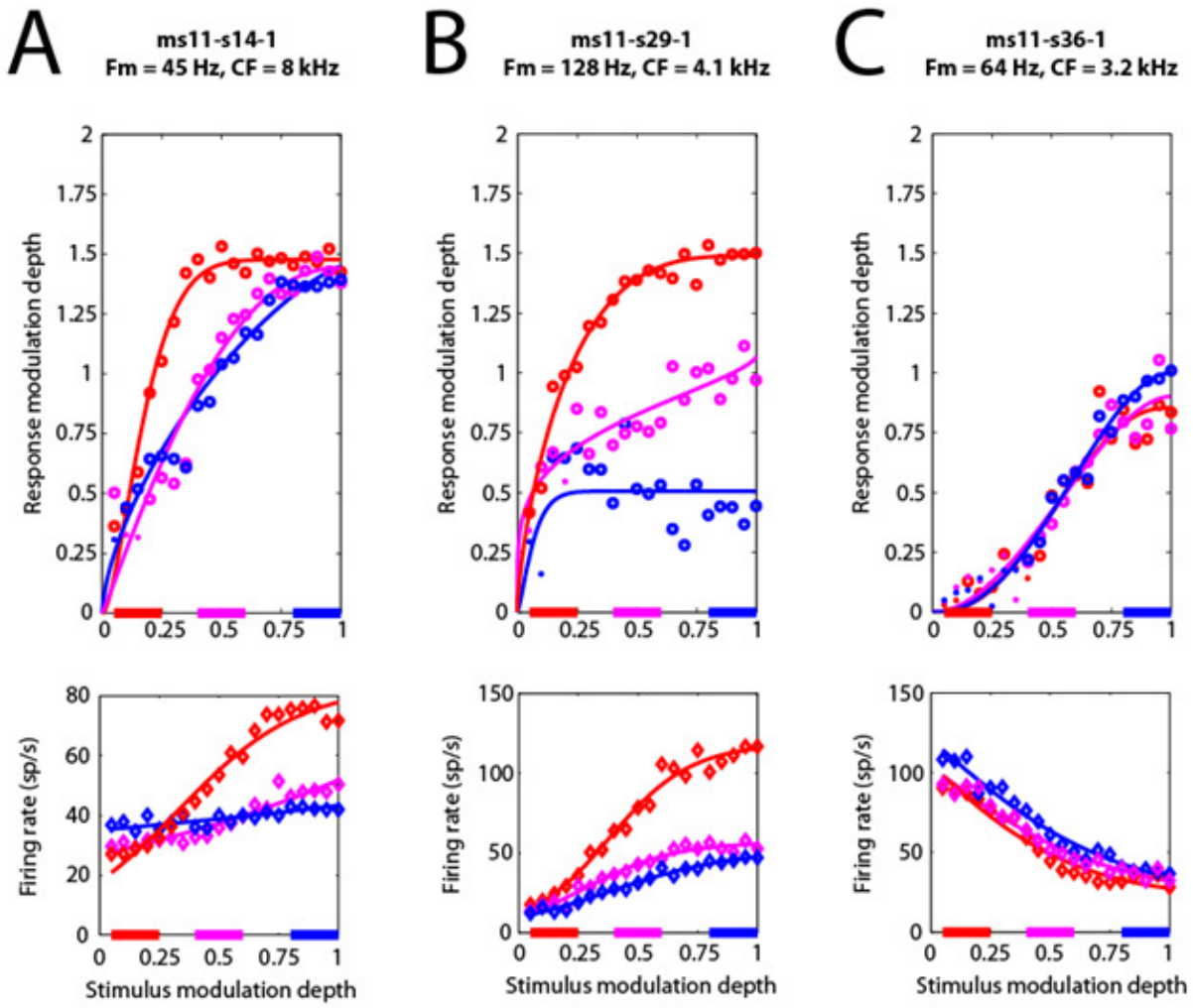

Figure 4.2 Example neural responses to non-switching dynamic stimuli
Modulation Input/Output Functions (MIOFs, top panels) and Rate-Depth Functions (RDFs, bottom panels) in example neurons measured with dynamic stimuli for which anechoic-like and reverberant-like conditions were presented separately. Response modulation depths are represented with an open circle when they are significant (Rayleigh test of uniformity, $\alpha<0.05$), and with a dot when they are not significant. Red: Strong reverberant-like condition. Magenta: Moderate reverberant-like condition. Blue: Anechoic-like condition. Regions of high probability of the distributions of stimulus modulation depths are highlighted on the x axis ($0.05 \leq m \leq 0.25$ for strong reverberant-like condition, $0.4 \leq m \leq 0.6$ for moderate reverberant-like condition, and $0.8 \leq m \leq 1$ for anechoic-like condition). **A.** Example neuron (f_m=45 Hz, CF=8 kHz) in which anechoic-like MIOF was significantly different from strong reverberant-like MIOF, but not from moderate reverberant-like MIOF. **B.** Example neuron (f_m=128 Hz, CF=4.1 kHz) in which anechoic-like MIOF was significantly different from both strong reverberant-like and moderate reverberant-like MIOFs. **C.** Example neuron (f_m=64 Hz, CF=3.2 kHz) for which all 3 dynamic conditions produced similar MIOFs.

rate were largest in the strong reverberant-like condition. These changes in RDF across stimulus condition are consistent with simple firing rate adaptation: As this neuron is more responsive for large stimulus modulations, firing rate tends to adapt more in the anechoic-like condition, for which stimulus modulation depths are more often in the more responsive region, than for the strong reverberant-like condition, for which stimulus modulation depths are more often in the least responsive region.

The example of Figure 4.2 B shows a very strong effect of stimulus statistics that consisted in a vertical rather than horizontal shift of the MIOF: In the anechoic-like condition, phase-locking to the modulation was rather poor, and RMD did not substantially vary with *m* except near 0. In contrast, the moderate reverberant-like condition produced MIOFs with greater gains and steeper slopes compared to the anechoic-like condition. Following the same trend, the strong reverberant-like MIOF had even higher gains and steeper slopes, aside from the saturation at high *m*. In all 3 conditions, the MIOF had steeper slopes in the lowest range of *m*. The effect on firing rate was similar to that in example A, and consistent as well with firing rate adaptation.

The third example in Fig. 4.2 C shows a neuron in which stimulus statistics did not have a significant effect on the MIOF: All 3 conditions had the same sigmoidal shape with quantitatively similar RMDs. The firing rates were also similar across conditions, although the anechoic RDF was somewhat higher for small modulation depths. This small effect on firing rate is also consistent with firing rate adaptation: In this case, RDFs were monotonically decreasing; Therefore the most responsive region was the region of small modulation depths, which the anechoic-like stimulus did not frequently visit.

Adaptation to modulation depth statistics with the switching paradigm

To ascertain that the effects observed in the previous section were not due to long-term changes in the operating range of the neuron unrelated to the history of stimulation (i.e. for example a slow degradation or improvement in spike quality over time during which the unit is in contact), and to characterize the time course of these effects, we used a switching paradigm for all subsequent experiments (40 well isolated single units): Because the switching dynamic stimulus is alternating between anechoic and strong reverberant-like conditions (Fig. 4.3 A), any long-term change in the operating range of the neuron would apply to both conditions, and therefore would not cause the differences seen across conditions.

Figure 4.3 B shows data from an example neuron in which the MIOFs, measured with a switching stimulus, were significantly affected by modulation depth statistics. In this example, the RMD was significantly larger in the reverberant-like condition than in the anechoic-like condition at most stimulus modulation depths. As a result, the MIOF seems to shift horizontally toward the region of high probability: The slopes of the MIOF (top panel in Fig. 4.3 C) were largest for low m in the reverberant-like condition, while the largest slopes occurred at relatively higher m in the anechoic-like condition. Although the point of maximum slope occurred in the high probability region in the reverberant-like condition, this was not quite the case in the anechoic-like condition, where it occurred near $m = 0.6$ rather than in the anechoic-like high probability region ($m \geq 0.8$).

To assess the sensitivity of a neuron to changes in stimulus modulation depth m, the slope of the MIOF is not sufficient, as it does not take into account the variability in the responses. To take into account variability across repetitions of m, we estimated the standard deviation of the RMD as a function of m with a bootstrapping procedure (Methods). In the example of Figure

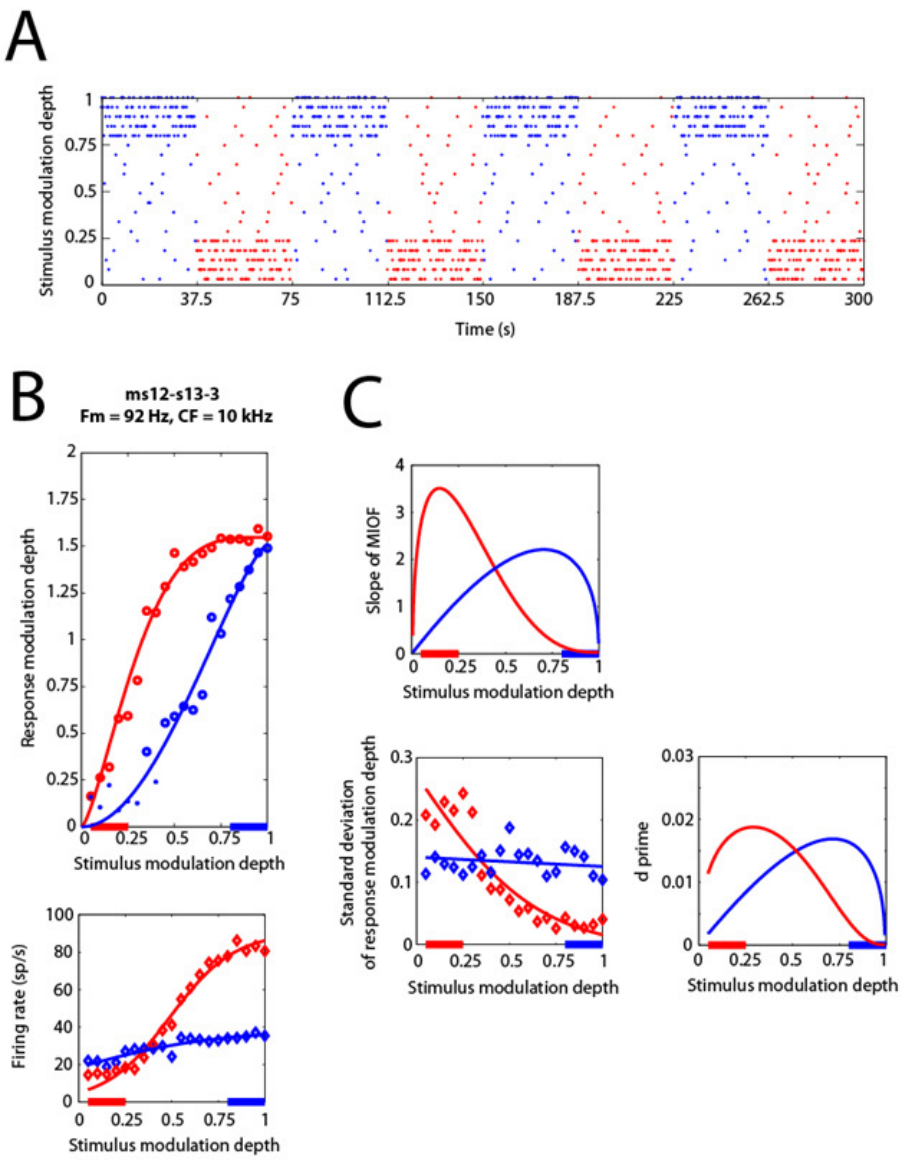

Figure 4.3 Example neural response and discriminability analysis for the switching paradigm
A. Modulation depth as a function of time in the first 300 seconds of a switching dynamic stimulus. Modulation depths are drawn every 250 ms from the anechoic-like (blue dots) or strong reverberant-like (red dots) distributions. Distributions are switched every 37.5 seconds. **B.** Modulation Input/Output Functions (MIOFs, top) and Rate-Depth Functions (RDFs, bottom) measured in an example neuron (f_m=92 Hz, CF=10 kHz) in response to a switching stimulus. Response modulation depths are represented with an open circle when they are significant (Rayleigh test of uniformity, $\alpha<0.05$), and with a dot when they are not. Blue: Anechoic-like condition. Red: Reverberant-like condition. Regions of high probability of the distributions of stimulus modulation depths are highlighted on the *x* axis. In this example, response modulation depths were larger in the reverberant-like condition than in the anechoic-like condition at most stimulus modulation depths, and the MIOF appears to shift towards the region of high probability. **C.** *Top:* Slope of the fitted MIOFs as a function of modulation depth. Here, slopes were largest for small stimulus modulations in the reverberant-like condition, and largest for high stimulus modulations in the anechoic-like condition. *Bottom, left:* Standard deviation of response modulation depths across repetitions, estimated with a bootstrap procedure, and fitted with a sigmoidal function (Methods). *Bottom, right: d'* discriminability index (Methods) computed from the fitted MIOFs (in B) and the fitted standard deviations. In this neuron, sensitivity to modulation depth improved in the region of high probability.

4.3, the standard deviation (C, bottom left panel) decreased with increasing *m*, although at a faster rate and from a larger value in the reverberant-like condition than in the anechoic-like condition. These differences in standard deviation are consistent with the differences in firing rate (Fig. 4.3 B, bottom), as the standard deviation of RMD decreases with increasing spike count (Mardia and Jupp, 1999). We fitted the dependence of standard deviation on *m* with a sigmoidal function (Methods) and used the fitted standard deviations and MIOFs to compute a *d'* discriminability index for changes in *m* (right panel). In this neuron, sensitivity to stimulus modulation depth assessed with *d'* followed a similar trend to the slopes of the MIOF. In particular, the maximum *d'* seemed to shift toward the region of high probability, hence providing better differential sensitivity for the most frequent modulation depths.

The effects of stimulus statistics on sensitivity were very different in other neurons. In the majority of neurons studied with the switching stimulus (27 out of 40), modulation depth statistics did not produce significant effects on the MIOF. Figure 4.4 A shows one such neuron. In this case, the maximum *d'* occurred in the mid-range of modulations, and did not differ much between anechoic-like and reverberant-like conditions. Despite the lack of effect on the MIOF, firing rates were significantly affected by stimulus conditions: The slope of the RDF was larger (less negative) for the reverberant-like condition than for the anechoic-like condition. Again, this effect is consistent with firing rate adaptation.

In other neurons, the effects of stimulus statistics were significant, but small. For example, in Figure 4.4 B, changes in the MIOF were mostly in its overall shape: While the anechoic-like case was characterized by large slopes for small *m*, and a saturation of RMD at high *m*, the reverberant-like MIOF was more sigmoidal, with the steepest slope occurring in the mid-range, near $m = 0.5$. In this example, the point of highest discriminability was shifted away

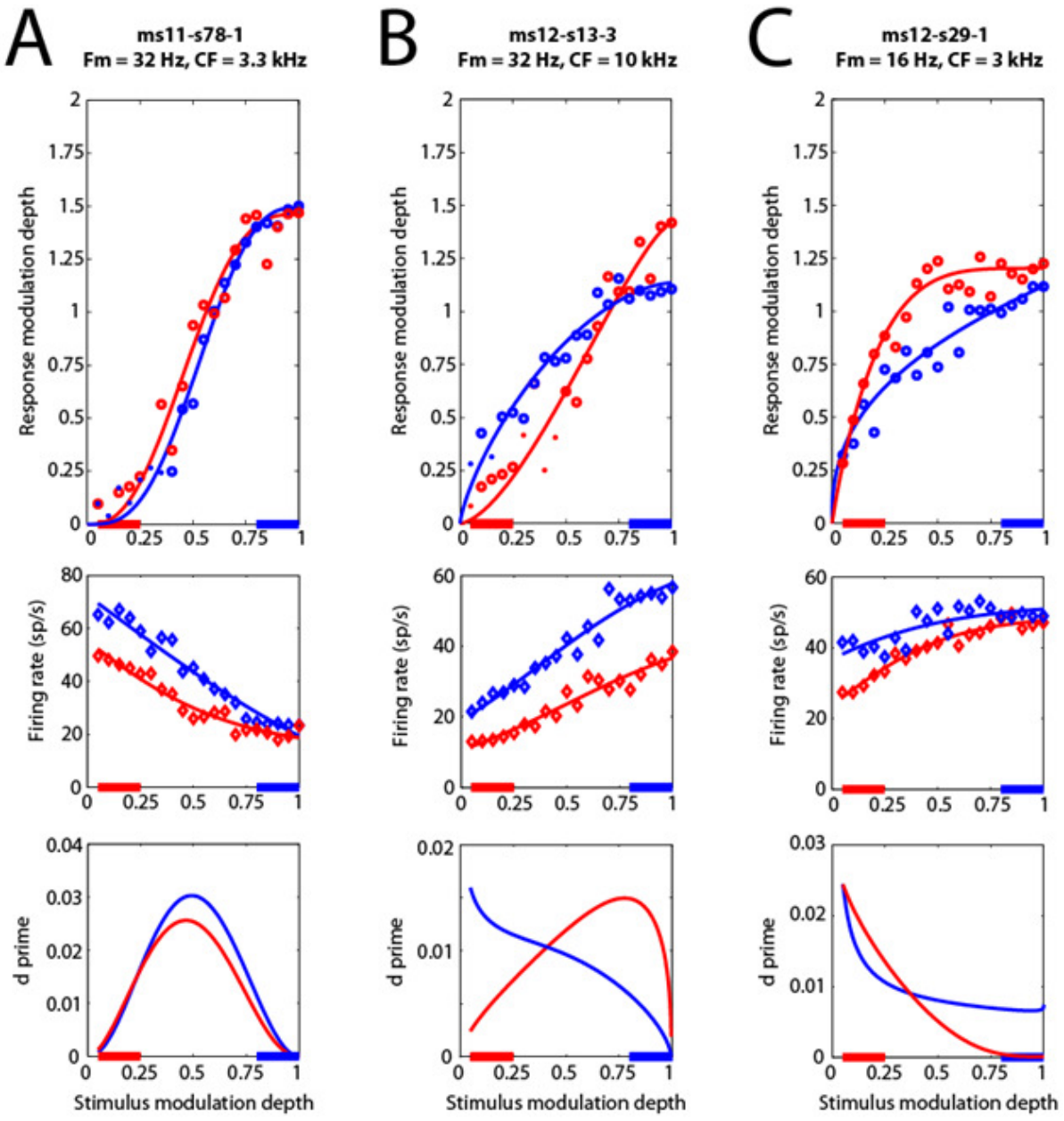

Figure 4.4 Example neurons showing the variety of effects observed
Modulation Input/Output Functions (MIOF, top panels), Rate-Depth Functions (RDF, middle panels) and *d'* sensitivity index (bottom panels) in three example neurons, measured in response to switching dynamic stimuli. Blue curves: Anechoic-like condition. Red curves: Reverberant-like condition. Regions of high probability of the distributions of stimulus modulation depths are highlighted on the *x* axis. **A.** Example neuron (f_m=92 Hz, CF=10 kHz) for which there was no significant effect of stimulus statistics on the MIOF, although firing rates significantly changed. **B.** Example neuron (f_m=32 Hz, CF=10 kHz) for which response modulation depths in the reverberant-like condition were larger at high stimulus modulation depths, and smaller at small stimulus modulation depths than in the anechoic-like condition. In this neuron, sensitivity was largest away from the region of high probability. **C.** Example neuron (f_m=16 Hz, CF=3 kHz) for which response modulation depths in the reverberant-like condition were larger at most stimulus modulation depths than in the anechoic-like condition. Sensitivity was largest in the region of high probability in both conditions.

from the region of high probability in both conditions. Moreover, the mean *d'* for $m \leq 0.25$ was larger in the anechoic-like condition than in the reverberant-like condition, and, similarly, the mean *d'* for $m \geq 0.8$ was larger in the reverberant-like condition than in the anechoic-like condition. This is inconsistent with the hypothesis that the neuron shifts its MIOF to optimize coding of changes in *m* in the region of high probability.

Figure 4.4 C shows another example for which the reverberant-like RMD was larger than the anechoic-like RMD at most stimulus modulation depths. The resulting reverberant-like MIOF had a steeper slope for small *m*, and saturated at a lower *m* than the anechoic-like MIOF. In this example, *d'* was largest for the lowest range of stimulus modulation depths in both conditions. Moreover, *d'* was clearly higher in the anechoic-like condition than in the reverberant like condition for large *m*, and somewhat higher in the reverberant-like condition than in the anechoic-like condition for small *m*. This particular neuron's behavior is consistent with the hypothesis that AM coding is optimized for the region of high probability. In this neuron, the effect of *m* distribution on firing rate was small.

Across neurons, the reverberant-like condition increased RMD but effects on discriminability of changes in stimulus modulation depth were mixed

The MIOF characterizes the transformation of stimulus modulation depth into neural modulation depth. Sensitivity to changes in stimulus modulation depth *m* were assessed with *d'*, which depends on the slope and variability of the MIOF. However, even in cases when the MIOF is nearly flat in some range of *m*, resulting in poor sensitivity to changes in *m*, the MIOF can still carry important information about the modulation frequency f_m of the stimulus, as long as RMDs are significant. Effects of modulation depth statistics on the MIOF can therefore be of

two kinds: (1) Effects on the coding of modulation frequency f_m, assessed by the RMD of the MIOF (a larger RMD means a stronger neural representation of f_m); (2) Effects on the coding of changes in modulation depth m, assessed by the d' (a larger d' means a higher sensitivity to changes in m).

We compared the RMDs of the MIOF averaged over the entire range of m between anechoic-like and reverberant-like conditions across the neural population (Fig. 4.5 A, middle panel). On average, the mean RMD was significantly larger in the reverberant-like condition than in the anechoic-like condition. This suggests that envelope frequency representation improved in the reverberant-like condition, compared to the anechoic-like condition (Paired t-test, p=0.01). To investigate the effects on sensitivity to changes in m across the population, we compared the d' averaged over the entire range of m between conditions (Fig. 4.5 B, middle panel). Although the mean d' was larger in the anechoic-like condition than in the reverberant-like condition, these differences were not significant (p=0.1), suggesting that, on average, sensitivity to changes in m do not systematically depend on m distribution across the population.

Instead of looking at average effects over the entire range of m, an interesting question is whether the neurons optimize coding of AM in the region of high probability. To test this hypothesis, we compared the effect of stimulus statistics in 3 ranges of m: the low range $m \leq 0.25$ (corresponding to the region of high probability in the strong reverberant-like condition), the mid-range $0.25 < m < 0.8$ (corresponding to a region of low probability in both stimulus conditions), and the high range $m \geq 0.8$ (corresponding to the region of high probability in the anechoic-like condition). A one-way analysis of variance (ANOVA) on the differences between reverberant-like and anechoic-like RMDs averaged in the three ranges revealed no effect of m (p=0.12), suggesting that on average, there is no improvement of coding of f_m in the regions of

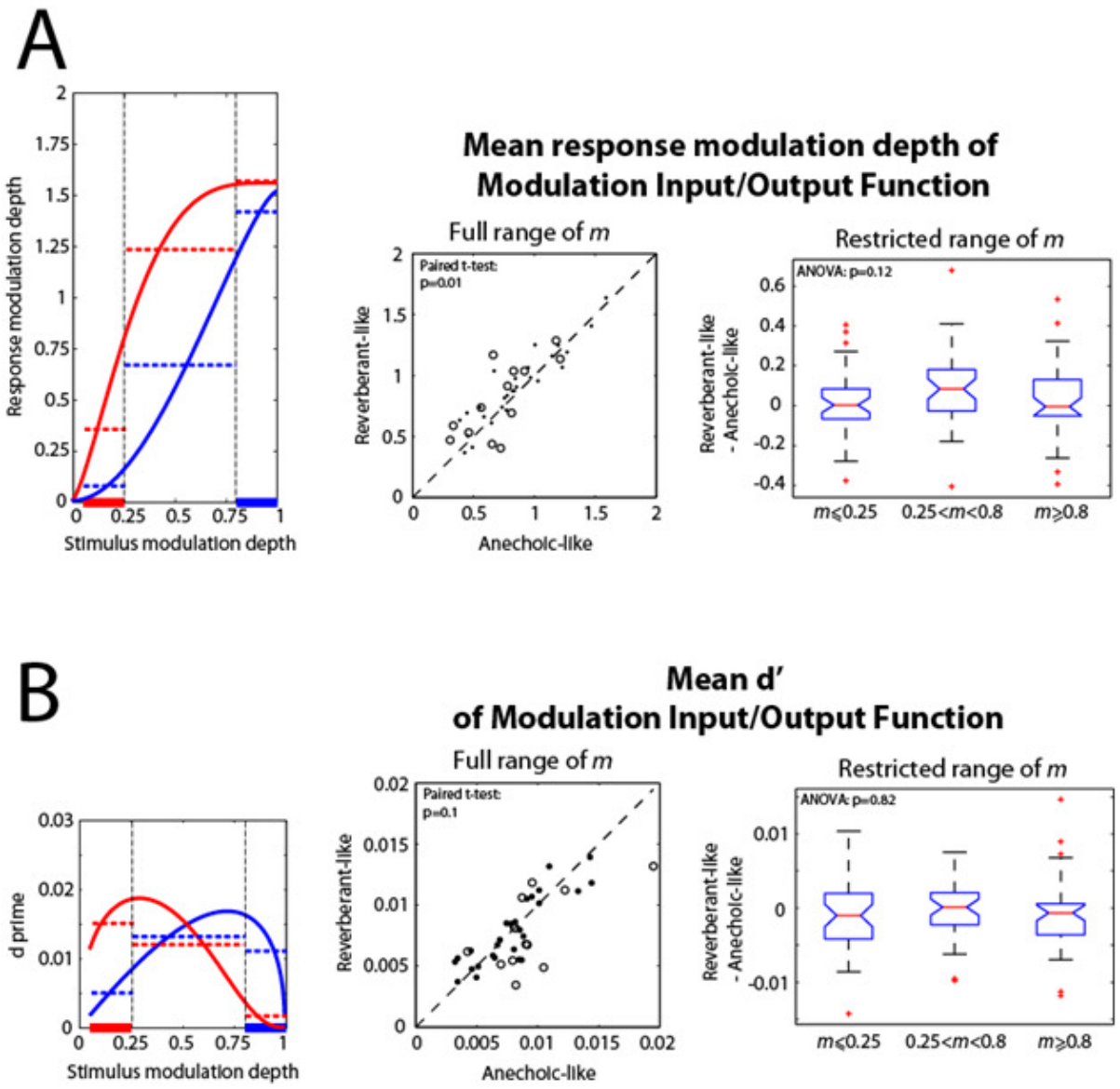

Figure 4.5 Summary of responses to switching dynamic stimuli across population
A. Mean response modulation depth was significantly larger (p=0.01) in the reverberant-like condition when averaged over the entire range of stimulus modulation depths (middle panel). Neurons for which changes in modulation depth statistics significantly affected the MIOF are represented with open circles. One-way analysis of variance (ANOVA) showed no significant differences in effect of stimulus statistics on RMD across three ranges of stimulus modulation depths. **B.** d' sensitivity index was not significantly different between anechoic-like and reverberant-like conditions (p=0.1). No significant differences were found in the effect of stimulus statistics on mean d' across ranges of stimulus modulation depths (one-way ANOVA, p=0.82).

high probability (Fig. 4.5 A, right panel). Similarly, a one-way ANOVA on the differences between reverberant-like and anechoic-like *d'* averaged in the three ranges defined above was not significant (p=0.82), suggesting that on average, there was no improvement of sensitivity to changes in *m* in the regions of high probability (Fig. 4.5 B, right panel).

Overall, the reverberant-like stimulus tended to increase mean RMD while no systematic effect on neural sensitivity was found. For both metrics, there were large variations in effects across neurons. On average, temporal coding of AM was not optimized for the region of high probability.

Time course of the adaptation to modulation depth statistics greatly varied across neurons

The time course of the adaptation to modulation depth statistics was studied by performing Likelihood Ratio Tests as a function of time following a switch in *m* distribution (Methods). For each neuron studied with the switching paradigm, and for which modulation depth statistics had a significant effect (13 neurons), we estimated the likelihood of each spike's phase at the stimulus modulation depth *m* for each distribution (reverberant-like or anechoic-like), assuming a von Mises distribution (Fig. 4.6 B). We summed the log likelihood ratios (LR) of all spikes occurring in a given time window relative to switch onset (2.5 second span, 98% overlap). Figure 4.6 C shows the smoothed log LR as a function of time in an example neuron. In this neuron, log LR was negative up to τ_a = 13.25 seconds after a switch from reverberant-like to anechoic-like distribution, indicating that, during this time period, the temporal pattern of neural firing was more consistent with the previously presented reverberant-like distribution, than with the ongoing anechoic-like distribution. In contrast, the log LR stayed positive for only τ_r = 0.55 seconds after the onset of the reverberant-like condition in the same neuron. This

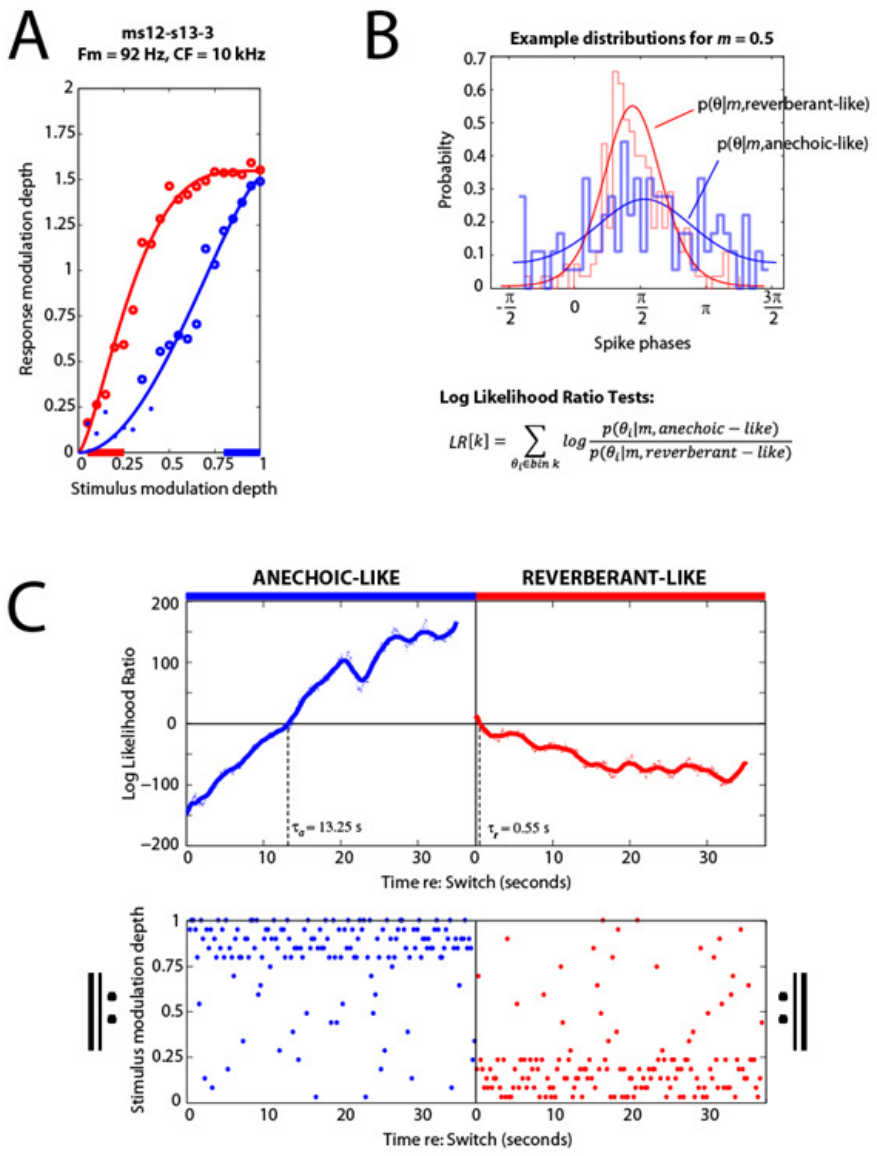

Figure 4.6 Time course of the effect of modulation depth statistics in an example neuron
A. MIOFs in an example neuron (f_m=92 Hz, CF=10 kHz) for which there was a significant effect of stimulus statistics, measured in response to a switching stimulus. **B.** Spike times were binned across switches to construct period histograms (40 bins/period) in the anechoic-like (blue, shaded stairs) and reverberant-like (red, shaded stairs) conditions. Von Mises distributions (blue and red, solid curves) were fitted to the histograms (Methods) to describe the likelihood of a spike phase for each stimulus modulation depth and dynamic condition. Logarithmic Likelihood Ratio Tests were performed as a function of time by summing the log likelihood ratios of all spikes occurring in a given time bin (Methods). **C.** *Top:* Time course of the log likelihood ratios. Positive log likelihood ratios meant a greater resemblance to the anechoic-like distribution. Spike times were binned across switches in time windows relative to switch onset. Time bins were 2.5 second long, with 98% overlap (50 ms delay between the onset of two time bins). Log likelihood ratios from each bin (dots) were smoothed with a moving average filter (solid lines) with a 2.5 second span. Time constants were defined as the time needed for the log likelihood ratio to cross the *x* axis (time for which the spike times begin to be more consistent with the distribution from which the stimuli were actually drawn). In this example, the effect was very slow in the anechoic-like condition (τ_a = 13.25 s), and relatively faster for the reverberant-like condition (τ_r = 0.55 s). *Bottom:* Distribution was switched every 37.5 seconds. Only one full switch is represented, but sequence of modulation depths was random across switches.

indicates that, in this neuron, adaptation to modulation depth statistics was slow (of the order of seconds), and significantly slower when going from the reverberant-like to the anechoic-like distribution, than from the anechoic-like to the reverberant-like distribution.

Figure 4.7 A-B provides two other examples showing the diversity of RMD adaptation time constants we encountered. In the first example (Fig. 4.7 A), RMDs were larger in the reverberant-like condition than in the anechoic-like condition. The time course of RMD adaptation for this neuron was faster than in the previous example: In the anechoic-like condition, the log LR was already positive in the first 2.5 second bin, and in the reverberant-like condition, the log LR became negative for $\tau_r = 0.25$ sec. Figure 4.7 B shows data from the same neuron as in Figure 4.6, but at a different modulation frequency (f_m = 32 Hz instead of 92 Hz). The time course of RMD adaptation in the anechoic-like condition was very slow for both f_m (τ_a = 8.9 seconds at 32 Hz). In the reverberant-like condition, however, the effect was faster than in the anechoic-like condition (τ_r = 4.05 seconds), but slower than previously measured at 32 Hz (τ_r was 0.55 seconds).

We compared the time constant of RMD adaptation in the reverberant-like condition to the time constant in the anechoic-like condition across the population of neurons that showed a significant effect of modulation depth statistics (Fig. 4.7 C). There was no significant difference between the time constants across conditions (p=0.17). In 5 cases, the time constant in one of the conditions was greater than 3 seconds (up to ≈13 s). In 7 cases, the time constants measured in both conditions were less than 1 second. Overall, the time course of the effect of modulation depth statistics varied greatly across neurons.

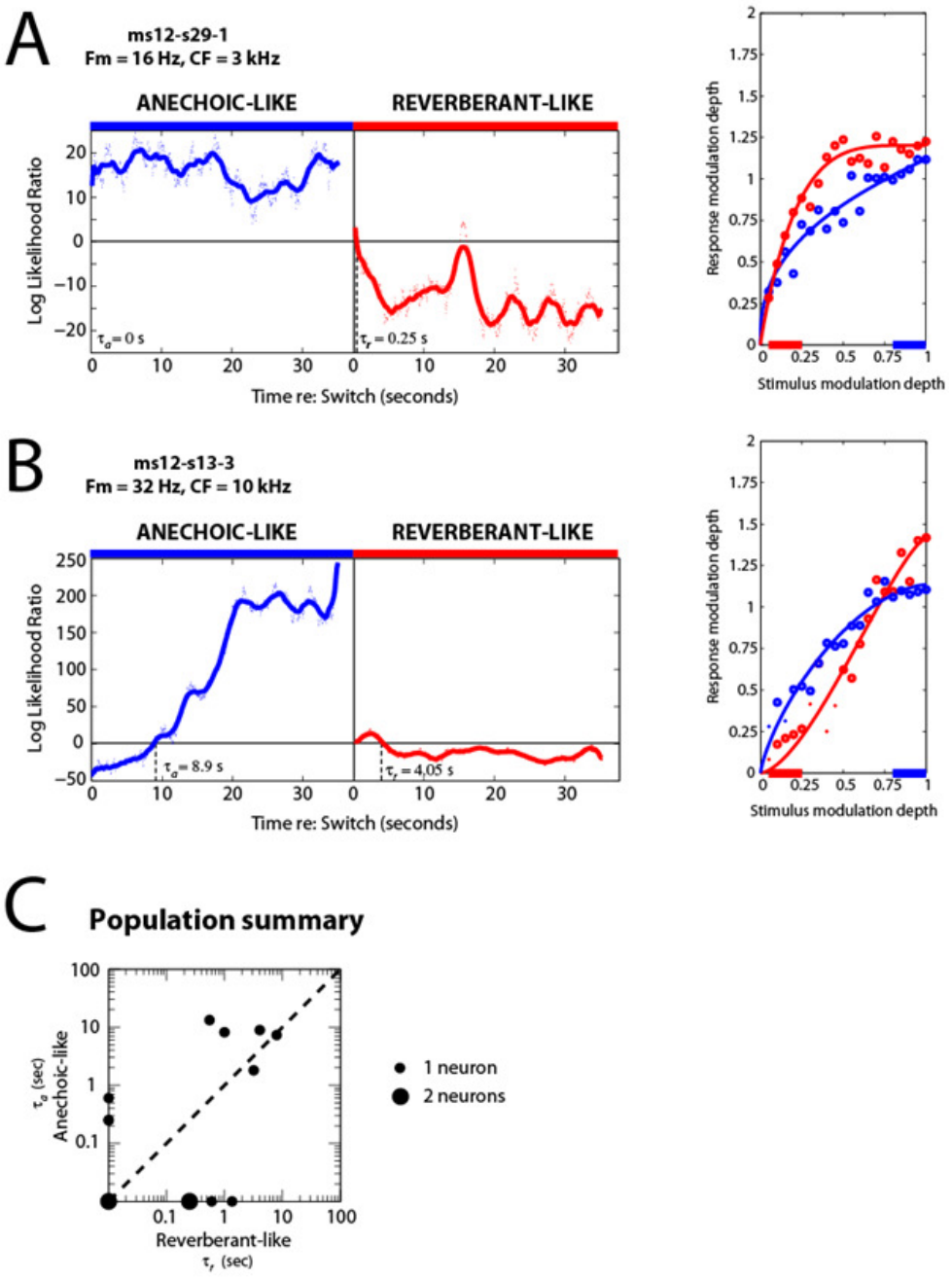

Figure 4.7 Time course of the effect of modulation depth statistics varied greatly across neurons
A. Example neuron with a fast effect of stimulus statistic ($\tau_a = 0$ s, $\tau_r = 0.25$ s). **B.** Example neuron with a slow effect of stimulus statistic ($\tau_a = 8.9$ s, $\tau_r = 4.05$ s). Time bins were 2.5-second long with 98% overlap. **C.** Scatter plot summary in the 13 neurons (out of 40) that showed a significant effect of stimulus statistic on response modulations. Time constants of 0 were displayed on the X and Y axis. In 5 cases, the time constant in one condition was greater than 3 seconds. In 7 cases, the time constant in both conditions was less than 1 second. There was no significant difference between the time constants across condition (p=0.17).

Discussion

We recorded from single units from IC of unanesthetized rabbit in response to Sinusoidally Amplitude Modulated (SAM) broadband noise stimuli whose modulation depth was dynamically drawn from distributions designed to imitate the properties of anechoic and reverberant environments. We found that modulation depth statistics had a significant effect on the Modulation Input/Output Functions (MIOFs) in a subset of neurons (13 of 40), which therefore showed a novel form of adaptation. Overall, response modulation depths (RMD) were higher in the reverberant-like condition than in the anechoic-like condition. In contrast, discriminability as assessed with the *d'* index did not change systematically across conditions. The time course of adaptation was slow in some neurons (≈ 3 – 13 seconds) and faster in other neurons (< 3 seconds).

Significance of the adaptation of a temporal property of neural firing to stimulus statistics

In recent studies examining adaptation to the statistics of a distribution of sound levels in IC of anesthetized guinea pig (Dean *et al.*, 2005; Dean *et al.*, 2008), in auditory nerve of anesthetized cat (Wen *et al.*, 2009), or adaptation to the statistics of a distribution of interaural level differences in IC of anesthetized ferrets (Dahmen *et al.*, 2010), investigators reported changes in neural firing rate. We also observed changes in firing rate with changes in the mean of the distribution of modulation depths. These changes were usually consistent with simple firing rate adaptation. Theoretically, if firing rate increases monotonically with *m*, the anechoic-like stimulus is 80% of the time in the most responsive region of the neuron, and should thereby decrease its firing in this region compared to the strong reverberant-like condition. This effect should result in a decrease in the slope of the RDF in the anechoic-like condition. We compared

the mean slopes of the RDFs across conditions, and found a very significant decrease in slope ($p=0.001$) in the anechoic-like condition, compared to the reverberant-like condition, consistent with simple firing rate adaptation.

In addition to changes in average firing rate, we observed, in a subset of neurons, changes in the temporal pattern of firing induced by changes in the mean of the stimulus distribution. In these neurons, changes in RMD could be very large (corresponding to increases or decreases of up to 10 dB), and we ascertained that they were not artifacts of long term fluctuations in the operating range of the MIOF by using a switching paradigm.

Effect of stimulus statistics on a temporal property of neural firing has been reported in another study of amplitude modulation in IC of anesthetized cat (Kvale and Schreiner, 2004). These authors modulated sinusoidal carriers with a stepwise rectangular function (800 steps/sec) whose amplitude was drawn from a Gaussian distribution, and investigated the effect of the variance of the distribution on both firing rate and temporal receptive fields. They found that the latency of the receptive fields increased slightly in most units (of the order of 1-2 msec) after a transition from a low variance to a high variance distribution, and that the amplitude of the temporal receptive fields changed by up to 10-20 %. The time course of the effect was of the order of hundreds of milliseconds. It is difficult to compare our study to theirs as we examined the effect of the mean of the distribution rather than the variance, in unanesthetized rabbit rather than anesthetized cats, and with SAM broadband noise carriers with AM frequencies $\leq$ 128 Hz, rather than stepwise modulations of sinusoidal carriers with 800 Hz modulation frequency. The mechanisms involved may be different in the two studies, for several reasons: (1) We found a significant effect in only about 1/3 of the neurons, while they found an effect in the majority of their units; (2) The time course of the effect was sometimes very slow in our case (of the order of

seconds), while they report time constants of the order of hundreds of milliseconds; (3) The magnitude of the effect was often much larger in our study (see e.g. Fig. 4.3).

Possible explanations

A possible explanation for the effect of stimulus statistics on the MIOF is a change in temporal precision due to firing rate adaptation of the neuron. As described earlier, firing rate depended monotonically (more often increasing but sometimes decreasing) on stimulus modulation depth *m*, and changing the stimulus statistics usually decreased the slope of the RDF in the anechoic-like condition, consistent with firing rate adaptation. Since the mean RMD was lower, on average, in the anechoic-like condition, a possibility is that firing rate adaptation caused a decrease in the precision of spike timing. We found a moderate correlation ($r=0.42$, $p=0.01$) between the effect of modulation statistics on the mean RDF slope and the effect on the mean RMD of the MIOF, suggesting that firing rate adaptation partly contributes to the effect. However, there are clear counter-examples to this hypothesis: For example, in Figure 4.4 A, changes in firing rate were significant, and consistent with firing rate adaptation, whereas there were no significant differences between the anechoic-like and reverberant-like MIOFs.

Another possibility is that the effect is related to forward masking of AM. In a recent study in IC of unanesthetized rabbit, Wojtczak *et al.* (2011) presented 1 second stimuli consisting of a 500 ms fully modulated or unmodulated masker tone preceding a partially modulated target tone embedded in an unmodulated tone. They found that, in a subset of neurons, the temporal pattern of the response to the target modulated tones was affected by the preceding masker. However, they did not provide a quantitative evaluation of the effect on RMD across population. In the examples they provide, the fully modulated masker tended to degrade phase-locking to the

target signal. It is difficult to compare the results of their study to ours, as the maskers they used were always fully modulated or unmodulated, and the stimuli were presented with interstimulus intervals greater than 1 second. Nevertheless, it is possible that in our experiments, the observed changes in RMD were induced by complex forward masking interactions between stimulus segments with different modulation depths. For example, the overall decrease in RMD in the anechoic-like condition, relative to the reverberant-like condition, may be caused by masking from the preceding frequent high modulation depths. However, the time constant of forward masking in Wojtczak *et al.* (2011) was of the order of tens or hundreds of milliseconds, whereas some of our neurons had much longer time constants (3 – 13 seconds), suggesting a more complex phenomenon.

In general, corticofugal mechanisms could contribute to changes in the MIOF following changes in modulation depth statistics. Effects of corticofugal connections on the coding of various sound parameters (e.g. frequency, duration, level) have been demonstrated in the IC of several species. For example, in unanesthetized big brown bat, Yan and Suga (1998) showed that electrically stimulating the auditory cortex caused the best frequency of IC neurons to shift toward that of the stimulated cortical neurons. Although effects of corticofugal feedback on AM coding have not been reported, it is possible that coding of modulation frequency in reverberant-like environments is enhanced through a corticofugal mechanism.

Implications for AM coding in reverberant environments

We found that RMDs tend to be higher in the reverberant-like condition than in the anechoic-like condition. This effect of modulation depth statistics may help robustly encode

envelope frequencies in reverberant situations, in which the listener is typically exposed to small modulation depths, due to the smearing effect of reverberation on envelopes.

If this effect of modulation depth statistics seems beneficial for coding in reverberant environments, it can also be seen as problematic for anechoic environments, for which envelope coding would be relatively less robust. Moreover, although the reverberant-like condition had, on average, a coding advantage over the anechoic-like condition, the effect on RMD was not significantly different across different ranges of *m*. Therefore, the neurons, on average, did not optimize the temporal pattern of their firing to the statistics of our stimuli.

We also assessed the sensitivity to changes in *m* using a *d'* metric. We did not find any significant difference, on average, between anechoic-like and reverberant-like *d'* averaged over the entire range of *m*, nor did we find differences across different ranges of *m*. Again, this finding is inconsistent with the hypothesis that neurons optimize temporal coding to adapt to stimulus statistics.

However, the population statistics we report do not reflect the high variability across neurons. In some neurons, the effect of stimulus statistics on the MIOF was consistent with an improvement of discriminability in the region of high probability, hence resulting in more sensitive temporal coding of AM for the most frequent modulation depths, both in the anechoic-like and the reverberant-like conditions. Other neurons seemed to optimize their MIOF to the statistical characteristic for one condition, but not the other. It is possible that these neurons constitute different functional populations, each optimized for a different type of environment, and that the information from these different populations is emphasized or suppressed as needed by a decoding mechanism in higher stages of the auditory system.

Overall, our findings are consistent with other studies in pointing to the importance of considering past stimulation history in investigating temporal coding of AM in IC neurons.

Chapter 5

General conclusions and discussion

In this thesis, we investigated the coding of amplitude envelope in reverberant environments by recording from single units in the Inferior Colliculus (IC) of unanesthetized rabbit in response to Sinusoidally Amplitude Modulated (SAM) broadband noise in simulated anechoic and reverberant environments. In Chapter 2, we find that reverberation degrades rate and temporal coding of Amplitude Modulations (AM), but that the compressive shape of the nonlinear transformation from stimulus modulations to neural modulations (the Modulation Input Output Function or MIOF) helps compensate for the neural degradation. We further find in Chapter 3 that in a subset of neurons, reverberant stimuli have a coding advantage over anechoic stimuli with the same modulation depth, and that binaural features of the reverberant stimuli may be partly responsible for this advantage. Finally, in Chapter 4, we report that temporal coding of AM frequency is not static, but depends dynamically on past stimulation in a subset of neurons, and that on average, AM coding is more robust in conditions when low modulation depths predominate, as in reverberant environments.

Importance of the Modulation Input Output Function (MIOF) for AM coding in reverberation

In Chapters 2 and 3, we compare the Response Modulation Depth (RMD) measured in response to reverberant stimuli to the RMD measured in response to anechoic stimuli with the same modulation depth. Although we find a subset of neurons in which reverberant RMD is significantly larger than depth-matched anechoic RMD, the population statistics supports the

view that stimulus modulation depth *m* is the primary determinant of temporal coding both in anechoic and reverberant conditions: For example, in Chapter 3, reverberant and depth-matched anechoic RMD were highly correlated across the population ($r=0.75$, $p<0.001$). Moreover, the RMD to the depth-matched anechoic stimuli were highly correlated to the RMD to any other depth-matched condition, no matter which additional features of the reverberant stimuli were matched. Therefore, a good prediction of the temporal response to SAM stimuli can be provided by the MIOF, which we assumed in Chapter 2 to be a function of *m* only. This is consistent with psychophysical studies of speech intelligibility in reverberation: For example, the Speech Transmission Index (STI) reliably predicts speech intelligibility for a variety of noise and reverberation conditions, and is computed only from the attenuation of *m* in different frequency bands (Houtgast *et al.*, 1980; Steeneken and Houtgast, 1980).

Our description of MIOFs measured in the IC provides a more detailed picture of temporal coding of AM than available in the literature. In particular, we find in Chapter 2 that a majority of MIOFs have a compressive shape (i.e. a slope < 1 in a log-log plot), and that this compressive shape may help counteract the attenuation of *m* due to reverberation by limiting the degradation in temporal coding.

Dependence of the MIOF on binaural features of the stimulus

Although the MIOF is primarily a function of *m*, we find in Chapter 3 that in a subset of neurons, binaural features of the stimulus such as Interaural Cross-Correlation (IACC) and Interaural Envelope Disparities (IEDs) can have an effect on RMD. Although effects of binaural features of the stimulus were significant for only subsets of neurons, it seems important to consider these features for studying AM coding with stimuli more relevant to realistic

communication situations. Moreover, we focused on binaural listening situations with the sound source located in front of the receivers, which is the most likely situation for human communication, but future work should investigate the effect of varying azimuth on AM coding. It is possible that temporal coding of AM is stronger for a specific range of azimuths. Further, we compared diotic and dichotic binaural conditions, but did not focus on monaural listening. This would be an interesting case to investigate and to compare to psychophysical studies, as binaural listening in reverberation has been reported to help speech intelligibility over monaural listening (e.g. Nabelek and Robinson, 1982).

Finally, we show in Chapter 3 that the mean change in IACC and IEDs introduced by reverberation only partially account for the coding advantage of reverberant stimuli over depth-matched anechoic stimuli. We hypothesize that oscillations in IACC at the modulation frequency introduced by reverberation may enhance phase-locking to the modulation frequency. Future work should test this hypothesis by synthesizing anechoic stimuli matching both the modulation depth and IACC oscillations of reverberant stimuli, and comparing the RMD they elicit in IC neurons. In synthesizing such stimuli, attention should be paid to the binaural and spectral frequency tuning characteristics of the neurons recorded from, as a neuron will only "see" the IACC of the band-pass stimulus at its best delay. Moreover, the phase difference between IACC oscillations and AM should be carefully controlled, as effects of IACC oscillations and AM may enhance phase-locking for some values of phase shift, but suppress it for other values. For example, if in a hypothetical neuron, an increase in IACC tends to decrease instantaneous firing rate, while an increase in amplitude tends to increase instantaneous firing rate, an in-phase relationship between IACC oscillations and AM in the stimulus will reduce phase-locking relative to the diotic case, as the two effects go in opposite directions.

Alternatively, if IACC oscillations and AM have a 0.5 cycle phase shift, phase-locking would be enhanced in such a neuron.

Dependence of the MIOF on modulation depth statistics

The finding that MIOFs depend dynamically on preceding stimulation in a subset of neurons (Chapter 4) is significant and suggests that the study of AM coding should take into account stimulus statistics to be more relevant to everyday communication situations. Future work should replicate the experiments of Chapter 4 at other processing stages of the auditory system. An important first step will be to perform these experiments in the Auditory Nerve (AN), in order to elucidate whether there is a contribution of the auditory periphery. In a study on sound level coding in the AN of anesthetized cat, Wen *et al.* (2009) show that dynamic range adaptation to mean sound level occurs in the auditory periphery, although the peripheral effects are not as marked as they are in the midbrain (Dean *et al.*, 2005; 2008). In our case, it is also possible that adaptation to the modulation depth statistics of the stimulus originates in the AN. However, AM coding in the AN and in the IC have different characteristics. For example, modulation gains are typically larger in the IC than in the AN (Joris *et al.*, 2004). Moreover, the dependence of average firing rate on stimulus modulation depth m is often monotonically increasing in the IC, while it tends to be flat in the AN (Joris and Yin, 1992). Since such a major transformation in AM coding occurs between the periphery and the IC, it is likely that adaptation to modulation depth statistics arises beyond the AN.

Neural coding of speech envelopes in reverberant environments

This thesis investigated the neural coding of sinusoidal envelopes in realistic reverberant environments. We chose to focus on sinusoidal modulations for simplicity, although the amplitude envelopes of speech signals are usually more complex, with a low-pass modulation spectrum (e.g. Attias and Schreiner, 1997). The experiments presented in this thesis could easily be extended to the study of speech, although the effect of cochlear filtering should be carefully controlled as speech envelope depends on frequency band. The strength of speech envelope representation in auditory neurons could be assessed by the peak cross-correlation between response spike train and envelope waveform, and compared in anechoic and reverberant conditions, similar to Chapter 2. To test the influence of binaural features of reverberation on the coding of speech envelope, experiments such as those carried out in Chapter 3 could be done as well, by synthesizing anechoic stimuli matching the envelope and selected binaural features of reverberant stimuli. Further, the effect of preceding stimulation on envelope frequency representation demonstrated in Chapter 4 should also be investigated with speech signals.

The possibility raised in Chapter 3 that fluctuations in IACC occurring at the modulation frequency of reverberant SAM stimuli may enhance phase-locking compared to depth-matched anechoic stimuli may also apply to speech. In speech signals, bursts of energy are produced, especially at low modulation frequencies corresponding to the syllabic rate. These bursts of energy are less affected by the decorrelation introduced by reverberation, compared to low amplitude segments of the signal, which are dominated by reverberant energy. Therefore, it is possible that variations in IACC coupled to variations in amplitude could enhance speech envelope representation in reverberation, and therefore partially counteract the attenuation of AM.

Finally, another extension of this study could focus on effects of hearing impairment on speech envelope representation in reverberation, as speech intelligibility is substantially altered in reverberation for the hearing-impaired, even if they do well in quiet. The effect of reverberation on the neural coding of speech envelope could be compared in normally hearing and hearing impaired animals, and would be an important step toward making hearing aids and cochlear implants perform better in everyday communication situations.

References

Adams J. C. (**1979**). Ascending projections to the inferior colliculus. *J. Comp. Neurol*. **183**:519-538.

Albeck Y. and Konishi M. (**1995**). Responses of neurons in the auditory pathway of the barn owl to partially correlated binaural signals. *J Neurophysiol.* **74**:1689-1700.

Allen J. B. and Berkley, D. A. (**1979**). Image method for efficiently simulating small room acoustics. *J. Acoust*. *Soc. Am*. **65**:943-950.

Apoux F. and Bacon S. P. (**2004**). Relative importance of temporal information in various frequency regions for consonant identification in quiet and in noise. *J. Acoust*. *Soc. Am*. **116**:1671-1680.

Apoux F. and Bacon S. P. (**2008**). Differential contribution of envelope fluctuations across frequency to consonant identification in quiet. *J. Acoust*. *Soc. Am.* **123**:2792–2800.

Attias H. and Schreiner C. E. (**1997**). Temporal low order statistics of natural sounds. In *Advances in Neural Information Processing Systems 9,* ed. M. C. Mozer, M. I. Jordan, and T. Petsche, pp. 103-109. Cambridge, MA: MIT Press.

Bacon S. P. and Viemeister N. F. (**1985**). Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. *Audiol*. **24**:117-134.

Borg E., Engström B., Linde G., and Marklund K. (**1988**). Eighth nerve fiber firing features in normal-hearing rabbits. *Hear. Res.* **36**:191-201.

Baumann S., Griffiths T. D., Sun L., Petkov C. I., Thiele A., and Rees A. (2011). Orthogonal representation of sound dimensions in the primate midbrain. *Nat. Neurosci.* **14**:423-425

Caicedo A and Herbert H. (**1993**). Topography of descending projections from the inferior colliculus to auditory brainstem nuclei in the rat. *J. Comp. Neurol.* **328**:377-392.

Coffey C. S., Elbert C. S. Jr., Marshall A. F., Skaggs J. D., Falk S. E., Crocker W. D., Pearson J. M., and Fitzpatrick D. C. (**2006**). Detection of interaural correlation by neurons in the superior olivary complex, inferior colliculus and auditory cortex of the unanesthetized rabbit. *Hear. Re*s. **221**:1-16

Culling J. F., Colburn H. S., and Spurchise M. J. (**2001**). Interaural correlation sensitivity. *J. Acoust*. *Soc. Am*. **110**:1020-1009.

Dahmen J. C., Keating P., Nodal F. R., Schulz A. L., and King A. J. (**2010**). Adaptation to stimulus statistics in the perception and neural representation of auditory space. *Neuron*. **66**:937-948.

Dean I., Harper N. S., McAlpine D. (**2005**). Neural population coding of sound level adapts to stimulus statistics. *Nat. Neurosci*. **8**:1684-1689.

Dean I., Robinson B. L., Harper N. S., and McAlpine D. (**2008**). Rapid neural adaptation to sound level statistics. *J Neurosci.* **28**:6430-6438.

Devore S. and Delgutte B. (**2010**). Effects of reverberation on the directional sensitivity of auditory neurons across the tonotopic axis: Influences of ITD and ILD. *J. Neurosci.* **30**:7826-7837.

Devore S., Ihlefeld A., Hancock K. E., Shinn-Cunningham B., and Delgutte B. (**2009**). Accurate sound localization in reverberant environments is mediated by robust encoding of spatial cues in the auditory midbrain. *Neuron* **62**:123-134.

Dreschler W. A. and Leeuw A. R. (**1990**). Speech reception in reverberation related to temporal resolution. *J. Speech Hear. Res.* **33**:181-187.

Drullman R., Festen J. M., and Plomp R. (**1994**). Effect of reducing slow temporal modulations on speech reception. *J. Acoust*. *Soc. Am*. **95**:2670–2680.

Eggermont J. J. (**1991**). Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hear. Res.* **56**:153-167.

Ewert S. D. and Dau T. (**2004**). External and internal limitations in amplitude-modulation processing. *J. Acoust*. *Soc. Am*. **116**:478–490.

Griffin S. J., Bernstein L. R., Ingham N. J., and McAlpine D. (**2005**). Neural sensitivity to interaural envelope delays in the inferior colliculus of the guinea pig. *J Neurophysiol.* **93**:3463-3478.

Hancock K. E., Noel V., Ryugo D. K., and Delgutte B. (**2010**). Neural coding of interaural time differences with bilateral cochlear implants: effects of congenital deafness. *J Neurosci.* **30**:14068-14079.

Harris R. W. and Swenson D. W. (**1990**). Effects of reverberation and noise on speech recognition by adults with various amounts of sensorineural hearing impairment. *Audiol*. 2**9**:314-321.

von Helmholtz H. L. F. (**1863**). *On the Sensation of Tones as a Physiological Basis for the Theory of Music*. New York: Dover.

Houtgast T., Steeneken H. J. M., and Plomp R. (**1980**). Predicting speech intelligibility in rooms from the Modulation Transfer Function. I. General room acoustics. *Acustica.* **46**:60-72.

Irwin R. J. and McAuley S. F. (**1987**). Relations among temporal acuity, hearing loss, and the perception of speech distorted by noise and reverberation. *J. Acoust*. *Soc. Am*. **81**:1557–1565.

Joris P. X., van de Sande B., Recio-Spinoso A., and van der Heijden M. (**2006**). Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation. *J. Neurosci.* **26**:279-289.

Joris P. X., Schreiner C. E., and Rees A. (**2004**). Neural processing of amplitude-modulated sounds. *Am. Physiol. Soc*. **84**:541-577.

Joris P. X. and Yin T. C. T. (**1992**). Responses to amplitude-modulated tones in the auditory nerve of the *cat*. *J. Acoust*. *Soc. Am*. **91**:215-232.

Kiang N. Y. S. and Moxon E. C. (**1974**). Tails of tuning curves of auditory nerve fibers. *J. Acoust. Soc. Am.* **55**:620-630.

Kim D. O., Bishop B., and Kuwada S. (**2010**). Acoustic cues for sound source distance and azimuth in rabbits, a racquetball and a rigid spherical model. *J Assoc Res Otolaryngol.* **11**:541-557.

Krebs B., Lesica N. A., and Grothe B. (**2008**). The representation of amplitude modulations in the mammalian auditory midbrain. *J. Neurophysiol.* **100**:1602-1609.

Krishna B. S. and Semple M. N. (**2000**). Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J. Neurophysiol.* **84**:255-273.

Kuwabara N. and Zook J. M. (**2000**). Geniculo-collicular descending projections in the gerbil. *Brain Res*. **878**:79-87.

Kuwada S. and Batra R. (**1999**). Coding of sound envelopes by inhibitory rebound in neurons of the superior olivary complex in the unanesthetized rabbit. *J Neurosci.* **19**:2273-2287.

Kuwada S., Stanford T. R., and Batra, R. (**1987**). Interaural phase sensitive units in the inferior colliculus of the unanesthetized rabbit. Effects of changing frequency. *J. Neurophysiol.* **57**:1338-1360.

Kvale M. N. and Schreiner C. E. (**2004**). Short-term adaptation of auditory receptive fields to dynamic stimuli. *J. Neurophysiol.* **91**:604-612.

Lacher-Fougère S. and Demany L. (**1998**). Modulation detection by normal and hearing-impaired listeners. *Audiology*. **37**:109-121.

Lorenzi C., Berthommier F., Apoux F., and Bacri N. (**1999**). Effects of envelopeexpansion on speech recognition. *Hear. Res.* **136**:131-138.

Mardia K. V. (1972). *Statistics of Directional Data.* London and New York: Academic Press.

Mardia K. V. and Jupp P. E. (1999). D*irectional Statistics.* New York: Wiley.

Merzenich M. M. and Reid M. D. (**1974**). Representation of the cochlea within the inferior colliculus of the cat. *Brain Res.* **77**:397-415.

Miller L. M., Escabi M. A., Read H. L., and Schreiner C. E. (**2001**). Functional convergence of response properties in the auditory thalamocortical system. *Neuron* **32**:151–160.

Møller A. R. (**1974**). Responses of units in the cochlear nucleus to sinusoidally amplitude-modulated tones. *Exp. Neurol.* **45**:105-117.

Nelken I., Rotman Y., and Bar Yosef O. (**1999**). Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*. **397**:154-157.

Nelson P. C. and Carney L. H. (**2007**). Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *J. Neurophysiol.* **97**:522-539.

Neuman A. C., Wroblewski M., Hajicek J., and Rubinstein A. (**2010**). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear.* **31**:336-44.

Payton K. L., Uchanski R. M., and Braida L. D. (**1994**). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *J. Acoust. Soc. Am.* **95**:1581-1592.

Plomp R. and Duquesnoy A. J. (**1980**).Room acoustics for the aged. *J. Acoust. Soc. Am.* **68**:1616-1621.

Poissant S. F., Whitmal N. A. 3rd, and Freyman R. L. (**2006**). Effects of reverberation and masking on speech intelligibility in cochlear implant simulations. *J. Acoust. Soc. Am.* **119**:1606-1615.

Rabinowitz N. C., Willmore B. D., Schnupp J. W., and King A. J. (**2011**). Contrast gain control in auditory cortex. *Neuron*. **70**:1178-1191.

Rees A. and Møller A. R. (**1983)**. Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hear. Res.* **10**:301-330.

Rees A. and Palmer A. R. (**1989)**. Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea-pig inferior colliculus, and their modification by broadband noise *J. Acoust. Soc. Am.* **85**:1978-1994.

Richards D. G. and Wiley R. H. (**1980**). Reverberations and amplitude fluctuations in the propagation of sound in a forest: implications for animal communication. *Am. Nat.* **115**:381-399.

Rhode W. S. (**1994**). Temporal coding of 200% amplitude modulated signals in the ventral cochlear nucleus of cat. *Hear Res.* **77**:43-68.

Saldaña E. and Merchàn M. A. (**2005**). Intrinsic and commissural connections of the Inferior Colliculus. In *The Inferior Colliculus*, ed. C. A. Winer and C. E. Schreiner, pp. 155-181. New York: Springer.

Sato H., Sato H., Morimoto M., and Ota R. (**2007**). Acceptable range of speech level for both young and aged listeners in reverberant and quiet sound fields. *J. Acoust. Soc.Am.* **122**:1616-1623.

Schreiner C. E. and Langner G. (**1988**). Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J. Neurophysiol.* **60**:1823-1840.

Schroeder M. R. (**1981**). Modulation transfer functions: Definition and measurement. *Acustica* **49**:179:182

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

Shannon R. V., Zeng F.-G., and Wygonski J. (**1998**). Speech recognition with altered spectral distribution of envelope cues. *J. Acoust. Soc.Am.* **104**:2467-2476.

Shinn-Cunningham B. G., Desloge J. G., and Kopco N. (**2001**). Empirical and modeled acoustic transfer functions in a simple room: Effects of distance and direction. Proceedings of the 2001 IEEE Workshop of Applications of Signal Processing to Audio and Acoustics, New Pfaltz, NY, 19-24 October 2001, 419-423.

Sinex D. G., Henderson J., Li H., and Chen G. D. (**2002**). Responses of chinchilla inferior colliculus neurons to amplitude-modulated tones with different envelopes. *J. Assoc. Res. Otolaryngol.* **3**:390-402.

Steeneken H. J. and Houtgast T. (**1980**). A physical method for measuring speech-transmission quality. *J. Acoust. Soc. Am.* **67**:318-326.

Ter-Mikaelian M., Sanes D. H. and Semple M. N. (**2007**). Transformation of temporal properties between auditory midbrain and cortex in the awake mongolian gerbil. *J. Neurosci.* **27**:6091-6102.

Tollin D. J., Populin L. C., and Yin T. C. T. (**2004**). Neural correlates of the precedence effect in the inferior colliculus of behaving cats. *J. Neurophysiol*. **92**:3286-3297.

Vater M. and Feng A. S. (**1990**). Functional organization of ascending and descending connections of the cochlear nucleus of horseshoe bats. *J. Comp. Neurol.* **292**:373-395.

Viemeister N. F. (**1979**). Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am.* **66**:1264-1380.

Wakefield G. H. and Viemeister N. F. (**1990**). Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise. *J. Acoust. Soc. Am.* **88**:1367-1373.

Wen B., Wang G. I., Dean I., and Delgutte B. (**2009**). Dynamic range adaptation to sound level statistics in the auditory nerve. *J. Neurosci.* **29**:13797-13808.

Wenstrup J. J. and Grose C. D. (**1995**). Inputs to combination-sensitive neurons in the medial geniculate body of the mustached bat: the missing fundamental. *J. Neurosci.* **15**:4693-4711.

Winer J. A., Larue D. T., Diehl J. J., and Hefti B. J. (**1998**). Auditory cortical projections to the cat inferior colliculus. *J. Comp. Neurol.* **400**:147-174.

Wojtczak M., Nelson P. C., Viemeister N. F., and Carney L. H. (**2011**). Forward masking in the amplitude-modulation domain for tone carriers: psychophysical results and physiological correlates. *J. Assoc. Res. Otolaryngol.* **12**:361-373.

Wong M. Y. (**1989**). Likelihood estimation of a simple linear regression model when both variables have error. *Biometrika.* **76**:141-148.

Woolf N. K., Ryan A. F., and Bone R. C. (**1981**). Neural phase-locking properties in the absence of cochlear outer hair cells. *Hear. Res.* **4**:335-346

Yan W. and Suga N. (**1998**). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nat. Neurosci.* **1**:54-58.

Yang W. and Bradley J. S. (**2009**). Effects of room acoustics on the intelligibility of speech in classrooms for young children. *J. Acoust. Soc. Am.* **125**: 922-933.

Yin T. C., Kuwada S., and Sujaku Y. (**1984**). Interaural time sensitivity of high-frequency neurons in the inferior colliculus. *J. Acoust. Soc. Am.* **76**:1401-1410.

Yin T. C., Chan J. C., and Carney L. H. (**1987**). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J. Neurophysiol.* **58**:562-583.

Zahorik P. (**2009**). Perceptually relevant parameters for virtual listening simulation of small room acoustics. *J. Acoust. Soc. Am.* **126**: 776-791.