ON THE MECHANISM OF RESPONSE LATENCIES IN AUDITORY NERVE FIBERS

by

Ethan I. Huang, MD

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Ethan I. Huang

It was defended on

November 30, 2011

and approved by

J. R. Boston, PhD

Catherine V. Palmer, PhD

Sheila R. Pratt, PhD

Dissertation Advisor: John D. Durrant, PhD

Copyright © by Ethan I. Huang

2011

ON THE MECHANISM OF RESPONSE LATENCIES IN AUDITORY NERVE FIBERS

Ethan I. Huang, MD

University of Pittsburgh, 2011

Despite the structural differences of the middle and inner ears, the latency pattern in auditory nerve fibers to an identical sound has been found similar across numerous species. Studies have shown the similarity in remarkable species with distinct cochleae or even without a basilar membrane. This stimulus-, neuron-, and species- independent similarity of latency cannot be simply explained by the concept of cochlear traveling waves that is generally accepted as the main cause of the neural latency pattern.

An original concept of Fourier pattern is defined, intended to characterize a feature of temporal processing—specifically phase encoding—that is not readily apparent in more conventional analyses. The pattern is created by marking the first amplitude maximum for each sinusoid component of the stimulus, to encode phase information. The hypothesis is that the hearing organ serves as a running analyzer whose output reflects synchronization of auditory neural activity consistent with the Fourier pattern.

A combined research of experimental, correlational and meta-analysis approaches is used to test the hypothesis. Manipulations included phase encoding and stimuli to test their effects on the predicted latency pattern. Animal studies in the literature using the same stimulus were then compared to determine the degree of relationship.

The results show that each marking accounts for a large percentage of a corresponding peak latency in the peristimulus-time histogram. For each of the stimuli considered, the latency

predicted by the Fourier pattern is highly correlated with the observed latency in the auditory nerve fiber of representative species.

The results suggest that the hearing organ analyzes not only amplitude spectrum but also phase information in Fourier analysis, to distribute the specific spikes among auditory nerve fibers and within a single unit.

This phase-encoding mechanism in Fourier analysis is proposed to be the common mechanism that, in the face of species differences in peripheral auditory hardware, accounts for the considerable similarities across species in their latency-by-frequency functions, in turn assuring optimal phase encoding across species. Also, the mechanism has the potential to improve phase encoding of cochlear implants.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

The seed of this work was sowed about 5 years ago, while I was meditating on how the implant users' speech perception in noise can be improved. I would like to thank Dr. John Durrant, my mentor, for his leading me from the cycle of clinical operations and otolaryngology to audiology since that time, and further into auditory physiology and basic hearing science. I will always be appreciative of his guiding me to become a thinker. I would like to especially thank Dr. J. R. Boston for his encouraging me in the world of engineering and giving me the technical pen for science. I am grateful also to Dr. Catherine V. Palmer for her careful reading and constructive comments, and to Dr. Sheila R. Pratt for her professional guidance and continuous support. This work is made possible also by Dr. Mario A. Ruggero and Dr. Christine Koppl, who kindly provided their raw data of the animal recordings. Lastly, I want to thank my family—Sophie, Sarah, and Eric—for their constant love.

1.0 INTRODUCTION

The sense of hearing is the ability to perceive sound, which is a series of mechanical waves created by a vibrating source [1]. A sound can be decomposed into sinusoid components. A sinusoid can be completely described by three parameters: frequency, intensity (amplitude) and phase. These three kinds of information enable one to reproduce the sinusoid. Putting all the sinusoid components together regenerates the original sound signal.

Hearing is one important sense that some people such as Helen Keller even value more than vision [2]. However, about one to three people in a thousand in the total human population are congenitally deaf or profoundly sensory neural hearing impaired [3, 4]. They cannot take advantage from current hearing aids and lose this sensation. Currently, the only effective treatment is multiple-electrode cochlear implantation [5]. The cumulative number of implants in the world has exceeded 120,000, which is higher than the number for all other types of neural prostheses combined [6].

Nevertheless, current cochlear implants fail to deliver, if not totally discard, phase information [6-10]. In order to prevent inter-electrode interaction [11], current cochlear implants deliver the extracted envelope information to the electrode array with a constant temporal offset [10] (see 2.3.2.2). In a processing cycle, this replaces the variety of original phases (e.g., the various patterns of first-spike latencies for different tone bursts) by a non-informative constant offset (see 2.3.2.2). Although the change between the normal phase dispersion and the constant

temporal offset is on the scale of a few milliseconds, precision coarser than a few milliseconds in the auditory cortex might significantly reduce the encoded information [12], and thus reduced the performance of implantees.

Current cochlear implant users encounter significant difficulties in adverse situations. For example, cochlear implant users might achieve almost zero correct score for open-set sentence recognition in noise when signal-to-noise ratio was zero, and gradually improved to about 30% (still poor) when signal-to-noise ratio was 20 dB [13] (also see 2.3.3). The failure of delivering phase information to cochlear implant users is also relevant to many other poor (close to chance performance) auditory perceptions such as music perception and lexical tone perception (see 2.3.3.2 and 2.3.3.3).

In previous studies, it was suggested that additional low-frequency acoustic hearing may improve speech recognition in noise (e.g., [13], also see 2.3.3.1). The residual acoustic hearing had rich phase information, although the residual acoustic hearing was so bad that it did not contribute (0% correct) to speech perception alone. Also, studies have shown that rhythm (a time code) and additional low-frequency acoustic hearing may improve music perception (e.g., [14, 15], also see 2.3.3.2).

Therefore, it is an important topic as to how phase information can be preserved for the current cochlear-implant users and possible implant candidates in the 0.3% of total human population who will pursue implantation.

The normal hearing organ preserves phase information well. At the outer ear, the hearing organ continuously reflects the original sound without relative phase shift across all frequencies (see 2.1.2.1). It transmits (mostly passively) the signal through the middle ear (see 2.1.3). The

hearing organ then encodes the signal in the peripheral auditory neurons with precise all-or-none spikes (action potentials).

The monaural response latencies usually are shorter in high characteristic-frequency neurons and longer in low characteristic-frequency neurons. In mammals, evidence supports that this latency pattern is related to the displacement and speed of the basilar membrane [16, 17], and the mechanics of cochlear traveling waves [18]. In the conventional view, different positions of the cochlear partition are accessed via the delay line imposed by traveling-wave propagation [19] (Figure 1). The frequency-place mapping of the cochlea results in the basal neurons discharging sooner than the apical neurons, because the basal neurons are closer to the middle ear. Thus, the latency pattern is affected and determined by any structures between the ear drum and the neural axon [20] (also see 2.1.3, 2.1.4, and 2.2.7).

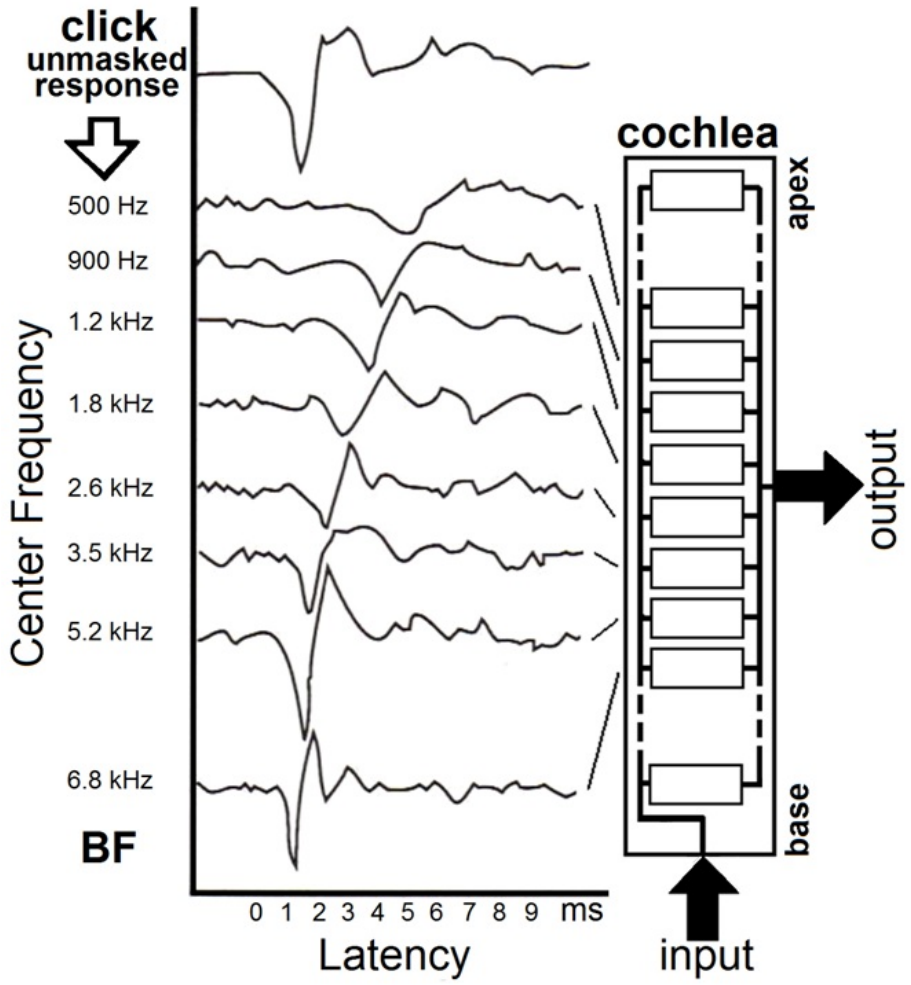


Figure 1. In the conventional view, the neural latency pattern in mammals is accounted for by a delay line imposed by traveling-wave propagation. Derived auditory brainstem responses with various center frequencies in humans show different latencies that form a latency pattern. Inset: block diagram of filter bank due to the frequency-selective response of each place along the basilar membrane. These (in effect) band-pass filters have various best frequencies (BFs) from base to apex and are accessed along the cochlear partition via the delay line imposed by traveling-wave propagation. Reprinted with permission from Durrant, J.D. and Feth, L.L. (in press). Hearing Science--The Bases. Pearson, Boston.

Based on the frequency response characteristics of the middle ear and the frequency-place mapping of the cochlea, remarkable differences are evident among samples of animals

across phyla [21] (Figure 2). Similarly, the latency pattern of primary neural responses to an identical sound is expected to differ correspondingly (Figure 3).

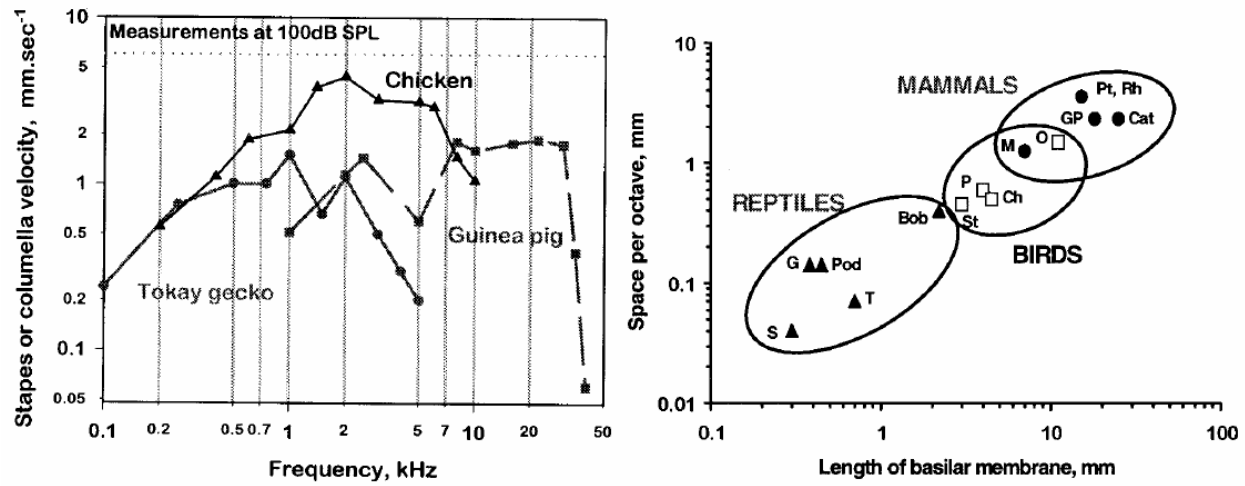


Figure 2. Variation of middle and inner ears. Left: The Tokay gecko is a sensitive lizard, the chicken represents birds, and the guinea pig represents mammals. The main difference observed is not in sensitivity, but in the upper frequency limits and the high-frequency flanks. Right: Lizards: G, alligator lizard, Gerrhonotus; T, turtle; Pod, European lizards of the genus Podarcis; S, the granite spiny lizard Sceloporus; Bob, Australian Bobtail lizard Tiliqua. Birds: St, starling, Sturnus; Ch, chicken, Gallus; P, pigeon, Columba; O, barn owl, Tyto. Mammals: M, mouse, Mus; GP, guinea pig, Cavia; Cat, house cat, Felis; Pt, Rh, two bat species, Pteronotus, the mustached bat, and Rhinolophus, the horseshoe bat. Reprinted with permission from Figs. 2 and 9, Manley, G. A. (2000). "Cochlear mechanisms from a phylogenetic viewpoint." Proc Natl Acad Sci U S A 97(22): 11736-11743. Copyright 2000, the National Academy of Sciences of the United States of America.

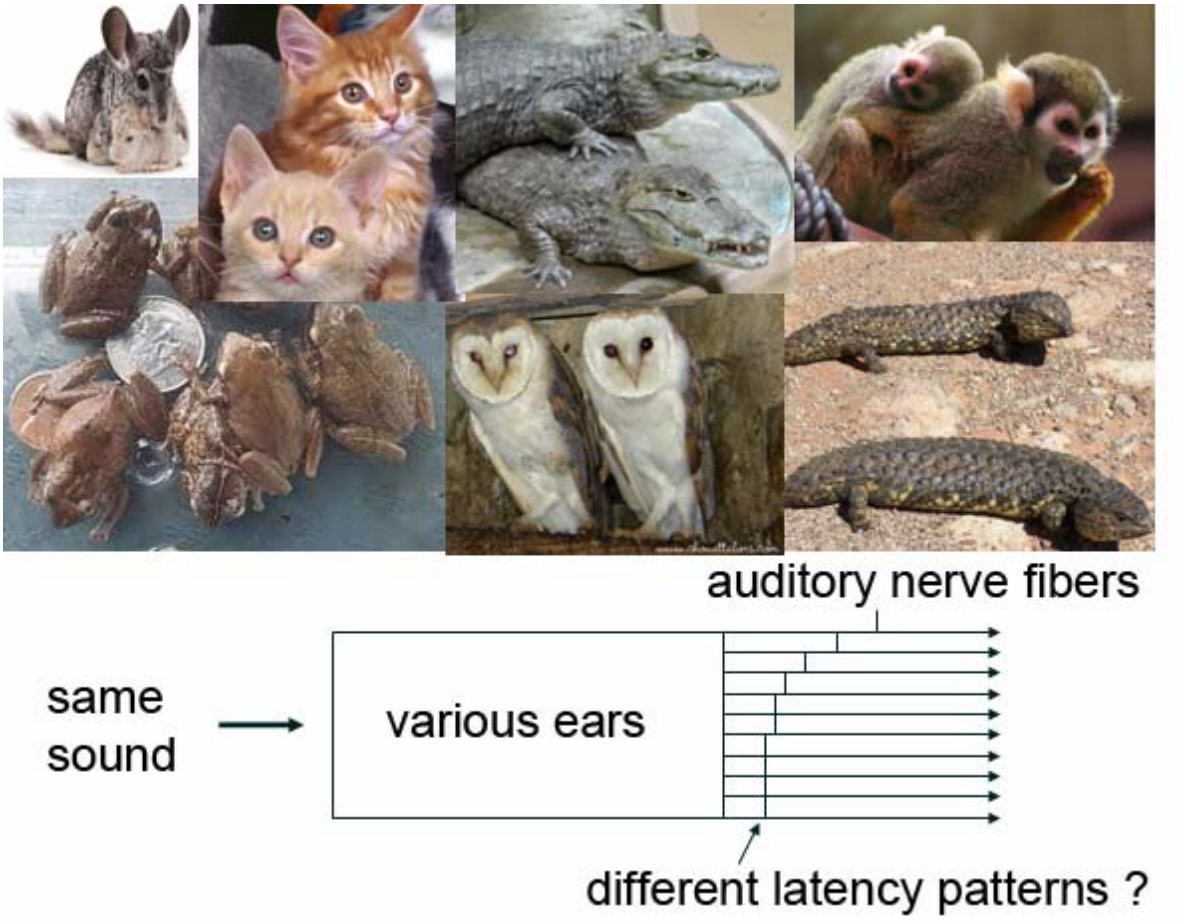


Figure 3. Does the latency pattern differ among various species and individuals? Based on the structural variation such as the content and sensitivity of the middle ear, and the length and frequency-place mapping of the cochlea, the latency pattern to an identical sound is expected to be different in different species, or even to be different between two individual animals of an identical species.

However, previous animal studies show the first-spike latency pattern to various tone bursts from different individual animals and even different species are similar (with considerations of the duration of an action potential and choice of the regression model) [22] (see 2.2.1) (Figure 4). Also, numerous studies have shown the phase-locking latency patterns to tone bursts are similar across species (e.g., [23-26], also see 2.2.3). Moreover, the similarity is demonstrated in the first-spike latency pattern to a rarefaction click [22] (see 2.2.4) and also in

classical click responses across numerous species [27] (see 2.2.5). Specifically, the compound peristimulus-time histogram shows multiple peaks separated by the inverse of the characteristic frequency, and the peaks from rarefaction and condensation clicks neatly interleave.

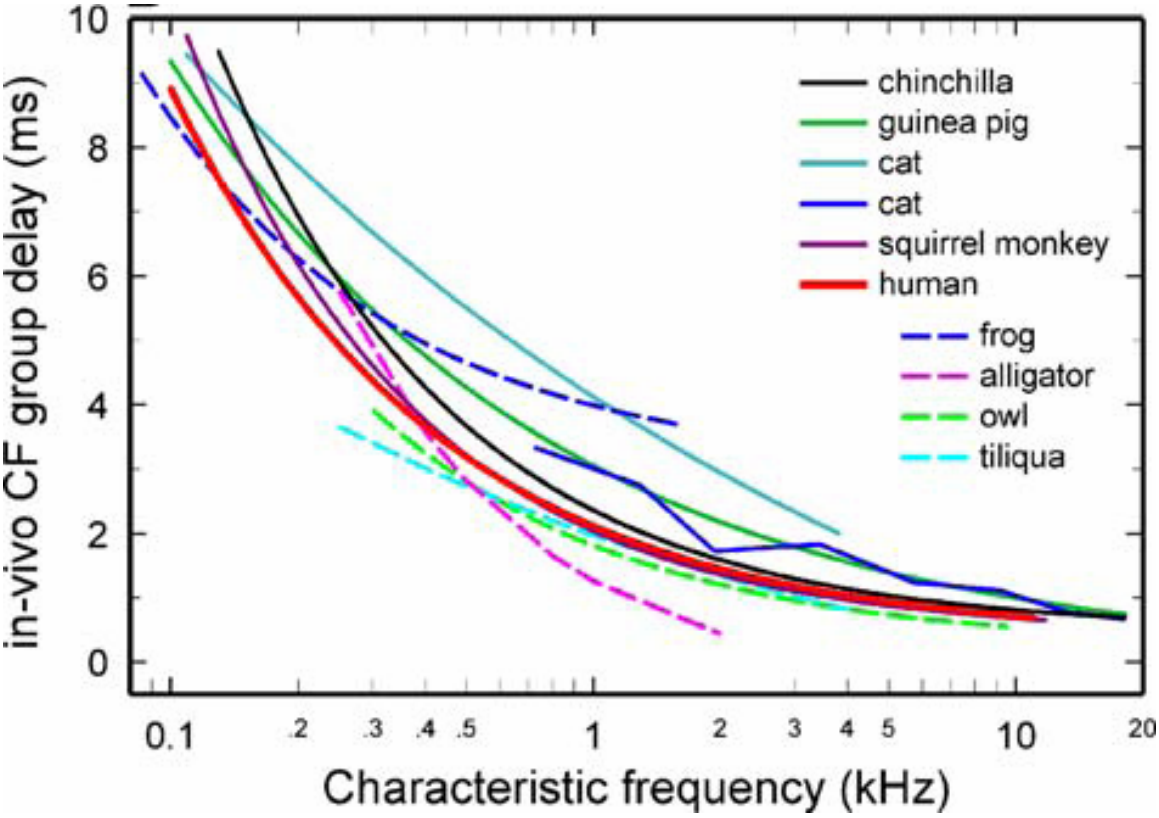


Figure 4. Comparison of mean first-spike latencies to various tone bursts in the hearing organs across species and individual animals. With kind permission from Springer Science+Business Media: Figure 8B, Ruggero, M. A. and A. N. Temchin (2007). Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods. J Assoc Res Otolaryngol. 8: 153-166.

These animal studies include three remarkable species. The first is barn owls, whose frequency representation of 5–10 kHz occupies more than half of the papillar length, so the latency is expected to show an unusually fast increase from 10 to 5 kHz [25] and thus differ from other species, but it does not. The second is bobtail lizards, Tiliqua, whose basilar membrane

does not support a traveling wave [28]. The third is tree frogs, Eleutherodactylus coqui, whose hearing organs lack a basilar membrane altogether [24].

These similarities (in the face of dramatic structural differences) suggest that the hearing organs across different species probably operate a common mechanism to decompose the input sound on the way to eliciting neural spikes. Since the similarity holds in animals that do not even have a basilar membrane, etc., such a common mechanism should comprise more than the mechanics of cochlear traveling waves.

It is contended herein that the shared attributes require a different interpretation than in the past, and proposed herein is an alternative model of Fourier analysis by the hearing organ. The conventional view of cochlear function, for example, implies that the hearing organ acts predominantly as a spectrum analyzer [29], converting a sound into components of different frequencies and magnitudes (two of the fundamental kinds of information). Phase encoding tends to be treated as an after-thought, having been long-embroiled in differing opinions of relevance and perceived limitations of synchronization abilities of auditory neurons (e.g., see 2.2.3.2).

Phase, nevertheless, affects the relationship among various auditory variables and broadly impacts auditory functions (Figure 5). The next chapter (Chapter 2, review of literature) provides in-depth background on phase encoding into auditory neural activity, wherein these auditory variables in both acoustic and electric hearing are explored. Also, the chapter will present the concept of a Fourier analyzer applicable to an across-species mechanism. The chapter ends with a set of hypotheses that were tested (Chapter 4 to 6) using a novel combination of meta-analysis and empirical designs and mathematical modeling (Chapter 3), over a vast body of information on neural response latency across species.

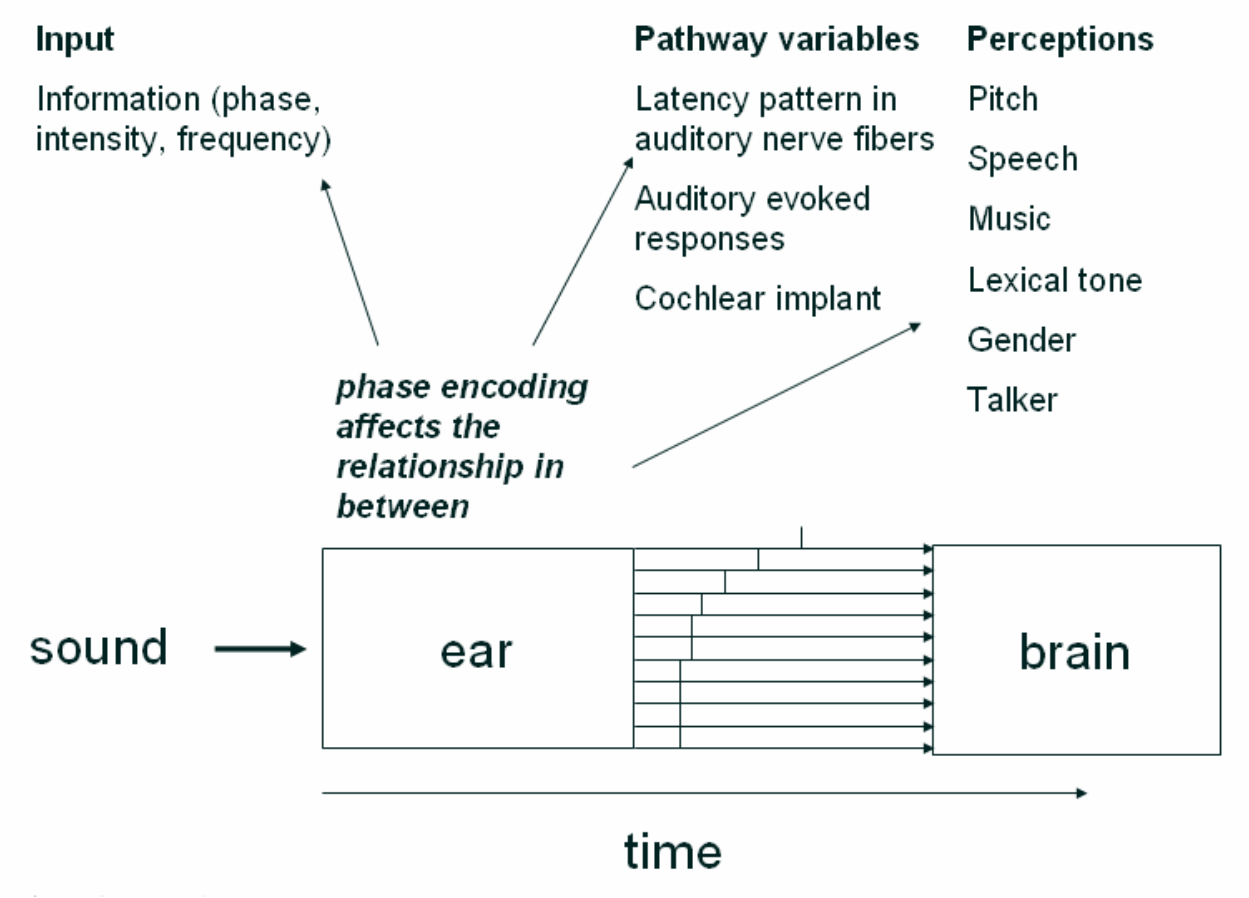


Figure 5. Phase affects the relationship among various auditory variables and broadly impacts auditory functions.

2.0 REVIEW OF LITERATURE

This literature review is devoted to exploring today's knowledge of phase encoding in acoustic hearing, the representation in auditory neural activity, and in electric hearing. It includes the parts of the whole auditory periphery, auditory neural responses to acoustic and electric stimulation, and phase processing of cochlear implants.

With this knowledge, it is possible to evaluate the idea of improving the electric hearing by providing an ever-closer approximation to the normal latency patterns of neural discharge at the periphery.

This section explores phase encoding in the auditory periphery from basic science and demonstrates that after nearly a century of relevant research in auditory encoding, there remain issues to be resolved.

2.1 ACOUSTIC HEARING AND PHASE

2.1.1 Sound waves

2.1.1.1 What is phase?

Sound is a series of mechanical waves created by a vibrating source in a medium [1], including solid, liquid or gas. Whenever an object in air vibrates, it causes longitudinal or

compression waves traveling through the air. These waves thus move away from the object in the medium as a disturbance called “sound”.

Sound waves are characterized by pressure change in the medium over time. This is why sound waves are usually recorded by measuring sound pressure over time. The two main variables, pressure and time, embed rich information of speech or music and give a sound wave three main features: frequency, intensity and phase. Frequency reflects how quickly pressure changes in the medium. Intensity reflects how much pressure is changed. Phase, which reflects a frequency dependent time difference, is the main topic here across all sections.

Frequency refers to how many cycles are completed per unit of time. It also is called temporal frequency since it is calculated on the axis of time. A directly related phenomenon to frequency is period, which is the duration of one cycle in a repeating event, so the period is the reciprocal of the frequency. Frequency is measured in cycles per second. The unit of this measurement is Hertz (Hz) in honor of the physicist Heinrich Hertz [1].

Intensity, at a given distance, is the sound energy that depends on the amplitude of the waves. For a pure tone, amplitude is the displacement between the equilibrium pressure and the peak excursion pressure. The most common approach to sound intensity measurement is to use the decibel scale with the unit dB.

Phase reflects a frequency dependent time delay when a periodic process or physical phenomenon is a time signal. In order to further understand what phase is, the section starts with a sound’s fundamental unit, a sinusoid. A sinusoid can either be generated from a sine function or a cosine function. The cosine function will be used as an example to connect concepts of phase in the sections below, and to avoid a sine-to-cosine phase conversion.

Any specific cosine curve can be completely shown in the equation below (Figure 6).

$$x(t) = A \cdot \cos(2\pi ft + \phi)$$

Figure 6. Equation of a simple cosine function.

where A is the amplitude of the cosine function and reflects its intensity, f is frequency, and ϕ is the starting phase of the cosine function. One way to predict a system's response to a sound, namely in a linear time-invariant system, is to investigate how the system processes a single sinusoid [30]. Although the auditory system is well known to be neither linear nor time-invariant, it may be reasonably evaluated as such, at least, in a piecewise manner or under reasonable constraints as characterize the considerations to follow and subsequently methods of study to be applied in this work.

The solid line in Figure 7 shows a 100-Hz cosine function with a period of 10 ms. The dashed line presents a reference with the same frequency and intensity but different phase. Since this is a time signal, the horizontal axis (time) can be converted to radians by plotting one cycle (of 10 ms) as 2π radians. So the phase according to the reference is the difference in radians between the signal and the reference. In Figure 7, the phase difference is $\frac{3\pi}{2}$ radians.

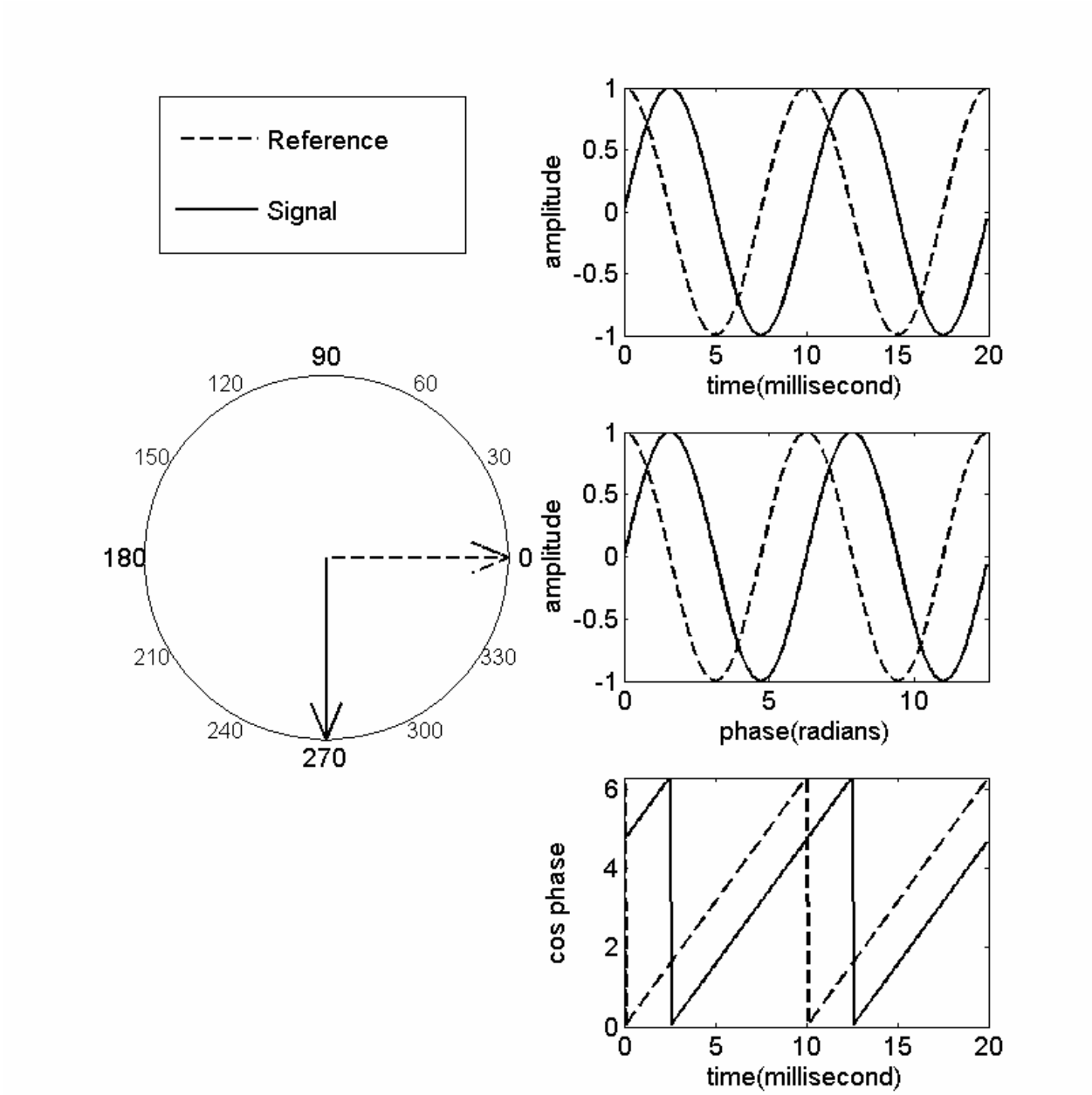


Figure 7. Phase in a cosine wave. Phase reflects the time delay between a reference and a signal (right upper panel), or the radian difference by plotting one cycle as 2π radians (right middle panel). The phase of the solid wave is 3π/2 (about 4.71 radians) which equals 7.5 ms. The continuous phase change of a simple cosine wave looks like a function of sawtooth (right lower panel). Phase is often shown in polar expression (left panel shows the starting phase).

In this example (Figure 7), because the frequency and intensity are the same, phase is the only fundamental information that makes these two sinusoids different.

If phase is viewed as a frequency dependent "time delay", there is an issue about which wave happens first. In Figure 7, since the first peak of the dash wave comes earlier than the solid wave, the dash wave is said to lead the solid wave. However, this may be controversial when it is considered from another point of view.

Many mathematical or physical theories and their applications are based on infinite sinusoids. If both sinusoids in Figure 7 are viewed as infinite sinusoids, the former statement is right by saying that the dash wave leads the solid wave for 2.5 ms. It is also correct to say that the dash wave lags the solid wave for 7.5 ms. Referring to the left panel, it may be found the dash leads the solid by $\frac{\pi}{2}$, or the solid leads the dash by $\frac{3\pi}{2}$.

There have been various names and technical terms that are closely related to phase. Researchers in different areas use different terms to describe phase related phenomenon. Learning the possible names of phase helps in investigating the literature and in interpreting the role of phase in different areas. Table 1 lists various names and technical terms that are closely related to phase, as well as their brief descriptions.

Table 1. Different names and descriptions that are closely related to phase. Among the names, temporal fine structure is frequently used in psychoacoustic studies, and Fourier phase is usually used for signals in the frequency domain with short-term Fourier transforms.

Name	Description
Phase	A particular point in a periodic process or phenomenon
Temporal phase	A frequency dependent time delay
Cosine phase	The starting phase of a waveform generated by cosine function
(Temporal) fine structure	The rapid oscillations in the structure of a signal
Fourier phase	The angle of a complex number (i.e.,frequency term) of a Fourier transform to the real axis in polar expressions

Among the diverse terms, the name “phase” gives the general concept of a specific point in a periodic phenomenon, which can be a wave varying on any kind of axis, such as time or space. For time signals, it is named temporal phase, which is a frequency dependent time delay. Cosine or sine phase indicates the starting phase of a waveform generated by cosine or sine functions, respectively. These two need to be clarified when trying to calculate the phase of a simple sinusoid.

Temporal phase and cosine phase usually describe concepts in the time domain. For the other representation of the same physical signal in the frequency domain, one popular name used to describe phase is Fourier phase. It is usually defined and widely used in engineering related areas. Briefly, Fourier phase is the angle of a frequency term, one output component of a Fourier transform, which is expressed in polar coordinate. For example the Fourier phase of $1+1i$ is the angle of this complex number in polar expression, i.e., $\frac{\pi}{4}$.

The above-mentioned names are derived from either phase itself or its characteristics, such as over what axis the wave is traveling, or how it is generated or calculated. However, although the term frequently used in psychoacoustic studies, “temporal fine structure” (TFS), is closely related to phase, the term itself does not tell how it derives from phase.

It is called temporal fine structure because of the observation that when plotting a sound wave over time, there are two components having different changing rates. One, envelope, is the usually slowly varying component that covers the peaks of the sound wave and represents its instantaneous intensity (Figure 8, middle panel). The other, temporal fine structure is the rapid oscillation in the structure of the signal (Figure 8, lower panel). When the sound is band-pass filtered, the oscillating rate is close to the center frequency of that filter band.

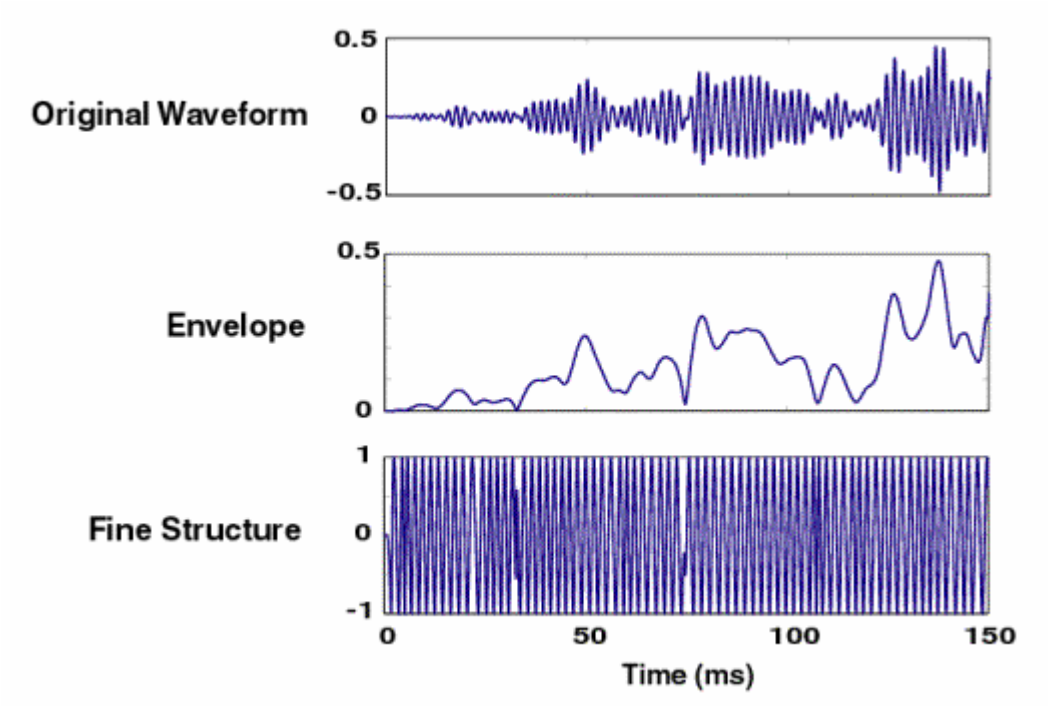


Figure 8. Envelope and temporal fine structure. The name "temporal fine structure" describes the "fine structure", a fast moving component, of a sound. It is highly related to phase. Adopted from research.meei.harvard.edu, with permission from Dr. Bertrand Delgutte.

Due to the controversies such as the definition and the unclear relationship between phase and temporal fine structure, the concept of temporal fine structure will not be used in the present study.

2.1.1.2 Measurement of phase

Calculating from period or frequency

For a sinusoid, the measurement of phase for a cosine wave can be calculated from its period (or frequency, i.e., period = 1/frequency), when the reference is given. For instance, when the reference phase angle is zero, the starting phase of the solid wave in Figure 7 is $\frac{3\pi}{2}$. The (instantaneous) phase angle of any following point at a certain time position t can be defined as the sum of $2\pi ft + \frac{3\pi}{2}$, which has been shown inside the parenthesis of the equation for a simple cosine wave in Figure 6. Referring to Figure 7, for these two 100 Hz cosine waves, the phase (to the reference signal with starting phase 0) at time 10 milliseconds (0.01 second) is $2\pi \times 100 \times 0.01 + \frac{3\pi}{2} = 10.9956$ radians. Because it is out of the range of 0 to 2π , it is an unwrapped phase. Note this unwrapped phase is equal to $10.9956 - 2\pi = 4.7124$ radians, or $\frac{3\pi}{2}$, as shown in the right lower panel of Figure 7. The phase difference between these two cosine waves at a time t is always $\frac{3\pi}{2}$, because $(2\pi ft + \frac{3\pi}{2}) - (2\pi ft + 0) = \frac{3\pi}{2}$.

It is straightforward to calculate the phase difference between two sinusoids with the same frequency and intensity (the other 2 fundamental kinds of information). It may become another issue in complex sounds. In one complex sound, there may be phase interference among

the tone components (sinusoids) and a complex sound usually does not have clear repeating cycles.

By Fourier transform

A complex sound (or any signal) can be decomposed by Fourier transform into sinusoids [31, 32]. The operation transforms the real time function (the sound) into another complex frequency function. In the application of sound processing, the domain of the original sound is time and is accordingly called the time domain. The domain of the new function is called the frequency domain, and the new function itself is called the frequency-domain representation of the original sound.

These two interchangeable functions represent the same physical phenomenon. In other words, the three kinds of fundamental information (intensity, frequency and phase) in the original sound are represented in the three kinds of fundamental information in the sinusoids. For each component, Fourier phase is the angle of a frequency term (a complex number) to the real axis in its polar expression.

Figure 9 demonstrates the two domains (upper panel and the lower set of sinusoids) and the Fourier phases of the first three sinusoid components (2^{nd} to 4^{th} frequency terms, right polar expressions). The original signal is a 5-ms 1000 Hz (chosen arbitrarily for demonstration) tone burst sampled at a rate of 40,000 Hz (chosen arbitrarily for demonstration). By use of a 400-point (chosen arbitrarily for demonstration) fast Fourier transform, the original sound can be decomposed and represented by a set of sinusoids with various frequencies, intensities (magnitudes) and Fourier phases.

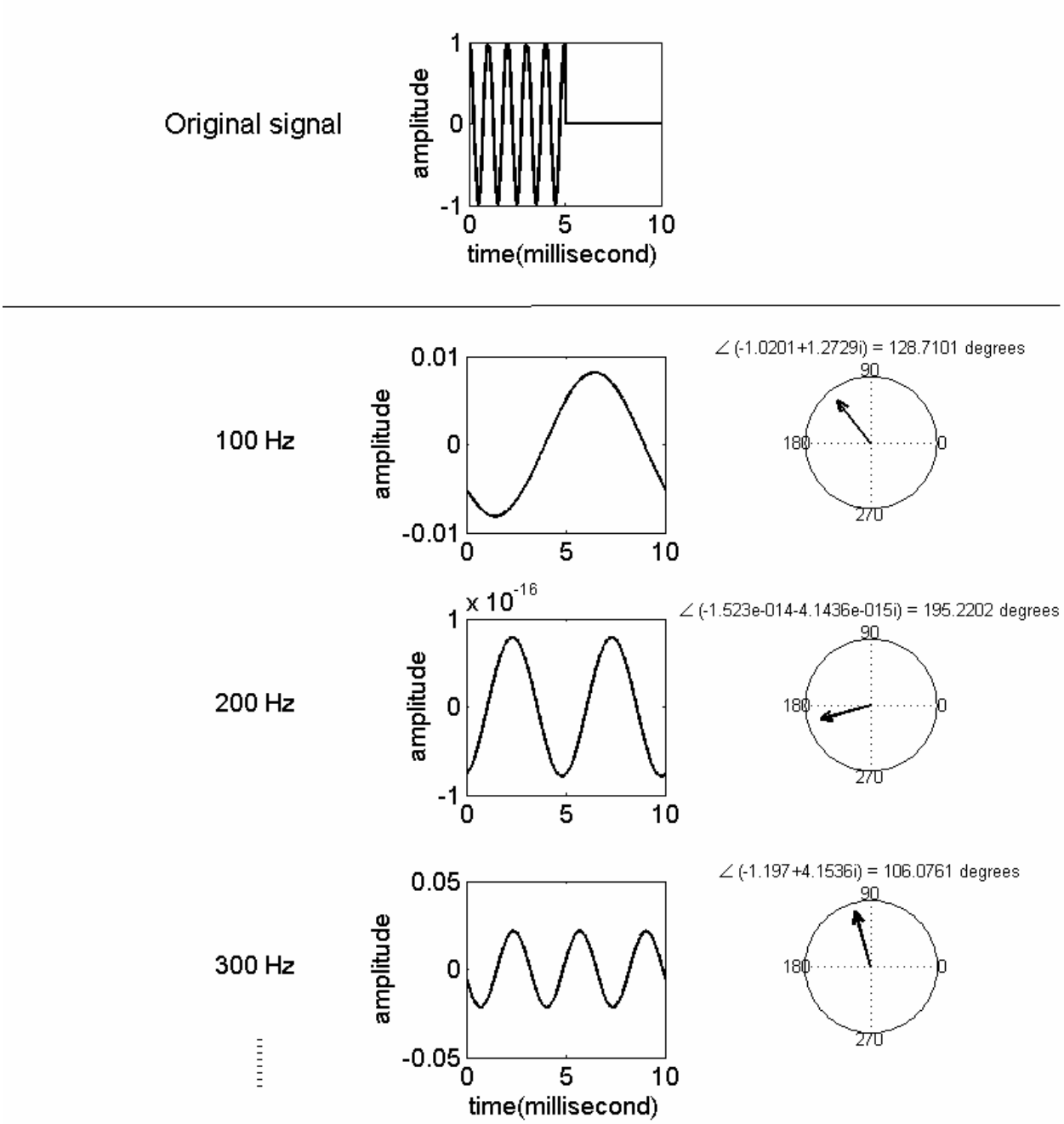


Figure 9. Fourier phase. A sound can be decomposed by Fourier transform into sinusoids. The original signal is a 1000 Hz (chosen arbitrarily for demonstration) tone burst sampled at a rate of 40,000 Hz, from the equation x(t) = cos(2 · π · 1000 · t), where t = n/40,000, 1≦n≦200. By use of a 400-point fast Fourier transform, the original sound can be decomposed and represented by a set of sinusoids (every 40k/400=100 Hz) with various frequencies, intensities (magnitudes) and Fourier phases. For example, the

Fourier phase of the 100 Hz sinusoid component is calculated from the inverse tangent (arctangent) for the 100 Hz frequency term, which is arctan (-1.0201 + 1.2729i) and equals to 129 degrees (or 2.2464 radians). Complex numbers shown on the tops of right polar coordinates are the frequency terms of the fast Fourier transform.

For instance, the Fourier phase of the 100 Hz sinusoid component, 129 degrees, is calculated from the inverse tangent (arctangent) for the 100 Hz frequency term (complex numbers shown on the tops of right panels), which is arctan (-1.0201 + 1.2729i) and equals 129 degrees (or 2.2464 radians), when $0 \leq \text{Fourier phase} < 2\pi$. The defined range must be considered while calculating the phase difference between a constant reference and the Fourier phase (e.g., see the step 3 in 3.2.1).

Since the decomposed components are all simple sinusoids, Fourier phase usually simplifies the interpretation when the effect of phase is investigated.

Furthermore, in a series of Fourier transformed sinusoids, intensity information is encoded by the amplitude of each sinusoid and frequency information is encoded by the serial number of the component (i.e., frequency term, whose frequency equals the frequency of the sinusoid).

If one is asked to reproduce a set of various sinusoid components, given the information of their frequencies and amplitudes, a way to encode phase information is to "mark" the first amplitude maximum for each sinusoid. Marking each sinusoid component with the first amplitude maximum creates a specific pattern among the sinusoids.

For example, marking the first amplitude maximum for each sinusoid component in Figure 9, such as those dots in the right upper panels in Figure 10, forms a specific pattern (Figure 10, lower panel) for the original tone burst (see 3.2 for details). This specific pattern

created by marking the first amplitude maximum for each sinusoid component is defined as a Fourier pattern. This concept is unique to this work, intended to characterize a feature of temporal processing—specifically phase encoding—that is not readily apparent in more conventional analyses.

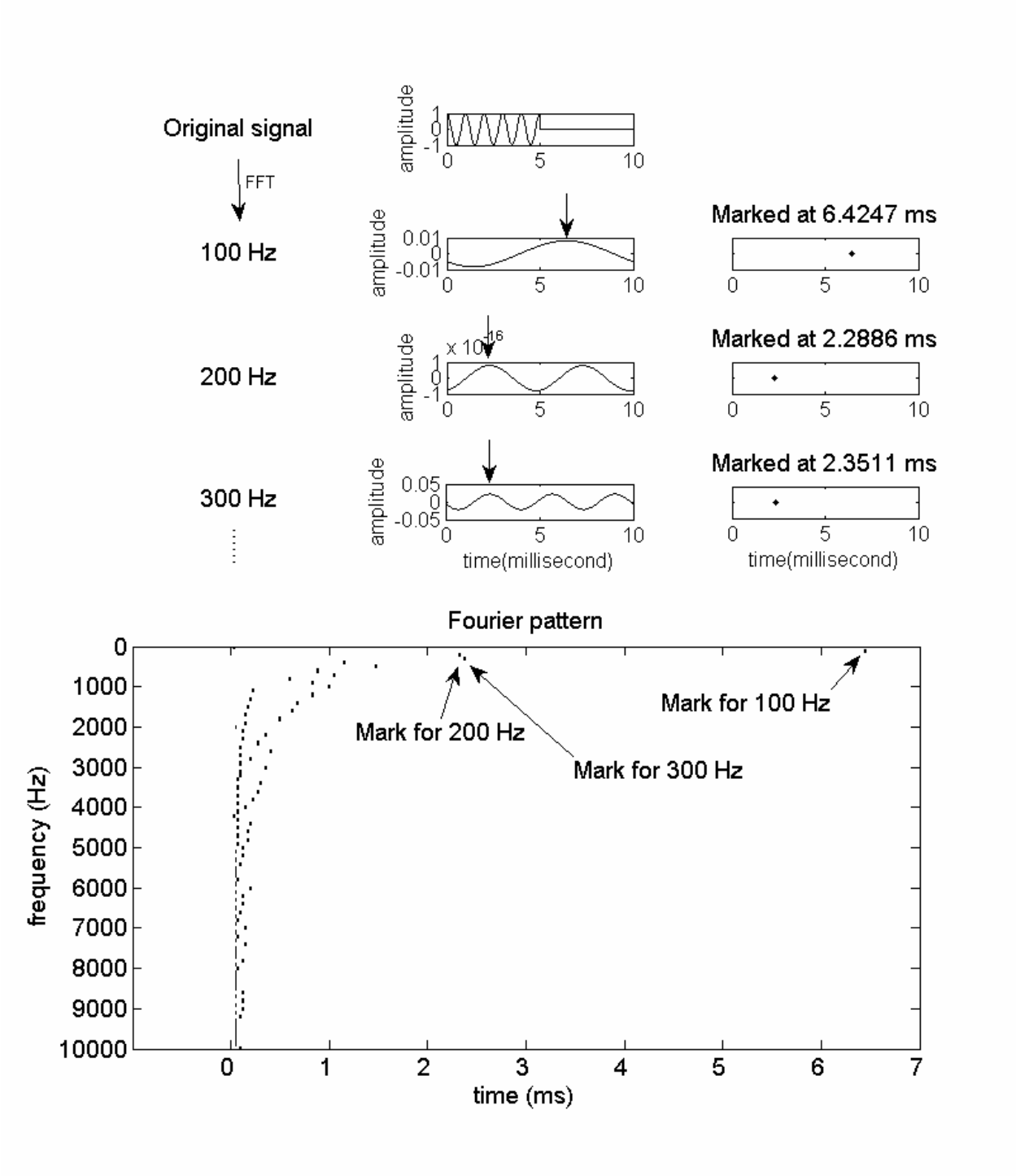


Figure 10. Fourier pattern. Marking the first amplitude maximum for each sinusoid creates a Fourier pattern among the sinusoids. Arrows indicate the marking times that are shown in right panels and the panel below. Since phase is defined as a frequency-dependent time delay and the Fourier pattern is a

frequency-dependent temporal pattern, it is reasonable to view the Fourier pattern as a "phase code" (encoding of phase information).

The latency of the marking in each sinusoid component can be calculated according to the phase difference between 2π and its Fourier phase, by the time-shift property of Fourier transform (see 3.2.1 for details).

Because phase is defined as a frequency-dependent time delay and the Fourier pattern is a frequency-dependent temporal pattern, it is reasonable to view the Fourier pattern as a "phase code" (that encodes phase information).

In addition to providing sinusoids as the basic units (to make phase investigation and interpretation easier), and creating a phase encoded Fourier pattern, short-term Fourier transforms are currently employed in processing strategies of cochlear implants, including advanced combination encoder (ACE)[33] and HiRes with Fidelity 120 (HiRes 120)[34]. Understanding the Fourier analysis seems to be especially important for phase investigations in electric hearing.

2.1.1.3 Significance of phase

Since temporal phase differences reflect frequency dependent time delays, phase processing in our auditory system is related to frequency dependent temporal processing, which may happen everywhere starting from the outer, middle and inner ears, to the nerve and central auditory cortex.

Time scales

Although these are neither absolute nor universally accepted numbers, among various auditory perceptions, different perceptions require different temporal resolution on the time scale: microseconds (one millionth (10^{-6}) of a second) for the processing of binaural cues, milliseconds (one thousandth (10^{-3}) of a second) for processing of temporal synchrony, tens of milliseconds for the processing of speech transients and voicing information, and hundreds of milliseconds to seconds for the processing of prosodic and suprasegmental cues [35].

Perception of phase

Perception of phase and how phase is processed in the auditory system are not well known so far. There seem to be rare direct and systematic studies for perception of phase. Among different listening situations, languages and subjects, phase has been involved in speech perception (especially in noise), lexical-tone perception, and music perception in the literature.

Speech perception

Oppenheim and Lim investigated the effect of phase on sentence perception in the Fourier representation [36]. They showed the importance of phase in speech perception in two experiments. In each experiment, they decomposed sentences by Fourier transform into phase and magnitude.

In one experiment, they showed three spectrograms: an original sentence, a phase-only sentence and a magnitude-only sentence. The phase-only sentence had the Fourier phase and unity magnitude. The magnitude-only sentence had the Fourier magnitude and zero phase. As suggested by the spectrograms and confirmed by normal listeners, they suggested intelligibility was lost in the magnitude-only reconstruction but not in the phase-only reconstruction.

In the other experiment, they decomposed two sentences A and B by Fourier transform into phase and magnitude. The authors showed the spectrogram of the sentence obtained from the phase of sentence A and the magnitude of sentence B most closely resembled the spectrogram of sentence A. The spectrogram of the sentence obtained from the phase of sentence B and the magnitude of sentence A most closely resembled the spectrogram of sentence B. In other words, the spectrogram of the component-exchanged sentence was more similar to that of the sentence providing phase.

Instead of calculating phase by Fourier transform, a speech sample's or its filtered component's instantaneous phase also can be calculated by Hilbert transform. There were numerous studies investigating the effect of temporal fine structure on speech perception by use of Hilbert transform.

Because of the controversial relationship between phase and temporal fine structure (which contains phase and at least part of frequency information), it seems inappropriate to interpret the effect of temporal fine structure on speech perception as the direct effect of phase on speech perception. In addition to this indirect relationship, there are some potential problems when researchers used temporal fine structure to study speech perception.

Current studies usually generate "TFS-speech" by decomposing a speech sample with Hilbert transform and removing the envelope cue [37]. A potential problem with TFS-speech was identified by Ghitza [38]. His study showed, although envelope cues were removed from TFS-speech, they may be automatically reconstructed by cochlear filters. This was especially noticeable when only a few broad filter bands were used in the processing [39].

Considering the possible problems when investigating TFS in speech perception, Hopkins et al tried a new approach. They band-pass filtered signals into channels. They set a cut-

off channel to remove TFS information above the cut-off channel, and then combined all the channels. Their data showed more TFS information resulted in better speech reception thresholds. This study suggested that TFS, and phase information, is related to speech reception thresholds [40].

It is also generally accepted that TFS information plays an important role in speech perception in noise [41]. The ability also was called "listening in the background dips". Although the mechanism was not clearly explored, some studies showed perceptual segregation of mixtures of sounds might be involved, such as that by Drennan et al in 2003 [42].

Lexical tone perception

Xu and Pfingst followed the method of generating auditory chimera [9] to exchange the envelope and TFS components among 4 lexical tones of a same Mandarin Chinese monosyllable [43]. Their study showed subjects based their identification of the lexical tone significantly and mainly on the fine structure rather than the envelope. The percentage of responses that were consistent with the fine structure of the chimeric stimuli (that had envelope of one lexical tone and temporal fine structure of another) could be as high as 92, 90 and 85% in 4, 8 and 16 frequency bands, respectively. These results suggested the contribution of envelope information in lexical tone recognition was only about 10%. Their results supported the significance of phase information in lexical-tone perception.

As it plays an important role in speech perception, phase is likely even more important in lexical tone recognition than envelope information.

In summary, temporal phase as defined as a frequency-dependent time delay is one of the three fundamental kinds of information (intensity, frequency and phase) in a sinusoid, which is the

basic unit of a sound. These three kinds of information enable one to reproduce the original sinusoid components. Putting all the sinusoid components together regenerates the original sound signal.

Phase information can be measured by calculating from a sinusoid's period or frequency, by Hilbert transform, and by Fourier transform (Fourier phase).

In Fourier analysis, one way to encode phase information is to "mark" the first amplitude maximum for each sinusoid. The "markings" creates a pattern among the sinusoids and is defined as a Fourier pattern, which is a frequency-dependent temporal pattern and is therefore a phase code. This concept is unique to this work, intended to characterize a feature of temporal processing—specifically phase encoding—that is not readily apparent in more conventional analyses.

Phase information is involved in many auditory perceptions such as pitch perception, speech perception (especially in noise), and lexical-tone perception. With the knowledge of the definition, measurement and significance of phase, the next issue to be reviewed is how the peripheral auditory system processes phase information.

2.1.2 Head and pinna effect

2.1.2.1 Phase in monaural hearing

In air, the speed of sound propagation is the same across all frequencies. The speed is only affected by stiffness and density of the air, or generally by temperature and air pressure.

The basic role of the external ear is to collect impinging sound waves and direct them toward the inner ear's sensory cells. In addition, the study of Batteau suggested that the external ear can be considered as a continuous reflecting and delaying system [44]. The study showed

there was no relative phase shift of components but an amplitude change among components of different frequencies.

These differences in the amplitude versus frequency profiles may result from the shape of the external ear. The amplitude differences play a role in spatial hearing, especially in vertical locating in monaural hearing. The amplitude differences also give interaural spectral difference (ISD) in binaural listening.

The concept remains today that the pinna is a continuous reflecting and delaying system. In recent literature about head related transfer function (HRTF, a filtering process using a complex frequency response function) or head related impulse response (HRIR), the concept remains that pinna has no effect on phase information. The time delay across frequency only changes when the speed of sound changes (e.g., [45]).

2.1.2.2 Phase in binaural hearing

The relative positions of the two external ears on the head allow comparison of two samples of incoming sounds, which provide at least two categories of important information, interaural time difference (ITD) and interaural intensity difference (IID) [1].

When a sound is produced by a source close to one ear and because of the sound's finite speed, it reaches the ear first. The time difference information provided by two different samples is an ITD. Under an identical circumstance, the ITD is the same across all frequencies. Another difference regarding phase, interaural phase difference (IPD), may exist because of the same head effect. Although ITD and IPD result from the relationship of two-ear locations, they may give different information. How the brain uses IPD together with ITD is not fully understood.

The second kind of information given by two ears at different locations is the interaural intensity difference (IID). It exists because a sound reaches one ear first and must diffuse around the head to reach the far ear, losing amplitude in the process.

The information provided by the head and the external ears (ITD, IPD, IID and ISD (see 2.1.2.1)), may be further processed by the central auditory pathway and the auditory cortex, and contribute to various perceptions such as three-dimensional spatial hearing (or sense of sound space) [1], speech perception in noise, and perceptual segregation of sounds [42].

In summary, in normal monaural hearing, the pinna is a continuous reflecting and delaying system. The time delay across all frequencies is usually constant. It only changes when the stiffness and density of the air changes. These changes are usually impacted by temperature.

For cochlear-implant users, because the microphone locates outside the pinna, implant users are not provided with the differences in the amplitude versus frequency profiles and therefore the ISD. If the technology advance ultimately overcomes the difficulty of delivering phase, bilateral implants may give approximated ITD, IPD and IID. It is reasonable to reconsider the location of the microphone in order to provide the differences in the amplitude versus frequency profiles and the ISD.

2.1.3 Role of the middle ear

The middle ear bridges sound transmission between the external ear and the cochlea. The human middle ear comprises the tympanic membrane, three tiny ossicles and their supporting structures. All these structures may have different phase response at various frequencies. In addition, middle ear structures (including mass) may vary between two different species (e.g., fusion of malleus

and incus) ,or even two individuals of the same species (e.g., small and big gerbils). Furthermore, in vivo, it has been suggested that the middle ear is not a passive structure but one that is under continuous central control [30]. Thus, phase response in this dynamic system is more complicated than that in the external ear, a continuous reflecting and delaying system [44, 45].

Reports of the phase response of the middle ear seemed to differ among previous animal studies. This might result from the nature that phase response is different in different animals, but might also be affected by the measurement method of the ultra-short delays, the conditions of the animals such as age, and the sensitivity of the measuring equipment.

Overstreet and Ruggero anesthetized gerbils with ketamin HCl and sodium pentobarbital, hydrated with Pedialyte, placed the gerbils upon a vibration-isolation table within a sound-insulated chamber and controlled the rectal temperature at 39 ℃ to measure stapes responses [46]. Having vibration, sound and temperature controlled, their results still showed large variation in phase response.

In addition, they reviewed some studies investigating phase-response of the gerbil middle ear by various methods such as measuring stapes velocity [47], umbo velocity [48, 49], and scala-vestibuli pressure [50]. Figure 11 shows the variation of results of phase response of gerbils' middle ear among different measuring methods.

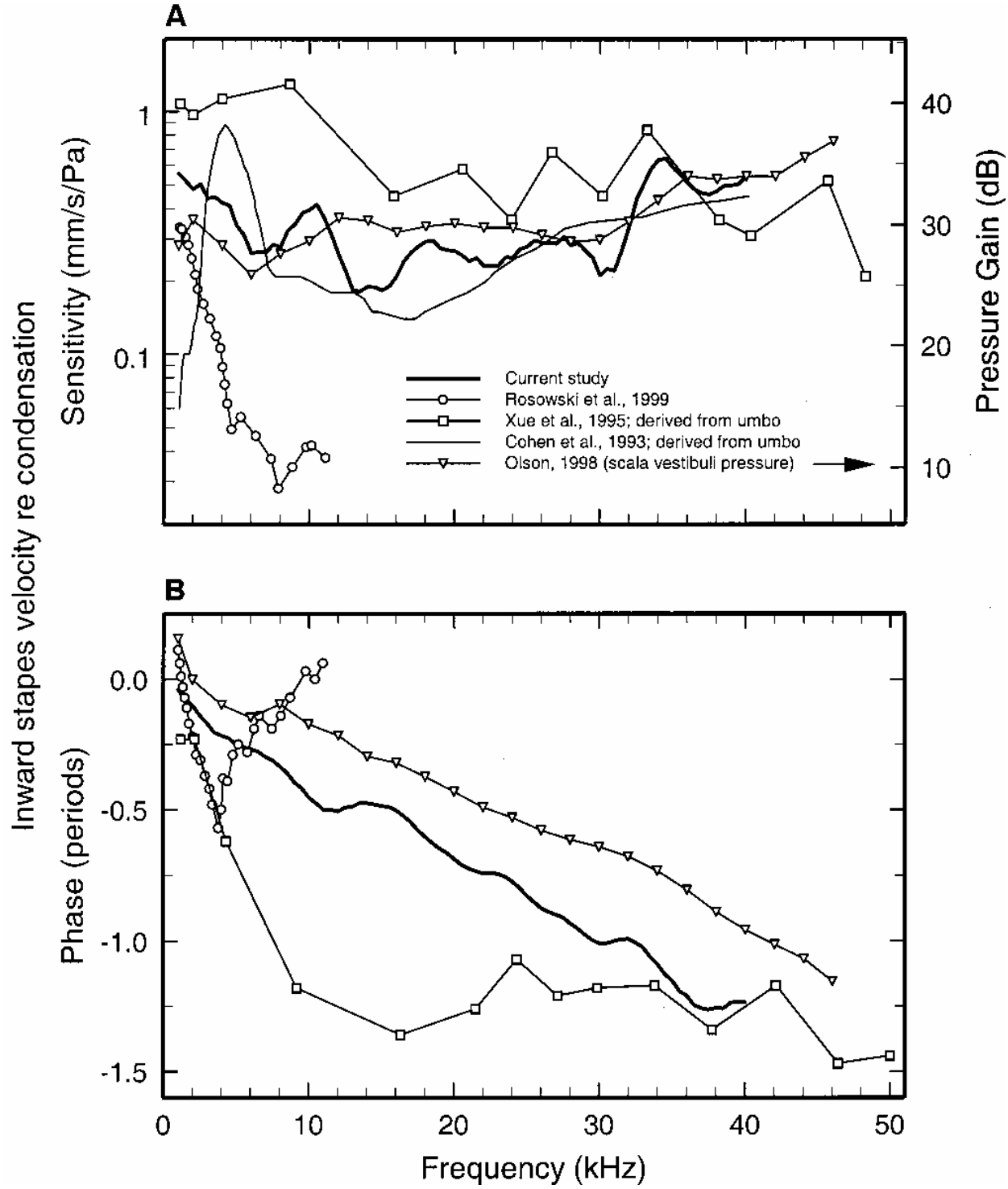


Figure 11. Phase responses of gerbil's middle ear. The reviewed figure showed that phase response of gerbil's middle ear might be a frequency-dependent function. But it also showed the large variation among different measuring methods, or even between two studies using the same measuring method such as stapes velocity. Reprinted with permission from Fig. 7, Overstreet, E. H., 3rd and M. A. Ruggero (2002).

"Development of wide-band middle ear transmission in the Mongolian gerbil." J Acoust Soc Am 111(1 Pt 1): 261-270. Copyright 2002, Acoustic Society of America.

Puria and Allen anesthetized healthy mature cats with chlorpromazine and sodium pentobarbital, and calibrated temperature to measure middle-ear impedance near the tympanic membrane [51]. Their results suggested the tympanic membrane was a lossless transmission line with a frequency-independent delay of 36 microseconds, which was estimated by converting the impedance data to reflectance and analyzing the reflectance group delay. Their study also suggested that the ossicles were represented as lumped-parameter elements for all three conditions they tested, including intact ossicles, drained cochlea, and disarticulated stapes.

According to studies with anesthetized animals, it remains unclear whether there is a similarity of middle ear delays among various frequencies across different species, or even whether these delays are frequency-independent (i.e., constant) or not.

One generally accepted observation is that these middle-ear delays at all frequencies are very short [46, 51, 52], less than 0.1 ms. For instance in Figure 11, the phase difference between 0 and 4k Hz for the gerbil was about 0.2 cycle, which was only about 0.05 ms (i.e., $0.2 \times \frac{1000}{4000}$).

In summary, although the middle-ear delays at all frequencies are very short (less than 0.1 ms), the phase responses may vary among different species or even different individuals of the same species. It seems to be unanswered that whether each middle ear is individualized for the hearing organ (so it does not affect the accuracy of phase encoding), or whether the variety in that kind of short time scale has no significant effect on hearing.

2.1.4 Cochlear delays

A sound is transmitted via the ossicles to the cochlea. The human cochlea is a complicated system covered by a seashell-shaped outer bone. It contains a spiral of 3 main channels with cochlear fluid and the sensory tissue known as the cochlear partition. Motion of the stapes, the input of the cochlea, sets the cochlear fluid in motion, which in turn sets the cochlear partition near the stapes in motion. The initial signal in the cochlea is the continuous pressure change of perilymph close to the stapes. The output signal is the set of discrete and time-specific spikes in each auditory nerve fiber. Cochlear mechanics, between the two, play an important but not completely understood role in phase processing.

2.1.4.1 Cochlear mechanics

Any structure or its function between the input and output of the cochlea may contribute an effect on phase information, e.g., changes of the fluid pressure [53], fluid flow [54, 55], relative movement of tectorial membrane [54, 56, 57], shape deformations of the organ of Corti [58], excitation of outer hair cells [16, 59-65] and inner cells [16, 59-61, 65-67], and the basilar membrane displacement which is highly related to the cochlear traveling waves (see sections below). They all take some time and they usually interact with one another (e.g., [55, 57, 59, 64]). The fundamental elements of cochlear traveling waves are fluid pressure, fluid motion, partition forces and partition motion [53].

As mentioned above, the initial signal in the cochlea is the pressure change of perilymph close to the stapes in scala vestibuli. Olson inserted a pressure sensor in basal scala tympani, scala vestibuli (where the characteristic frequency was about 30k Hz) and turn-one scala tympani (where the characteristic frequency was about 20k Hz) in young adult gerbils. The stimuli were

tone bursts ranging from 15k to 40k Hz. Comparing fluid pressures at the same location to different tone bursts and fluid pressures at the two locations to an identical tone burst, she demonstrated that the phase of the scala-tympani pressure relative to the scala-vestibuli pressure was both stimulus-frequency and cochlear-position dependent [53]. In other words, at the same cochlear location, her results showed that the delay of pressure change decreased when the stimulus frequencies decreased.

Changes of fluid pressure set the cochlear fluid and the cochlear partition (i.e., the vertical structure of sensory tissue) in motion. For example, boundary displacements and pressure change in scala media drive the flow of endolymph in the inner spiral sulcus and the subtectorial space [54]. Direct evidence of the oscillatory motion of the medial olivocochlear fibers also suggested that this fiber motion was caused by the oscillatory fluid flow in the tunnel of Corti [55]. Since the sensory stereocilia on the inner hair cells are "fluid coupled" to mechanical stimuli, and traveling of the fluid flow takes time, it is believed the cochlear fluid motion (a 3-dimentional flow) contributes to processing of phase information. However, how this 3-dimentional fluid motion encodes the specific spike timings is not fully understood.

In addition to fluid motion, changes of fluid pressure also set the cochlear partition on the basilar membrane in motion. It displaces and changes its shape passively, following physical laws, or positively by somatic motility of outer hair cells in the mammalian cochlea. Outer hair cells generate voltage-dependent length changes at acoustic rates [63, 64], which deform the cochlear partition.

Dallos et al. [16] investigated the roles of the basilar membrane displacement and the speed of the displacement by recording cochlear microphonic potentials (i.e., the electric response of the cochlear hair cells to acoustic stimulation) to a triangular stimulus in guinea pigs.

They showed the cochlear microphonics were approximately proportional to the displacement of the basilar membrane. After kanamycin treatment, the displacement microphonic disappeared in the first cochlear turn where the outer hair cells were destroyed, but remained in the third cochlear turn where outer hair cells were remained.

The work of Dallos et al. [16] is sufficient for a working hypothesis that outer hair cells respond to basilar membrane displacement and the inner hair cells to its displacement speed, the time derivative of the displacement [17]. If the outer hair cells respond to displacement and the inner hair cells to speed, these responses should reflect in auditory neural discharges. If there is no interaction between the inner and outer hair cells, most units should show a speed response, and a few, a displacement response. In the event of interaction, both responses could be present in the same fibers. Therefore, Zwislocki and Sokolich [59] recorded action potentials in Mongolian gerbils to trapezoidal and triangular stimuli with a fundamental frequency of 40 Hz. Comparing the acoustic signal, the cochlear microphonic potentials and the neural discharge rate, their results indicated that nearly all auditory neurons respond to both displacement and displacement speed of the basilar membrane. This suggested that the outer and inner hair cells interact.

Investigating the role of outer hair cells in phase-lock phenomenon, Woolf et al. examined phase-locking across frequency in auditory nerve fibers in chinchillas. They found selective outer-hair-cell loss exhibited a marked reduction in the frequency range over which they could phase-lock [61]. All these results together suggest that outer hair cells play an essential role in phase encoding of auditory neural discharges.

Phase processing of the cochlear mechanics is still a mystery. Even to a simple stimulus like a tone burst, the interaction among the 4 main elements in cochlear mechanics (fluid

pressure, fluid motion, partition forces and partition motion) is unclear. Although some studies (e.g., [65, 67]) reported that the phase difference could be as much as 90 degrees between the receptor potential from an individual inner hair cell and that from the organ of Corti fluid space (reflecting outer hair cell receptor currents), the underlying mechanism of the amount and the theoretical meaning remain unknown.

The directionality of hair cell stimulation combined with the vibration of the basilar membrane causes the auditory nerve fiber action potentials. In the normally functioning mammalian cochlea, the vibration of the basilar membrane supports a mechanism known as the cochlear traveling wave. In humans, these waves were first described by von Bekesy in human cadavers, progressing from the oval window area to the apex, that is from effectively the sound input from the middle ear and traveling along the inner ear to its apical terminus [18].

The excitatory waves of motion elicited by sound energy thus travel from high-frequency-sensitive areas to low-frequency-sensitive areas. As this is a time-consuming process, the traveling-wave mechanism transmits information to the brain about high-frequency components earlier than that pertaining to low-frequencies, as waves travel (propagation) along the hearing organ, like sounds in air.

2.1.4.2 Traveling-Wave Speed (Velocity)

Physically speaking, velocity is the rate of change of position. It is a vector quantity. However in the literature, traveling-wave velocity usually indicated traveling-wave speed, the scalar absolute value of velocity. Regardless of the three-dimensional structure and moving direction, traveling-wave speed can be measured in meters per second (m/s). If the time is known for how long it takes for one wave to “travel” between two known positions in the cochlea, it is possible to calculate the mean speed between these two positions. Measuring speed became one

of the major interests to investigate traveling waves, after von Bekesy discovered the traveling-wave motion on the basilar membrane and provided a map for characteristic frequency vs. cochlear position.

Zerlin reported a psychoacoustic study to measure traveling-wave speed, based on the following theory [68]:

"When two tone bursts of equal intensity but somewhat different frequency are simultaneously presented to both ears, a single coherent tonal image can be heard lateralized toward the higher-frequency side. When the higher-frequency signal is delayed by an appropriate amount, the sound image is brought to the midline. According to place theory, the two different frequencies are now arriving simultaneously at their designated places on the cochlear partition."

Five tone-burst pairs were used to obtain the "traveling time" between the two frequencies. The five pairs were: 400 and 500; 800 and 1000; 1600 and 2000; 3200 and 4000; 5000 and 6300 Hz. The "traveling distance" of one pair was calculated based on the cochlear map provided by von Bekesy [18]. So Zerlin calculated the traveling distance divided by the traveling time for each pair to get the traveling-wave speed at that frequency.

Another method of measuring traveling-wave speed is made possible by using derived response of ABR [69-72]. Gould and Sobhy conducted a study using the typical derived response technique. With a click stimulus, the derived band ABR was achieved by 3 steps: 1) they masked out a region of cochlear partition to get ABR A. 2) they masked out a larger region of cochlear partition to get ABR B (i.e., a smaller region was stimulated). 3) they subtracted ABR B from ABR A [69]. Then they calculated the traveling time by using the latency values of wave V in the derived responses. For example, they determined the two mean latencies of wave V for the

derived response centered at 2000 and 4000 Hz, and the traveling time was calculated as the latency difference between them.

The traveling distance was obtained in their study according to von Bekesy's cochlear map as well. Therefore, having the cochlear map and the data of the wave V latencies for the derived ABRs for 1000, 2000, 4000 and 7500 Hz, they were able to calculate the traveling-wave speed at the cochlear partition for 1000, 2000 and 4000 Hz.

Although there may be an argument that their results from click stimuli should not be compared to previous results using tone bursts such as the study by Zerlin [68], Figure 12 shows their result, Zerlin's result, and the results of some other previous studies. They noticed the traveling-wave speeds they measured by derived band ABRs were faster than all the others. They interpreted that it might be because of slight differences in filter cutoff frequencies between theirs and previous studies.

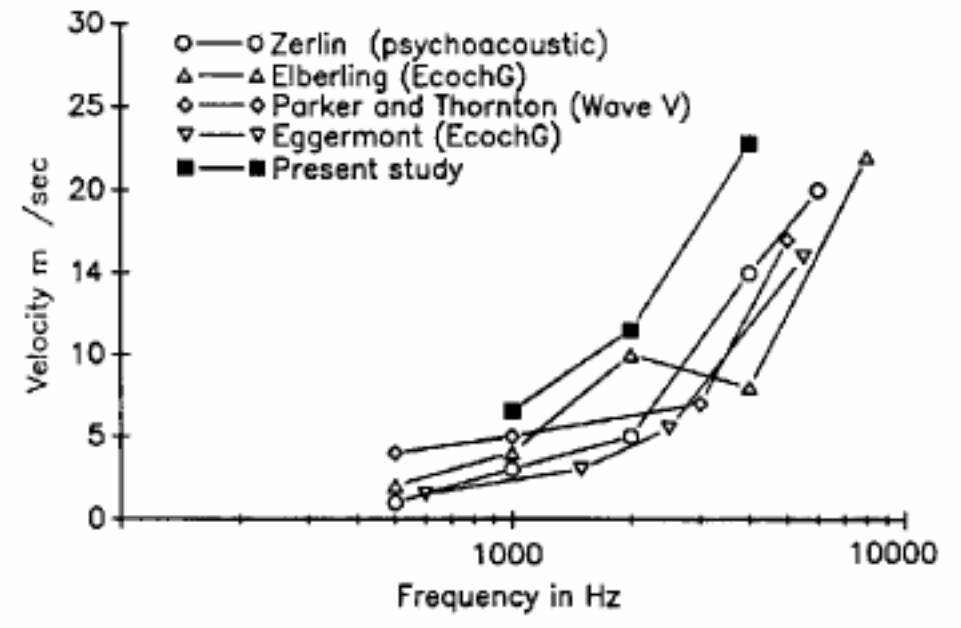


Figure 12. Traveling-wave speed of five previous studies. Traveling-wave speed was calculated by derived ABR technique. Reprinted with permission from Gould, H. J. and O. A. Sobhy (1992). "Using the

derived auditory brain stem response to estimate traveling wave velocity." Ear Hear 13(2): 96-101. Copyright 1992, Wolters Kluwer Health.

Don et al. [71] implemented the value of effectively unwrapping the putative cochlear phase dispersion, namely via their stacked-band analysis of the ABR, in clinical application. They showed that a smaller wave V of the stacked derived-band ABR amplitude is sensitive to the presence of small intracanalicular tumors in patients and has excellent specificity for the absence of tumors in normal-hearing individuals [73, 74].

The traveling-wave speed also was investigated as a potential diagnostic method for detecting Meniere's disease. It was proposed based on the assumption that, if endolymphatic hydrops is associated with endolymphatic hypertension, the increased pressure will result in a stiffer basilar membrane and, hence, an increase in the traveling-wave speed along the basilar membrane [72, 75-77].

It remains controversial whether ABRs can be used to detect Meniere's disease. For instance, Murray et al. using tone burst ABRs showed that wave V latencies and estimates of cochlear travel time cannot be used to distinguish Meniere's disease from other forms of cochlear hearing loss or from normal-hearing ears [78]. However, the results from Don et al. [79] showed that using high-passed noise masked ABRs successfully detected Meniere's disease. What needs attention here is that the technique used to detect small tumors is stacked ABR while the one used for Meniere's test is only high-passed noise masked ABRs to clicks.

2.1.4.3 Group delay of cochlear traveling-waves

In addition to traveling-wave speed, several similar concepts were developed in order to investigate the group delay of human cochlear traveling waves at different frequencies. Various

stimuli such as tone bursts and clicks were used with some non-invasive measurements such as auditory brainstem responses (ABR) and otoacoustic emissions (OAE). Among reports using tone bursts as input materials, two typical studies could be the ABR study by Neely et al. in 1988 and OAE study by Shera et al. in 2002.

Neely et al. used tone bursts at ten frequencies (0.25 – 8 kHz) and nine intensities (20 to 100 dB SPL) to measure ABR V waves among 20 normal-hearing young adults. The lengths of the tone bursts they used had frequency-dependent durations in order to keep at least 3 cycles at low frequencies. Their results are shown in Figure 13. They recognized that their basilar membrane delay estimates were much longer than most previously reported values, and interpreted this as being the result of human subjects having longer, more compliant cochleas and, therefore longer cochlear travel times than most experimental animals. Cochlear travel-time estimates derived from measurement on laboratory animals may not be good estimate of cochlea travel time in the human cochlea [80].

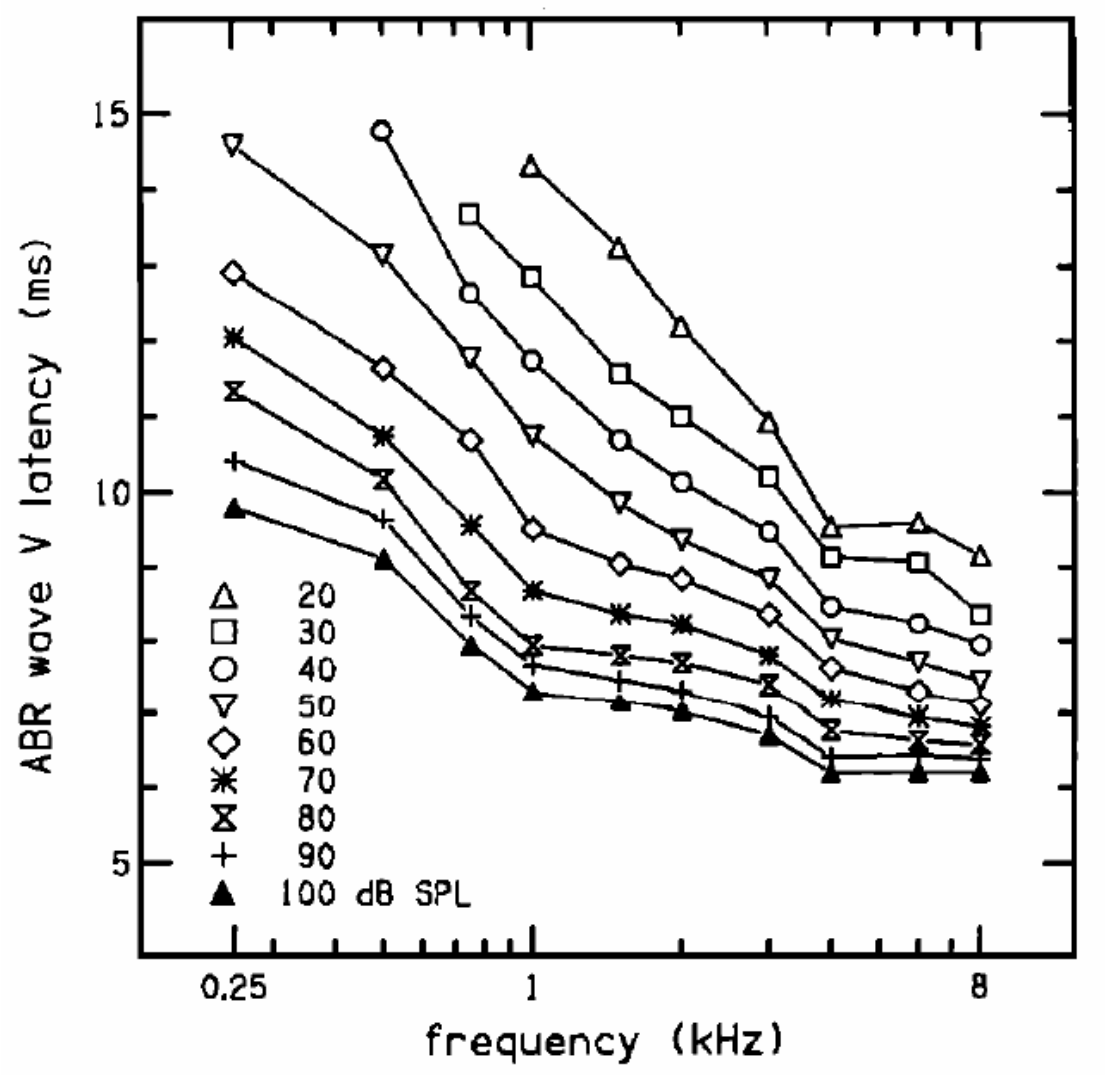


Figure 13. ABR V latencies at different frequencies. Neely et al. estimated basilar membrane delays in humans from the wave-V latencies of ABRs to tone bursts. The results, especially that of 20 dB SPL, were much longer than most previously reported values from aminal studies. Reprinted with permission from Fig. 1, Neely, S. T., S. J. Norton, et al. (1988). "Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli." J Acoust Soc Am 83(2): 652-656. Copyright 1988, Acoustic Society of America.

Shera et al. measured the group delays of human stimulus-frequency otoacoustic emissions (SFOAEs) to estimate the group delays of the basilar membrane. They also compared

SFOAEs in cats and guinea pigs to show that basilar membrane delays are much longer in humans than in common experimental animals (Figure 14). They suggested that this was because "the group delay of SFOAEs is equal to twice the group delay of the basilar-membrane mechanical transfer function, evaluated at the cochlear location with characteristic frequency equal to the stimulus frequency".

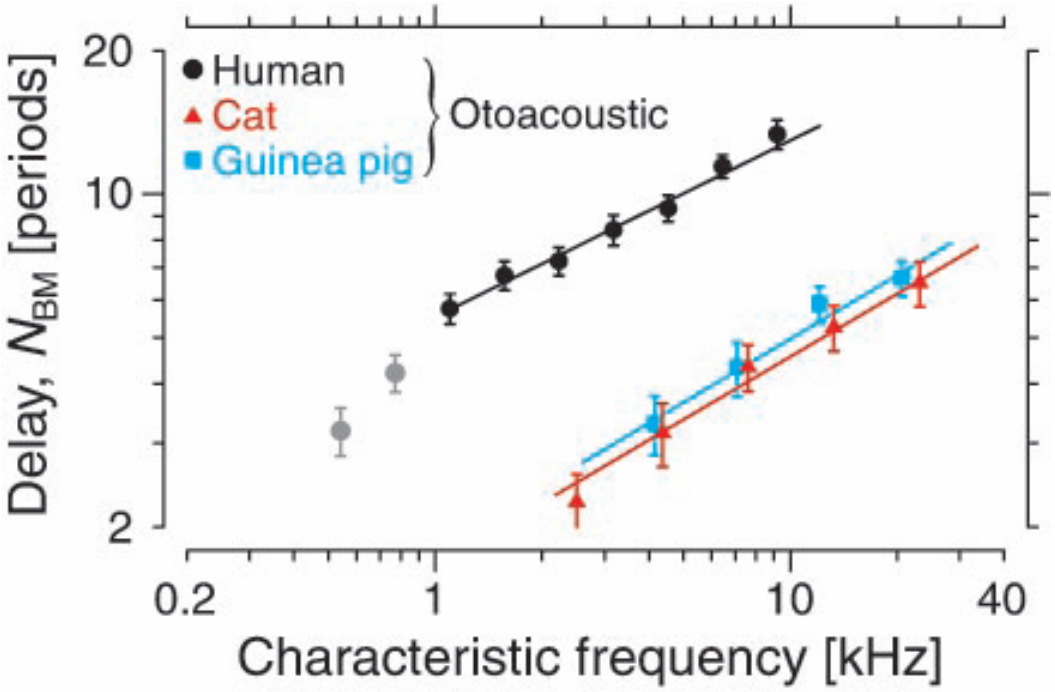


Figure 14. Delays of stimulus-frequency otoacoustic emissions in human, cat and guinea pig. Human SFOAE delays were longer than those in cat and guinea pig. Reprinted with permission from Fig. 2, Shera, C. A., J. J. Guinan, Jr., et al. (2002). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements." Proc Natl Acad Sci U S A 99(5): 3318-3323. Copyright 2002, the National Academy of Sciences of the United States of America.

Some studies pointed out that the over-estimated traveling-wave delays in humans from ABRs [80] actually resulted from flawed methodology and interpretation [22], and the delays from SFOAEs [81] were based on a proven wrong assumption [82]. Based on multiple animal

studies and human cadavers, a review study showed that the mean first-spike latencies in human auditory nerves should be similar to those of some other species [22] (See the section "Mean first spike latency to various tone bursts" below).

2.1.4.4 Effect of mean traveling-wave delays on speech perception

The cochlea works pretty much the same throughout life, with only minor shifts or changes in place-code mapping during early development. Therefore, the central auditory system is accustomed to this non-simultaneous pattern of traveling-wave delays. The frequency-dependent latency of the neural code thus is the norm. A reasonable hypothesis is that, if human brains get used to temporally non-simultaneous transmission at different frequencies provided by traveling waves, nullifying this temporal feature could somehow degrade speech recognition ability, at least under circumstances that deprive the brain from deducing word identity via redundancy and/or linguistic cues. This issue is not merely of basic science interest, but potentially of clinical value as prosthetic hearing devices have been developed with little/no regard to this factor, such as cochlear implants.

Huang et al. (unpublished) used the modified Rhyme Test [83] to investigate the effect of mean traveling-wave delays on perception of the original, transient extracted and quasi-steady-state extracted speeches in noise. They generated reverse-transformed speech words via a computer program according to the time-table of the reverse-latency function (the reverse-latency of a specific frequency was constant). The formula they used in the study for estimating human TW delays [84] was close to the estimates of Neely's [80]. The intensities of the reverse transformed and control stimuli were normalized in pairs. The target words were randomly selected to show on a monitor from six 50-word lists and monaurally delivered to 10 normal

listeners' right ears in a sound isolation booth. Subjects were instructed to press a button as soon as they heard the word they thought matches the target word.

Although there seemed to be a similar trend that recognition scores in all their 3 reverse traveling-wave transformed groups were slightly worse than those in the paired control groups, their results showed that reversing the traveling-wave delays did not significantly diminish speech perception in noise (Figure 15). Their results suggested that it may not significantly affect speech perception for cochlear implant users whether the mean traveling-wave delays are implemented or not.

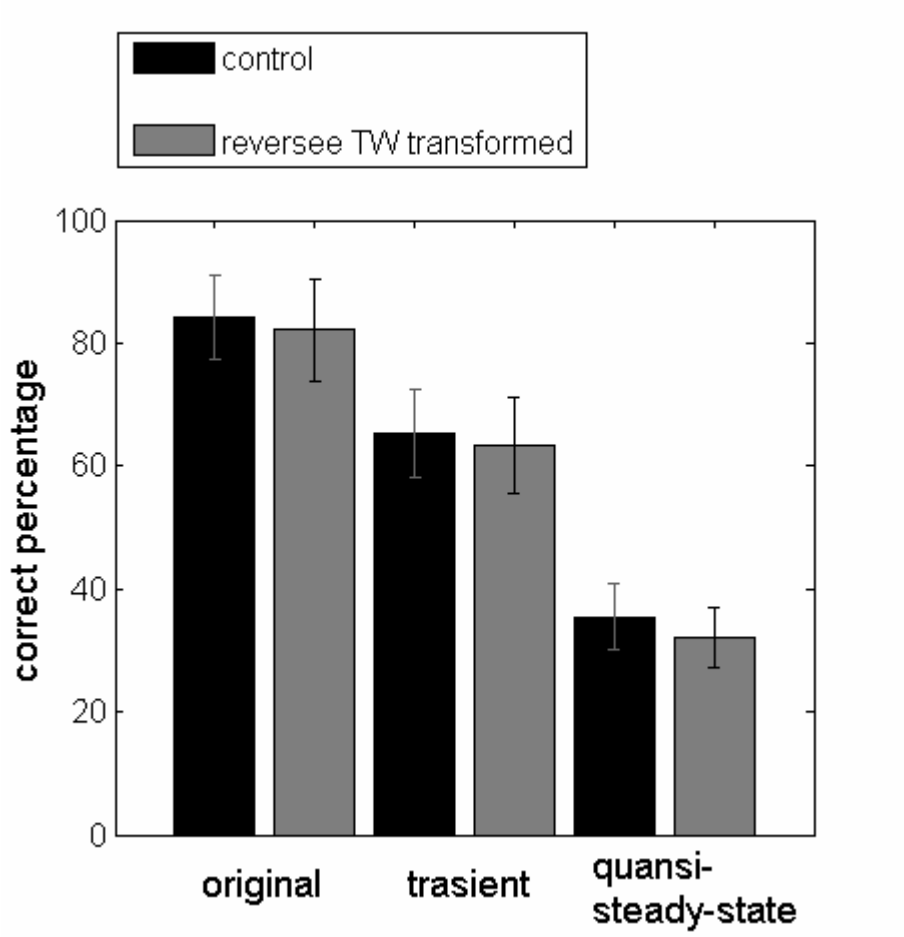


Figure 15. Effect of constant traveling-wave delays on speech perception in noise for original, transient-extracted and quasi-steady state-extracted speech. Bar graphs indicate mean score and one

standard error. The recognition scores of modified Rhyme Test in all of the 3 reverse-traveling-wave transformed groups are not significantly different from the paired control groups. The reverse transform was based on the group-delay formula obtained from auditory brainstem responses by Elberling et al., 2007.

In contrast, Taft et al. reported a significant group improvement ($p = 0.027$) when they investigated sentence perception in quiet with 8 cochlear implantees using advanced combination encoder (ACE) sound-processing strategy with additional TW delays [33]. However, they reported non-significant effects on other tests, for example sentences in noise. They suggested that their results are encouraging for cochlear implant sound processing because across-frequency delays can be incorporated easily and efficiently into available sound-processing strategies.

Both of these pioneering studies were investigating the effect of mean traveling-wave delays. In other words, the delay at a specific frequency was constant (no matter what frequency and phase information the original signal contains), which was not the result measured in previous animal studies. In addition, there is not a generally accepted gold standard for estimating human traveling-wave delays. The delays used in both of the studies [70, 84] seemed to be much longer than those measured from mean first-spike latencies in animal auditory nerve fibers [22].

Thus, future studies are required to investigate the effect of auditory neural first-spike latency (which changes when the stimulating tone burst's frequency or phase changes) on speech perceptions. Then one can start to evaluate its incorporation in cochlear implants.

2.1.5 Summary and discussion

Temporal phase is defined as a frequency-dependent time delay, and is one of the three fundamental kinds of information (intensity, frequency and phase) in a sinusoid. A sinusoid is the basic unit of a sound. These three kinds of information enable one to reproduce the original sinusoid components. Putting all the sinusoid components together regenerates the original sound signal.

Phase information can be measured by calculating from a sinusoid's period or frequency, by Hilbert transform (instantaneous phase), or by Fourier transform (Fourier phase).

In Fourier analysis, one way to encode phase information (again) is to "mark" the first amplitude maximum for each sinusoid. When marking the first amplitude maximum for each sinusoid component, the "markings" create a pattern among the sinusoids and is defined as a Fourier pattern, which is a phase code of Fourier analysis. Phase information is involved in many auditory perceptions such as pitch perception, speech perception (especially in noise), and lexical-tone perception.

The role of the outer ear for phase information in monaural hearing is a continuous reflecting and delaying system. The time delay across all frequencies only changes when the stiffness and density of the air change. Phase responses of the middle ears vary in different animal studies. The variety may be due to structural (such as size and mass) differences among different species and individual animals.

The inner ear is believed to be the main place of phase processing at auditory periphery. Approaches to estimate human cochlear traveling delays include psychoacoustic methods, stimulus-frequency otoacoustic emissions, wave V latencies of auditory brainstem responses at different frequencies, and derived response (or stacked-band) auditory brainstem responses.

Phase processing of the cochlear mechanics is still a mystery. Even to a simple stimulus like a tone burst, the interaction among the 4 main elements in cochlear mechanics (fluid pressure, fluid motion, partition forces and partition motion) is unclear.

Considering different stimuli, the effort to elucidate the interaction among the 4 main elements in cochlear mechanics to tone bursts may only explain the processing of tone burst stimuli. The observation or interpretation may not apply to another stimulus. For example, there have been studies suggesting that auditory-nerve-fiber responses to clicks are due to the combination of multiple drives in the cochlea [85, 86], and some motion other than the traveling wave must produce the first-spike responses of auditory nerve fibers to clicks [87].

Considering various species, an established model of mammalian cochlear mechanics may only be good for mammalians but not another species. For example, a basilar-membrane based cochlear model that successfully predicts the neural responses to a click in mammals can not explain the similar responses in tree frogs which lacks the cochlear partition motion.

Based on the frequency response characteristics of the middle ear and the frequency-place mapping of the cochlea [21], remarkable differences are evident among samples of animals across phyla. Similarly, the latency pattern of primary neural responses to an identical sound is expected to differ correspondingly. However, previous animal studies show recordings stimulated by various tone bursts from different individual animals and even different species yield a similar averaged delay at a characteristic frequency and form a similar latency pattern [22]. This motivates us to further look into the pattern of neural spikes, especially how these spikes encode phase information.

2.2 SIMILARITY IN NEURAL LATENCY PATTERN

2.2.1 Mean first spike latency to various tone bursts

In order to measure the latency pattern in auditory nerve fibers, a typical experimental setting was a miniature earphone inserted into the animal's ear canal (very close to the tympanic membrane) and a glass microelectrode placed in the targeted auditory neuron (e.g., [25]). The mean first-spike latency of an auditory nerve fiber is the mean latency computed from multiple recordings to various tone bursts around the neuron's characteristic frequency. Some studies termed this as "mean response delays" or "group delays".

In previous studies, the mean first-spike latency of an auditory unit to tone bursts is calculated from the slope of the unit's phase-versus-frequency function (e.g., [23-26]), which is derived from the most stable mean response phase [25] (or the mean phase angle [24]) in period histogram to various tone bursts around the neuron's characteristic frequency (also see 2.2.3.2). Usually, period histograms are generated through the use of the positive zero crossing of the stimulus waveform to trigger a fast data acquisition unit, which marks the occurrence of each action potential with reference to a single cycle of the stimulus frequency (e.g., [24, 25]). Assuming the phase-versus-frequency function is a linear function, researchers usually use the slope of the regression line for calculation (e.g., [24, 26]). Because the latency changes while the starting (or reference) phase of the stimulating tone burst varies, phase calibration (and reference) is important for recording the phase-versus-frequency function.

In addition to frequency or phase change of the stimuli, the mean first-spike latency stimulated by a tone burst may be changed by any structure within the pathway, such as outer ear and middle ear, cochlear traveling wave, neural transmission, and synaptic transmission.

Because of the above-mentioned structural differences such as the content and sensitivity of the middle ear and the length and frequency-place mapping of the cochlea [21], the mean first-spike latency at a characteristic frequency is expected to be different in different species, or even to be different between two individual animals of the same species. Mean first-spike latencies have been recorded in many different animals, such as chinchillas [88], cats [89], alligators [90], barn owls [25], monkeys [23], lizards (Tiliqua rugosa) [26] (whose basilar membrane does not support traveling waves [28]), and frogs [91](whose hearing organs lack a basilar membrane altogether [24]). Despite the structural differences among these species, which reasonably would be expected to cause variety in first-spike latencies, Ruggero and Temchin published findings of a comparative analysis of results from previous animal studies demonstrating mean first-spike latencies across species to be compellingly similar, as was presented in Introduction [22] (Figure 4). In these data obtained (again) from primary neurons excited by tone bursts, the latencies (delays) across species showed comparable dependencies on characteristic frequency.

Also, they studied the effect of death on mean first-spike latencies. They found the ratio between in-vivo characteristic frequency and postmortem best frequency was close to being a constant (about 1.5). Assuming the effect of death on a human produces first-spike latencies similar to that in animals, they concluded that the estimated mean first-spike latencies in humans are similar to those in the auditory nerve fibers of other animals across species.

In summary, based on the frequency response characteristics of the middle ear and the frequency-place mapping of the cochlea [21], remarkable differences are evident among samples of animals across phyla (Figure 2). Similarly, the latency pattern of primary neural responses to an identical sound is expected to differ correspondingly (Figure 3). However, previous animal

studies show recordings from different individual animals and even different species yield a similar averaged delay at a characteristic frequency and form a similar latency pattern [22], including recordings from the bobtail lizard (Tiliqua), whose basilar membrane does not support a traveling wave [28], and tree frogs, whose hearing organs lack a basilar membrane altogether [24].

This similarity suggests that the hearing organs across different species probably operate a common mechanism to decompose sound into neural spikes whose mean first-spike latencies are similar. Since the similarity holds in animals that do not support traveling waves or do not even have a basilar membrane, this common mechanism should comprise more than the mechanics of cochlear traveling waves. There seems to be no species- and individual-independent sharing of common attributes that explain this similarity, except for the assertion that the hearing organ acts as a Fourier analyzer [29].

As defined in Chapter one, temporal phase is a frequency-dependent time delay. The mean first-spike latency at a characteristic frequency is a frequency-dependent time delay. It is thus reasonable to view this similarity as a common phase pattern.

There have been numerous studies examining the mean first-spike latency stimulated by ***various*** tone bursts of different frequencies, as some of them are cited in this section. However, reports about the first-spike latencies among different auditory neurons stimulated by a ***single*** (same-frequency and same-phase) tone burst are not so numerous in the literature.

2.2.2 First spike latency to a same-frequency tone burst

The first-spike latencies among different auditory neurons stimulated by a single tone burst are not numerous in the literature. And the descriptions and conclusions made in available studies

are vague. Here the available studies will be reviewed first, followed by discussion of possible theoretical basics. The original description and conclusion will be quoted. My personal interpretations will be discussed.

The first study providing direct evidence of the first-spike latencies among auditory neural fibers by a single speech signal was Shamma's [92] in 1985. Shamma examined the latency pattern stimulated by a single stimulus /da/ in large populations of auditory-nerve fibers in the cat. The first-spike latencies to the /da/ created a curvy spatiotemporal curve.

Looking at the spatiotemporal representation, Shamma noticed that the fiber section can be roughly subdivided into two regions:

"(1) a region basal to the point of resonance of the harmonic where the fiber peristimulus-time (PST) histograms accumulate only small delays (or phase shifts) relative to each other reflecting the fast speed of propagation of the traveling wave, and (2) a region at or very near the point of resonance where the responses exhibit drastic relative phase shifts owing to the sudden slow down of the traveling wave and the consequent rapid accumulation of phase shifts." [92] (Figure 16)

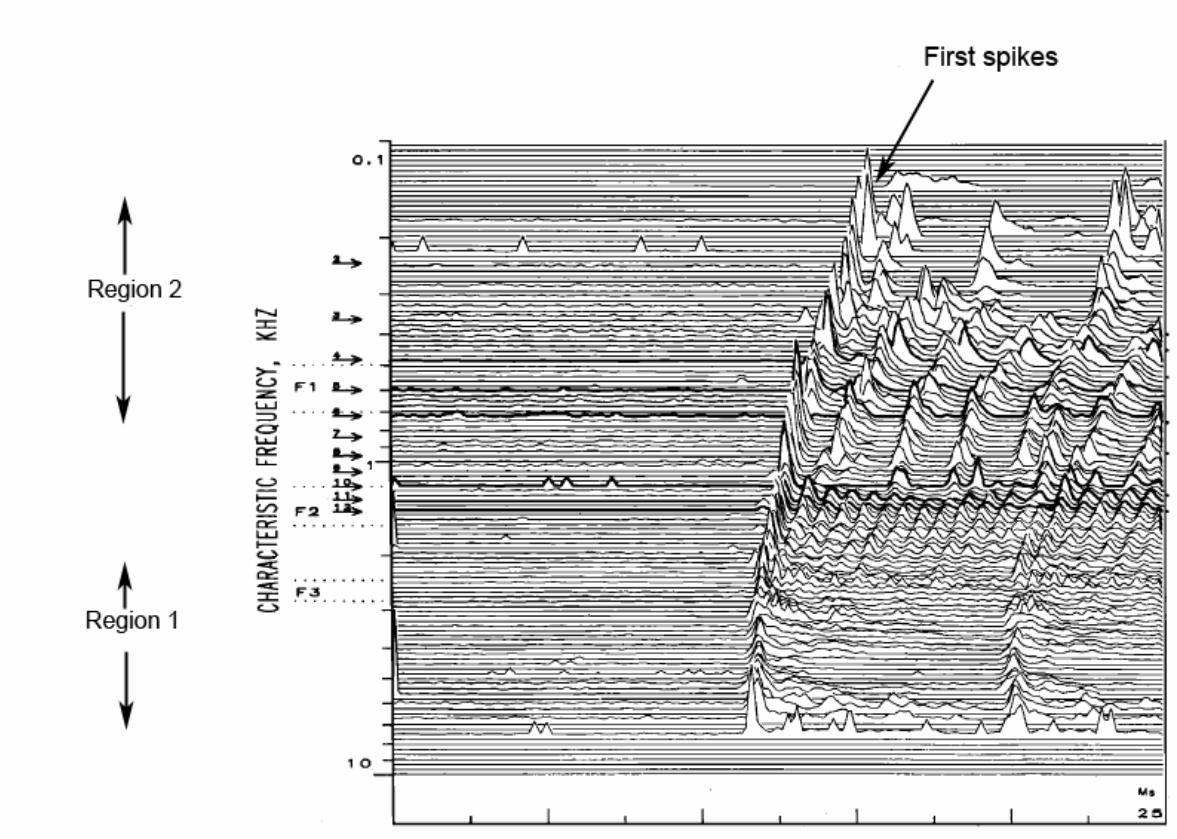


Figure 16. Spatiotemporal responses of the auditory nerve evoked by the stimulus /da/. Personal interpretation at left indicates possible region 1 and 2 in the author's conclusion: "(1) a region basal to the point of resonance of the harmonic where the fiber peristimulus time (PST) histograms accumulate only small delays (or phase shifts) relative to each other reflecting the fast speed of propagation of the traveling wave, and (2) a region at or very near the point of resonance where the responses exhibit drastic relative phase shifts owing to the sudden slow down of the traveling wave and the consequent rapid accumulation of phase shifts." Reprinted with permission from Figure 5B, Shamma, S. A. (1985). "Speech processing in the auditory system. I: The representation of speech sounds in the responses of the auditory nerve." J Acoust Soc Am 78(5): 1612-1621. Copyright 1985, Acoustic Society of America.

Variety of the first-spike latencies to a single tone burst were demonstrated in Loeb's review article [7]. Loeb assumed the first-spike latency was caused by cochlear traveling waves. He described that:

"at the characteristic place (Figure 17A, highlighted critical band), a tonal stimulus would result in a convergence of high levels of neural activity (to be decoded as place pitch) and maximal phase acceleration (to be decoded as phase-locked pitch)… A much lower pitched tonal stimulus (Figure 17D) might produce a much flatter phase gradient…"

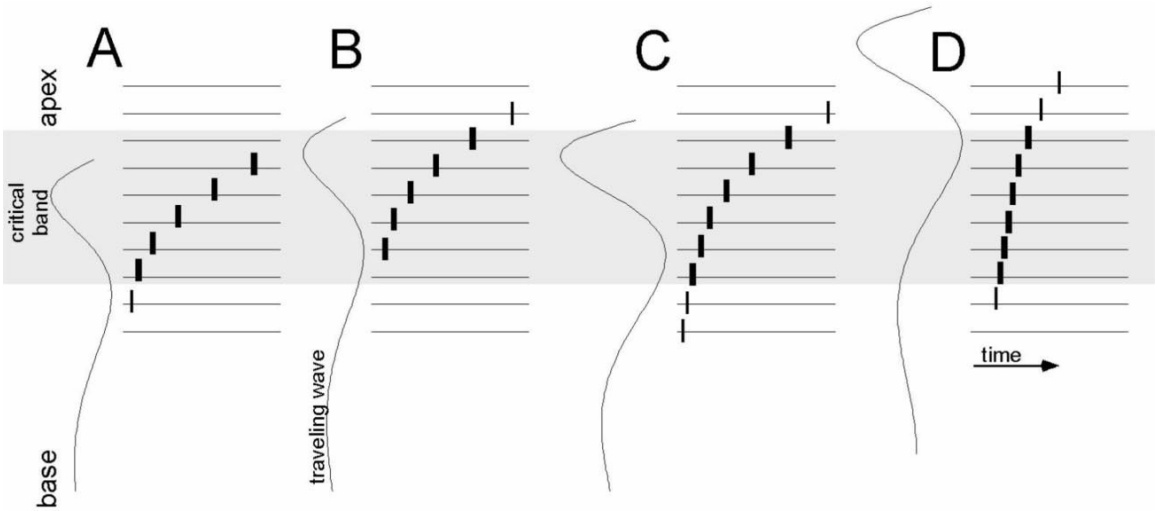


Figure 17. Schematic representation of first-spike latencies to a single tone burst. When a single tone is presented at sufficient intensity, the trains of action potentials across auditory nerve fibers will exhibit a coherent spatial gradient. At the same location, a tone around the characteristic frequency generates a curvy and rapidly changing spatial gradient (A, B), but a tone with much lower frequency shows a much shallower and more constantly-changing phase gradient (D). Stimulus in C has the same frequency as in B but higher intensity. Reprinted with permission from Fig. 1, Loeb, G. E. (2005). "Are cochlear implant patients suffering from perceptual dissonance?" Ear Hear 26(5): 435-450. Copyright 2005, Wolters Kluwer Health.

This concept or even the description itself is not clear. The bottom line is that these "gradients" are frequency-dependent time patterns. Since the hearing organ acts as a Fourier analyzer [29] and the Fourier pattern (see 0) is frequency-dependent, it is investigated here by

looking into the Fourier pattern created by a single tone burst to seek possible cues that imitate the first-spike latencies to a single tone burst as those in Figure 17.

Applying Loeb's concept by using three tone bursts: a selected tone burst, the 2nd tone burst with a slightly lower frequency, and the 3rd tone burst with a much lower frequency, Figure 18 shows the Fourier patterns created by a 505 Hz, a slightly lower 495 Hz, and a much lower 125 Hz long tone bursts. All signals are sampled at 20000 Hz (chosen arbitrarily for demonstration) and the length of fast Fourier transform is 400 points (chosen arbitrarily for demonstration) to show a frequency resolution of 50 Hz.

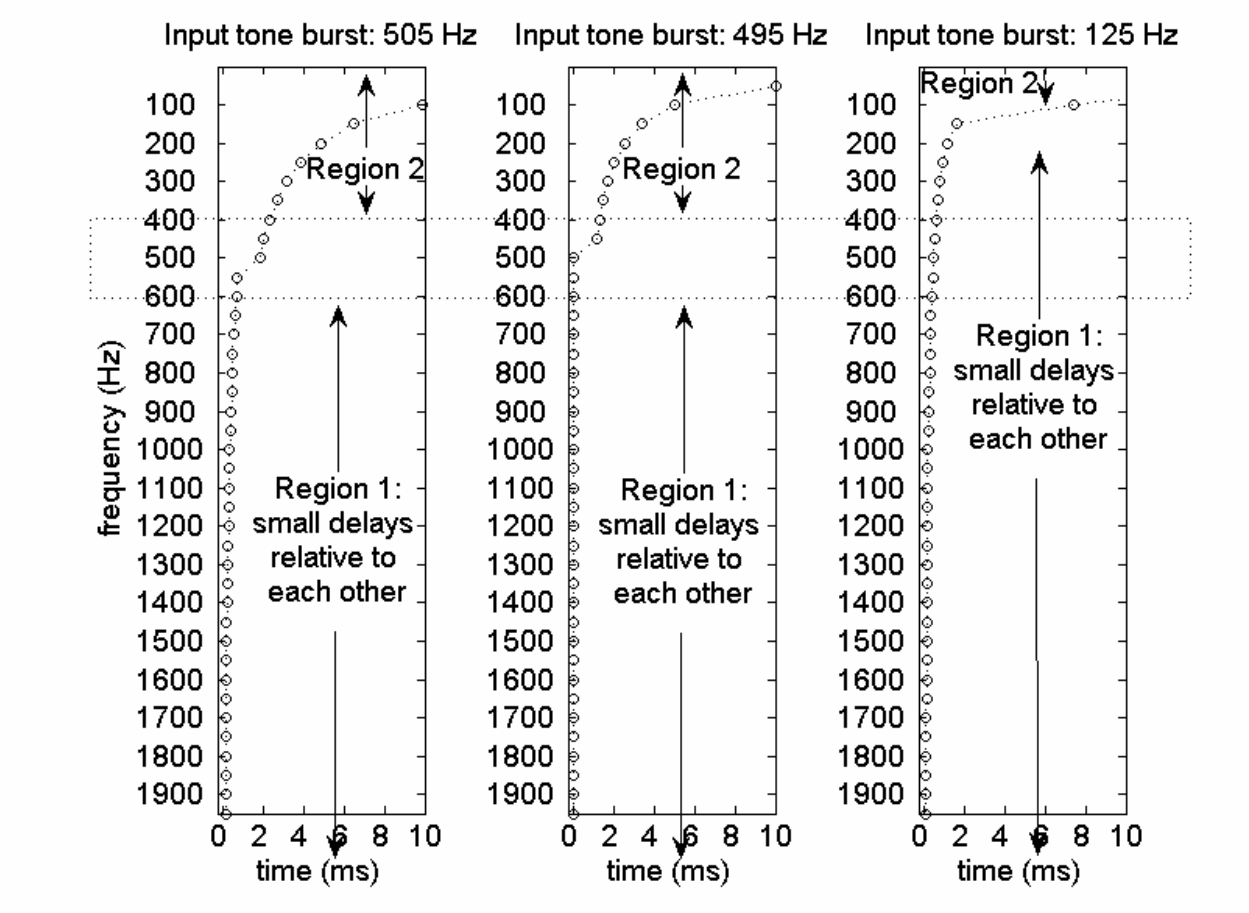


Figure 18. The similar features in the Fourier pattern created by a single long tone burst for Loeb's and Shamma's concepts. Regarding the concept of Loeb, 2005: when a selected place is examined (the dash

rectangle, the characteristic place around 500 Hz) and a tone is produced around the characteristic frequency (505 or 495 Hz), the generated Fourier pattern will be curvy and rapidly changing. In contrast, a tone with much lower frequency (125 Hz) generates a much shallower and more constantly-changing Fourier pattern (right panel, within the dash rectangle). Regarding the concept of Shamma, 1985: The frequency of the tone burst separates the created Fourier pattern into two regions -- in the one (region 1) containing high frequencies, the Fourier pattern accumulates only small delays (or phase shifts) relative to each other; in the other (region 2) which contains low frequencies, the Fourier pattern exhibits large relative time delays.

Compared to Loeb's concept and looking at the dashed rectangle in Figure 18 (the characteristic place around 500 Hz), a tone around the characteristic frequency (505 or 495 Hz) generates a curvy and rapidly changing Fourier pattern (Figure 18, left two panels), but a tone with much lower frequency (125 Hz) shows a much shallower and more constantly-changing Fourier pattern (Figure 18, right panel).

Compared to Shamma's description and looking at the two regions separated by the stimulating frequency in Figure 18 (although Shamma's stimulus was a complex sound /da/), in the one (region 1) containing high frequencies, the Fourier pattern accumulates only small delays (or phase shifts) relative to each other; in the other (region 2) which contains low frequencies, the Fourier pattern exhibits large relative time delays.

The similar features in the Fourier pattern created by a single tone burst suggest again the hearing organ may operate a common mechanism, which also generates Fourier-pattern-like response delays. Future studies are necessary to demonstrate direct recordings from animal auditory neurons stimulated by a single tone burst (or repeatedly stimulated by a single tone burst).

The stimulus-frequency dependent pattern change in Loeb's scheme and in the Fourier pattern also suggests that there may be a pattern-only pitch perception. That is, since stimulus

frequency determines the neural pattern, different patterns may give different pitch perceptions. Future studies are required to first validate the existence of the Loeb described feature in animal auditory neural responses, and then to test whether there is a pattern-only pitch perception.

In summary, Shamma and Loeb demonstrated that the first-spike latencies to a given tone burst may form a specific pattern that represents the tone burst's frequency and phase information. The Fourier pattern created by a given tone burst shows the similar features demonstrated by Shamma and Loeb. There are very few reports in the literature showing direct evidence of this stimulus-specific first-spike latency pattern and its impact. Future studies are required to first validate the existence of the feature described by Loeb in animal auditory neural responses, and then to test its effect on auditory perceptions.

From another point of view, after each first spike in single auditory neuron, the neural spikes evoked by a tone burst that follow are grouped in a specific way. This is the temporal pattern to the tone burst for the auditory neuron. The pattern is determined by frequency of the tone burst and the characteristic frequency of the neuron.

2.2.3 Latency pattern in a single unit to tone bursts

Since temporal phase is a frequency-dependent time delay, the phase information represented in a single neural fiber at least should reflect frequency or phase information of the stimulus. Responding to a tone burst, a single auditory neuron tends to discharge at phase-locked timings to the tone burst. The interval of two adjacent spikes in the single nerve tends to be the period of

the stimulus (or integral multiples). In other words, this encoding at least reflects the stimulus frequency.

Researchers in numerous studies (e.g., [93-96]) assumed that the neural latency pattern to a tone burst is accounted for by a mechanism of half-wave rectification during the mechano-electrical transduction process in the cochlea. In this hypothesis, the neural latency pattern is synchronized to one side (positive or negative) of the tone's waveform.

The phase-lock phenomenon has been reported in auditory nerve fibers with characteristic frequencies up to 10 kHz in barn owls [25]. Also, the phase-lock phenomenon has been reported in animal studies across different species, including squirrel monkeys[23, 97-99], barn owls[25], bobtail lizards[26], chinchillas[88], caimans[90], frogs[91], and cats [100, 101].

The neural activity to a tone burst has been frequently reported by several typical ways to demonstrate the phase-specific timing, such as interspike interval (ISI) histogram, period histogram, peristimulus-time (PST) histogram, or by calculating the synchronization index [102-106].

2.2.3.1 Interspike interval histogram

Plotting number of intervals as a function of time, Figure 19 demonstrates a typical interspike interval histogram to a tone burst. Each histogram was composed by several modals (grouped spikes at a specific phase to the tone burst). Each modal distribution is approximately normal. Each modal population tends to be larger than the one which has a larger interspike interval. The peak of each partial distribution is usually quite close to an integral multiple of the period of the stimulating tone burst.

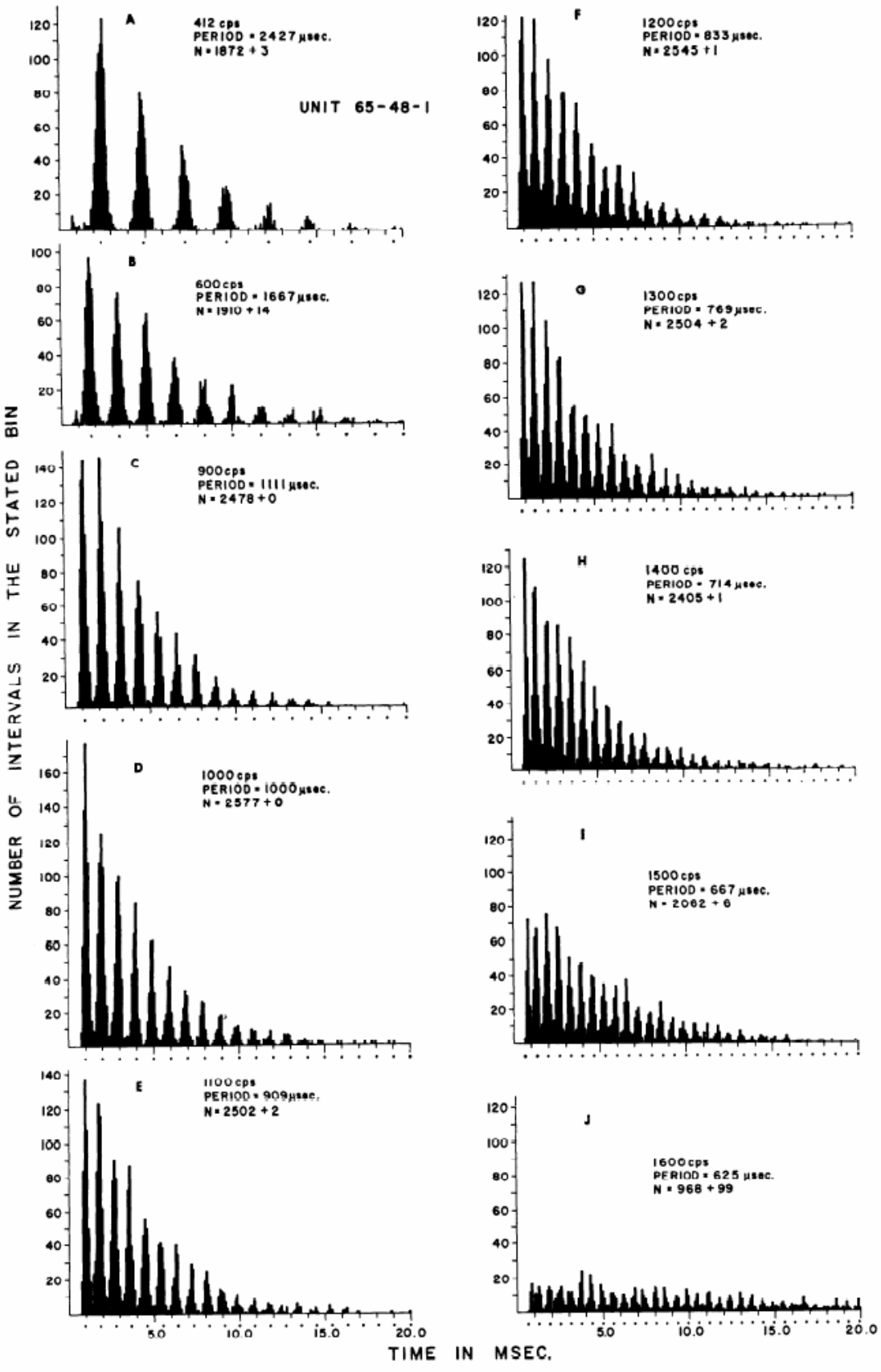

UNIT 65-48-1
A
412 cps
PERIOD = 2427 µsec.
N = 1872 + 3
B
600 cps
PERIOD = 1667 µsec.
N = 1910 + 14
C
900 cps
PERIOD = 1111 µsec.
N = 2478 + 0
D
1000 cps
PERIOD = 1000 µsec.
N = 2577 + 0
E
1100 cps
PERIOD = 909 µsec.
N = 2502 + 2
F
1200 cps
PERIOD = 833 µsec.
N = 2545 + 1
G
1300 cps
PERIOD = 769 µsec.
N = 2504 + 2
H
1400 cps
PERIOD = 714 µsec.
N = 2405 + 1
I
1500 cps
PERIOD = 667 µsec.
N = 2062 + 6
J
1600 cps
PERIOD = 625 µsec.
N = 968 + 99
NUMBER OF INTERVALS IN THE STATED BIN
TIME IN MSEC.
5.0
10.0
15.0
20.0

Figure 19. Interspike interval histogram to stimuli of different frequencies. An auditory nerve fiber of the squirrel monkey with characteristic frequency of about 1100 Hz tended to discharge at a rate of the stimulus frequency. The inter-peak-intervals were multiples of the period of the stimulus frequency. Stimuli for each histogram were 10 one-second tone bursts of various frequencies as shown in each panel. Reprinted with permission from Fig.1, Rose, J. E., J. F. Brugge, et al. (1967). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey." J Neurophysiol 30(4): 769-793. Copyright 1967, The American Physiological Society.

Rose et al. [98] reported the period was that of the stimulating tone burst. For all stimuli which lie within the response area, they found it was the stimulus frequency and not the best frequency of the nerve fiber that determined how the time-code information was transmitted.

Also, using six 10-second 100 dB tone bursts, they reported three interesting findings of a neuron (whose best frequency was 400 Hz) stimulated by low-frequency tone bursts [98] (Figure 20): 1. The dispersion around the mean becomes larger as the period lengthens (i.e., decreasing stimulating frequency). 2. There may appear another highly skewed distribution with a peak at about 1 ms. As shown in Figure 20, the interspike interval of the largest number of intervals by stimuli of 120, 100 and 80 Hz was 1 ms. In other words, when stimulated with these low-frequency tone bursts, there were many adjacent spikes whose interval was about 1 ms. 3. Even when the period was as long as 12.5 ms, as shown in F panel, the nerve still failed to respond to every cycle since there was a small but distinct second peak at intervals around 25 ms with a suggestion of even a third cluster at 37.5 ms.

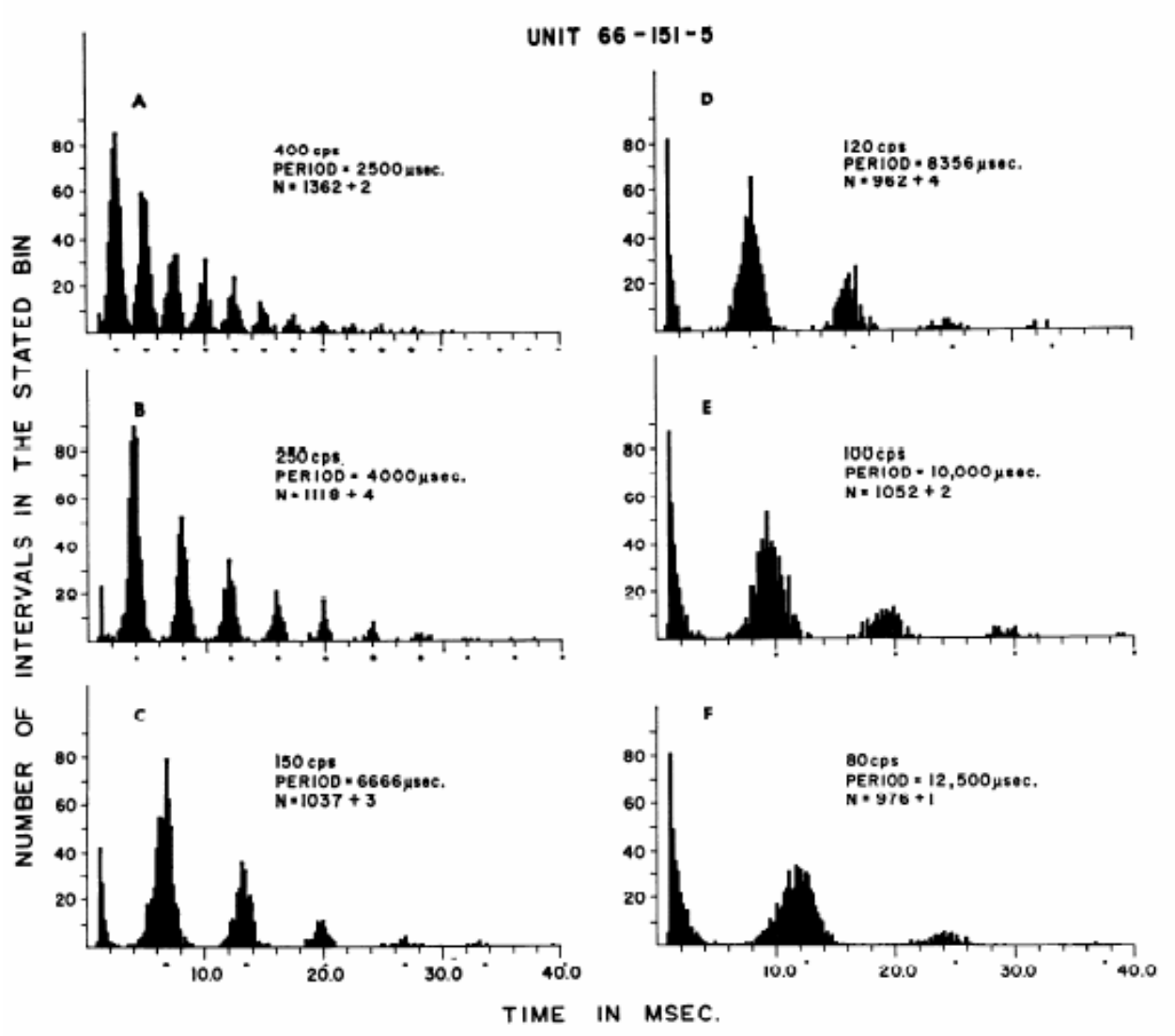


Figure 20. Periodic distributions of interspike intervals in a nerve of best frequency of 400 Hz. Stimuli were six 10-second 100 dB tone bursts. The dispersion around the mean becomes larger as the period lengthens. There may appear another highly skewed distribution with a peak at about 1 msec. Even when the period is as long as 12.5 msec, as shown in F panel, the nerve still fails to respond to every cycle. Reprinted with permission from Fig. 4, Rose, J. E., J. F. Brugge, et al. (1967). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey." J Neurophysiol 30(4): 769-793. Copyright 1967, The American Physiological Society.

The shortest interspike interval to low-frequency tone bursts in Figure 20 is believed to be the refractory period of the auditory neuron. That is, responding to low-frequency tone bursts, an auditory neuron may discharge spike immediately after recovering from the refractory period, within the tone's one cycle.

Interspike interval histograms have been continuously used for analyzing auditory neural discharges in various animal studies up to today [107-110]. From the interspike intervals generated by a tone burst, the following conclusions are generally accepted: 1. The discharges are spaced at intervals grouped around the integral multiples of the period of the stimulating tone. It holds for all frequencies activating the neuron regardless of the best frequency of the unit. 2. The discharges distribute symmetrically in each group. 3. The dispersion around the mean becomes larger as the period lengthens, i.e., frequency decreases. 4. An auditory neuron tends to fail to respond to every cycle despite of stimulus frequency and intensity.

Since an auditory neuron tends to discharge periodically, another way to show the periodicity is to plot all the discharge within one period to show the level of synchronization. This is a period histogram.

2.2.3.2 Period histogram

A period histogram plots the number of spikes as a function of phase (0 to 2π) or cycle (0 to 1). With a selected reference, the mean response phase in a period histogram indicates the phase difference to the reference.

Period histograms have been used to investigate temporal processing in neural representations [23, 25, 104, 109, 111-113] such as phase response of an auditory neuron, and strength of phase-locking. In one typical phase related application, Temchin and Ruggero [113]

showed the period histograms (Figure 21) of a chinchilla auditory neuron (characteristic frequency = 3281 Hz) responding to 80-dB SPL tone burst of 12 different frequencies (100 to 4100 Hz, shown at left upper corner of each panel). Compared to the reference (100 Hz stimulus panel), it was possible to calculate the relative delay (phase difference) to the reference.

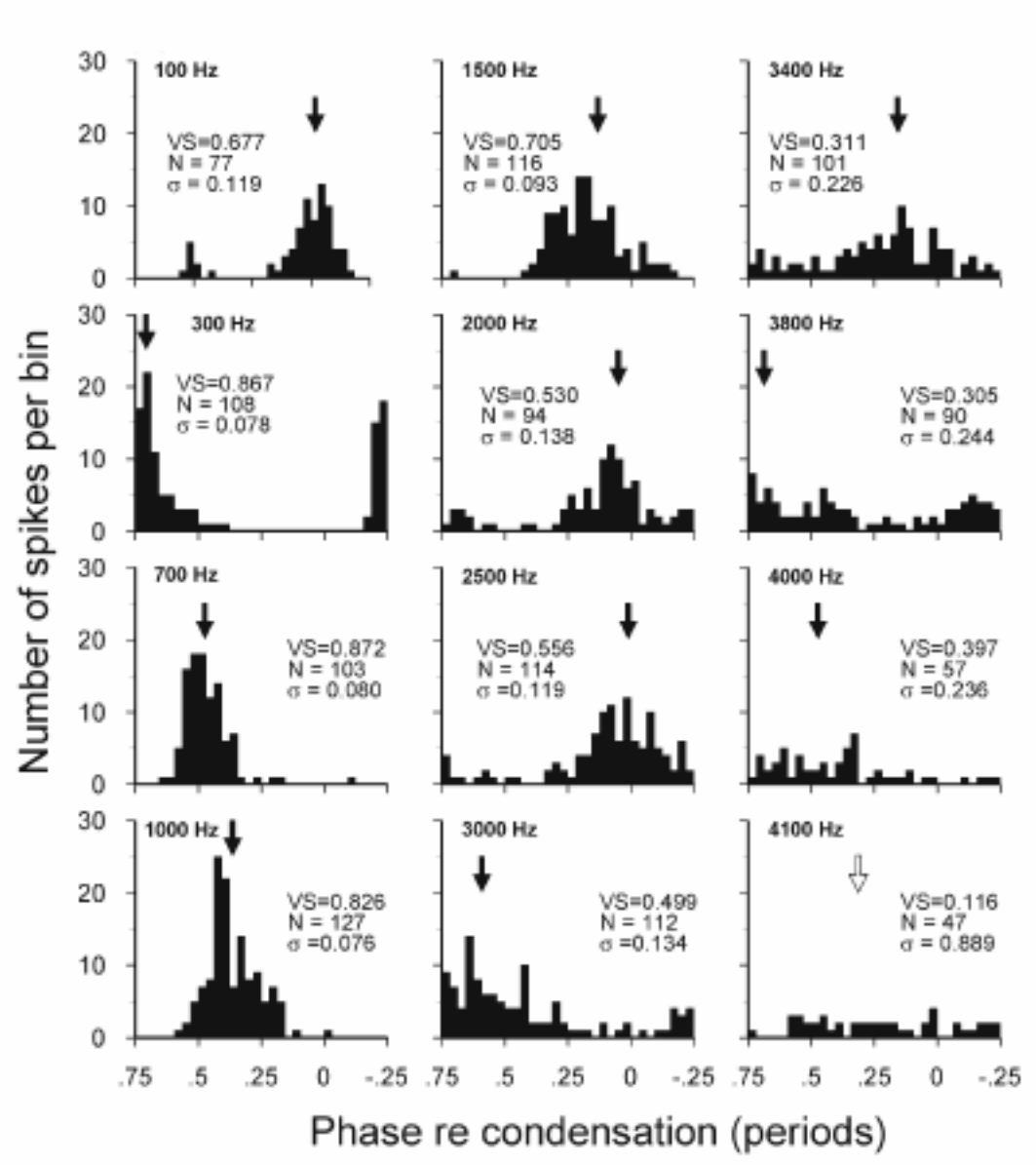


Figure 21. Period histograms of a chinchilla auditory neuron responding to 80-dB SPL tone burst of 12 different frequencies. The characteristic frequency was 3281 Hz. Vertical arrows indicate mean response

phase. Vector strength (VS), number of spikes (N), and standard deviation of the phase (σ) are indicated. With kind permission from Springer Science+Business Media: Fig 1A, Temchin, A. N. and M. A. Ruggero (2010). "Phase-locked responses to tones of chinchilla auditory nerve fibers: implications for apical cochlear mechanics." J Assoc Res Otolaryngol 11(2): 297-318.

Period histogram is a good way to show the level of synchronization. However, when using period histograms to analyze not only the level of synchronization (e.g., time delays), one common issue for these animal studies of auditory neural responses was that the starting phase of the stimulating tone burst (or what instantaneous phase was used as the reference among windows) was not controlled (or described in the articles how they calibrated). When the starting phase or the reference is not constant, the relative phase difference to two different-frequency tone bursts might be interpreted incorrectly. For example, it could be the phase difference between a 500 Hz tone burst with a mean starting phase of X, and a 100 Hz tone burst with a mean starting phase of Y, but interpreted as the constant phase difference between 500 and 100 Hz tones.

Otherwise, period histograms of previous animal studies showed auditory neurons with low characteristic frequencies showed steady phase-locking. This synchronization was even evidenced in some animals when stimulus frequency was as high as 10k Hz [25].

2.2.3.3 Peristimulus-time histogram

Plotting discharge rate (spikes/second) or number of spikes as a function of time, a peristimulus-time histogram shows the raw data that an auditory neuron responds to the stimulus. A peristimulus-time histogram (some studies termed poststimulus-time histogram) illustrates the time relationship between spikes and the onset of the stimulus. It is usually used to show the

spike structures over time [94, 104, 109, 111, 114-116], such as synchronized modals (spike groups) and the overall process (onset, adaptation, offset and recovery).

Findings in peristimulus-time histograms to tone bursts did not change significantly in the past several decades. For low-frequency tone bursts, the discharges are grouped at the period timing of the stimulating tone. The discharge-distribution width in each group is similar, if not the same. The dispersion around the mean becomes larger as the period lengthens, i.e., stimulus frequency decreases.

In summary, starting with the specific first-spike latency to the tone burst, discharges of an auditory neuron represent spike groups with the following characteristics: 1. Spike modals are synchronized at the period timing of the stimulating tone. 2. The duration of a spike group is similar. 3. The dispersion of a group becomes larger as the period lengthens (or when the difference between stimulus frequency and the neural characteristic frequency is bigger). 4. An auditory neuron fails to respond to every cycle despite stimulus frequency and intensity.

In other words, phase information of a stimulus tone seems to be encoded by a unique latency pattern among all auditory neurons. In one auditory neuron, the main variables include the specific first-spike latency, the width or duration of each modal, and the period of the repeating modals. The stimulus frequency and the neural characteristic frequency together affect all the main variables above. The underlying mechanism of how they interact is not fully understood.

In addition to the stimulus frequency and the neural characteristic frequency, the starting phase of the tone burst should affect the first-spike latency and thereby determine the following

spike modals. However, it seemed not to be well controlled or described how the starting phase was calibrated in most of current available studies.

The similarity of latency pattern has not only been found to tone bursts but also to a click. The next two sections will review the similar latency patterns of the first spikes and the classical responses to rarefaction and condensation clicks across numerous species.

2.2.4 Mean first-spike latency to a click

The first-spike latency (or sometimes termed as onset latency or signal-front delays) to a click is usually obtained from the peristimulus-time histogram of responses to a intense repeatedly presented click (e.g., [88]). The latency is the direct measurement of the time difference between the click onset and the action potential detected subtracted by the acoustic transmission to the ear drum. In other words, the latency is the delay between the time when the stimulus starts to vibrate the eardrum and that the action potential discharged by the corresponding eardrum vibration.

In addition to the latency pattern to tone bursts (first spikes and phase locking), comparative studies also have been carried out previously on the first-spike latency to a rarefaction click and thus also were summarized by Ruggero and Temchin [22] (Figure 22). At first glance, the similarities across species seem less striking, comparing Figure 22 (below) and Figure 4 (see Introduction), than in the case of tone-burst stimulation. However, this is simply a distraction from presumably some overall inefficiency of the tree frog's peripheral auditory system and the choice of regression model among other species. Nevertheless, the overall trends of the functions tell similar stories of the dependencies of first-spike latencies on characteristic frequency, despite remarkable structural differences among their middle and inner ears. The

animals studied include barn owls, whose frequency representation of 5–10 kHz occupies more than half of the papillar length and so the latency is expected to show an unusually fast increase from 10 to 5 kHz [25] and thus differ from other species, and tree frogs (Eleutherodactylus coqui), whose hearing organs lack a basilar membrane altogether [24].

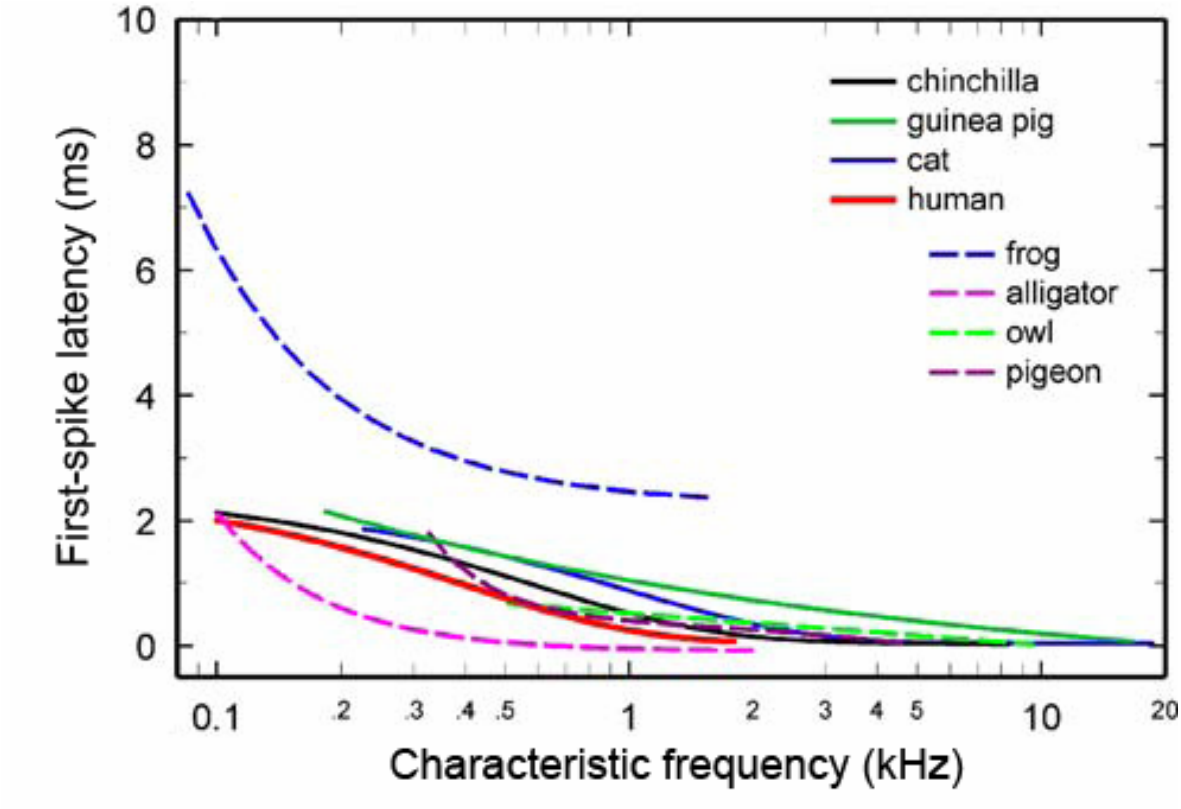


Figure 22. Across-species comparison of first-spike latencies in auditory nerve fibers to a rarefaction click. The 1-ms neural and synaptic transmission was subtracted. With kind permission from Springer Science+Business Media: Figure 8A, Ruggero, M. A. and A. N. Temchin (2007). Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods. J Assoc Res Otolaryngol. 8: 153-166.

It thus is argued that there seems to be a common mechanism that effectively minimizes the potential effects of anatomical diversity, yielding comparable latency patterns across species.

2.2.5 Latency pattern in a single unit to a click

The latency pattern in the peristimulus-time histogram of a single auditory nerve fiber to a click is similar across numerous species (e.g., among guinea pigs, chinchillas and cats [27]). At least for low-characteristic-frequency(CF) fibers, the classical responses in an auditory nerve fiber to a click in peristimulus-time histograms show multiple peaks separated by 1/CF, with a large initial peak and smaller subsequent peaks, and the peaks from rarefaction and condensation clicks neatly interleave [85, 86, 117-121] (Figure 23). Studies also showed these two characteristics are generally intensity invariant [27, 122-124].

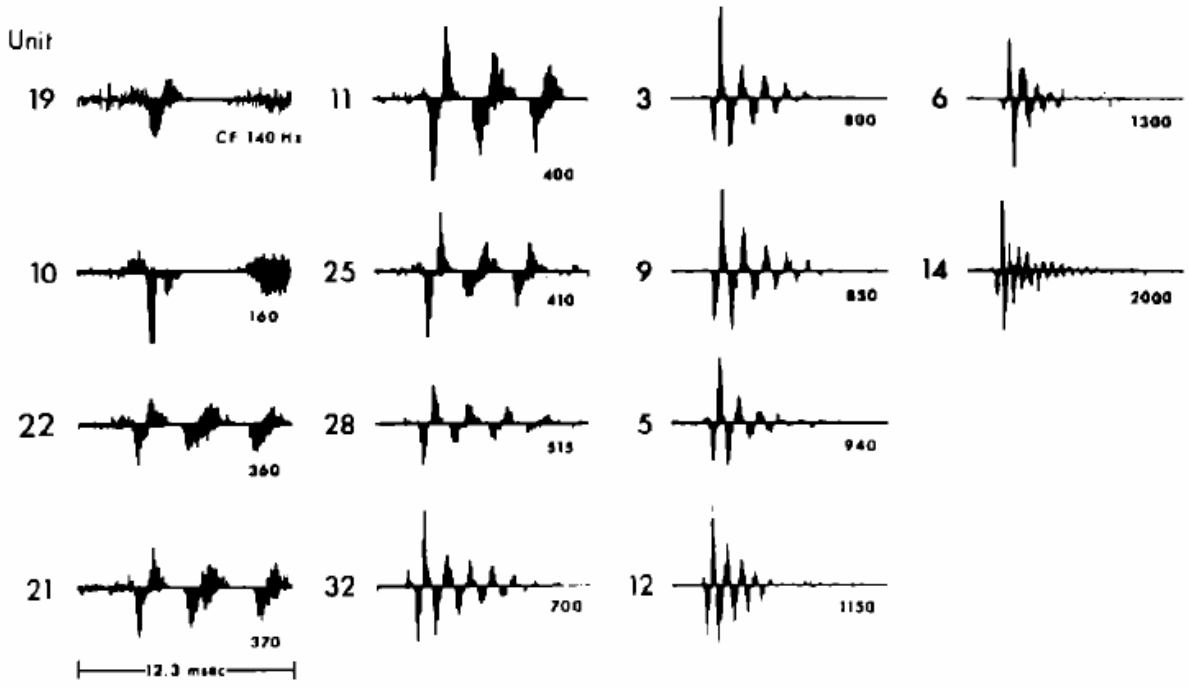


Figure 23. The classical responses in 14 auditory nerve fibers to a click. In peristimulus-time histograms (PST) the classical responses show multiple peaks separated by 1/characteristic-frequency(CF), and the peaks from rarefaction and condensation clicks neatly interleave. The compound histogram was formed by combining rarefaction-click peristimulus-time histograms with positive values (i.e., plotted upward) and the condensation-click PST histograms with negative values (i.e., plotted downward). Reprinted with

permission from Fig. 4, Pfeiffer, R. R. and D. O. Kim (1972). "Response patterns of single cochlear nerve fibers to click stimuli: descriptions for cat." J Acoust Soc Am 52(6): 1669-1677. Copyright 1972, Acoustic Society of America.

In mammals, the classical click responses are generally thought to be due to a resonant response of the basilar membrane at the neuron's characteristic frequency that is half-wave rectified by the inner-hair-cell synapse (e.g., [85, 93, 118]). The vibration of the basilar membrane stimulated by a click is considered as the vibration of a plucked violin string [85].

This hypothesis of basilar-membrane resonance implies that the latency pattern is determined by the half-wave rectification of the oscillatory excitation. In any event, this basilar-membrane-based interpretation fails to explain the responses in animals whose hearing organs lack a basilar membrane altogether [24].

2.2.6 Post-stimulus synchronized spikes

In quiet, peripheral auditory neurons discharge spontaneously at various rates [1, 30]. When a sound breaks the silence, the auditory neurons start to discharge specifically to the sound. The duration of the active discharge pattern usually represents the duration of the stimulus as well. However, a click that has an extremely short duration such as 25 to 100 microseconds (10^{-6} s) can be presented in the auditory nerve fibers for as long as 25 milliseconds (10^{-3} s) (e.g., [121]). The big duration discrepancy is remarkable.

As can be summarized from background information presented earlier, the post-stimulus neural latency pattern to a click is generally thought to be due to a resonant response of the sound transducer, filtered by the combined outer and middle ear, but thereafter dependent upon the

basilar membrane at the neuron's characteristic frequency that is half-wave rectified by the inner-hair-cell and synapse system (e.g., [85, 93, 118]). Therefore, the click responses in peristimulus-time histogram show multiple peaks separated by the inverse of the neuron's characteristic frequency, and the peaks from rarefaction and condensation clicks neatly interleave [85, 86, 117-121] (Figure 24A).

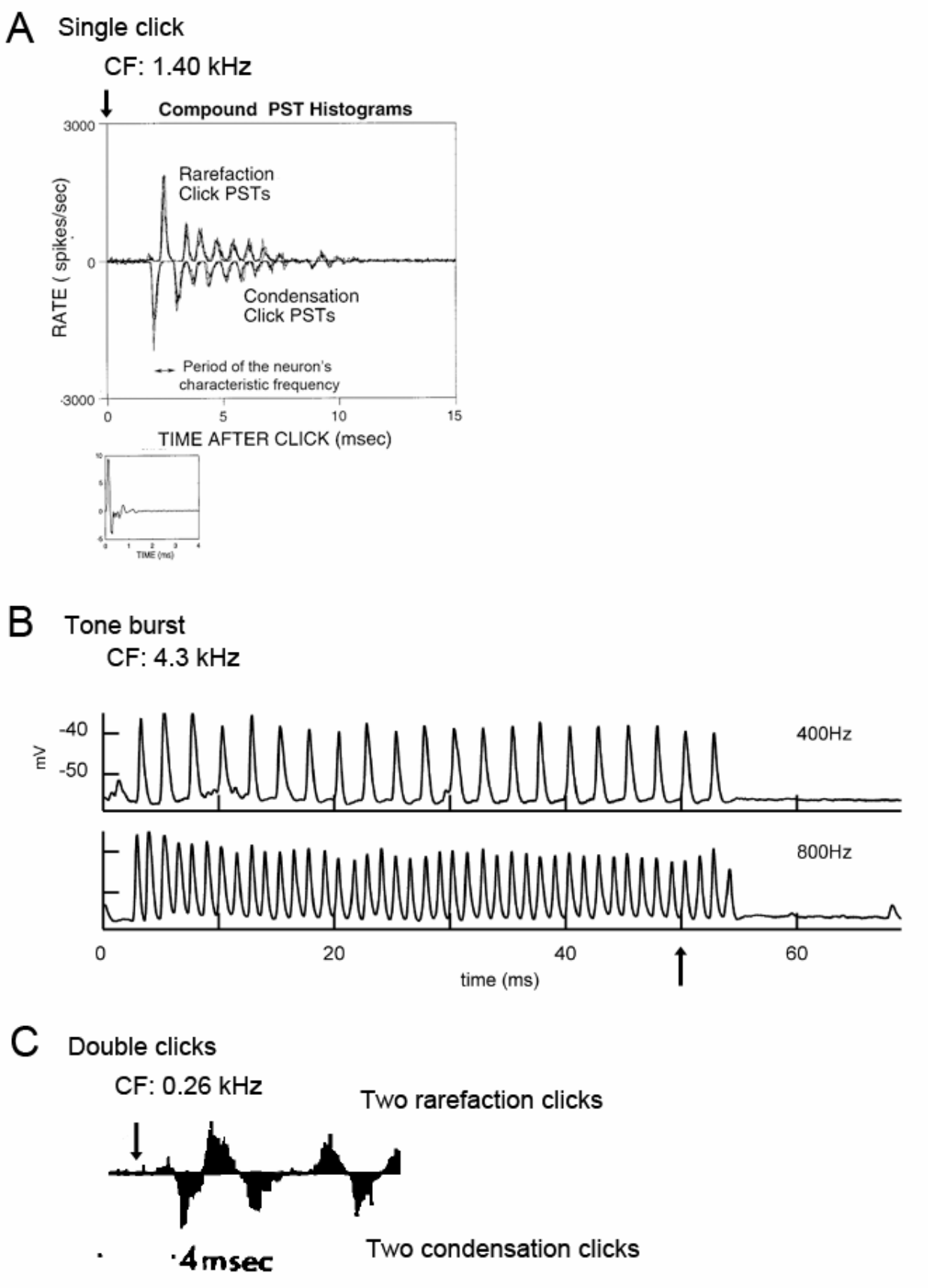

A
Single click
CF: 1.40 kHz
Compound PST Histograms
3000
0
-3000
RATE (spikes/sec)
Rarefaction
Click PSTs
Condensation
Click PSTs
Period of the neuron's
characteristic frequency
0
5
10
15
TIME AFTER CLICK (msec)
TIME (ms)
B
Tone burst
CF: 4.3 kHz
mV
-40
-50
400Hz
800Hz
0
20
40
60
time (ms)
C
Double clicks
CF: 0.26 kHz
Two rarefaction clicks
Two condensation clicks
4msec

Figure 24. A single auditory neuron continues to respond specifically to the stimulus after the stimulus ends. Arrows indicate the time when the stimulus is gone. The latency pattern changes when the stimulus varies. A. Compound peristimulus-time histogram to rarefaction and condensation clicks. CF: 1.40 kHz. The inter-peak interval is 1/CF. Reprinted with permission from Figs. 1 and 2, Lin, T. and J. J. Guinan, Jr. (2000). "Auditory- nerve-fiber responses to high-level clicks: interference patterns indicate that excitation is due to the combination of multiple drives." J Acoust Soc Am 107(5 Pt 1): 2615-2630. Copyright 2000, Acoustic Society of America. B. Peristimulus-time histograms of a cochlear-nucleus neuron to two 50-ms tone bursts, 400 and 800 Hz, respectively. The inter-peak interval is the inverse of the stimulus frequency. CF: 4.3 kHz. Threshold: 10 dB SPL Spontaneous rate: 16.5 spikes/second. Reprinted from Fig. 7, Neuroscience 154(1): 87-98. Rhode, W. S. "Response patterns to sound associated with labeled globular/bushy cells in cat", Copyright (2008), with permission from Elsevier. C. Compound peristimulus-time histogram of a single auditory nerve fiber to click-pair stimuli (a complex sound) with click separation of 1.96 ms. The responses to both click pairs are different from those to single clicks and are different in between the polarities. The inter-peak interval is irregular after the stimulus ends (when the second click is gone). CF: 260 Hz. Reprinted with permission from Fig. 6, Goblick, T. J., Jr. and R. R. Pfeiffer (1969). "Time- domain measurements of cochlear nonlinearities using combination click stimuli." J Acoust Soc Am 46(4): 924-938. Copyright 1969, Acoustic Society of America.

However, the duration discrepancy also can be found in auditory neural responses to other stimuli such as tone bursts [112, 116] or complex sounds (e.g., click pairs [118]). The duration of active neural responses is longer than that of the stimulus.

To tone bursts, most studies in the literature used a tone burst of the same frequency as the neuron's characteristic frequency to obtain the best synchronization of responses. It is impossible to interpret whether the post-stimulus responses are specific to the stimulus or not, because the stimulus frequency is the same as the neuron's characteristic frequency.

There are not many studies focusing on and demonstrating post-stimulus neural latency pattern in various units to an identical tone burst. Carney and Yin transformed a tone burst into a click in several steps and recorded the latency patterns to these stimuli in a single auditory nerve fiber in cats [94]. Their results showed when the stimulus was varied from a tone burst to a click, the dominant component of the synchronized response moved from the frequency of the stimulus toward the characteristic frequency of the fiber.

Rhode recorded a cat's single cochlear-nucleus neuron to various 50-ms tone bursts. The duration of active responses after the stimulus ended was about 5 ms (Figure 24B). The results suggested that the neuron was more likely to respond to the tone burst's frequency after the stimulus ends than to its characteristic frequency [116] (which is 4.3 kHz).

To complex sounds, taking the study of Goblick and Pfeiffer [118] for an example, their results showed that a single auditory neuron responded to a click-pair stimulus (two same-polarity clicks with the same intensity, separated by 1.96 ms) with a special latency pattern. Although the single unit responded to a single click similarly with the classical pattern in Figure 24A, it responded to the click pair in a specific pattern (Figure 24C) in which the peaks are not separated by the inverse of the neuron's characteristic frequency. The responses to the rarefaction-click pair were also different from those to the condensation-click pair. The auditory neuron continued to respond with stimulus-specific latencies after the second click ends, lasting for at least about 10 ms.

The examples in Figure 24 together show that an auditory neuron may continue to respond with different patterns to different stimuli, for about 5 to 10 ms (in these examples) after the stimulus ends. Thus in conventional view, the post-stimulus vibration of basilar membrane in

mammals must have different modes for different stimuli. But the underlying mechanism is unclear.

All of the above unique latency patterns come from the total response of pinna, middle ear, cochlear mechanics, and maybe the contribution from synaptic and neural transmission. Next, this final step will be reviewed. Synaptic and neural transmission may change the contribution proportion of parts of the pathway in the first spike latency.

2.2.7 Synaptic and neural transmission

Synaptic transmission here between a hair cell and a postsynaptic auditory nerve fiber [125-127] is an extremely high rate of exocytosis from a hair cell active zone to drive the auditory nerve fiber to discharge spikes [128, 129]. Still, it takes time. If delays resulting from this process are constant among different characteristic frequencies over time, it has no effect on encoding of phase information. On the other hand, if it is frequency dependent or time variant, it should have contributed in part to the specific neural patterns mentioned above. It is always true that the latency pattern in all animal studies contains the delay generated by synaptic and neural (the travel time along the axon) transmission.

Although there are not many reasons to suppose that the delay of neural and synaptic transmission varies with cochlear location, Ruggero and Rich [20] pointed out the first possible cause of variation is axon length. In cats, fibers originating in the 3- to 6-kHz characteristic frequency region are shortest (about 4.5 mm), whereas both basal and apical fibers are longer (apical about 6.5 mm) [130]. This length difference should imply that an action potential travel time is longer for apical fibers than for fibers in the 3- to 6-kHz characteristic frequency region,

assuming the fiber diameters are the same. The travel time may amount to 0.39 ms for the apical-most fibers, compared to 0.27 ms for those innervating the 3- to 6-kHz characteristic place [130], based on a general relation between axon diameter and conduction speed proposed by Arbuthnott et al [131], using a value of 2.54 μm for axonal diameter and a conduction speed of 16.8 m/s [130]. This suggests a maximum difference in travel time of 0.12 ms. The actual difference in cats and chinchillas is probably smaller [20]. There thus seems to be a mechanism of fiber diameter that apical fibers appear to be thicker than other fibers [130, 132, 133], to compensate the axon-length difference and maintain the delay of neural transmission constant.

The second possible cause of delay variation is from synaptic transmission. In mammals, each auditory nerve fiber receives input from only one inner hair cell synapse [134]. Thus, this single synapse determines the temporal precision of the fiber. There is growing interest in investigating the temporal mechanism, such as vesicle pools and ribbons, of synaptic exocytosis [128, 129, 134-136]. In brief, hair cells employ a whole set of mechanisms to achieve the extraordinary temporal acuity of sound coding with chemical synaptic transmission: short membrane time constants, Ca2+ nano-domain control of vesicle exocytosis, large readily releasable pool with parallel exocytosis of multiple vesicles and rapid gating of postsynaptic receptors [134]. All together these features minimize the delay added by synaptic transmission.

Studies on synaptic variation mainly reported differences of ribbon synapse morphology (spherical in apical and ellipsoid in basal inner hair cells). Johnson et al. showed the existence of tonotopic differences in the Ca^{2+} dependence of exocytosis and vesicle pool replenishment at ribbon synapses with high-frequency cells being more indefatigable than low-frequency cells [137]. In other words, when repetitively stimulating adult inner hair cells, vesicle pool refilling

could become rate limiting for vesicle release, with high-frequency inner hair cells able to sustain greater release rates.

Including Johnson et al, there seemed to be no literature indicating delay differences from synaptic transmission at different cochlear locations. Ruggero and Rich [20] concluded that, the neural and synaptic delay may be taken as equal to the latency asymptote in the onset of high-characteristic-frequency cochlear nerve responses to rarefaction clicks (where cochlear delay is the minimum). Their latency asymptote has a value of 1 ms. The value of 1 ms for synaptic and neural transmission in the chinchilla is consistent with other independent estimates from data for guinea pig and cat.

First, Palmer and Russell [138] measured synaptic delay in the basal turn of the guinea pig cochlea. Adding their estimate (0.8 ms) to an action potential travel time of 0.15-0.18 ms, calculated on the basis of morphological data of Brown [139] and conduction velocity estimates of Arbuthnott et al.[131], the delay for synaptic and neural transmission equals 0.95-0.98 ms. Second, van den Honert and Stypulkowski [140] measured responses of cat cochlear afferents to electrical stimulation. The latency was 0.9-l .3 ms.

To date, the constant value of 1 ms has been used as the putative delay for synaptic and neural transmission across different cochlear locations [22, 88, 113]. This implies that the synaptic and neural transmission has no effect on phase encoding. Even if experimental technology advances in the future and changes the invariant, any possible variation among different cochlear locations should be smaller than 1 ms. Moreover, this will change only the delay proportions contributed by middle ear, cochlear mechanics, and synaptic and neural transmission. The final unique spatiotemporal pattern in auditory neural population remains unchanged as it is.

2.2.8 Summary and discussion

Based on the frequency response characteristics of the middle ear and the frequency-place mapping of the cochlea [21], remarkable differences are evident among samples of animals across phyla. Similarly, the latency pattern of primary neural responses to an identical sound is expected to differ correspondingly. However, previous animal studies show recordings from different individual animals or even different species yielding a similar mean first-spike latency pattern to a rarefaction click and to tone bursts [22]. Also, the phase-locking responses to tone bursts and the classical responses to a click are similar across numerous species. The similarity reveals in remarkable animals such as the bobtail lizard (Tiliqua), whose basilar membrane does not support a traveling wave [28], and tree frogs, whose hearing organs lack a basilar membrane altogether [24].

These similarities (in the face of dramatic structural differences) suggest that the hearing organs across different species probably operate a common mechanism to decompose the input sound on the way to eliciting neural spikes. Since the similarity holds in animals that do not even have a basilar membrane, etc., such a common mechanism should comprise more than the mechanics of cochlear traveling waves. It is contended herein that shared attributes require a different interpretation than in the past. There seems to be no species- and individual-independent sharing of common attributes that explain this similarity today, except for the assertion that the hearing organ acts as a Fourier analyzer [29].

The assertion usually implies that the hearing organ as a spectrum analyzer converts a time-domain sound into frequency-domain components of different frequencies and intensities (two of the fundamental kinds of information). Even assuming all hearing organs are operating

Fourier analysis, it remains unanswered what specific mechanism in Fourier analysis makes the across-species similarity possible.

There have been numerous studies examining the mean first-spike latency stimulated by various tone bursts of different frequencies, as some of them are cited in this section. However, reports about the first-spike latencies among different auditory neurons stimulated by a single (same-frequency and same-phase) tone burst, or even a single stimulus, are not so numerous in the literature.

Shamma and Loeb demonstrated that the first-spike latencies to a given tone burst may form a specific pattern which represents the tone burst's frequency and phase information. The Fourier pattern created by a given tone burst in Fourier transform shows the similar features demonstrated by Shamma and Loeb. This mathematical-like encoding of sound by the auditory periphery (in effect, Fourier analysis) is proposed to incorporate a mechanism of phase pattern capture that, in the face of species differences in peripheral auditory hardware, accounts for the considerable similarities across species in their latency-by-frequency functions, in turn assuring optimal phase encoding across species.

Although the Fourier pattern varies on the scale of a few milliseconds, previous work has shown that onset responses to isolated brief sounds can be millisecond-precise and carry acoustic information [141-143].

For example, Heil and Irvine [143] demonstrated the striking precision of the first spike in a total of 77 auditory nerve fibers in cats. The fibers had characteristic frequencies ranging from 0.6 to 35.5 kHz. Ten fibers (13%) were low spontaneous rate (SR). Twenty-seven (35%) were medium-SR and 40 (52%) were high-SR. The stimuli were 200-ms tone bursts at each

stimulated neuron's characteristic frequency. They found the standard deviation of the first-spike latency could be equal to or smaller than 1 ms in medium-SR and high-SR fibers.

The precise response timing not only prevails at peripheral stages of the auditory system, but is a general principle of auditory coding throughout the sensory pathway. For example, Kayser et al. [12] recorded single neurons' responses of 41 neurons in the caudal auditory cortex of passively listening macaque monkeys, to a sequence of pseudorandom tones and a continuous 52-second sequence of natural sounds. Using information theoretical analysis, they demonstrated that neurons in primate auditory cortex can indeed use spike times with a precision of a few milliseconds to carry information about complex and behaviorally relevant sounds.

Furthermore, trained rats were shown to be able to report millisecond differences in patterns of ***electrical*** microstimulation applied to auditory cortex [144], demonstrating that auditory cortex can access temporally precise patterns of local activity. Modern cochlear implants now permit precise timing of stimulation. For example, the CLARION CII provides 1 μ sec resolution for arbitrary waveforms including biphasic pulse rates > 100 kpps [7]. However, current implant processors only deliver envelope information. The envelope detection process replaces variety of original phases by a non-informative constant offset of time.

As Kayser et al. demonstrated that registering spikes at a precision coarser than a few milliseconds significantly reduced the encoded information [12], the non-informative phase pattern distorts the encoded information and thus reduce the performance of implantees. Next, literature will be reviewed, about the status and issues of phase encoding in electric hearing established by current cochlear implants.

2.3 ELECTRIC HEARING AND PHASE

2.3.1 Neural representation

The stimulating electrode array of a cochlear implant is usually implanted within the scala tympani surrounded by a layer of conductive perilymph. Beyond perilymph is a layer of bone. The stimulating electrodes can activate an auditory neuron by using the electrodes as bipolar pairs, where current is delivered by one electrode and returned through an adjacent electrode, or as monopolar electrodes, where the current return is to a distant ground embedded under the temporalis muscle.

Since there are various media and barriers between an electrode and the auditory neuron to be activated, it is essential to learn the real activity of the auditory neurons in electric stimulation. When phase in electric hearing is investigated, it is essential to understand the temporal responses of the auditory neurons to electric stimulation.

2.3.1.1 Single auditory neuron response in single electric stimulation

Liang et al. [145] recorded anesthetized cats' single-auditory-nerve response to electrical stimulation delivered to the cochlea through an implant with an eight-electrode array similar in construction to the Hochmair electrode [146], but scaled appropriately for the cat cochlea. The electrode array was inserted into a cat's scala tympani through the round window. The return electrode was placed in a neck muscle ipsilateral to the scala tympani electrode.

They used a (100 μ s/phase) charge balanced biphasic pulse as a stimulus. Charge balanced biphasic pulses are mostly used in current cochlear implants to prevent damaging electrochemical reactions [147]. They showed the single auditory neuron response by subtracting

the reference (no response) from the sum of the large artifact and the single auditory neuron response [145]. Figure 25 shows a typical response they recorded. The action potential almost immediately starts in the middle of the electric stimulus and peaks in 0.2 ms to above 5 mV.

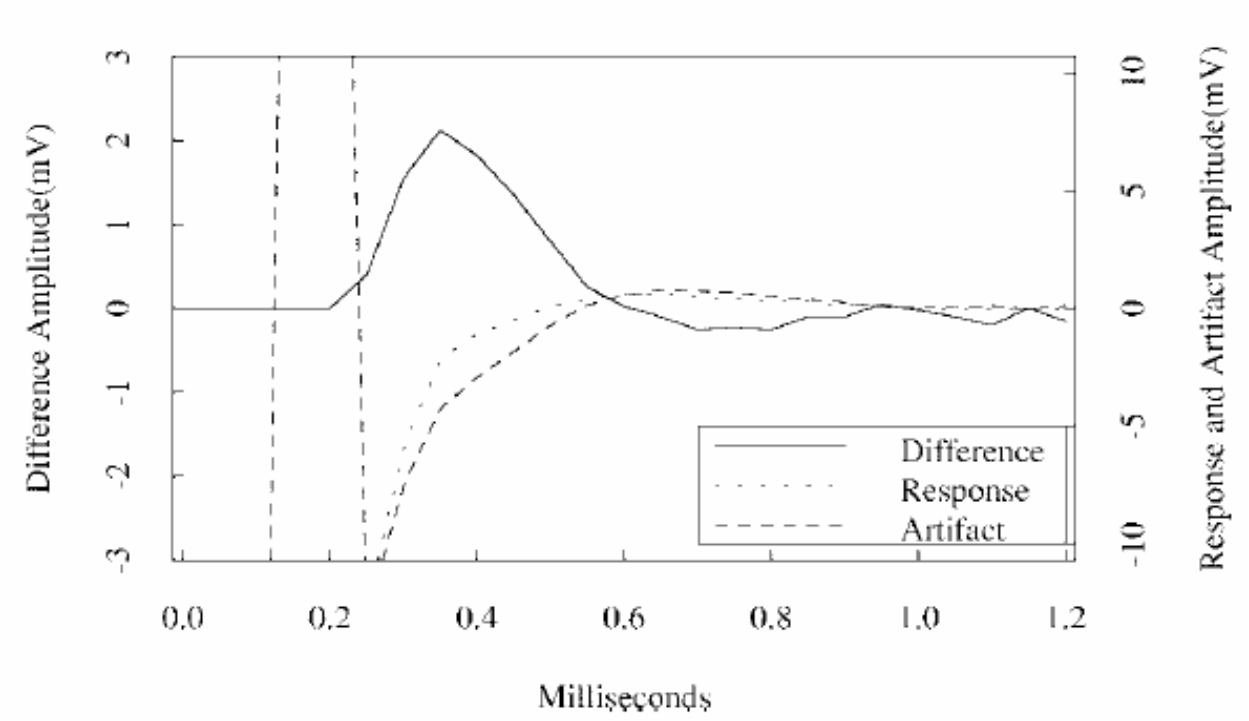


Figure 25. Anesthetized cat's single auditory neuron response in electric stimulation. The stimulus was a (100 μs/phase) charge balanced biphasic pulse. The large artifact is shown by the dashed line and the sum of the artifact and the single-unit response is given by the dotted line. The difference, which is the single-unit response in isolation, is given by the solid line. Reprinted with permission from Fig. 2, Liang, D. H., H. S. Lusted, et al. (1999). "The nerve- electrode interface of the cochlear implant: current spread." IEEE Trans Biomed Eng 46(1): 35- 43. Copyright 1999, IEEE.

Animal studies demonstrate that auditory-nerve neurons comply with basic neurophysiologic principles in their responses to electrical stimulation [148]. Studies using monophasic or charge-balanced biphasic current pulses have shown that the auditory neuron

response is dominated by a highly synchronized short-latency (~0.5 ms) response [149-152]. Although the latencies between electric stimulation and neuron response may differ in different auditory neurons [152, 153], for an identical neuron stimulated by charge-balanced biphasic current pulses, the response latency to the same pulse strongly holds constant [152, 153]. Figure 26 showed some examples of responses to four different charge-balanced biphasic current pulses.

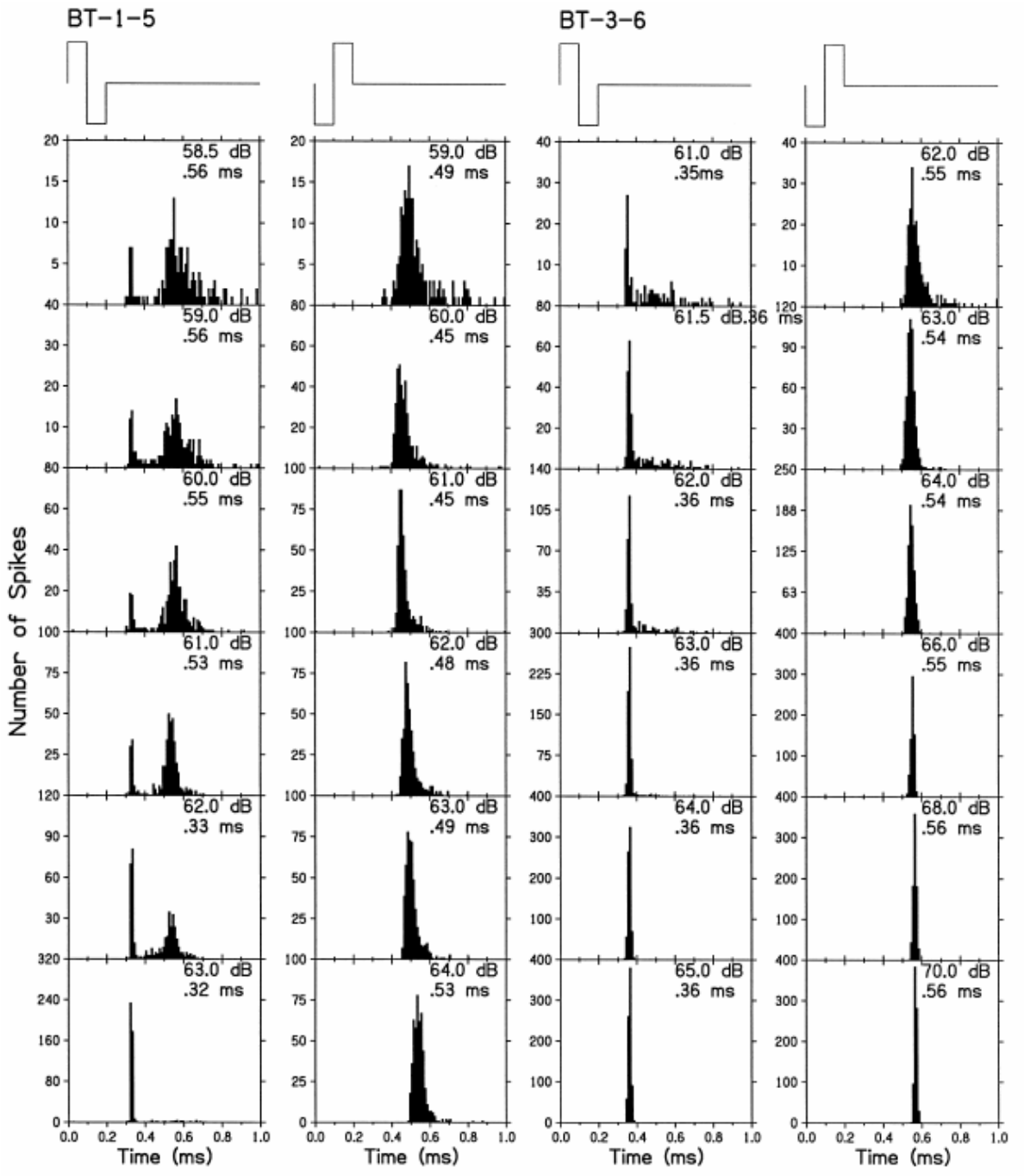


Figure 26. Period histograms of 2 (cat) auditory neurons responding to charge-balanced biphasic current pulses. The responses evoked by the anodic leading phase of BT-1-5 illustrates an example of peak splitting, where the lagging phase appeared to evoke the near threshold activity while the leading phase dominated the response as the neuron approached saturation. Reprinted from Fig. 3, Hear Res 130(1-2): 171-

188. Shepherd, R. K. and E. Javel "Electrical stimulation of the auditory nerve: II. Effect of stimulus waveshape on single fibre response properties", Copyright (1999), with permission from Elsevier.

Whether the latency between the electric pulse and the response is constant will be further discussed in the section “Latency between pulse and response (by interspike interval analysis)”. This is an essential factor to replicate a stimulating pattern in spike responses.

2.3.1.2 Single auditory neuron response and stochastic effect

A high-rate (e.g., 5000 pulse-per-second, pps) desynchronizing pulse train was proposed to better represent natural auditory activity by modulating the original electric stimulation [154, 155], due to the fact that current cochlear implant processors poorly represent sound waveforms in the temporal discharge patterns of auditory-nerve fibers. This strategy is based on the observation that high rate pulse trains may produce random spike patterns which are statistically similar to those produced by spontaneous activity in the normal cochlea [156].

Litvak et al. showed that the temporal representation of sinusoidal stimuli can be improved in a majority of auditory nerve fibers by encoding the stimuli as small modulations of a sustained 5 kpps desynchronizing pulse train [155]. Litvak and associates also calculated, in anesthetized cats, the correlation coefficient between the period histograms of stochastic-threshold-model responses to vowel modulators and those of auditory-nerve-fiber responses as a function of pseudo-spontaneous discharge rate. With high correlation, they suggested that using high-rate desynchronizing pulse trains may improve neural representation of vowels in electric stimulation [154].

However, as concluded by Litvak et al. [154], key unanswered questions included whether a desynchronizing pulse train could be safe but stimulate a sufficient number of auditory

nerve fibers to make a difference, and whether the central nervous system would be able to make use of the information provided by a desynchronizing pulse train. Another key issue, raised by Rubinstein et al. [156], was that the pulse train was perceptible to the cochlear implant user when they presented at levels sufficient to desynchronize the nerve. As Rubinstein and associates said, it is reasonable to hypothesize that an electrical stimulus which evokes pseudospontaneous activity should not be perceptible over the long term as long as the electrically-evoked activity mimics normal spontaneous activity.

2.3.1.3 Single auditory neuron response to a pulse train

Miller et al. reported their auditory nerve recordings in 8 anesthetized deafened cats to 300 ms electric pulse trains of 3 different constant rates (250, 1k and 5k pps) [151]. They examined the effects of pulse trains on auditory nerve fiber responses across the duration of such stimuli.

Latency between pulse and response (by interspike interval analysis)

The inter-pulse interval of a 250 pps electric pulse train is 4 ms. It is longer than the refractory period of an auditory neuron (given an absolute refractory period of 0.5 ms [1]). However, the inter-pulse interval of a 5 kpps electric pulse train is 0.2 ms, which is shorter than an absolute refractory period. Figure 27 shows a constant latency between pulse and response when the stimulating rate was 250 pps, and increased "randomness" on latency when the stimulating rate was 5 kpps. The figure also shows that increasing level of current will help to synchronize the latency between pulse and response. However, even with high-level current, the latency will not be the inter-pulse interval of the pulse train when the interval is close to or shorter than the refractory period (e.g., 1k and 5k pps) [151].

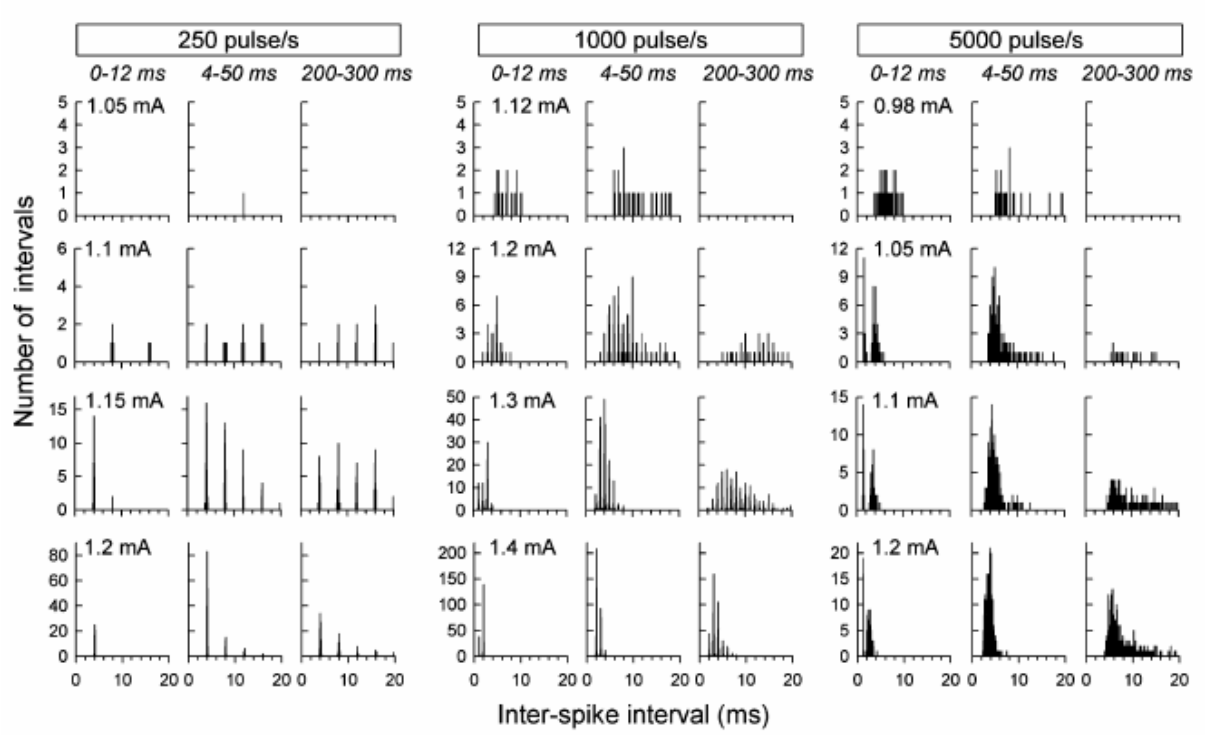


Figure 27. Interspike interval showed effect of electric pulse rate and level on latency between pulse and response. Interval histograms were obtained from single auditory neurons of anesthetized deafened cats. The inter-pulse interval of a 250 pps electric pulse train is 4 ms. It is longer than the refractory period of an auditory neuron. When a low-rate pulse train whose inter-pulse intervals are longer than an auditory neuron's refractory period, the auditory neuron tends to respond to a pulse in a constant latency. Therefore, the auditory neuron discharged with the same pattern as the electric pulse train (but failed to catch every pulse). With kind permission from Springer Science+Business Media: Fig. 1, Miller, C. A., N. Hu, et al. (2008). "Changes across time in the temporal responses of auditory nerve fibers stimulated by electric pulse trains." J Assoc Res Otolaryngol 9(1): 122-137.

This provided evidence that, at least when a low-rate pulse train whose inter-pulse interval is longer than an auditory neuron's refractory period, the auditory neuron tends to respond to a pulse in a constant latency. Therefore, the auditory neuron discharged with the same pattern as the electric pulse train (but failed to catch every pulse).

In contrast, interspike intervals evoked by 5k pps trains were dominated by refractory and adaptation effects. This suggests a desynchronizing effect of high-rate trains [151]. In other words, even although the pulse rate is constant, the high-rate pulse train may randomize discharges of an auditory neuron. One concern for the reliability of the desynchronizing effect at high rate is the temporal resolution of their experimental equipment. A constant latency at 250 pps (Figure 27) is confirmed because, given enough intensity, the main inter-spike interval is 4 ms. At least a temporal resolution of 2 ms (4 ms / 2) is required to tell that there was almost no spike between 4 ms and 8 ms. Similarly, at least a temporal resolution of 0.1 ms (0.2 ms / 2) is required to prove whether the latency is constant or not at 5k pps.

In addition, as reported by Miller et al., the use of artifact subtraction schemes requires care, as none of them work perfectly. It may result in distortion contributing to altered latency and, in some cases, increases in apparent spike synchrony [151] (that is why they visually inspected all waveforms).

In summary, animal studies show that an auditory neuron responds to electric pulses of an intracochlear electrode with a short latency. When the interpulse interval is longer than the refractory period (i.e., stimulating rate is low), neural spikes in an auditory neuron can precisely represent the stimulating pattern. When the stimulating rate is high (e.g., > 1000 pps), neural spikes can precisely represent the beginning [151] and ending [157] timing. During a high-rate pulse train, an auditory neuron can sustain the maximal discharge rate (every refractory period) for tens of milliseconds in the onset stage (Figure 27). These together suggest that the precise stimulating pattern can be replicated in auditory nerve fibers including the onset pulse, and at least the onset and offset of the stimulating pulse trains.

2.3.2 Phase processing of current devices

2.3.2.1 Compressed analogue (CA) strategy

Before currently employed continuous interleaved sampling strategy [10] or its derivatives, the standard processing strategy for cochlear implants was CA strategy [158]. Each of the channels in the standard CA processor consisted of a bandpass filter, a voltage-to-current converter, and an electrode. Since the CA processor simply bandpass filtered the input signal then separately converted to appropriate electric current for each electrode, it presented analogue waveforms simultaneously to all electrodes by voltage differences. In other words, phase information in each channel is supposed to be presented in the corresponding electrode.

However, according to physical laws, electric current occurs whenever there is a vector sum of the electric field among electrodes and the vector sum overcomes the resistance in between. There are multiple interactions among the 3-dimensional intracochlear electrodes employing CA strategy. These interactions both fail to deliver accurate place code and to preserve the original phase information within one channel, because of interfered current.

2.3.2.2 Continuous interleaved sampling (CIS) strategy and its derivatives

Most of current cochlear implant processors use CIS strategy or its derivatives. The CIS strategy was first reported in 1991 [10] and is used as the default strategy or as a processing option in all implant systems now worldwide [6]. It addresses the problem of channel interactions [11] through the use of interleaved non-simultaneous stimuli. Charge-balanced biphasic pulse trains are delivered to each electrode with temporal offsets that eliminate any overlap across channels. The envelopes of the bandpass outputs are extracted by rectification and

lowpass filtering. Finally, the envelopes of each channel are converted to electric amplitude and are compressed into the dynamic range appropriate for that channel.

The rectification or similar envelope detecting process fails to preserve the phase information, at least to a certain extent [6]. In addition, the non-informative temporal offsets further distort the original phase information. This is the well-known limitation that current cochlear implants do not deliver phase (or temporal fine structure) information [6-10].

In summary, most current cochlear implant processors use continuous interleaved sampling strategy or its derivatives. This processing successfully addresses the problem of channel interactions by use of interleaved non-simultaneous biphasic pulse with a constant temporal offset. However, as concluded by Wilson and Dorman [6], the only genuine phase information provided by current envelope-detection based processors is time-varying envelope information (sometimes referred to by other articles as temporal envelopes), if it is not totally removed as claimed by some researchers. As mentioned in previous sections, (Hilbert) envelope of a sinusoid is a straight line, which contains no phase information. Envelope and temporal fine structure of a complex sound together contain the three kinds of fundamental information (intensity, frequency and phase), but how much phase information is represented in envelope is unclear.

2.3.3 Auditory perceptions with current devices

Auditory perceptions by electric stimulation are affected by many well-known or not yet understood mechanisms. Many possible limitations lead to Wilson's phrase that "cochlear implant signals are sparse and distorted inputs" [6]. In addition to the above-mentioned failure of

genuine phase delivery, these limitations and weaknesses may include restricted channel number [8, 9], channel interaction [11], shifted frequency range and mismatched frequency mapping [8, 159-162], limited stimulating rate, bad presentation of the fundamental frequency [6], lacking coordination of rate with place of stimulation [6], and no replication of the latency fields that occur in normal hearing [6, 7].

It is not fully understood how much is affected by phase information in each electric auditory perception, and how different perceptions interact with each other. Generally accepted, temporal fine structure significantly affects at least speech recognition in noise, music perception and tone perception in tonal languages [6]. Therefore, here studies of most possible affected perceptions are reviewed.

Other auditory perceptions of cochlear implant users are measured in some studies. The performances varied and highly depended on the test designed, from poor (close to chance level) to moderate (but not as good as normal listeners). Relevant reports include gender identification [163-166], talker recognition [167-170], perception of fundamental frequency [171-180], and voice emotion recognition [181]. To date, the effect of phase processing on these electric auditory perceptions remains unclear.

2.3.3.1 Speech recognition in noise

As directly pointed out by Wilson and Dorman, even cochlear implant users with 100% correct scores in quiet have significant difficulty in noise [6]. There seems to be no generally accepted manipulation of delivering vs. non-delivering phase information to cochlear implant users. Therefore, the method of adding vs. not-adding phase-contained information was employed in some studies. Although it may not be the best approach, one possible way to

investigate the effect of phase information on speech recognition in noise for implantees is to look at the effect of additional non-contributing low-frequency acoustic hearing.

Kong et al. recruited 5 cochlear implant users with residual acoustic hearing in the non-implanted ear to elucidate the effect of additional low-frequency acoustic hearing on speech recognition in noise [13]. All subjects were implanted with regular long electrode devices. They manipulated the signal-to-noise ratio (0 to 20 dB) in the test for sentence recognition in noise. They presented a target sentence which has 5 key words with a fixed competing sentence by another talker, “Port is a strong wine with a smoky taste”. They then asked the subjects to type what they heard using a keyboard and the results were scored by Matlab software.

Figure 28 shows the correct score was almost zero when signal-to-noise ratio was zero, and gradually improved to about 30% (implant only) when signal-to-noise ratio was 20 dB. Although the residual low-frequency (<1000 Hz) acoustic hearing produced essentially no recognition for speech in noise, it significantly enhanced performance when combined with the electric hearing [13].

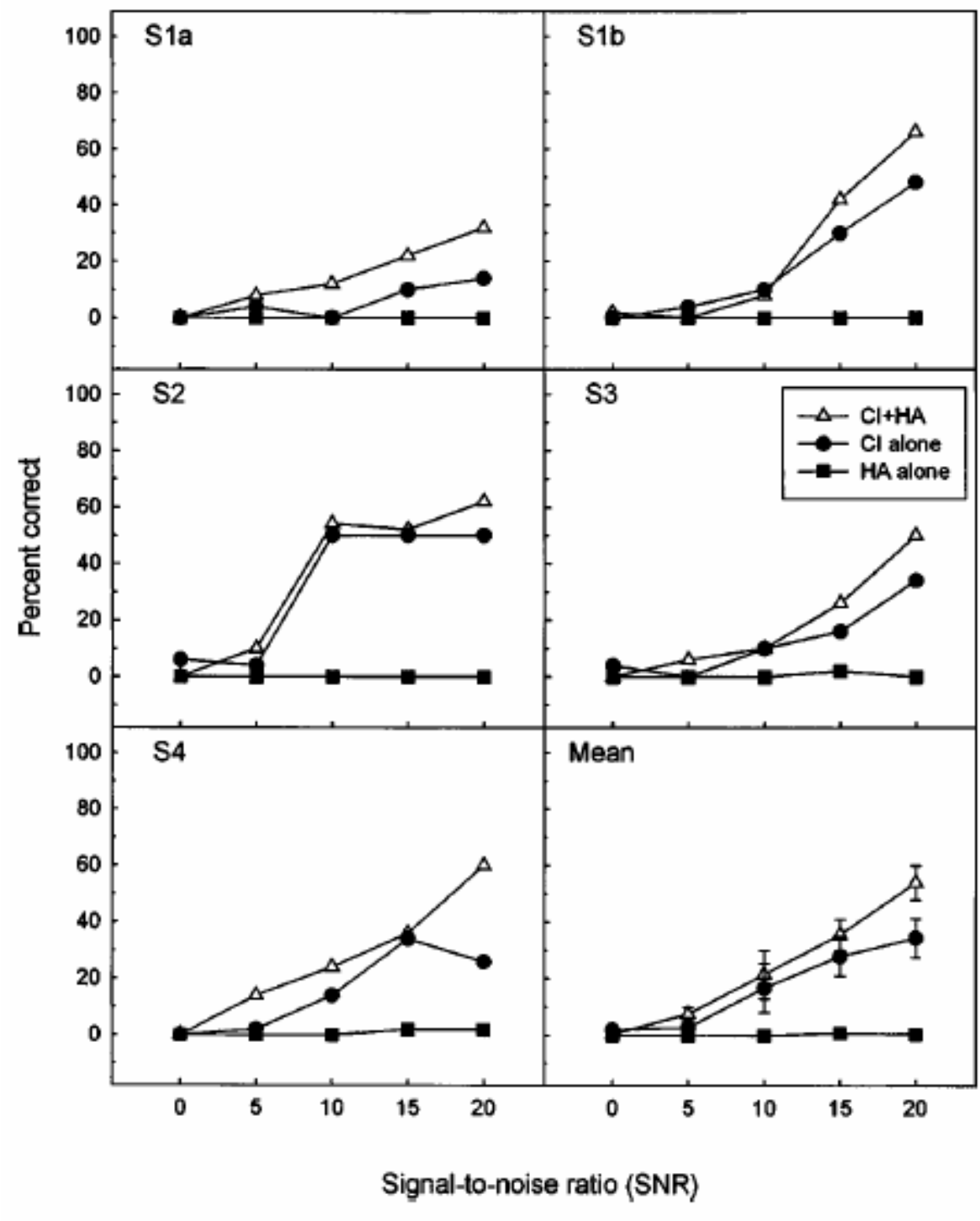


Figure 28. Individual and mean open-set sentence recognition scores in noise for cochlear implant (CI) users. The three functions in each panel represent the CI + hearing aid (HA) (open triangles), CI alone (closed circles), and HA alone (closed squares) conditions. A target sentence having 5 key words was presented with a same competing sentence by another talker, "Port is a strong wine with a smoky taste".

Subjects were then asked to type what they heard using a keyboard and the results were scored by Matlab software. Reprinted with permission from Fig. 2, Kong, Y. Y., G. S. Stickney, et al. (2005). "Speech and melody recognition in binaurally combined acoustic and electric hearing." J Acoust Soc Am 117(3 Pt 1): 1351-1361. Copyright 2005, Acoustic Society of America.

Balkany et al. [182] conducted a randomized, controlled, prospective, single-blind study using single-subject repeated measures at fourteen academic centers in the United States and Canada to evaluate hearing outcomes with the Nucleus Freedom cochlear implant. They used Consonant Nucleus Consonant (CNC) words and Hearing in Noise Test (HINT) on 71 severely/profoundly hearing impaired adult implantees. Mean scores for Consonant Nucleus Consonant words was 57%, Hearing in Noise Test (HINT) sentences in quiet 78%, and HINT sentences in noise 64%.

Nie et al. manipulated channel number (4 to 12) or stimulation rate (from 1000 to 4000 Hz per electrode on four pairs of fixed electrodes) by a Research Interface Box to test five MED-EL COMBI 40+ cochlear implant users' performances of vowels, consonants, and HINT sentences in quiet and noise [183]. For effect of channel number, they reported that the noise decreased the sentence recognition score by 30 percentage points with 4 electrodes (quiet 53% and noise 23% correct) and 32 percentage points with 12 electrodes (quiet 84% and noise 52% correct) ($p < 0.05$). For stimulation rate, increasing stimulation rate from 1000 to 4000 Hz only improved 4 % correct for sentence recognition in noise (non-significant).

Hong and Turner adapted the method of Robert et al. [184] to investigate the correlation between the ability of pure tone stream segregation and understanding of speech in noise with 8 cochlear implant users [185]. They presented 2 tone streams to the subjects and asked them to choose which tone sequence had an irregular rhythm. Each of the tone streams was composed of

two 60-ms tone burst in an overlapped AB pattern (tone burst A, tone burst B, tone burst A, tone burst B…). The tone stream with irregular rhythm was made by three parts: steady rhythm in the first 6 AB cycles, progressively irregular rhythm (delaying the onset time of tone burst B) in the next 4 cycles, and maintained irregular rhythm in the final 2 cycles (accumulated delay). Their results showed better stream segregation was associated with better understanding of speech in noise, and that the inability of some cochlear implant users to perform stream segregation may therefore contribute to their difficulties in noise backgrounds [185].

Although above research suggests precise time coding and additional phase-rich non-contributing low-frequency acoustic hearing may improve speech recognition in noise, there seems to be no literature that provides direct evidence. The key is the lack of a generally accepted manipulation of delivering vs. non-delivering phase information to cochlear implant users. The bottom line is that performances of speech recognition in noise for implant users are not good but improvable.

2.3.3.2 Music perception

Music perception has been investigated along many characteristics such as rhythm perception, chord discrimination, dissonance rating, emotion rating [186], pitch relationship between 2 tones (same, higher, or lower) [187], melodic recognition [13-15, 188-191], timbre [189, 192-195], direction of pitch change (higher or lower) [15]. Music perception also has been studied among different speech processing strategies [196] and different hearing paradigms (acoustic, electric and combined) [15, 186].

Kong et al. conducted one of the representative studies to investigate music perception for cochlear implant users [14]. Twelve popular and familiar melodies were used as test stimuli in four conditions (original rhythm, 1-band rhythm, original no-rhythm, and one band no-

rhythm). After a 5-to-10-minute practice session to familiarize all the stimuli (with visual feedback), the subjects were asked to choose what they heard from the 12 titles shown on a computer screen (chance level = 8.3%). Figure 29 shows the correct scores of both normal listeners and cochlear implant users. In the no-rhythm conditions (all melodies were played using notes of the same duration ,quarter notes with 350 ms in duration, with a silent period of 150 ms between notes), the performance of cochlear implant users was almost at chance level.

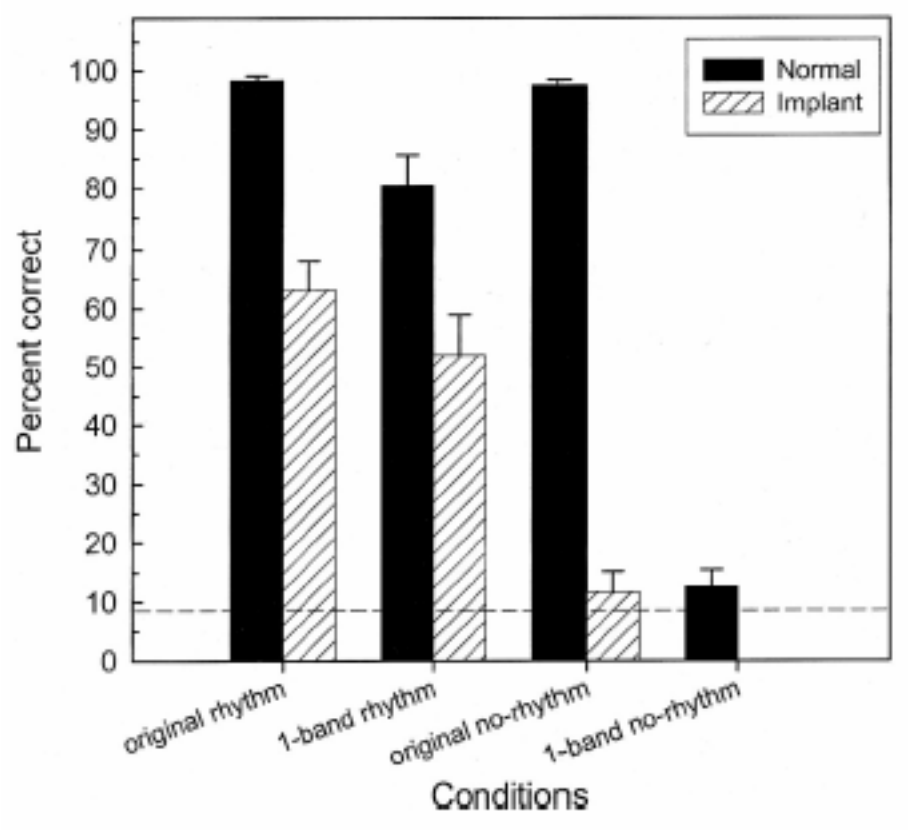


Figure 29. Close-set melody identification for normal listeners and cochlear implant users. After 5-10 minutes practice session to familiarize all the stimuli (with visual feedback), the subjects were asked to choose what they heard from the 12 titles shown on a computer screen. The horizontal dashed line indicates the mean chance performance. Reprinted with permission from Fig. 5, Kong, Y. Y., R. Cruz, et al. (2004). "Music perception with temporal cues in acoustic and electric hearing." Ear Hear 25(2): 173-185. Copyright 2004, Lippincott Williams.

Singh et al. [190] separated melodies into 3 different groups with fundamental frequencies of low (104-262 Hz), middle (207-523 Hz), and high (414-1046 Hz) ranges to investigate the effect of a melody's frequency range on the recognition score in 5 cochlear implant users. They also manipulated the number of active electrodes to measure melody recognition score. They found that frequency range significantly affected cochlear implant melody recognition, with higher frequency ranges producing better performance. Increasing the number of activated electrodes did not improve performance with low- and middle-frequency melodies but produced better performance with high-frequency melodies [190].

Gfeller et al. [15] recruited 101 monaural cochlear implant users, 13 hybrid device users and 21 normal listeners to investigate performance of pitch ranking among 3 groups, which suggested groups of electric hearing, high frequency electric + low frequency acoustic hearing (which contains rich phase information), and normal hearing. The stimuli in their study were tone bursts ranging from 131 Hz (one octave below middle C on the piano) to 1048 Hz (two octaves above middle C on the piano). The tones were 500 ms in duration, with 25 ms rise-fall ramps. The time interval between tones of each trial sequence was 300 ms. Each trial consisted of three tones presented sequentially (the first two tones were the same frequency and the third tone was higher or lower). Subjects were asked to indicate by touch screen whether the final tone was higher or lower in pitch than the first two [15]. Figure 30 shows that the long electrode recipients performed significantly poorer on pitch discrimination than the normal listeners and hybrid device users. There seemed to be a trend that the performances of long electrode recipients and hybrid device users became similar at pitch ranking test with a high base frequency. These together suggested the significant effect of the low frequency acoustic hearing on pitch ranking for implant users.

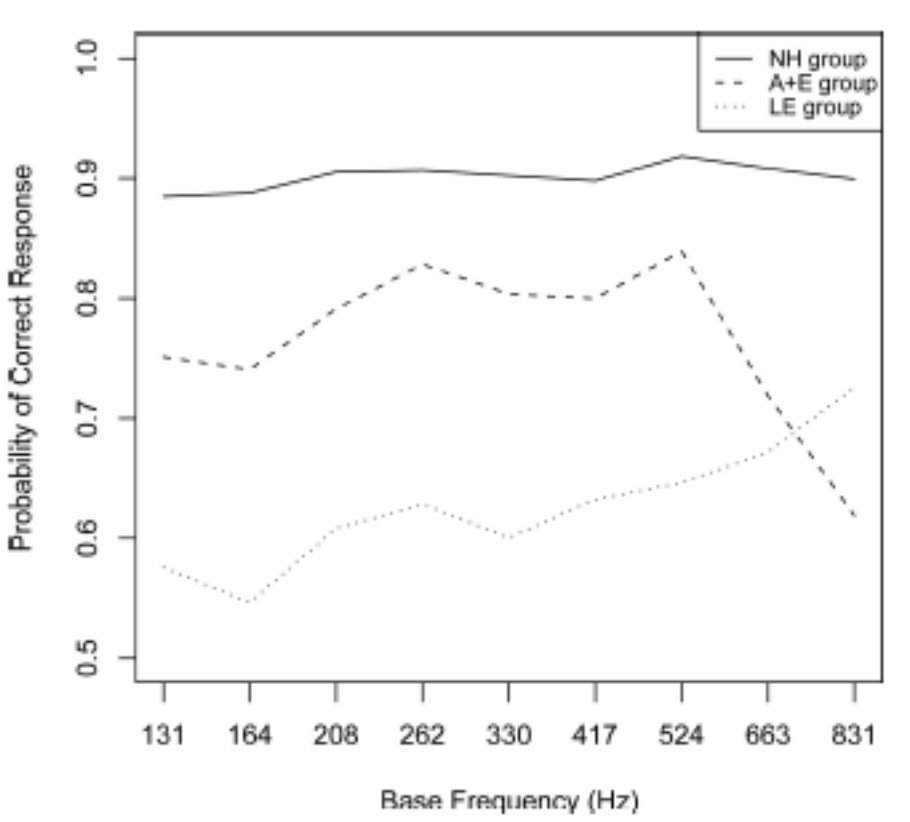


Figure 30. Probability of correct response of pitch ranking test as a function of base frequency (at interval size of 1 semitone). Subjects were asked to indicate by touch screen whether the final tone was higher or lower in pitch than the first two same tones. Chance level is 0.5. Normal-hearing (NH), short electrode users (A+E), and long electrode (LE) users are plotted separately. Reprinted with permission from Fig. 1, Gfeller, K., C. Turner, et al. (2007). "Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise." Ear Hear 28(3): 412-423. Copyright 2007, Lippincott Williams.

Again, although above researches suggested rhythm (a time code) and additional phase-rich low-frequency acoustic hearing may improve music perception, there seems to be no literature that provided direct evidence. The key is the lack of a generally accepted manipulation of delivering vs. non-delivering phase information to cochlear implant users. The results suggested that music perception for implant users is not good but probably improvable.

2.3.3.3 Tone perception in tonal languages

In tonal languages, a different fundamental-frequency contour of each monosyllable conveys a different lexical meaning. For example, there are 4 patterns of contour in Mandarin, and Cantonese is generally considered to have 6 tones. Information of lexical tones is conveyed by fundamental frequency and its harmonics. Therefore, how fundamental frequency and its harmonics are processed and delivered to cochlear implant users determines their electric lexical-tone perception.

With restriction of approach to the apex and mismatch of frequency range and mapping, cochlear implants represent less-than-resolute fundamental frequencies for complex sound [6]. Not representing the fundamental frequency well is an important issue for lexical-tone perception for cochlear implant users, but is not the only one.

Interestingly, as mentioned in the section “Lexical tone perception” above, evidence was shown in a lexical-tone study investigating the effect of exchanging envelope and temporal fine structure components among 4 lexical tones of an identical Mandarin Chinese monosyllable [43]. Subjects based their identification of lexical-tone significantly and mainly on the fine structure rather than the envelope. This result in normal listeners also suggested the significance of phase information in lexical-tone perception for cochlear implant users.

Wong and Wong reported poor outcome in lexical-tone perception in Cantonese implantees [197]. They recruited a total of 17 native Cantonese-speaking children with Nucleus cochlear implants between 4 to 9 years of age. They used a same/different paradigm to test tone discrimination (whose chance level is 50%) and a four-alternative, force-choice method to test tone identification (whose chance level is 25%) [197]. The result showed Cantonese-speaking implanted children could poorly discriminate and/or identify Cantonese lexical tones. Subjects'

performances in discrimination and identification tasks were both slightly above chance level (Figure 31).

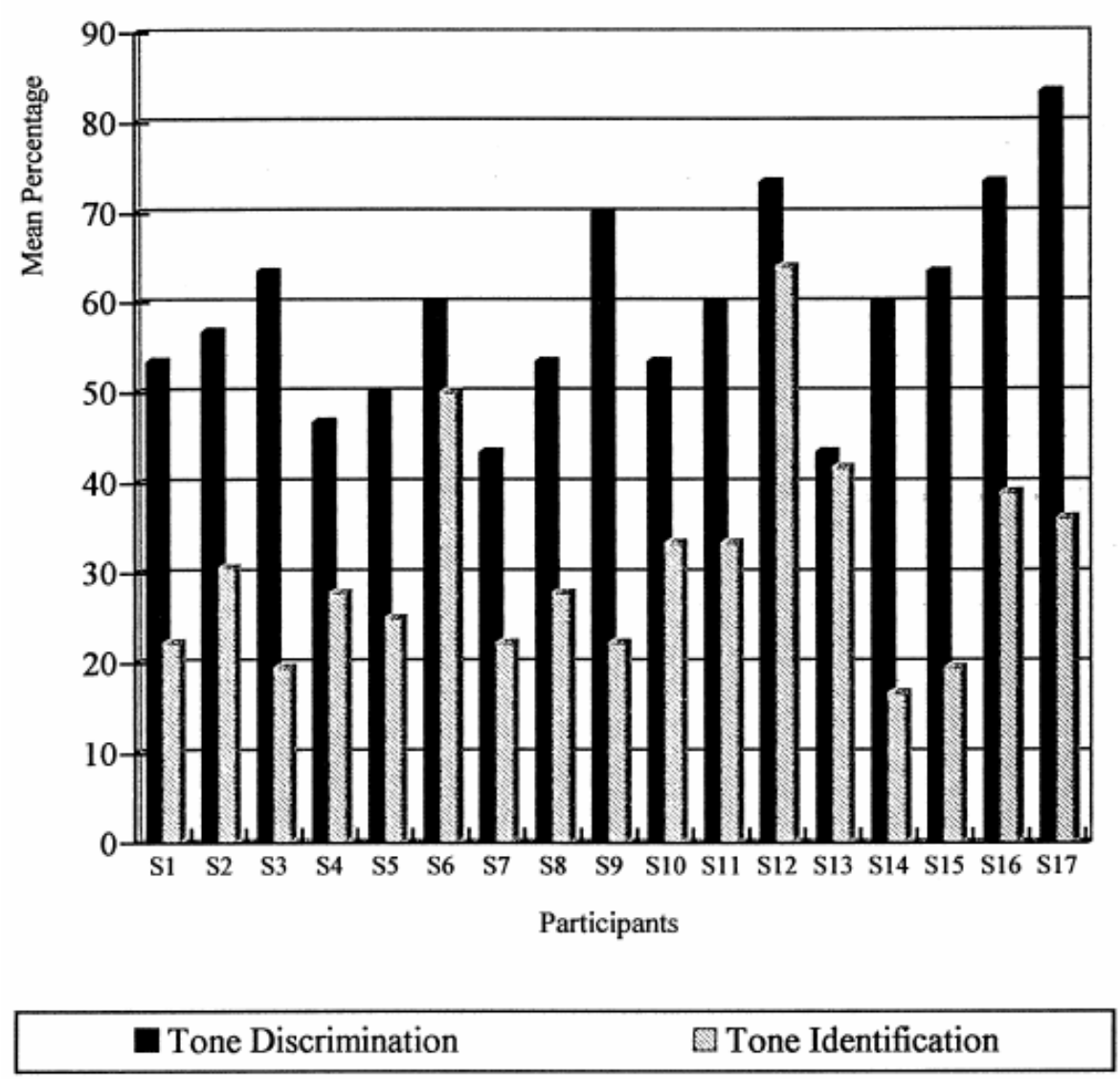


Figure 31. Cantonese lexical-tone perception in cochlear implant children. A same/different paradigm was used to test tone discrimination (whose chance level is 50%) and a four-alternative, force-choice method was used to test tone identification (whose chance level is 25%). Reprinted with permission from Fig. 2, Wong, A. O. and L. L. Wong (2004). "Tone perception of Cantonese-speaking prelingually hearing-impaired children with cochlear implants." Otolaryngol Head Neck Surg 130(6): 751-758.

Wei et al. further tested the effect of noise and pre- vs. post-lingually-deafened on lexical-tone perceptions (in Mandarin) [198]. The subjects were 7 pre-lingually-deafened cochlear implant users, 10 post-lingually-deafened cochlear implant users, and 5 normal listeners with a mean age of 22 years old (ranging from 10 to 49). The lexical-tone test was a four-alternative, forced-choice method [199] that the subjects were presented with a tone token in a random order and had to select the tone pattern (1, 2, 3, or 4) corresponding to the presented stimulus. The graphic interface presented both the phoneme-tone combinations and their corresponding Chinese characters. The average tone recognition of 17 cochlear implant users was 80% correct in quiet, which dropped to 55% at +10 dB signal-to-noise ratio and essentially chance performance at -5 dB signal-to-noise ratio. In comparison, the normal-hearing control subjects maintained essentially perfect performance over this signal-to-noise ratio range (at the ceiling even at -5 dB signal-to-noise ratio) [198] (Figure 32).

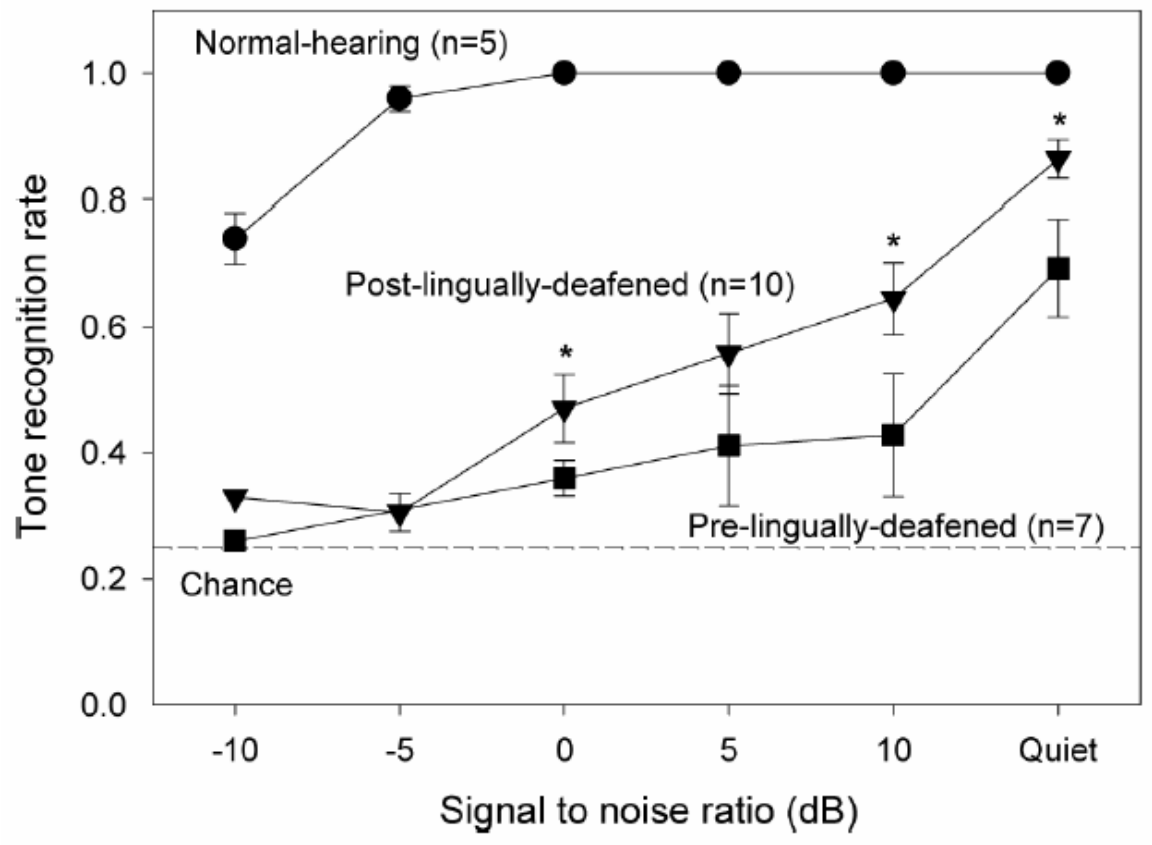


Figure 32. Mandarin tone recognition as a function of signal to noise ratio in cochlear implant users and normal listeners. Chance performance is 0.25. Asterisks represent a significant difference between groups at the tested conditions. Reprinted with permission from Fig. 3, Wei, C., K. Cao, et al. (2007). "Psychophysical performance and Mandarin tone recognition in noise by cochlear implant users." Ear Hear 28(2 Suppl): 62S-65S. Copyright 2007, Lippincott Williams.

Preliminary studies of the effect of newly developed and commercialized strategies, fine structure processing (FSP) [200] and HiRes 120 (current steering) [201], on lexical-tone perception both showed non-significant results [200, 201]. This suggests that the claimed temporal fine structure enhancement by zero-cross based technique and current steering failed to improve lexical-tone perception in cochlear implant users. (Please see the section below for technology details.)

In summary, current cochlear implant users encounter significant difficulties in speech recognition in noise, music perception and lexical tone perception. With adverse situations (such as decreasing speech to noise ratio to 0 dB or removing rhythm from the music), their performances can be as poor as 0% correct in open-set sentence recognition, or close-to-chance performance for music and lexical tone recognitions.

Studies suggested additional time information (such as rhythm) or additional phase-rich low-frequency acoustic hearing may improve speech recognition in noise and music perception. There seems to be no literature that provided direct evidence of the effect of phase information on these electric auditory perceptions. The key is the lack of a generally accepted manipulation of delivering vs. non-delivering phase information to cochlear implant users.

Future studies are necessary to develop a method which can at least deliver different amounts of phase information to cochlear implants. This method should not affect the other two kinds of fundamental information (frequency and intensity). Thus it may be achieved by somehow rearranging the stimulating time at each electrode precisely according to phase information. Or it may be achieved by simply recording the specific latency pattern in animal auditory nerve fibers to a stimulus, assuming human ears will do the same, and then delivering electric currents with or without the recorded pattern to test the effect on perception of the stimulus.

2.3.4 Current effort to preserve phase information

Two of the three principal manufacturers of implant systems recently introduced new processing strategies that are designed to increase the transmission of temporal fine structure information [202].

In one approach, the MED-EL company tried to present the phase-locked discharge pattern in auditory neurons by marking the zero-cross timing of the bandpass filtered signal to form “Channel Specific Sampling Sequences”. The timing of positive zero crossings in the bandpass filtered outputs with the lowest center frequency, or in the outputs of up to four bandpass filters with the lowest center frequencies, is "marked" with the presentation of a short pulse train for the corresponding channel and site of stimulation. The overall amplitude of the pulse trains for these special processing channels is determined by the magnitude of energy in the band for each channel, as in CIS. This strategy is called the fine structure processing (FSP) strategy [202]. In short, at each zero crossing in a channel's bandpass filter output, a series of stimulation pulses is started.

The starting time might reflect some phase information (how much and how accurate is unclear), but the pulse train reflects only the frequency allocation (or the corresponding characteristic frequency) without original phase information. Future studies are needed to develop an effective tool to measure and evaluate how much phase information is preserved and delivered by this “fine structure processing”.

Performances of the FSP users from preliminary reports varied from improvements in some speech and music test scores [203, 204] to no significant differences in speech perception [200] and channel number that is required in speech perception [205].

The other approach is to represent the fine frequency information within bands using a virtual channel technology termed “current steering”. For CIS and other derivatives, a single site is stimulated for each channel. The current steering technique creates multiple stimulating sites by “steering” the current with voltages of two adjacent electrodes [206-208]. This approach is a variation of HiRes (and CIS) and is called the HiRes with the Fidelity 120™ option (HiRes 120). It employs virtual channels as a way to increase the number of discriminable sites beyond the number of physical electrodes [202].

As reported by Koch et al., through simultaneous delivery of current to pairs of adjacent electrodes, the technique of “current steering” actually improves spectral resolution (place code) [208] instead of preserving phase information.

There has been growing interest in delivering temporal fine structure information [209-211] with similar zero-crossing hypothesis or other methods such as frequency modulation. Again, they seemed to fail to show evidence or extent of phase preservation and are not recognized by the principle implant manufacturers to commercialize.

2.3.5 Summary and discussion

Most of current cochlear implant processors use continuous interleaved sampling strategy or its derivatives. The only genuine phase information provided by current envelope-detection based processors is time-varying envelope information. However, envelope of a sinusoid is a straight line, which contains no phase information. Envelope and temporal fine structure of a complex sound together contain the three kinds of fundamental information (intensity, frequency and phase), but how much phase information is represented in envelope is unknown.

In a processing cycle, the envelope detection process replaces the variety of original phases by a non-informative constant offset. For example, implant processors replace the various patterns of first-spike latencies for different tone bursts by a constant temporal offset. Although the change between the normal phase dispersion and the constant temporal offset is on the scale of a few milliseconds, it was indicated that precision coarser than a few milliseconds significantly reduced the encoded information [12], and thus reduced the performance of implantees.

Current cochlear implant users encounter significant difficulties in speech recognition in noise, music perception and lexical tone perception. With adverse situations (such as decreasing speech to noise ratio to 0 dB or removing rhythm from the music), their performances can be as poor as 0% correct in open-set sentence recognition, or close to chance performance for music and lexical tone recognitions. Even using the claimed fine structure processing (FSP) strategy, the improvement of the performance in available literature was marginal.

Animal studies have shown that an auditory neuron responds to electric pulses of an intracochlear electrode with a short latency. When the interpulse interval is longer than the refractory period (i.e., stimulating rate is low), neural spikes in an auditory neuron can precisely represent the electric stimulating pattern. When the stimulating rate is high (e.g., > 1000 pps), neural spikes may only precisely represent onset and offset of stimulation. During a high-rate pulse train, an auditory neuron can sustain the maximal discharge rate (every refractory period) for tens of milliseconds in the onset stage. These together suggest that the stimulating pattern can be precisely replicated in auditory nerve fibers, at least for the timings of the first pulse and margins of the pulse trains.

Precise phase information in the stimulating pattern will be represented in auditory nerve fibers, and will be processed through the auditory pathway. The brain is a critical part for the outcomes of this prosthesis system. For patients with a compromised brain, it is unclear whether such inputs are optimal. For patients with a fully intact brain, as Wilson and Dorman [6] concluded, the bottom up (peripheral to central) approach to implant design probably is appropriate.

Scrutiny of the results in the literature leads to a compelling impression that the auditory system encodes timing and phase information well beyond what might have been expected, considering basic auditory neurons' limitations and extensive focus on the place code in the literature. A more comprehensive knowledge of just how well this encoding process works should contribute to a better understanding of such processes as speech perception and other complex signal processing ultimately by the brain. For such devices as cochlear implants, there also appears to be greater use of phase information possible than seems to have been considered to date, perhaps one of the keys to better performance in subsequent generations of these devices. It is useful to further explore the encoding of phase information in the auditory system and to determine if indeed there is a pervasive scheme by which, at least in effect, this is accomplished, perhaps virtually independently of species-specific specializations of acoustic input systems (like the mammalian cochlea and its intricate traveling-wave apparatus). To this end the following hypotheses were examined using the methods to follow.

2.4 HYPOTHESES

As concluded from the review of the literature, considerable evidence points to the notion that no species- or individual-independent sharing of common attributes explains the across-species similarity of latency patterns in auditory nerve fibers, except for the assertion that the hearing organ acts as a Fourier analyzer [29].

A way to encode phase information for all of the sinusoids in a Fourier analysis is to "mark" the first amplitude maximum for each sinusoidal component. The "markings" create a pattern among the sinusoids and is defined as a **Fourier pattern** (also see 3.2.1). This concept is unique to this work, intended to characterize a feature of temporal processing—specifically phase encoding—that is not readily apparent in more conventional analyses.

Because phase introduces a frequency-dependent time delay and the Fourier pattern reflects a frequency-dependent latency pattern, it is reasonable to view the Fourier pattern as manifest of a phase code (that is, manifest of how phase information is represented effectively in the nervous system). The demonstration of the salience of such a pattern in auditory neural activity, therefore, provides a test of the underlying system. If the results of that test are compellingly favorable, they may be taken as evidence that the auditory periphery is designed purposefully and specifically to be a phase analyzer.

2.4.1 Hypothesis 1: phase encoding among auditory nerve fibers

The hearing organ serves as a running analyzer whose output among auditory nerve fibers reflects synchronization of activity consistent with the Fourier pattern.

Although it is generally accepted that the hearing organ is described in the context of a Fourier analyzer, phase encoding tends to be an afterthought and tested only weakly by such tools as the period histogram of single-unit recordings. In contrast, the Fourier pattern, as defined, permits a critical and practical analytical method to test whether the hearing organ acts as a Fourier analyzer, as posited in this hypothesis. Namely, it is implied that the hearing organ is effectively a short-term Fourier analyzer. The underlying mechanisms effectively operate a Fourier transform and effectively mark "phase peaks" as described above. The output of this system manifests in the synchrony of the neural activity accordingly. If so, then a systematic pattern of activity appears as the "Fourier pattern". Otherwise, a comprehensive plot of markings, as used in this study, will appear simply as a noisy scatter plot.

This mathematical-like encoding of sound by the auditory periphery (Fourier analyzer, as described herein) is proposed to incorporate a mechanism of phase pattern capture that, in the face of species differences in peripheral auditory hardware, accounts for the considerable similarities across species in their latency-by-frequency functions, in turn assuring optimal phase encoding across species.

2.4.2 Hypothesis 2: phase encoding within an auditory nerve fiber

A single auditory nerve fiber expresses specific spike distribution consistent with the Fourier pattern created by, in effect, heavily overlapped Fourier analyses.

To appreciate the issue at hand, recall that the Fourier pattern is created among the sinusoid components. If the latency relationship *among* various auditory nerve fibers is

accounted for by the Fourier pattern, it must be asked, what accounts for the specific response latencies *within* a single auditory nerve?

The latency pattern in a single auditory nerve fiber to an identical stimulus is similar across species. Examples include the classical responses to single clicks (see 2.2.5) and the phase-locking phenomenon to single tone bursts (see 2.2.3). The specific spike distribution in a single auditory nerve fiber is the temporal code for the stimulating sound.

If the hearing organ operates as a Fourier analyzer per Hypothesis 1, a cycle of Fourier analysis generates a marking for each Fourier band. With multiple consecutive analyses, the marking distribution in a single Fourier band also will be a phase code at a given frequency.

It thus may be argued that, unless the encoder is operating to effect heavily overlapped Fourier analyses, it is unlikely that the hearing organ can accurately encode a sound's phase information that may vary all the time, hence Hypothesis 2.

The descriptor here, "heavily overlapped" is used to connote a process of moving-window Fourier analysis wherein the window is moved forward in a point-wise manner, i.e., calculating the transform repeatedly one point at a time while moving forward in time from the onset of the stimulus. The testing of this hypothesis thus goes beyond the focus of Hypothesis 1 on the latency relationship among various auditory nerve fibers (e.g, see 2.2.4 and 2.2.1) to explore more pervasive effects (e.g., see 2.2.5 and 2.2.3) of the putative Fourier-pattern analyzer within single units.

3.0 METHODS

3.1 OVERALL DESIGN OF STUDY AND RATIONALE

The hearing organ is hypothesized to act, in effect, as a Fourier analyzer whose output reflects synchronization of auditory neural activity consistent with the Fourier pattern (again, as defined herein). Consequently, each marking of the Fourier pattern predicts a latency counterpart in the peristimulus-time histogram of an auditory neuron whose characteristic frequency matches the Fourier band's frequency.

If this hypothesized mathematical-like mechanism indeed accounts for the across-species similarity of the latency pattern to one stimulus, the Fourier latency pattern should be close to, if not the same as, the measured latency pattern in auditory nerve fibers to that stimulus. The testing of the stated hypotheses presented a variety of technical and pragmatic issues for which the multimodal approach taken in this work was designed to address.

3.1.1 Method of study—a combined research approach

Since the hypothesis implies that the above-mentioned mechanism is stimulus-, neuron-, individual- and species-independent, the research design required a practical basis by which to study the latency patterns over a certain responding time via recordings in numerous auditory nerve fibers, ideally from different individual animals and species to multiple stimuli. This was

imperative to investigate the robustness of the hypothesis across different recording time sections, stimuli, neurons, individuals (and sexes), and species.

An in-depth empirical study meeting the study objectives would be extremely time-consuming, expensive (in both financial cost and cost of animal lives), and difficult, if not impossible in a single investigator's career. In addition, the similarity of neural latencies among different individuals and species has been shown in the literature (see 2.2.1, 2.2.3, 2.2.4 and 2.2.5) and thus does not require replication. Thus, the most efficacious approach is a meta-analysis wherein the predicted latency patterns are tested by effectively replicating numerous experiments of the previous well-conducted animal studies, each of which examines the neural latency pattern in numerous auditory nerve fibers of one or two representative species to at least one stimulus. Comparing the predicted latency and the counterpart latency recorded in various auditory nerve fibers of the representative animals to multiple stimuli thus permits statistical tests to reject or accept the hypotheses.

Because this approach describes and integrates multiple experiments, the goal was to describe the degree of relationship that exists between two variables, while manipulating (in effect) phase encoding, stimuli, number of Fourier analyses to test their effects on the predicted latency pattern. This research design represented a unique variation of meta-analysis, namely a correlational study, combined with an empirical research design, using mathematical modeling.

3.1.1.1 Meta-analysis

Meta-analysis is a term that was introduced by Glass to refer to the analysis of analyses, in the spirit of meta-evaluation [212]. When in an area where rarely a single study provides a complete answer to a research question, a number of studies must be conducted in order to obtain a sufficiently definitive answer. At some point, these studies need to be integrated and described.

Meta-analysis may get around the problems of subjective judgments, preferences and biases, because it uses a quantitative technique to analyze the results of the studies in question [213].

There are two main classes of meta-analysis techniques. One technique provides the answer of whether an overall significant result will emerge if the findings from numerous studies are combined in a statistical fashion [213]. For instance, meta-analysis can be used to compute the p values of three studies with non-significant results and three with significant results to yield one overall p value that would be representative of the combined effect of all six studies. The other meta-analysis technique deals with different estimates of effect sizes in various studies to come up with one overall effect size [213].

Thus, meta-analysis is usually a non-experimental quantitative research. It usually uses a quantitative technique to arrive at an overall effect size or significance. It is a descriptive type of research in which the goal is to attempt to give an accurate description or picture of a particular phenomenon with numbers.

However, the present study represents a unique variation of meta-analysis, namely a correlational study, combined with an empirical research design, using mathematical modeling. The variables of interest were manipulated (phase encoding, stimulus, and number of Fourier analyses) to test their effects on the predicted latency pattern, i.e., the effects of these treatments on the marking latency. The predicted latency patterns would then be compared with the descriptive numbers in numerous animal experiments using the same stimulus to determine the degree of relationship that exists between them. Therefore, here the descriptive approach of meta-analysis provides input regarding the effectiveness of the proposed hypotheses, as it is used to describe the status of a situation once a solution, suggested by experimental analysis, has been put into effect [213]. In fact, there is compelling precedence for this approach in this literature.

With numerous experimental analyses in hand, this kind of meta-analysis approach has been used by Ruggero and Temchin [22] to describe the status of across-species similarity in auditory-neural latency pattern to a rarefaction click and tone bursts, and by Shera [27] to describe the status of across-species similarity in frequency glides to a click.

3.1.1.2 Correlational study

A correlational study is a study that seeks to describe the degree of relationship that exists between two observed variables [213]. A correlational study can be incorporated into other quantitative research approaches. In this way, a descriptive index is computed that accurately and quantitatively presents the relationship between these two variables.

Pearson product-moment correlation was used in the present study (see 3.1.3). The Pearson correlation measures the degree and the direction of linear relationship between two variables. The weakness of the correlational approach is apparently the so-called third variable problem, which refers to the fact that two variables may be correlated not because they are causally related but because some third variable caused both of them. Nevertheless, even if there is a third-party variable that caused both the Fourier latency pattern and the recorded neural latency pattern, this variable must exist in the Fourier analysis by which the Fourier latency pattern is calculated. In this case, the cause of the across-species similarity is still a variable in the Fourier analysis, and probably also is a mathematical-like mechanism.

There are still other considerations for a correlational study. The correlation is related to the range of the data, i.e., a restricted range of data may produce incorrect correlation. Extreme data points (outliers) can dramatically impact the correlation coefficient. So the selected animal studies should have representative and adequate latency recordings. Scatter plots must be examined before interpreting Pearson correlation coefficient.

In addition to the direction and strength of the correlation coefficient, the coefficient of determination (r^2) also was calculated. Once the correlation coefficient was squared, this number indicated the proportion of variation in one of the variables that was accounted for by the other variable [213]. Therefore, r^2 showed what percentage of the variance in the neural latency pattern can be explained by the Fourier latency pattern.

3.1.1.3 Experimental study

Empirical research is a quantitative approach designed to discover the effects of presumed causes [213]. The key feature of this approach is that one variable is manipulated to see what happens to some other variable, or to examine the presumed causes. In the experimental research approach, one attempts to identify cause-and-effect relationships by deliberate control and manipulation.

The experimental approach has the disadvantage that the effects of nonmanipulated variables cannot be tested. It creates an artificial environment, and is frequently time consuming and difficult to design and carry out [213].

On the other hand, the experimental approach has at least two advantages. As mentioned above, the first one is that it permits researchers to infer causal relationships. However, it is easier to identify causal description, which describes the results of deliberately varying a treatment, than it is to achieve a causal explanation, which clarifies the mechanisms by which a causal relationship holds. A second advantage is that it controls for the influence of extraneous variables. Other advantages include that of producing lasting results, leading to new studies, and suggesting solutions to practical problems [213].

In summary, the present research is effectively a combination of experimental, correlational, and meta-analysis approaches. By the experimental research approach, the present study deliberately varied a treatment (such as marking randomly or by the Fourier pattern, manipulating the stimulus and number of Fourier analyses) to test their effects on the predicted latency pattern. By the correlational approach, the present study compared the predicted latency pattern to the recorded counterparts to attempt to clarify the mechanisms by which the causal relationship holds. The experimental and correlational procedures were carried out for various stimuli and different species to form with meta-analysis an integrated approach. The meta-analysis thus provided the backdrop for the implementation of the empirical and statistical methods, respectively, namely providing insights for forming and parameters by which to evaluate the hypotheses.

3.1.2 Material of study

The latency patterns in auditory nerve fibers were drawn from the literature, or by direct communication with other investigators. The predicted latency patterns were calculated in Matlab (see 3.2). Although it seems possible to replicate any animal studies that use various stimuli to test the hypothesis, there were several considerations for the material selection.

Since the latency pattern is on the scale of milliseconds, only studies providing adequate temporal resolution on that scale were considered. In general, the neural latency patterns were drawn from the selected studies, whose research design is adequately described and replicable, whose results of latencies can be easily and clearly identified for comparison (e.g., the regression curves or raw data are provided), whose measurement and description of the characteristic frequencies are reliable (since the latency is sensitive to the neuron's characteristic frequency),

whose sample size of auditory nerve fibers is adequate (i.e., covers a good frequency range), whose animal species is representative, and whose data collection method is appropriate and reliable (i.e., with valid measurement tools).

Because it may not be possible to obtain the original acoustic stimulus, studies that satisfy the above criteria were considered if they used simple signals such as clicks, or tone bursts. Using these simple signals reduces possible stimulus difference when reproducing the stimulus.

To test Hypothesis 1, two animal studies [25, 88] that comprised four experiments were selected because only these experiments best satisfied all of the above criteria. They provided original data with adequate temporal resolution among a large sample size of auditory nerve fibers. They clearly described how the starting phase of the tone bursts was controlled. The two species are representative, and the research design was adequately described and replicable.

Other animal studies that may partially meet the above criteria were not selected because of their similar results but smaller sample size or frequency range, which might cause a correlational study to be biased (see 3.1.1.2). Therefore, it is believed that replicating these animal studies would have been redundant due to the similarity or biased due to their small sample sizes.

Specifically, the predictions of latency in numerous auditory nerve fibers were tested virtually in numerous experiments, including the first-spike latency pattern (regression and raw data) to a rarefaction click in chinchillas [88] (that is a representative species in Figure 22), the first-spike latency pattern (regression and raw data) to a condensation click in chinchillas [88], the first-spike latency pattern (regression and raw data for both species) to various tone bursts in

chinchilla [88] (that is a representative species in Figure 4) and barn owl [25] (that is a representative species with significant phase locking at high frequencies up to 10 kHz [25]).

To test Hypothesis 2, the generalized conclusion of classical click responses (see 2.2.5) in single auditory nerve fibers was used to compare to the predicted latency pattern.

There seems to be no study that precisely reports the latencies to single tone bursts in auditory nerve fibers that have a wide range of characteristic frequencies. The generalized conclusion that a single auditory neuron tends to discharge at phase-locked timing to a tone burst (see 2.2.3) was used as a rough reference (for Hypothesis 2) while inspecting the combined Fourier pattern of various tone bursts.

The exact variables for individual experiment to test the hypotheses are detailed in the sections below.

3.1.3 Statistical design

The exact variables to test Hypothesis 1 included raw data or regression curves provided by the selected animal studies. The reported characteristic frequencies of the auditory nerve fibers were not linearly distributed as the band frequencies in a Fourier analysis. In order to make comparisons at each identical frequency, if a regression curve in the animal study was used, the neural latency pattern was defined as the points on the regression curve when using the Fourier band frequencies as the characteristic frequencies. If the individual points of the raw data in the animal study were used as the neural latency pattern, the predicted latency pattern was calculated as the points on the regression curve of the Fourier pattern when using the neurons' characteristic frequencies as the Fourier band frequencies.

In other words, there are always two latencies for each frequency. As mentioned above, if the hypothesized mathematical-like mechanism accounts for the across-species similarity of the latency pattern to one stimulus, the Fourier latency pattern should be close to, if not the same as, the measured latency pattern in auditory nerve fibers to that stimulus. In this case, the relationship between the pairs of measured versus predicted latencies should be linear. Thus, the Pearson correlation coefficient was calculated and statistically tested to show the strength and direction of the correlation between the measured and predicted latencies.

The statistical null hypothesis for testing the correlation coefficient was $H_0 : \rho = 0$ (there was no population correlation). A two-tailed t-test with $\alpha = 0.05$ was used to test whether or not a correlation exists in the population: $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$, where n is the number of pairs of latencies.

When neural latency pattern is demonstrated by latencies of the peaks in some peristimulus-time histograms with a generalized latency conclusion (e.g., the classical click response; see 2.2.5), the predicted latency pattern is examined by inspecting whether its exact numbers fit the conclusion.

3.2 CALCULATING THE FOURIER PATTERN

3.2.1 Latency pattern in a single Fourier analysis

As mentioned above, the Fourier pattern in a single analysis is created by marking the first amplitude maximum for each sinusoid component. Figure 33 illustrates how the Fourier pattern in a single analysis window is calculated by the following steps.

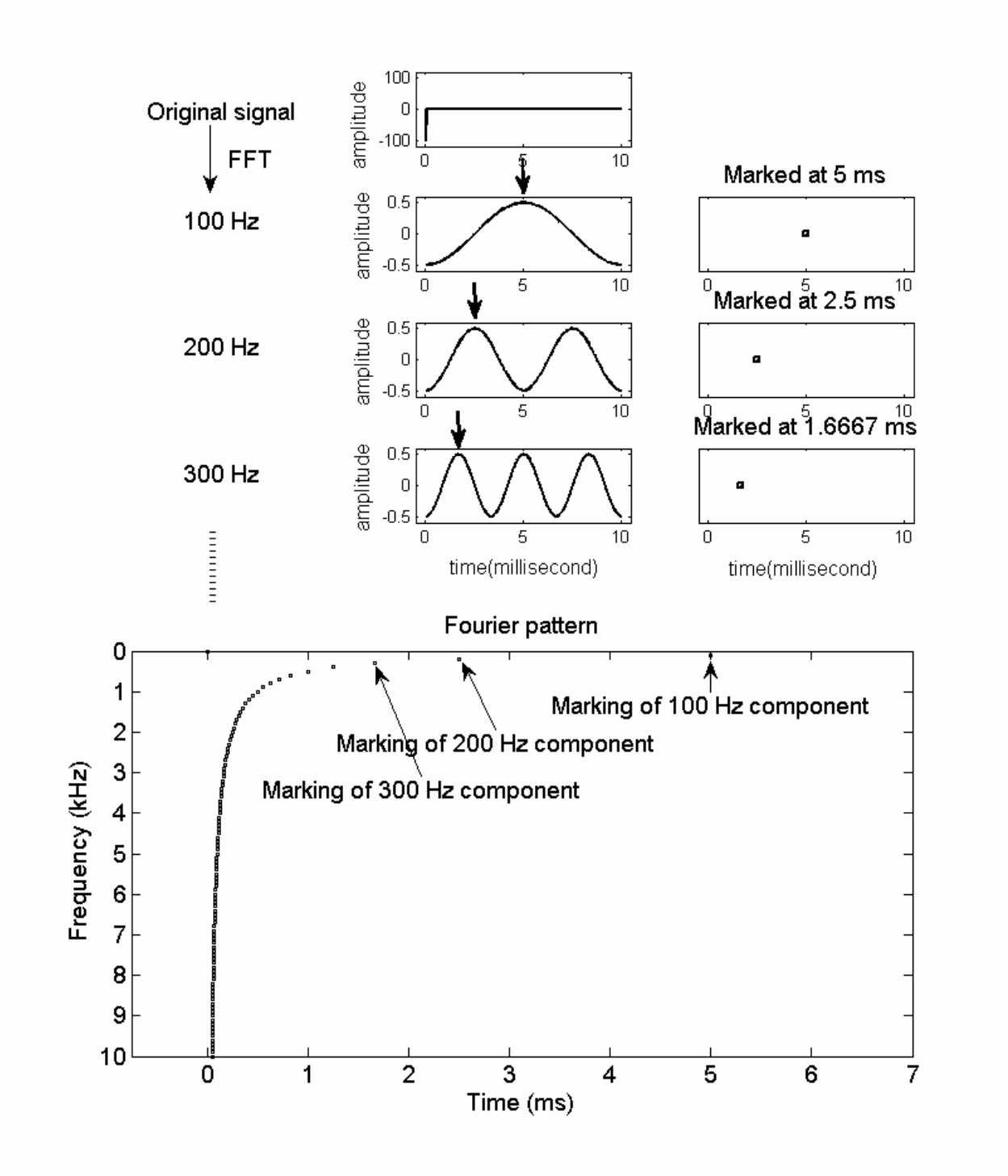


Figure 33. Fourier pattern of a rarefaction click. The sampling rate is 40000 Hz. A 400-point fast Fourier transform (FFT) is calculated to decompose the rarefaction click into sinusoid components (among which the first three are shown). For clarity, only the markings of components between 100 and 10000 Hz are shown.

1. $\boxed{x(n) \qquad 0 \le n \le N-1}$

Sample a signal at a rate of f_s Hz for N points to yield $x(n)$ in an N-point analysis window. The top panel of Figure 33 shows a rarefaction click ($n = 0, x = -100$; otherwise , $x = 0$), with f_s = 40000 Hz and N = 400 (chosen arbitrarily for demonstration).

2. $\boxed{x(n) \xrightarrow{\text{N-point FFT}} X(k) \qquad 0 \le k \le N-1}$

Calculate one N-point fast Fourier transform (FFT).

In Figure 33, the 400-point fast Fourier transform yields components every 100 Hz. Only the markings of components between 100 and 10000 Hz are illustrated to keep clarity.

The marking latency can then be calculated according to Fourier phase (the angle of a frequency term, a complex number, to the real axis in its polar expression). Also, it can be calculated by the phase of a sinusoid that is synthesized from that Fourier component. Fourier phase is used in the next step to skip unnecessary conversion and to simplify the process. The connection between Fourier phase and cosine phase (the starting phase of a waveform generated by cosine function) of the synthesized temporal signal is detailed in Appendix A.

3. $\boxed{d_k = \dfrac{2\pi - \angle X(k)}{\omega_k} \qquad 0 \le \angle X(k) < 2\pi}$

Calculate the latency of each marking d_k (seconds) by time-shift property of Fourier transform, according to the phase difference between the hypothesized reference 2π and the Fourier phase $\angle X(k)$. ω_k is the angular frequency of the band. Selection of the hypothesized reference is detailed in Appendix A.

For instance in Figure 33, all of the frequency terms equal -100. The Fourier phase, $\angle(-100)$, equals π in all components of the rarefaction click. The marking latency of the 100-

Hz component (d_1) equals $\frac{2\pi - \pi}{2\pi \cdot 100} = 0.005$ seconds (or 5 ms). The marking latency of the 200-Hz component (d_2) equals $\frac{2\pi - \pi}{2\pi \cdot 200} = 0.0025$ seconds (or 2.5 ms). The marking latency of the 300-Hz component (d_3) equals $\frac{2\pi - \pi}{2\pi \cdot 300} = 0.00167$ seconds (or 1.67 ms). Figure 33 demonstrates an example of the results for a single window.

All calculations were undertaken in Matlab 7.12.0.635 (R2011a) MathWorks (Natick, MA).

3.2.2 Latency pattern in running Fourier analyses

If a Fourier analyzer operates all the time, it continues to generate markings for all sinusoid components. The neural latency pattern is therefore hypothesized as the combination of the markings in all heavily overlapped windows.

Calculating the latency of a marking in one of the running Fourier analyses requires a reference time. In previous animal studies, the latency is the time difference between the onset of the stimulus and the corresponding action potential detected, subtracted by the acoustic transmission to the ear drum (see 2.2). If the action potential starts to be discharged at the same moment the stimulus starts to vibrate the eardrum, the latency is zero. Therefore, the moment when the stimulus just starts to vibrate the eardrum was used as the reference time in previous animal studies.

In the same manner, the time when the first point of the stimulus enters the window (as window 0 in Figure 34) is defined as the reference time in running Fourier analyses. Each bold bar in Figure 34 indicates one window that is a single Fourier analysis described in the previous section (3.2.1).

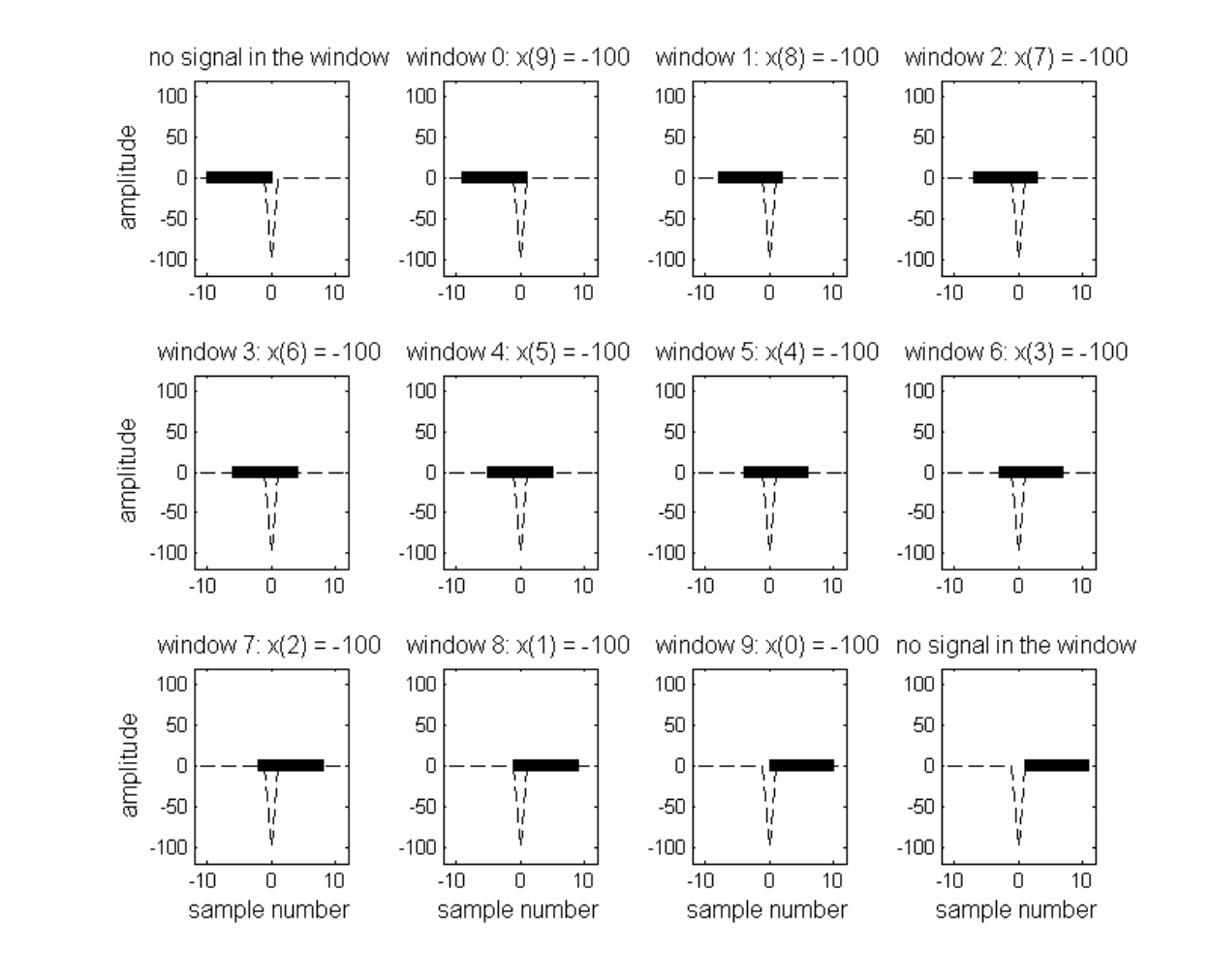


Figure 34. Time relationship between the stimulus and the window of the analyzer. The time when the first point of the stimulus enters the window (i.e., the time when window 0 occurs) is defined as the reference time in running Fourier analyses. In this example, the stimulus is a rarefaction click in which only one point equals -100. Window size: 10 points; Sampling rate: 40000 Hz (chosen arbitrarily for demonstration).

The relationship between the stimulus and the window of the analyzer yields different individual input signals. For example, Figure 35 shows the 400 individual input signals of the rarefaction click, while the 400-point window is moved forward in a point-wise manner. Above

each panel shows the window number and the time difference of how many milliseconds the window is away from the reference, window 0.

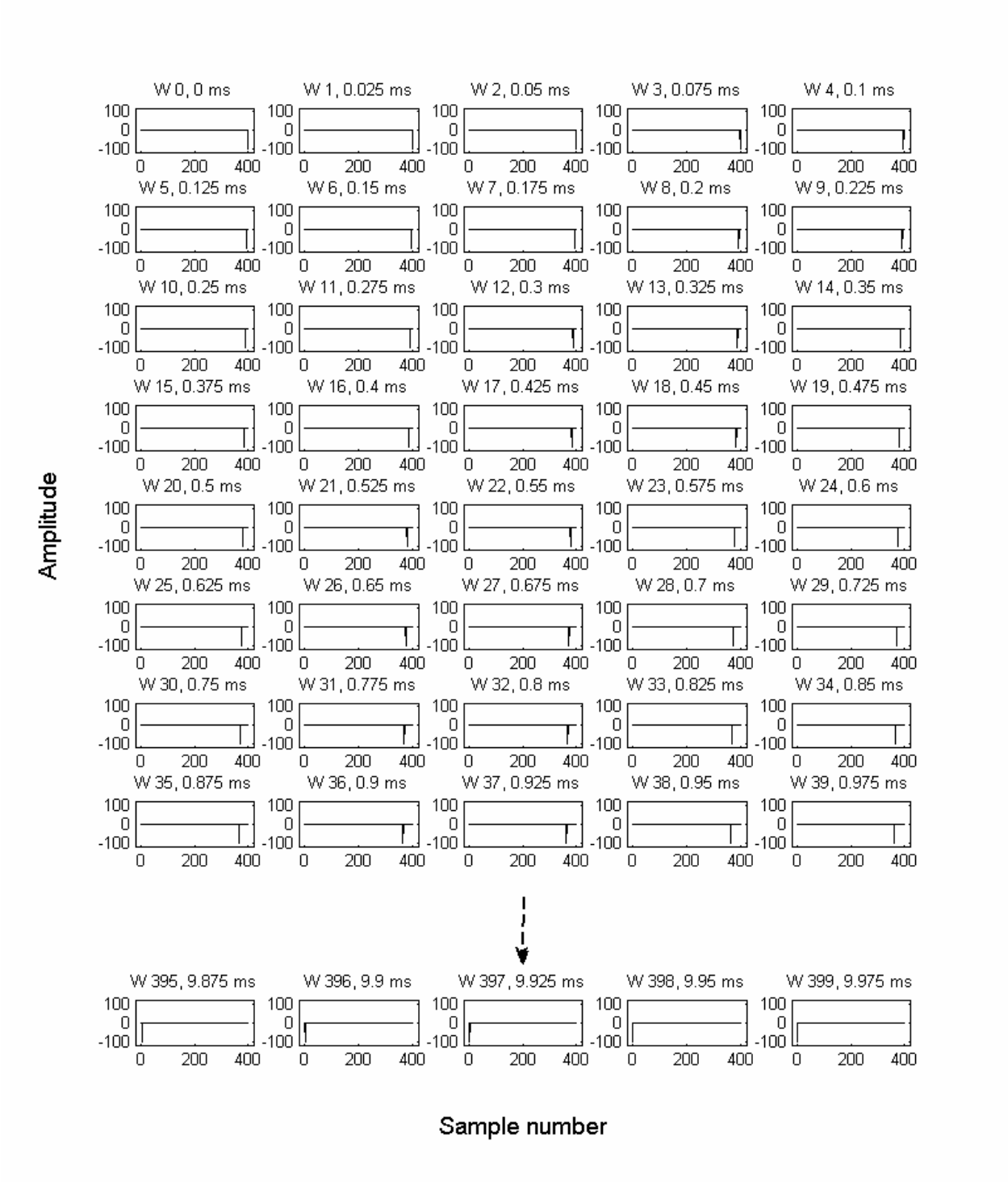


Figure 35. The individual input signals of the rarefaction click in moving windows. Above each panel shows the window number and the time difference of how many milliseconds the window is away from window 0 in the point-wise manner. Sampling rate: 40kHz. Window size: 400 points. The rarefaction click has a single non-zero point of -100. While the window is moved forward in the point-wise manner and the sampling rate is given, the time difference can be calculated for the predictions.

Computing the latency of each marking in an analyzing window is completed by the following step:

4. $\boxed{D(k,m) = d_k + t_m \qquad 0 \le m < N-1}$

Calculate the latency $D(k,m)$ of each marking in window m by adding the time difference t_m between window m and window 0, which equals $\frac{m-0}{f_s} = \frac{m}{f_s}$ (seconds), to d_k (seconds).

The time difference between window m and window 0 is determined by sampling rate and how the windows overlap. Because the window is moved forward in a point-wise manner, window m has m new points, comparing with those in window 0. Thus, window m is $\frac{m-0}{f_s} = \frac{m}{f_s}$ seconds away from window 0. The time difference can also be calculated by the location difference of a given stimulus point in these two windows, if it is in both of these windows. For example, the click is at sample number 400 and 300 in windows 0 and 100, respectively. The time difference equals $\frac{400-300}{f_s} = \frac{100}{40000}$ seconds.

The two examples below demonstrate the calculations in step 4.

Figure 36 shows results of the single analysis in window 0 (the first and the reference window). The marking latency is calculated from the Fourier phase $\angle X(k)$, the angular frequency of the band (ω_k), the window number m, and the sampling frequency f_s. According to steps 3 and 4, the marking latency of the 100-Hz component is :

$$D(k,m) = \frac{2\pi - \angle X(k)}{\omega_k} + \frac{m}{f_s}$$

Because the 100-Hz component ($k = 1$) is in window 0 ($m = 0$) and $X(k) = X(1) =$ -99.9877 - 1.5707i, the equation becomes:

$$D(1,0) = \frac{2\pi - \angle(-99.9877 - 1.5707\text{i})}{2\pi \cdot 100} + \frac{0}{40000}$$

$$= \frac{2\pi - 3.1573}{2\pi \cdot 100} = 0.004975 \text{ seconds (or 4.975 ms)}.$$ In the same manner, the marking latency of the 200-Hz component is:

$$D(2,0) = \frac{2\pi - \angle(-99.9507 - 3.1411\text{i})}{2\pi \cdot 200} + \frac{0}{40000}$$

$$= \frac{2\pi - 3.1730}{2\pi \cdot 200} = 0.002475 \text{ seconds (or 2.475 ms)}.$$ The marking latency of the 300-Hz component is:

$$D(3,0) = \frac{2\pi - \angle(-99.8890 - 4.7106\text{i})}{2\pi \cdot 300} + \frac{0}{40000}$$

$$= \frac{2\pi - 3.1887}{2\pi \cdot 300} = 0.0016417 \text{ seconds (or 1.6417 ms)}.$$

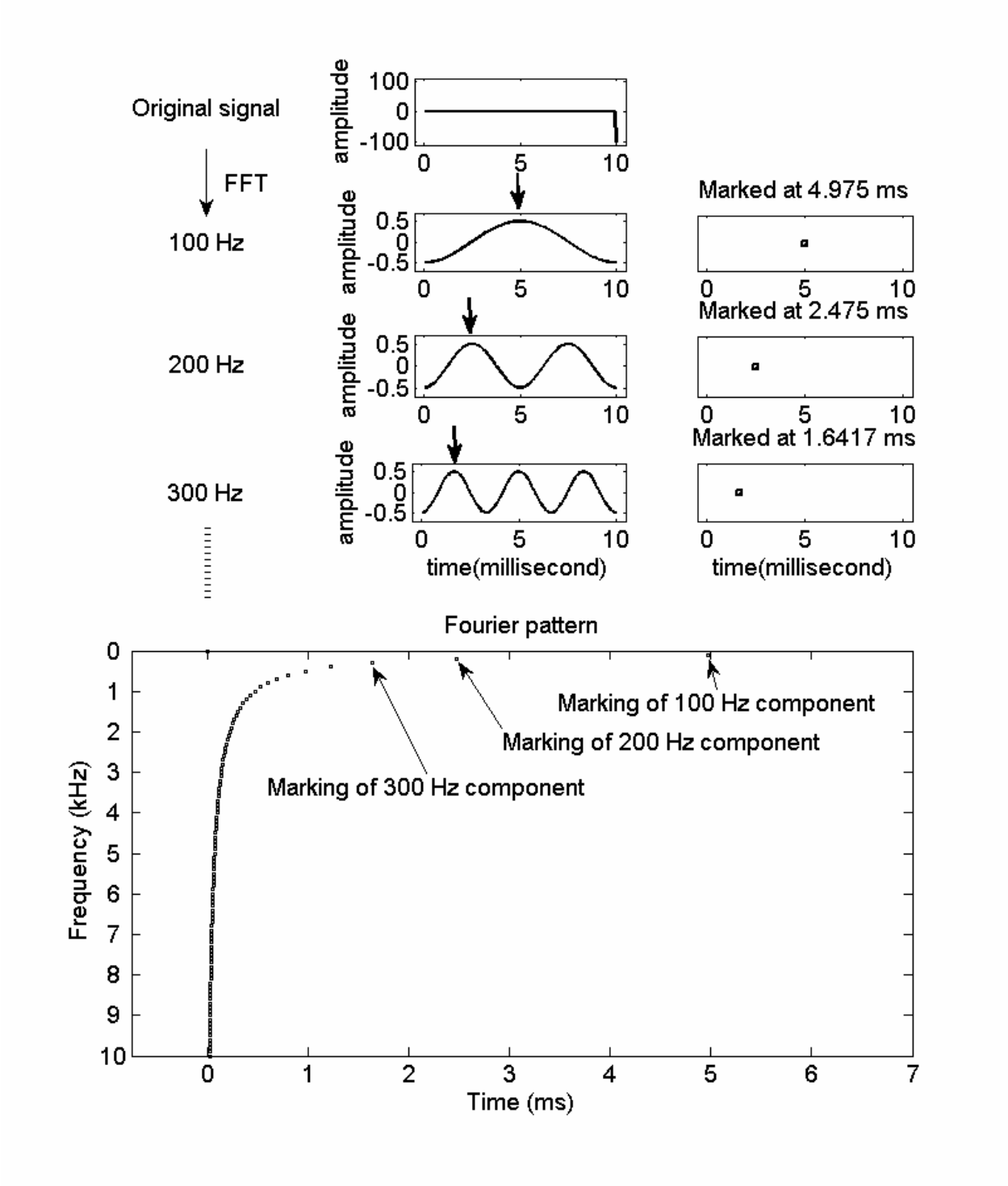


Figure 36. Fourier pattern in window 0 of the rarefaction click. Input x(n): x(399) = -100, otherwise x(n) = 0, 0 ≦ n ≦ 399.

For another example, Figure 37 shows results of the single analysis in window 30 (chosen arbitrarily for demonstration). According to steps 3 and 4, the marking latency of the 100-Hz component is:

$$D(k,m) = \frac{2\pi - \angle X(k)}{\omega_k} + \frac{m}{f_s}$$

Because the 100-Hz component ($k = 1$) is in window 30 ($m = 30$) and $X(k) = X(1) =$ -88.3766 -46.7930i, the equation becomes:

$$D(1,30) = \frac{2\pi - \angle(-88.3766 - 46.7930i)}{2\pi \cdot 100} + \frac{30}{40000}$$

$$= \frac{2\pi - 3.6285}{2\pi \cdot 100} + \frac{30}{40000} = 0.004225 + 0.00075 = 0.004975 \text{ seconds (or 4.975 ms). In the}$$

same way, the marking latency of the 200-Hz component is:

$$D(2,30) = \frac{2\pi - \angle(-56.2083 - 82.7081i)}{2\pi \cdot 200} + \frac{30}{40000}$$

$$= \frac{2\pi - 4.1155}{2\pi \cdot 200} + \frac{30}{40000} = 0.001725 + 0.00075 = 0.002475 \text{ seconds (or 2.475 ms). The}$$

marking latency of the 300-Hz component is:

$$D(3,30) = \frac{2\pi - \angle(-10.9734 - 99.3961i)}{2\pi \cdot 300} + \frac{30}{40000}$$

$$= \frac{2\pi - 4.6024}{2\pi \cdot 300} + \frac{30}{40000} = 0.0008917 + 0.00075 = 0.0016417 \text{ seconds (or 1.6417 ms).}$$

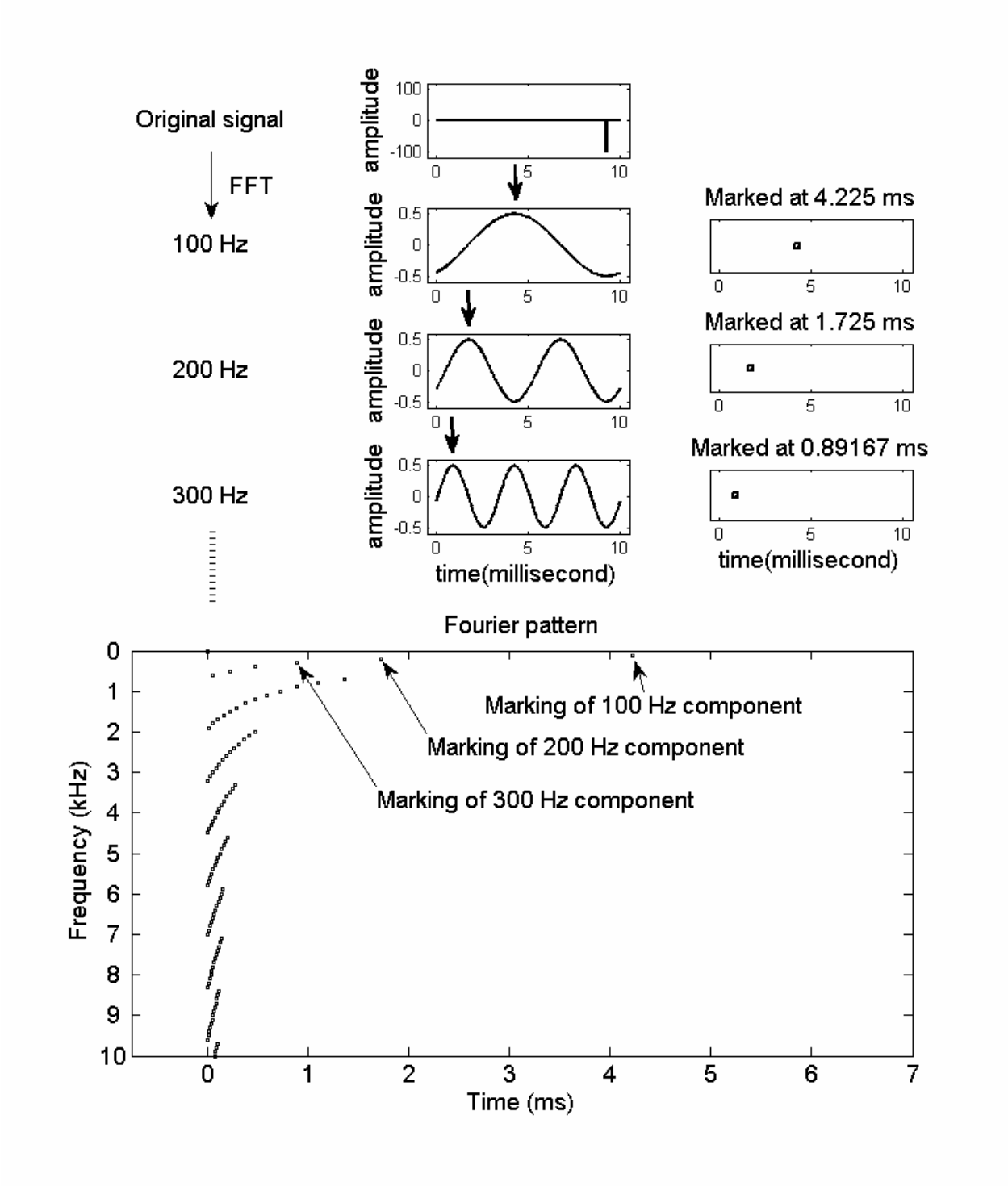


Figure 37. Fourier pattern in window 30 of the rarefaction click. Input x(n): x(369) = -100, otherwise x(n) = 0, 0 ≦ n ≦ 399. The marking latencies are results of a single window. The time difference between

window 0 and 30 needs to be added to the results. For example, the final latency of the 100-Hz marking is 0.004225+30/40000 = 0.004975 seconds (or 4.975 ms).

5. Combine all the markings by plotting them together, window by window, on a single figure to form the final latency pattern.

The markings that are from different windows but have the same latency $D(k,m)$ will superimpose. For example, the markings for the 100-Hz component from window 0 and 30 superimpose, so as those for the 200- and 300-Hz components from these two windows, because $D(1,0) = D(1,30) = 4.975$ ms, $D(2,0) = D(2,30) = 2.475$ ms, and $D(3,0) = D(3,30) = 1.6417$ ms.

3.2.3 Figure orientation and axis scale

In the literature, various plotting methods such as different orientations and axis scales were used. Figure 38 reveals the Fourier pattern (original vs. 1 ms added) of a *single* window demonstrated in Figure 33. The orientation and axis scale used in Figure 38A is the same as those in Figure 33. The figure also shows the Fourier pattern with additional 1 ms (the synaptic and neural transmission; see 2.2.7). To compare to previous animal studies and to improve clarity, the same plotting method in Figure 38B (rotated) and C (logarithmic frequency) are used throughout Results and Discussion.

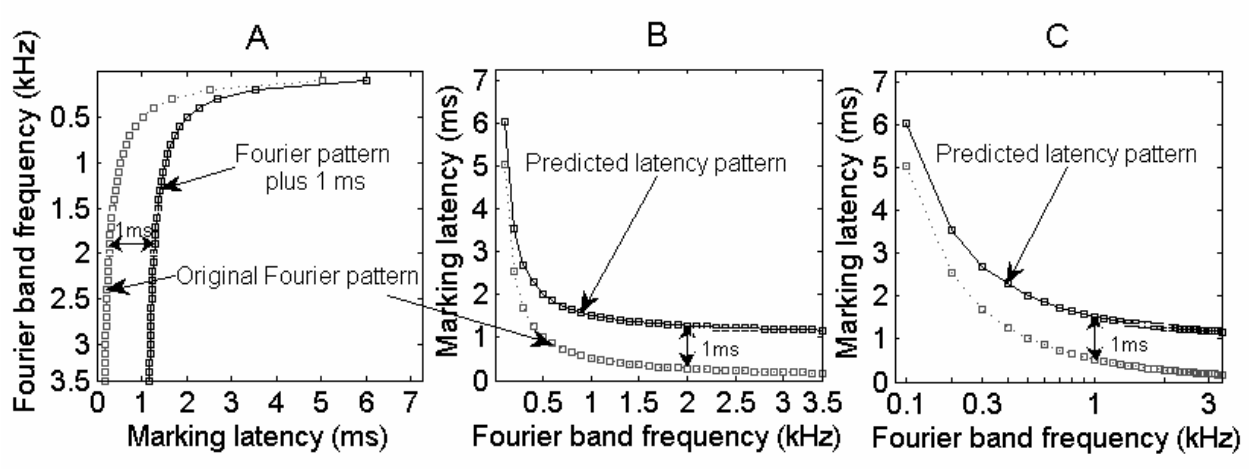


Figure 38. Orientation and axis scale. Gray markings: the original Fourier pattern of the rarefaction click. Black markings: the Fourier pattern with additional 1 ms. A. Plot oriented to the presentation used by Loeb, 2005 and in the section 3.2.1, in which the horizontal axis is time. B. Rotated plot to orientation with linear frequency on the horizontal axis. C. Similar plot to B. with logarithmic frequency and is usually used in animal studies of first-spike latencies. Plotting methods of B. and C. are used throughout the Results and Discussion.

3.2.4 Sampling rate and window size considerations

Sampling rate determines the frequency range of the Fourier bands, and also temporal resolution of the processing system. Window size, together with sampling rate, determines the frequency resolution. The fundamental frequency of a Fourier analysis is calculated by $\frac{\text{sampling rate (Hz)}}{\text{window size (points)}}$. The frequency of one Fourier band is a multiple of the fundamental frequency.

Sampling rate and window size should provide appropriate frequency range and resolution for comparison or demonstration, according to the data provided by the selected

animal studies. The characteristic frequencies of the auditory neurons in these animal studies range from about 0.3 to 10 kHz [25] and about 0.1 to 20 kHz [88].

The limitations of using any arbitrary sampling rate and window size include unmatched numbers between the present study and animal studies, due to the absence of information about these two parameters in animal hearing organs. The limitations also include not being to capture proper frequency or temporal changes within the set resolution.

To provide frequencies consistent with the characteristic frequencies of the auditory neurons reported in the selected animal studies and to clearly demonstrate the relationship between markings with adequate resolution, the sampling rate and window size were set to 40,000 Hz and 400 points (see 3.3.1.2, 3.3.2.2, and 3.4.1.2). These settings provide a frequency range of 0.1 to 20 kHz, which is consistent with the characteristic frequencies of the auditory neurons and the tested range. Although the provided resolution of 0.1 kHz may not be the resolution of the auditory system, this resolution is believed to be adequate to demonstrate the relationship of markings and the principles of encoding suggested in the present study.

3.3 TESTING PREDICTIONS OF FIRST-SPIKE LATENCY

Since there is no study that satisfies the selection criteria and reports the whole latency pattern, the predictions of latency in numerous auditory nerve fibers were evaluated in two main groups of experiments from animal studies. The first group of studies yielded reports of first-spike latency pattern in auditory nerve fibers to different stimuli. They included the first-spike latency pattern to a rarefaction click in chinchillas, latency pattern to a condensation click in chinchillas, and latency pattern to various tone bursts in chinchillas and barn owls. The second group of

experiments reported the latency pattern over a certain period of time after the first spikes (see 3.4).

3.3.1 Single rarefaction or condensation click

The animal study selected for testing the predicted latency pattern to both rarefaction and condensation clicks was conducted by Temchin et al [88]. The first-spike latency pattern to a click in this study is obtained from the peristimulus-time histograms of responses in 420 auditory nerve fibers of 29 chinchillas to an intense and repeatedly presented click. The latency, measured by the equipment whose temporal resolution was 1 microsecond, is the direct measurement of the time difference between the click onset and the action potential detected, subtracted by the acoustic transmission to the ear drum. Chinchilla is a representative species for the across-species similarity [22] to a rarefaction click in Figure 22.

3.3.1.1 Latency pattern in animal auditory nerve fibers

The neural latency patterns include the individual points of the raw data and regression curves. The raw data were 411 and 407 latencies of 420 auditory nerve fibers recorded in 29 chinchillas to a rarefaction click and a condensation click, respectively (Figure 39). The raw data were obtained by personal communication with the authors. The regression curve they reported to a rarefaction click is: $\tau\,(\text{ms}) = 1.244 + 2.402 \cdot e^{-1.699CF} + 0.083 \cdot e^{-0.269CF}$, where CF is characteristic frequency and $e = 2.7183$. The regression curve they used to a condensation click is: $\tau\,(\text{ms}) = 1.244 + 2.402 \cdot e^{-1.699CF} + 0.083 \cdot e^{-0.269CF} + 0.5 \cdot \frac{1000}{CF}$.

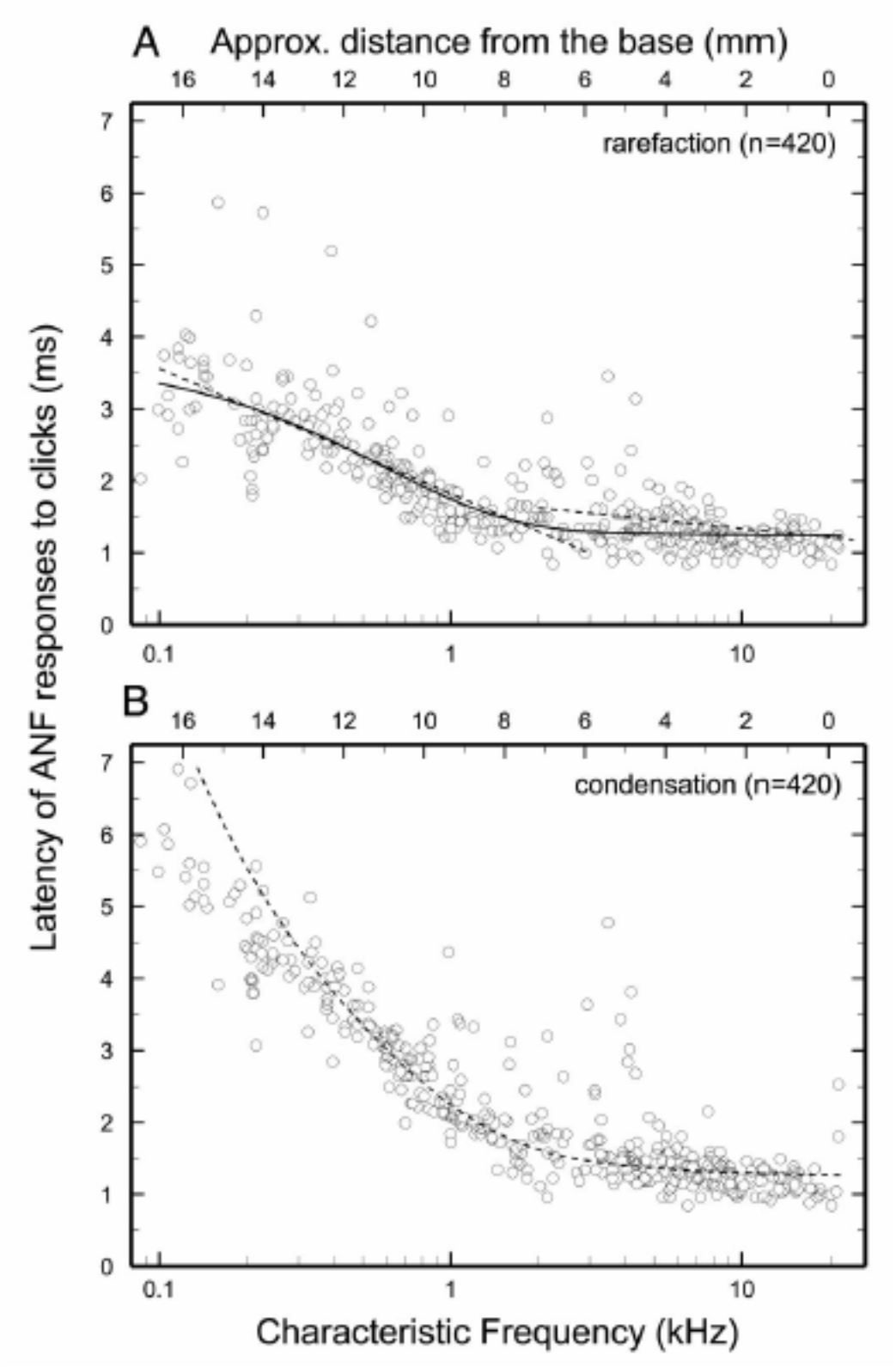


Figure 39. Neural latency patterns of first-spike responses to clicks. A. Latencies of 420 auditory nerve fibers recorded in 29 chinchillas to a rarefaction click. B. To a condensation click. Reprinted with

permission from Fig. 10, Temchin, A. N., A. Recio-Spinoso, et al. (2005). "Wiener kernels of chinchilla auditory-nerve fibers: verification using responses to tones, clicks, and noise and comparison with basilar-membrane vibrations." J Neurophysiol 93(6): 3635-3648.

As mentioned above (3.1.3), when the regression curve in the animal study was used, the neural latency pattern was defined as the points on the regression curve when using the Fourier band frequencies as the characteristic frequencies. Because the selected animal study does not show data below 0.1 kHz or above 20 kHz, a total of 98 Fourier band frequencies ranging from 0.2 to 9.9 kHz every 0.1 kHz were used to test the correlation with the regression curve (also see 3.3.1.2).

3.3.1.2 Latency pattern in Fourier analysis

The sampling rate and window size were set to 40,000 Hz and 400 points, to provide a frequency range of 0.1 to 20 kHz, the same range of the characteristic frequencies of the auditory neurons, with a resolution of 0.1 kHz (see 3.2.4).

The predicted latency pattern to a rarefaction or condensation click was calculated as the latencies of the first markings of the combined Fourier pattern (see 3.2) plus 1 ms (the synaptic and neural transmission; see 2.2.7).

The rarefaction click was generated as the signal: $x(n) = -100, \text{when } n = 0; x(n) = 0, \text{when } n \neq 0$. And the condensation click was generated as the signal: $x(n) = 100, \text{when } n = 0; x(n) = 0, \text{when } n \neq 0$.

A preliminary study of multiple-window analyses was undertaken, in which the windows overlapped every point. The result showed that, as expected (given the duration of the click signal), that all the first markings to either the rarefaction click or the condensation click are

from the first window (window 0, see 3.2.2). Therefore, the predicted latency pattern to a rarefaction or condensation click was obtained from the first window.

When the individual points of the raw data in the animal study were used as the neural latency pattern, a total of 411 and 407 neurons' characteristic frequencies were used to calculate the predicted latencies for the rarefaction and condensation clicks, respectively.

3.3.1.3 Markings with and without the Fourier pattern to a rarefaction click

The predicted latencies to a rarefaction click ***with*** the Fourier pattern were computed as described above. In order to investigate if the predicted latency pattern can be formed by chance, latencies ***without*** the Fourier pattern were calculated by randomly marking at an average rate of 60 Hz for 10 repetitions, to simulate the spike latencies in auditory nerve fibers of 60-Hz spontaneous rate (a physiologically high spontaneous rate [214], chosen arbitrarily for demonstration). Otherwise, all conditions were the same as those for calculating the predicted latencies with the Fourier pattern.

3.3.1.4 Correlation between Fourier and response latencies to a rarefaction click

The Pearson correlation coefficient was calculated and statistically tested to show the strength and direction of the correlation between the measured and predicted latencies to a rarefaction click. The coefficient of determination (r^2) was also calculated to show what percentage of the variance in the neural latency pattern can be explained by the Fourier latency pattern. When using the points on the regression curve as the neural latency pattern, n was 98. And when using the individual points of the raw data as the neural latency pattern, n was 411 (see 3.1.3).

3.3.1.5 Markings with and without the Fourier pattern to a condensation click

The same experimental procedures as for the rarefaction click (3.3.1.3) were undertaken except the *condensation* click was used as the input signal.

3.3.1.6 Correlation between Fourier and response latencies to a condensation click

The same statistical tests for the rarefaction click (3.3.1.4) were carried out, except the regression curve and raw data to a *condensation* click were used as the measured latency pattern. Also, when using the points on the regression curve as the neural latency pattern, n was 98. And when using the individual points of the raw data as the neural latency pattern, n was 407 (see 3.1.3).

3.3.2 Various tone bursts

Two animal studies were selected to test the predicted latency pattern to various tone bursts: one in chinchillas (Temchin et al. [88]) and the other in barn owls (Koppl [25]). Chinchilla is a representative species for the across-species similarity to tone bursts [22] in Figure 4, and was selected to represent the latency pattern at low frequencies (0.1 to 3 kHz). Barn owl is a species that demonstrates significant phase locking at high frequencies up to 10 kHz [25].

The mean first-spike latency of an auditory unit to tone bursts in these animal studies was calculated from the slope of regression fit of the unit's phase-versus-frequency function, which was derived from the mean phase in period histogram to various tone bursts around the neuron's characteristic frequency (see 2.2.1). Therefore, strong phase locking in period histograms is important and the latency pattern of barn owls was selected to represent the similarity of first-spike latencies at high frequencies (up to 9 kHz).

3.3.2.1 Latency pattern in animal auditory nerve fibers

The neural latency patterns included the individual points of the raw data and the regression curves in both chinchillas and barn owls. The raw data in chinchillas were 40 latencies of 40 auditory nerve fibers whose characteristic frequencies ranged from 100 to 2934 Hz. The raw data in barn owls were 40 latencies of 40 auditory nerve fibers whose characteristic frequencies ranged from 355 to 8895 Hz. The two sets of raw data were obtained by personal communication with the authors of the two animal studies. Although no regression curve was provided in the study in chinchillas, the power fit regression curve was calculated according to the raw data the authors provided. The regression curve in barn owls to tone bursts was: $\tau\,(\text{ms}) = 0.833 + 205 \cdot CF^{-0.706}$ (where CF is in Hz).

As mentioned above (3.1.3), when the regression curve in the animal study was used, the neural latency pattern was defined as the points on the regression curve when using the Fourier band frequencies as the characteristic frequencies. Due to the characteristic-frequency range of the raw data, a total of 27 Fourier band frequencies ranging from 0.2 to 2.8 kHz every 0.1 kHz were used to test the correlation with the chinchillas' regression curve, and a total of 71 Fourier band frequencies ranging from 1 to 8 kHz every 0.1 kHz were used to test the correlation with the barn owls' regression curve (also see 3.3.2.2), because there was only one datum below 1 kHz.

3.3.2.2 Latency pattern in Fourier analysis

The sampling rate and window size were set to 40,000 Hz and 400 points, respectively, to provide the frequency resolution of 0.1 kHz and the frequency range of 0.1 to 20 kHz to compare to the animal studies (see 3.2.4). These two parameters were the same as those used for

predicting the latency patterns to clicks in 3.3.1.2, to compare the predicted patterns to various stimuli (a rarefaction click, a condensation click, and tone bursts).

Input signals in Fourier analysis

In the animal studies, the above neural latency pattern to tone bursts (see 3.3.2.1) was calculated from period histograms, which were generated through the use of the positive zero crossing of the stimulus waveform [25] to trigger a fast data acquisition unit, which marked the occurrence of each action potential with reference to a single cycle of the stimulus frequency. Therefore, the input signals for calculating the predicted latency pattern should start with a positive zero crossing as in the animal studies. The tone bursts were therefore generated by the equation $x(n) = \sin(2\pi f \cdot \frac{n}{f_s})$, where f = 0.1, 0.2… 10 kHz, $1 \le n \le 400$, f_s (sampling rate) = 40,000. All these 100 tone bursts started with a positive zero crossing that was used as the reference in the animal studies.

The prediction of the first spike latency for a tone burst was calculated by the steps below. The frequency of 1 kHz is chosen arbitrarily for demonstration. Predictions for other frequencies can be calculated in the same manner.

The input signal is the 1-kHz tone burst shown in Figure 40. The input is sampled with a moving window process wherein the window is moved forward in a point-wise manner (see 3.2). There are totally 799 windows (i.e., 400×2-1) that contain at least one non-zero point of the 400-point stimulus.

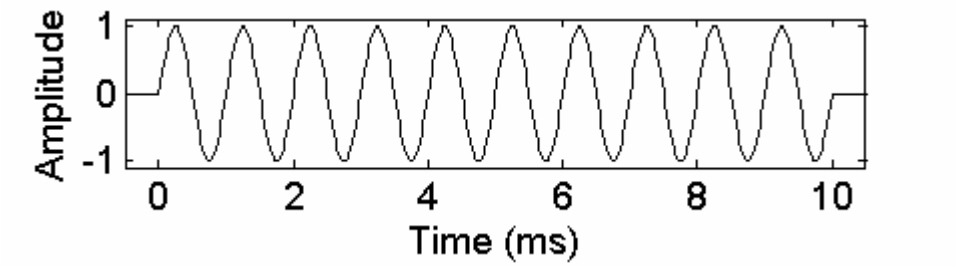


Figure 40. The 1-kHz tone burst

Calculate predicted latencies at the stimulus frequency

Calculate the predicted latency at the stimulus frequency (i.e., 1 kHz in this example) for each window and plot them together.

For example in the single window 0 (the first window, Figure 41A), the marking latency is $\frac{2\pi - 0.1571}{2\pi \cdot 1000} + \frac{0}{40000}$ = 0.000975 seconds (or 0.975 ms). Adding 1 ms (the synaptic and neural transmission; see 2.2.7), the predicted marking latency for window 0 is 1.975 ms. In the single window 99 (chosen arbitrarily for demonstration, Figure 41B), the marking latency is $\frac{2\pi - 1.7279}{2\pi \cdot 1000} + \frac{99}{40000}$ = 0.0032 seconds (or 3.2 ms). The predicted latency for window 99 is 3.2 + 1 = 4.2 ms. Calculate predicted latencies in all windows and plot them together (Figure 41E).

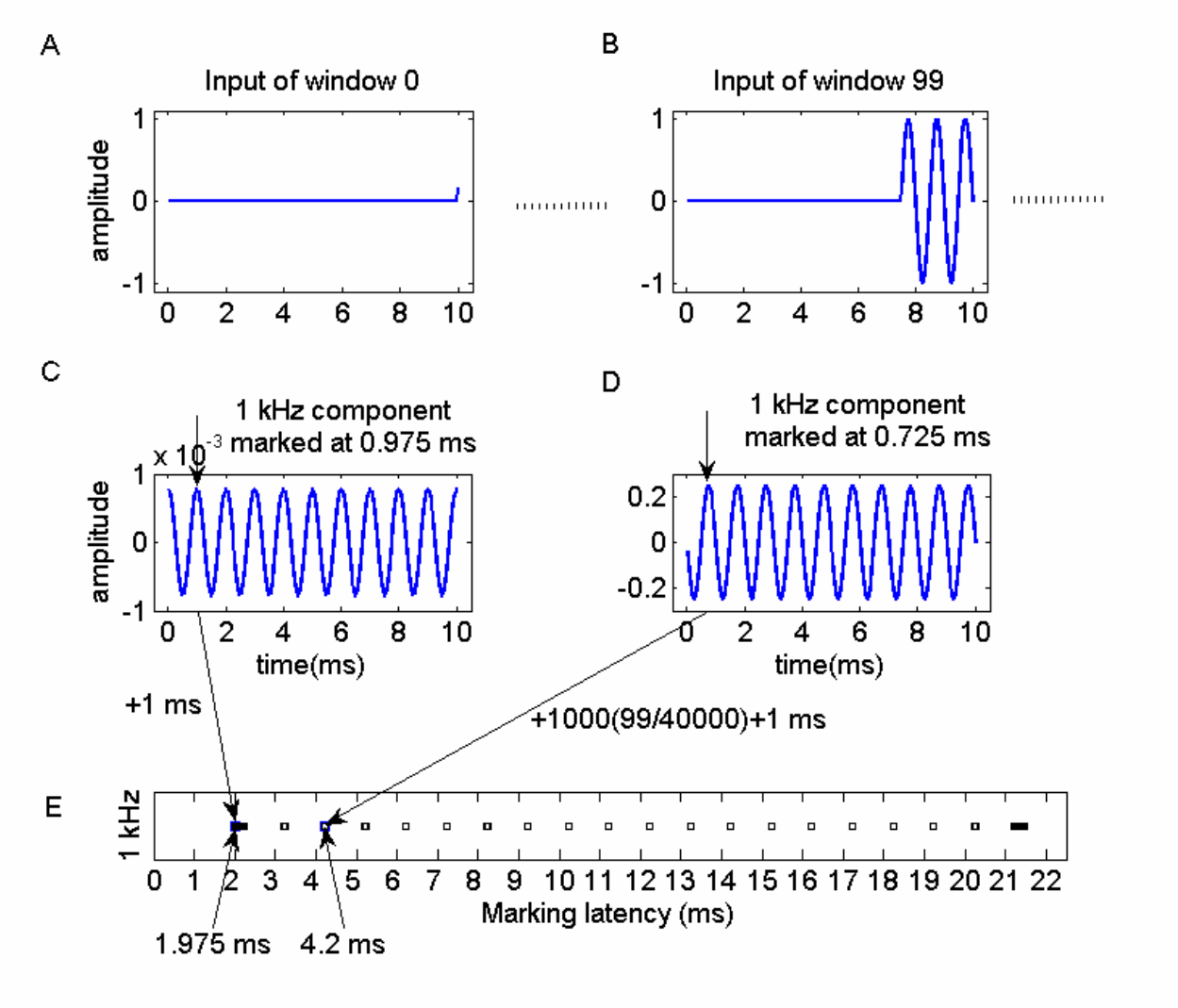


Figure 41. The predicted latencies at 1 kHz to the 1kHz tone burst. A predicted latency is the marking latency in a single window with the correction to reference (window 0) and the additional 1 ms. A. Input of window 0. B. Input of window 99. C. Marking of the first amplitude maximum of 1k-Hz component for window 0. D. Marking of the first amplitude maximum of 1k-Hz component for window 99. E. Distribution of all 799 markings at 1 kHz, including those from window 0 and 99 (indicated with big blue markings).

Figure 41E shows a periodic distribution that is related to the period of the stimulus (1 ms). So it is possible to calculate the theoretical points where these markings tend to be

distributed. The can be achieved by computing a modulus after dividing the latency by the stimulus period.

The two steps below are to compute the central tendency of the marking distribution, to find a theoretical first latency from the central tendency and then define the predicted first-spike latency.

Convert the predicted latencies to latencies between 0 and the period

Calculate the modulus after dividing the predicted latency by the stimulus period.

The modulus is defined as X - n×Y where X is the predicted latency, Y is the stimulus period, and n is the nearest integer less than or equal to $\frac{X}{Y}$. The predicted latencies are calculated as described in the previous step, as shown in Figure 41E. If the marking distribution in Figure 41E is viewed as a prediction of peristimulus-time histogram, this step is like "folding" the peristimulus-time histogram cycle-by-cycle to generate a period histogram.

For example, the converted latency for window 0 is the modulus after dividing 1.975 by 1, i.e., 0.975 ms. In the same way, the converted latency for window 99 is the modulus after dividing 4.2 by 1, i.e., 0.2 ms. Figure 42 shows the distribution of the converted latencies for all windows. The two big blue markings from windows 0 and 99 are indicated by arrows.

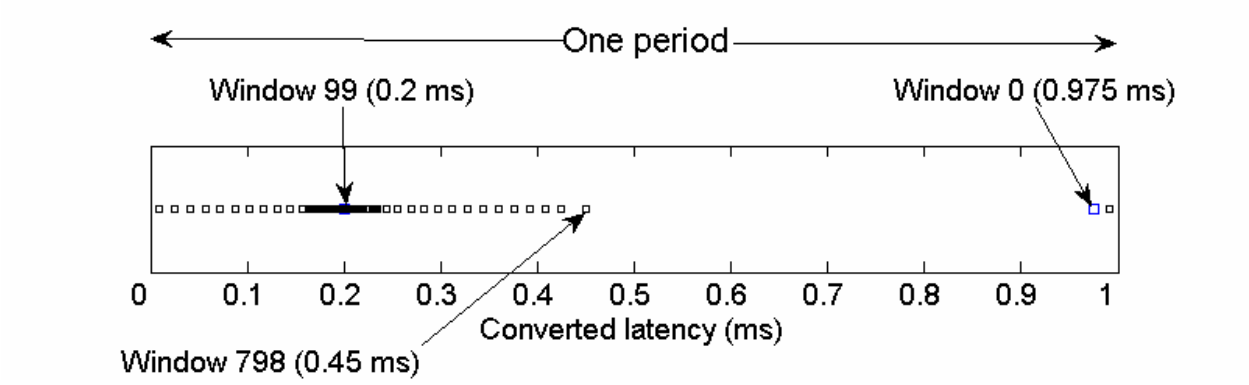


Figure 42. Distribution of the converted latencies between 0 and 1 ms to the 1-kHz tone burst. The two from windows 0 and 99 are indicated by big blue markings. If the markings are split (latency of window 0 is larger than that of window 798) as in this example, a cut line (average of these two latencies, i.e., 0.7125 ms) is set to move one part of the markings and merge all of them to calculate the mean latency.

Merge makings to calculate the mean converted latency

Assuming the markings are synchronized to a single point in one cycle, the mean converted marking latency should be calculated only when the markings are distributed in one group. If the markings are split (latency of window 0 is larger than that of window 798) as in this example, a cut line (average of these two latencies in the first and the last windows, i.e., 0.7125 ms) is set to move one side of the markings to the other and merge all of them to calculate the mean latency. Then make the mean converted latency range between 0 and the period (by adding or subtracting one period).

For instance, Figure 43 shows the merged markings made by moving the markings above 0.7125 ms in Figure 42 to the left (subtracting latencies by the period 1 ms). The mean converted latency is defined as the average of latencies of the merged markings, with the range between 0 and the period. It equals 0.20031 ms in this case.

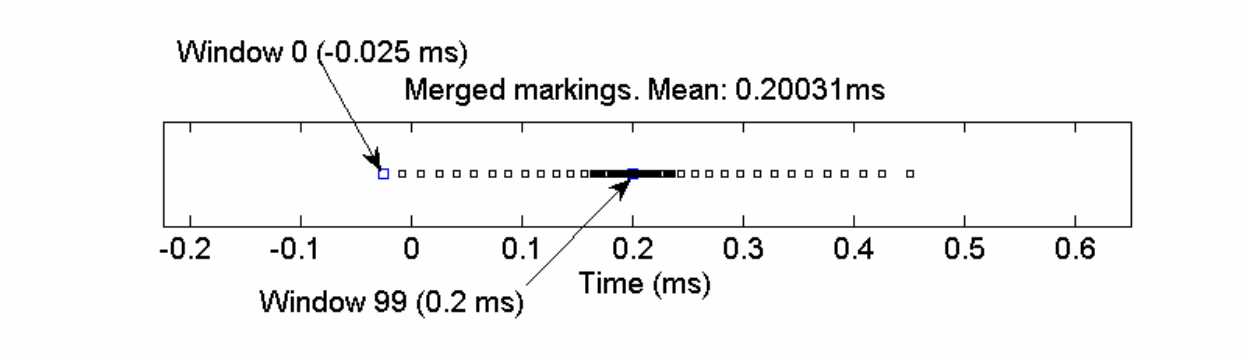


Figure 43. The merged markings made by moving the markings above the cutline to the left (subtracting latencies by the period). The mean converted latency is defined as the average of latencies of the merged markings, with the range between 0 and the period. The markings from window 0 and 99 are indicated by arrows and big blue symbols. The marking from window 0 is moved from 0.975 to -0.025 ms.

The mean converted latency of 0.20031 indicates that the predicted markings tend to be synchronized at theoretical points of 0.20031 + a multiple of the period (1 ms), i.e., 0.20031 ms, 1.20031 ms, 2.20031 ms, etc. These points are defined as predicted synchronization points for the 1-kHz tone burst.

Calculate the predicted first-spike latency to the tone burst

Comparing the predicted synchronization points (whose latencies are the mean converted latency + a multiple of the period, Figure 44B) with the predicted markings (Figure 44A). The latency of one of the predicted synchronization points (Figure 44B) whose latency is within the first marking group is defined as the predicted first-spike latency.

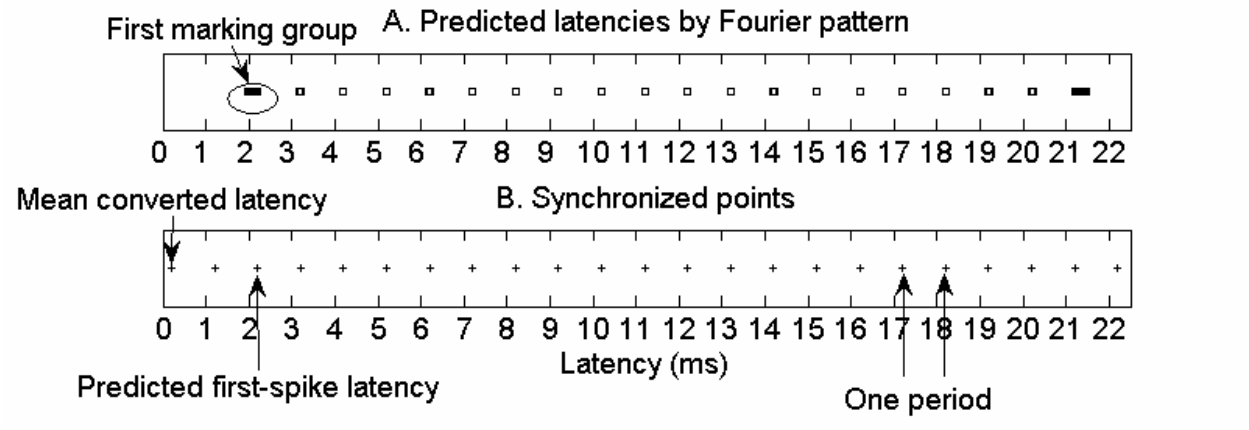


Figure 44. The predicted markings to the 1-kHz tone burst tend to be synchronized at specific points. The predicted first-spike latency is defined as the synchronization point whose latency is within the predicted first marking group. Arrows indicate the predicted first-spike latency and the first marking group.

There are no markings around the first and second predicted synchronization points in Figure 44A. The latency of the third predicted synchronization point (arrow in Figure 44B) is within the first marking group (arrow in Figure 44A). So it is defined as the predicted first-spike latency, which equals 2.20031 ms.

Calculate the predicted first-spike latency to all tone bursts

In the same manner, use tone bursts of 0.1, 0.2… 10 kHz and repeat the above steps to calculate the predicted first-spike latencies to all tone bursts.

The predicted latencies and the synchronized points for each stimulus were manually verified again that they matched well while calculating the predicted first-spike latencies. The predicted first-spike latencies would then be compared to response latencies in animal studies to investigate their correlation.

When the individual points of the raw data in the animal study are used as the neural latency pattern, a total of 80 neurons' characteristic frequencies were used to calculate the predicted latencies to compare with chinchillas' (40) and barn owls' (40) raw data.

3.3.2.3 Correlation between Fourier and response latencies to various tone bursts

The Pearson correlation coefficient was calculated and statistically tested to show the strength and direction of the correlation between the measured and predicted latencies to tone bursts. It was computed for both regression curves and raw data in both chinchillas and barn owls. The coefficient of determination (r^2) was calculated to show what percentage of the variance in the neural latency pattern can be explained by the Fourier latency pattern. When using the points on the regression curve as the neural latency pattern in chinchillas and barn owls,

n was 27 and 71, respectively (see 3.3.2.1). And when using the individual points of the raw data as the neural latency pattern in chinchillas and barn owls, n was 40 and 40, respectively (see 3.3.2.1).

3.4 TESTING PREDICTIONS AFTER FIRST SPIKES

As mentioned above (2.4.2), the predictions of latency in numerous auditory nerve fibers were evaluated in the second group of animal studies that yielded reports of the latencies after the first spikes, in addition to those in the first group, regarding first spike latencies. A single rarefaction click, a single condensation click and single tone bursts were used to test the hypothesis (Hypothesis 2) that a single auditory nerve fiber expresses specific spike distribution consistent with the Fourier pattern created by, in effect, heavily overlapped Fourier analyses.

3.4.1 Single rarefaction or condensation click

The latency predictions after first spikes were tested with the auditory-neural response latencies to rarefaction and condensation clicks, which are species-independent and have been precisely described and generally accepted (see 2.2.5).

3.4.1.1 Latency pattern in animal auditory nerve fibers

The latency pattern of the classical click responses is demonstrated by latencies of the peaks in numerous peristimulus-time histograms of auditory nerve fibers across many species (see 2.2.5) with a generally accepted conclusion that, at least for low-characteristic-frequency

fibers, there are multiple peaks separated by the inverse of the characteristic frequency, and the peaks from rarefaction and condensation clicks neatly interleave.

3.4.1.2 Latency pattern in consecutive Fourier analyses

The sampling rate and window size were set to 40,000 Hz and 400 points, to provide a frequency range of 0.1 to 20 kHz, with a resolution of 0.1 kHz (see 3.2.4). In order to test the relationship of markings between two Fourier analyses and whether the combined Fourier pattern accounts for the neural latency pattern after the first spikes, the Fourier pattern to a click in each of the consecutive analyses, in which the window moved forward one point at a time, and the combined Fourier pattern of all analyses were calculated, respectively.

Fourier pattern in each of the consecutive analyses

Due to space limitation, the latencies (plus 1 ms) of markings in each of the first 100 consecutive analyses to a rarefaction click were calculated to test if any single analysis or subset of analyses account(s) for the neural latency pattern. As before, the rarefaction click was generated as the signal: $x(n) = -100, \text{when } n = 0; x(n) = 0, \text{when } n \neq 0$.

Combined Fourier pattern of all analyses

Compared to the markings in a single analysis, the markings in all of the 400 analyses (because window size equals 400) were combined (plotted together) to test the effect of heavily overlapped analyses on latency pattern.

Combined Fourier pattern of both click polarities

The combined Fourier pattern (with additional 1 ms) to a *condensation* click ($x(n) = 100, \text{when } n = 0; x(n) = 0, \text{when } n \neq 0$) also was calculated to test the effect of click

polarity on latency pattern. The predicted latency pattern of clicks was defined as this combined Fourier pattern (with additional 1 ms) to both rarefaction and condensation clicks.

3.4.1.3 Comparison between predicted and recorded latencies

The predicted latency pattern was examined by inspecting whether its exact numbers fit the generally accepted conclusion for classical click responses. Specifically, the peaks in a peristimulus-time histogram are separated by the inverse of the characteristic frequency, and the peaks from rarefaction and condensation clicks neatly interleave.

3.4.2 Single tone bursts

Although auditory nerve fibers across various species demonstrate phase locking to tone bursts (see 2.2.3), there seems to be no study that precisely reports the latencies (such as the mean latency and the width of each spike group) to single tone bursts in auditory nerve fibers over a wide range of characteristic frequencies.

3.4.2.1 Latency pattern in animal auditory nerve fibers

The latency pattern to a single tone burst is usually reported by latencies of spike groups in the peristimulus-time histogram. Specifically, at least for low-characteristic-frequency fibers around the stimulus frequency, an auditory neuron tends to fire at a particular phase of a stimulating tone, and the inter-spike intervals tend to occur at integer multiples of the inverse of the stimulus frequency (see 2.2.3). This specific distribution will be used as the neural latency pattern for comparison.

3.4.2.2 Latency pattern in consecutive Fourier analyses

The combined Fourier patterns (plus 1 ms) to four tone bursts (0.5, 1, 2 and 4 kHz; chosen arbitrarily for demonstration) were used as the predicted latency patterns.

An example of calculating the combined Fourier pattern to the 1-kHz tone burst is detailed in **錯誤! 找不到參照來源。**.

3.4.2.3 Comparison between predicted and recorded latencies

The predicted latency pattern was examined by inspecting whether the distribution of the markings was consistent with the general description that an auditory neuron tends to fire at a particular phase of a stimulating tone, i.e., whether the markings were separated by the inverse of the tone burst's frequency.

4.0 RESULTS

4.1 PHASE ENCODING AMONG AUDITORY NERVE FIBERS

4.1.1 Rarefaction click

Results for the predicted latency pattern to the rarefaction click (3.3.1.2) are presented in Figure 45B. When using the points on the regression curve of the recordings as the neural latency pattern, correlation analysis reveals the predicted latencies to be related significantly to neural latencies, r = 0.97, n = 98, $p < 0.05$, two tailed test. The coefficient of determination (r^2) is 0.95. Ninety-five percent of the variance in the points on the regression curve of the neural pattern can be explained by the predicted latency pattern to a rarefaction click. The difference between predicted and animal latencies (predicted – animal) has a mean of -0.20 ms with a standard deviation (SD) = 0.079 ms, n = 98. On average, the predicted latencies are 0.2 ms shorter than the points on the regression curve of the neural pattern to a rarefaction click.

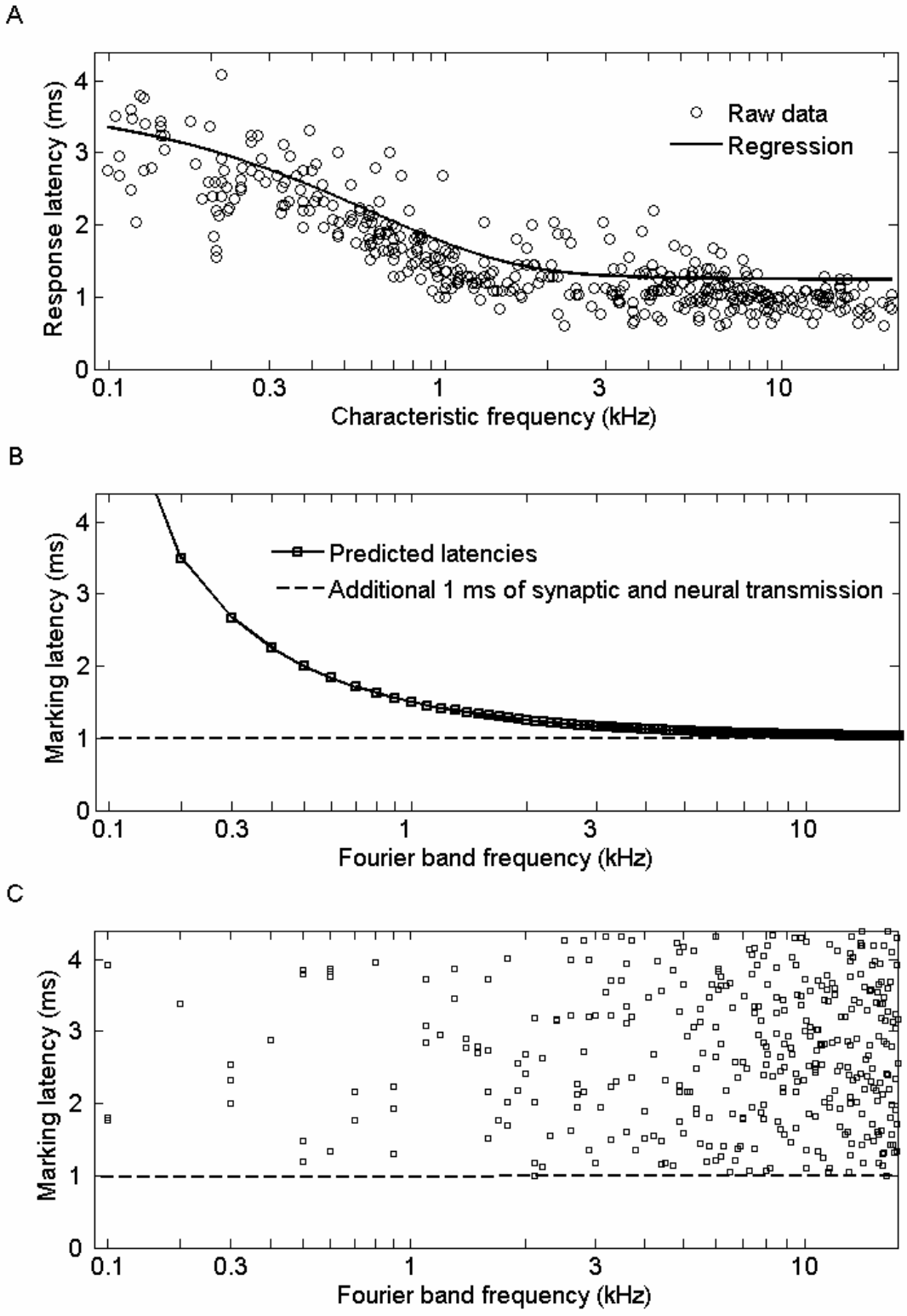

A
Response latency (ms)
Characteristic frequency (kHz)
Raw data
Regression
B
Marking latency (ms)
Fourier band frequency (kHz)
Predicted latencies
Additional 1 ms of synaptic and neural transmission
C
Marking latency (ms)
Fourier band frequency (kHz)

Figure 45. Predicted and recorded first-spike latency pattern in auditory nerve fibers to a rarefaction click. A. Individual recordings and regression of response latencies in chinchillas' auditory nerve fibers. Raw data obtained by personal communication. With permission from Temchin, A. N., A. Recio-Spinoso, et al. (2005). "Wiener kernels of chinchilla auditory-nerve fibers: verification using responses to tones, clicks, and noise and comparison with basilar-membrane vibrations." J Neurophysiol 93(6): 3635-3648. B. Fourier pattern pluses 1 ms. C. The predicted latencies without the Fourier pattern. Straight dash lines indicate the 1-ms synaptic and neural transmission. The predicted latency pattern (B) is significantly related to the neural latency pattern (A). The latencies without the Fourier pattern (C, which is calculated by randomly marking at an average rate of 60 Hz for 10 repetitions) are not related to the neural latencies.

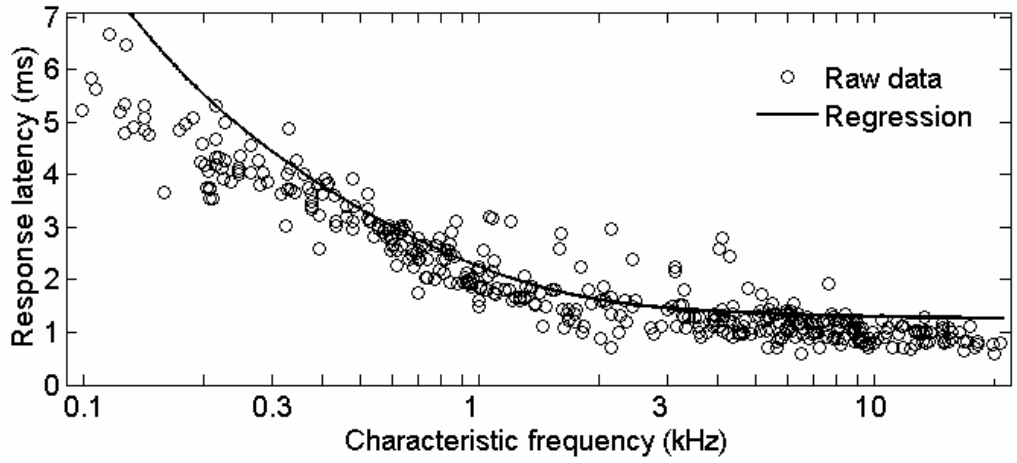
The marking latencies to the rarefaction click without the Fourier pattern are shown in Figure 45C. A correlation for the data reveals that the mean marking latencies of the 10 repetitions without the Fourier pattern are not related to the neural latencies, r = 0.031, n = 98, p = 0.66, two tailed test.

The power fit of the predicted latencies to a rarefaction click is $\tau\,(\text{ms}) = 0.975 + 500 \cdot f_{(Hz)}{}^{-1}$. When using the individual points of the raw data as the neural latency pattern, correlation analysis reveals the predicted latencies to be related significantly to neural latencies, r = 0.82, n = 411, $p < 0.05$, two tailed test. The coefficient of determination (r^2) is 0.68. Sixty-eight percent of the variance in the individual points of the raw data in the auditory nerve fibers can be explained by the predicted points to a rarefaction click. The difference between predicted and animal latencies (predicted – animal) has a mean of 0.10 ms with SD = 0.56 ms, n = 411. On average, the predicted latencies are 0.1 ms longer than the individual points of the raw data in the auditory nerve fibers to a rarefaction click.

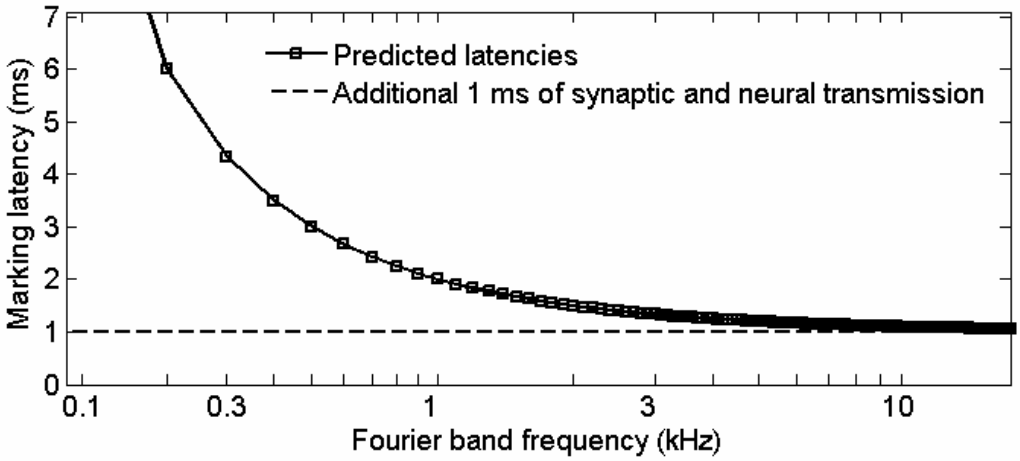
4.1.2 Condensation click

Results for the predicted latency pattern to the condensation click (3.3.1.2) are presented in Figure 46B. When using the points on the regression curve of the recordings as the neural latency pattern, correlation analysis reveals the predicted latencies to be related significantly to neural latencies, r = 0.99, n = 98, $p < 0.05$, two tailed test. The coefficient of determination (r^2) is 0.99. Ninety-nine percent of the variance in the points on the regression curve of the neural pattern can be explained by the predicted latency pattern to a condensation click. As the regressions for recordings of rarefaction and condensation clicks were identical, other than an additional 0.5 period for the latter (see 3.3.1.1), the mean and SD were not reported.

A


Response latency (ms)
Characteristic frequency (kHz)
Raw data
Regression
0
1
2
3
4
5
6
7
0.1
0.3
1
3
10

B


Marking latency (ms)
Fourier band frequency (kHz)
Predicted latencies
Additional 1 ms of synaptic and neural transmission
0
1
2
3
4
5
6
7
0.1
0.3
1
3
10

C

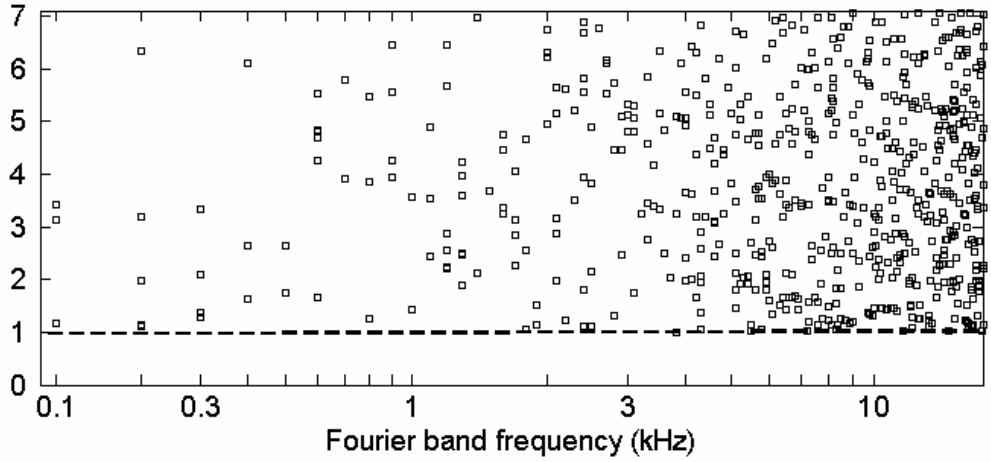

Fourier band frequency (kHz)
0
1
2
3
4
5
6
7
0.1
0.3
1
3
10

Figure 46. Predicted and recorded first-spike latency pattern in auditory nerve fibers to a condensation click. A. Individual recordings and regression of response latencies in chinchillas' auditory nerve fibers. Raw data obtained by personal communication. With permission from Temchin, A. N., A. Recio-Spinoso, et al. (2005). "Wiener kernels of chinchilla auditory-nerve fibers: verification using responses to tones, clicks, and noise and comparison with basilar-membrane vibrations." J Neurophysiol 93(6): 3635-3648. B. Fourier pattern pluses 1 ms. C. The predicted latencies without the Fourier pattern. Straight dash lines indicate the 1-ms synaptic and neural transmission. The predicted latency pattern (B) is significantly related to the neural latency pattern (A). The latencies without the Fourier pattern (C, which is calculated by randomly marking at an average rate of 60 Hz for 10 repetitions) are not related to the neural latencies.

The marking latencies without the Fourier pattern are shown in Figure 46C. A correlation for the data reveals that the mean marking latencies of the 10 repetitions without the Fourier pattern are not related to the neural latencies, r = -0.024, n = 98, p = 0.74, two tailed test.

The power fit of the predicted latencies to a condensation click is $\tau\,(\text{ms}) = 0.975 + 1000 \cdot f_{(Hz)}{}^{-1}$. When using the individual points of the raw data as the neural latency pattern, a correlation for the data reveals that the predicted latencies are significantly related to the neural latencies, r = 0.90, n = 407, $p < 0.05$, two tailed test. The coefficient of determination (r^2) is 0.82. Eighty-two percent of the variance in the individual points of the raw data in the auditory nerve fibers can be explained by the predicted points to a condensation click. The difference between predicted and animal latencies (predicted – animal) has a mean of 0.21 ms with SD = 0.90 ms, n = 407. On average, the predicted latencies are 0.21 ms longer than the individual points of the raw data in the auditory nerve fibers to a condensation click.

4.1.3 Tone bursts

Although no regression curve was provided in the study in chinchillas, the power fit regression curve is calculated according to the raw data the authors provided. The regression curve in chinchillas to tone bursts is: $\tau\,(\text{ms}) = 1 + 250.9 \cdot CF^{-0.71}$ (where CF is in Hz).

Results for the predicted latency pattern to the tone bursts (3.3.2.2) are presented in Figure 47B. When using the points on the regression curve of the recordings as the neural latency pattern, correlation analysis reveals the predicted latencies to be related significantly to neural latencies in both chinchillas and barn owls, ($r = 0.99$, $n = 27$, $p < 0.05$, two tailed test) and ($r = 0.997$, $n = 71$, $p < 0.05$, two tailed test), respectively. The coefficient of determination (r^2) is 0.99 and 0.99 in chinchillas and barn owls, respectively. In both animals, Ninety-nine percent of the variance in the points on the regression curve of the neural pattern can be explained by the predicted latency pattern to tone bursts. The means and standard deviations of latency difference (predicted – animal) to tone bursts in chinchillas and barn owls are -0.54 ± 0.20 ms (n = 27) and -0.15 ± 0.044 ms (n = 71), respectively. Thus, on average, the predicted latencies in chinchillas are 0.54 ms shorter than the points on the regression curve of the neural pattern to tone bursts, and are 0.15 ms shorter in barn owls.

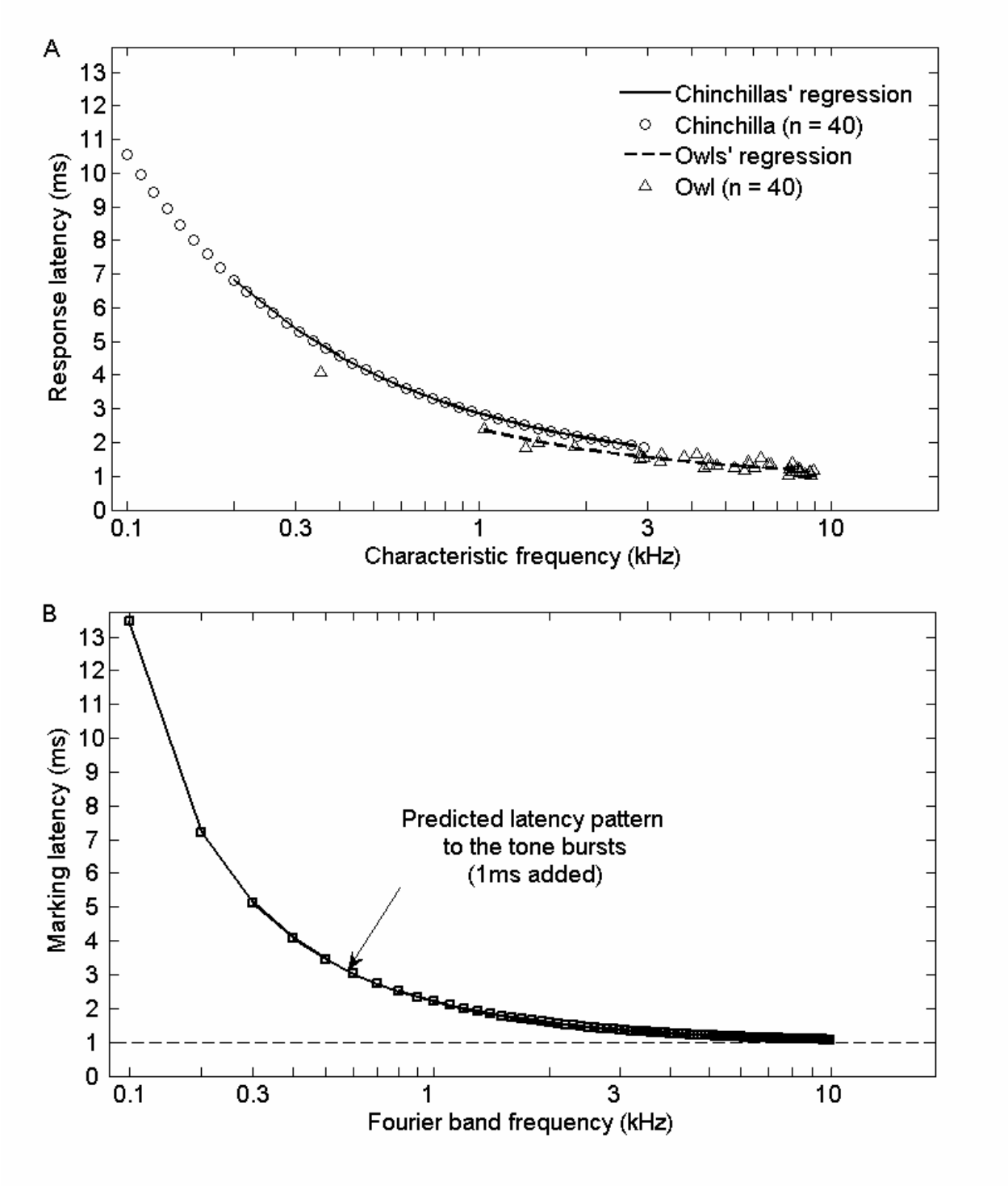


Figure 47. Predicted and recorded first-spike latency pattern in auditory nerve fibers to various tone bursts. A. Individual recordings and regression of response latencies in chinchillas' and barn owls' auditory nerve fibers. Raw data obtained by personal communication. With permission from Temchin, A. N., A. Recio-Spinoso, et al. (2005). "Wiener kernels of chinchilla auditory-nerve fibers: verification using responses

to tones, clicks, and noise and comparison with basilar-membrane vibrations." J Neurophysiol 93(6): 3635-3648, and from Koppl, C. (1997). "Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba." J Neurosci 17 (9): 3312-3321. B. Predicted latencies by Fourier patterns pluses 1 ms. Straight dash lines indicate the 1-ms synaptic and neural transmission. The predicted latency pattern (B) is significantly related to the neural latency patterns in chinchillas and barn owls, respectively (A).

The power fit of the predicted latencies to tone bursts is $\tau\,(\text{ms}) = 0.9501 + 1250 \cdot f_{(Hz)}^{-1}$. When using the individual points of the raw data as the neural latency pattern, correlation analysis reveals the predicted latencies to be related significantly to neural latencies in both chinchillas and barn owls, ($r = 0.99$, $n = 40$, $p < 0.05$, two tailed test) and ($r = 0.96$, $n = 40$, $p < 0.05$), respectively. The coefficient of determination (r^2) in chinchillas and barn owls is 0.99 and 0.93, respectively. Ninety-nine and ninety-three percent of the variance in the individual points of the raw data in the auditory nerve fibers can be explained by the predicted points to tone bursts in chinchillas and owls, respectively. The means and standard deviations of latency difference (predicted – animal) in chinchillas and barn owls are (-0.058 ± 0.94 ms, $n = 40$) and (-0.12 ± 0.15, $n = 40$), respectively. On average, the predicted latencies are 0.058 and 0.12 ms shorter than the individual points of the raw data in the auditory nerve fibers to tone bursts in chinchillas and owls, respectively.

4.1.4 Summary of correlation

The results show that all of the predicted latencies by Fourier patterns are significantly related to neural latencies across stimulus (a rarefaction click, a condensation click and various tone bursts),

neural data types (regression and raw data), and species (chinchillas, barn owls and to the extent they represent). Indeed, all measures of correlation proved to be quite high, as summarized in Table 2.

Table 2. Summary of correlation between the predicted and neural latencies. The predicted latencies by Fourier pattern are significantly related to neural latencies across stimulus (a rarefaction click, a condensation click and various tone bursts), neural data (regression and raw data), and species (chinchillas, barn owls and to the extent they represent).

Stimulus	Neural data	r	n	p	r^2
Rarefaction	Regression	0.97402	98	1.01×10^{-63}	0.95
click	Raw data	0.82301	411	1.64×10^{-102}	0.68
Condensation	Regression	0.99395	98	6.43×10^{-94}	0.99
click	Raw data	0.90440	407	6.85×10^{-152}	0.82
Tone bursts	Regression	0.99418	27	1.03×10^{-25}	0.99
(chinchilla)	Raw data	0.99421	40	1.98×10^{-38}	0.99
Tone bursts	Regression	0.99658	71	1.87×10^{-76}	0.99
(barn owls)	Raw data	0.96371	40	2.13×10^{-23}	0.93

4.2 PHASE ENCODING WITHIN AN AUDITORY NERVE FIBER

4.2.1 Rarefaction click

Results for the Fourier pattern in each of the first 100 consecutive analyses (plus 1 ms) to a rarefaction click are presented in Figure 48. The predicted latencies in each analysis usually present a saw-tooth pattern.

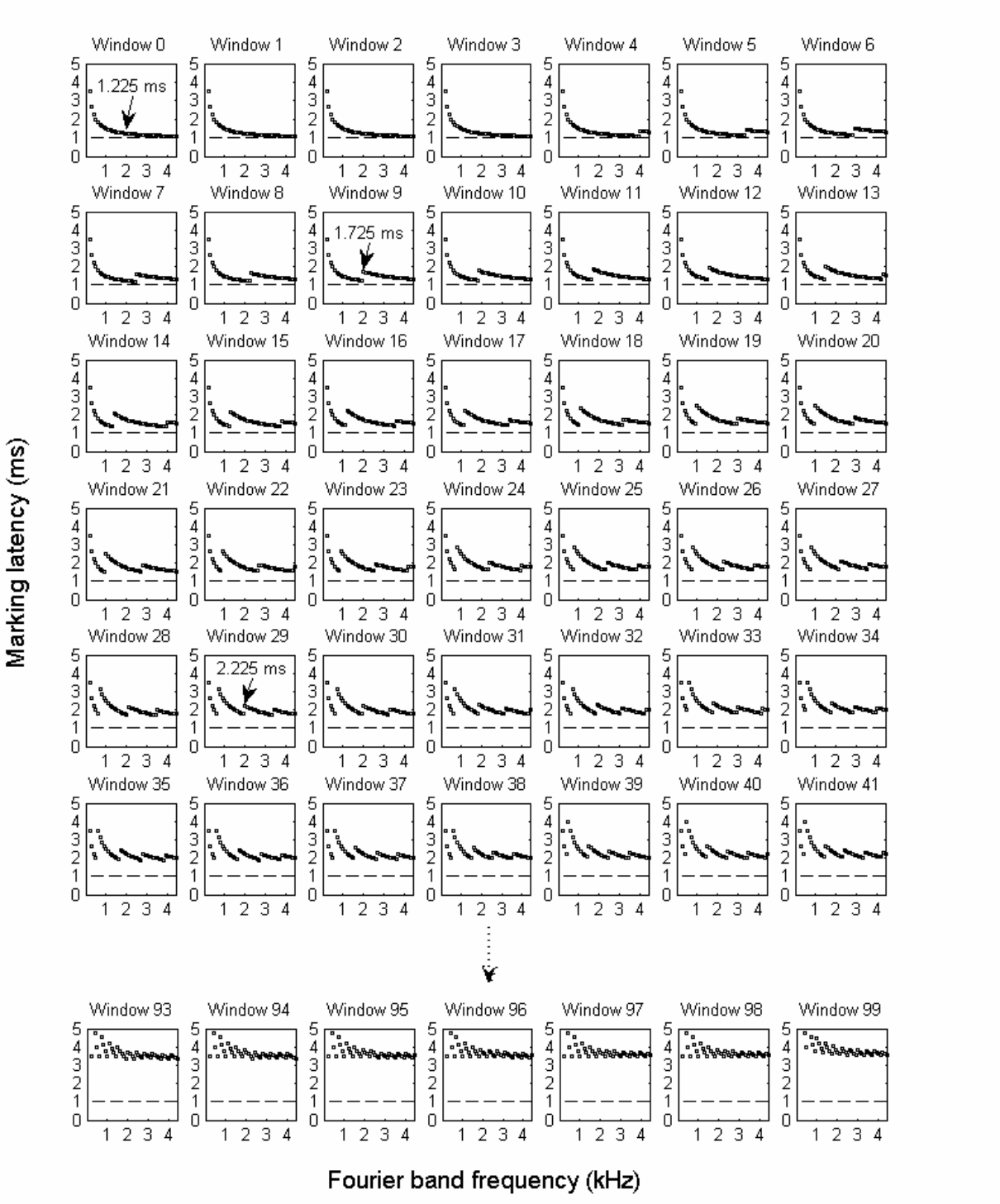


Figure 48. Predicted latency pattern in each of the consecutive windows to the rarefaction click. Markings beyond the time and frequency ranges are not shown. The markings usually present a saw-tooth pattern. There is always one marking at a frequency in one analysis. Arrows indicate the new values (latency

'jumps') at the frequency of 2 kHz in windows 0, 9 and 19. The rarefaction click is generated as the signal: x(n) = -100, when n = 0; x(n) = 0, when n ≠ 0. Sampling rate: 40 kHz. Window size: 400 points.

At a given Fourier band frequency, the predicted latency to the rarefaction click is usually the same for several consecutive analyses then jumps to a new latency that is separated from the previous one by the inverse of the Fourier band frequency. The marking stays at the new timing for several analyses before it changes again. For instance, tracing the markings at the same frequency of 2 kHz, as indicated by the arrows in Figure 48, the latency is 1.225 ms for the first 9 analyses (windows 0 to 8), 1.725 ms for the next 20 analyses (windows 9 to 28), etc. This indicates that, at a given frequency, the markings of all consecutive analyses are precisely (same numbers in Matlab) distributed at points that are separated by the inverse of that Fourier band frequency. In this example, because windows 0, 9, and 29 are the windows that have a new latency at 2 kHz, these windows are taken for examples later to illustrate the evolution of latency by frequency analysis.

The combined Fourier pattern (plus 1 ms) of all the 400 analyses is demonstrated in Figure 49 by plotting all the markings together. The saw-tooth pattern in each analysis is now embedded in the combined pattern. It again shows that, to the rarefaction click, the markings at a given frequency are precisely distributed at points that are separated by the inverse of that Fourier band frequency. For instance at 2 kHz, the 400 markings are distributed at 21 points that are separated by $\frac{1}{2\,\text{kHz}}\,\text{seconds} = 0.5\,\text{ms}$ (Figure 50A).

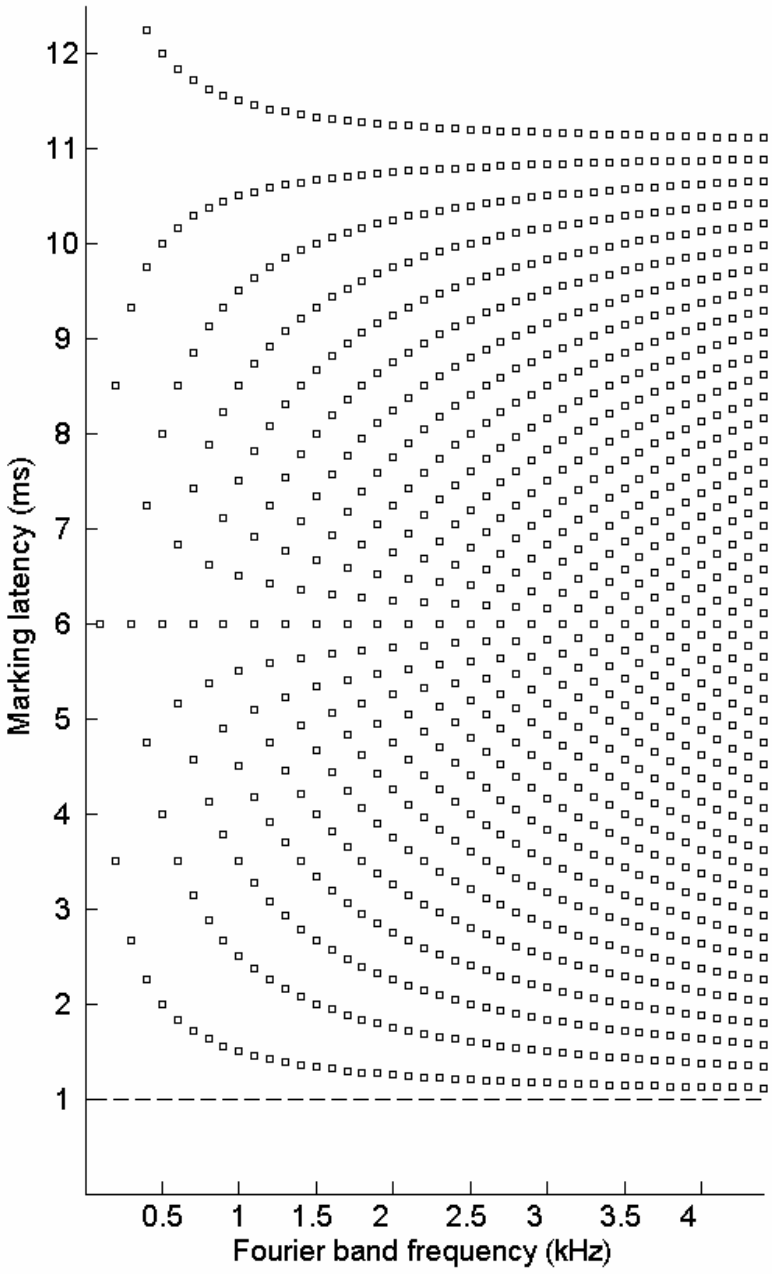


Figure 49. The combined Fourier pattern (plus 1 ms) of all the 400 consecutive analyses to the rarefaction click. Straight dash line indicates the additional 1 ms.

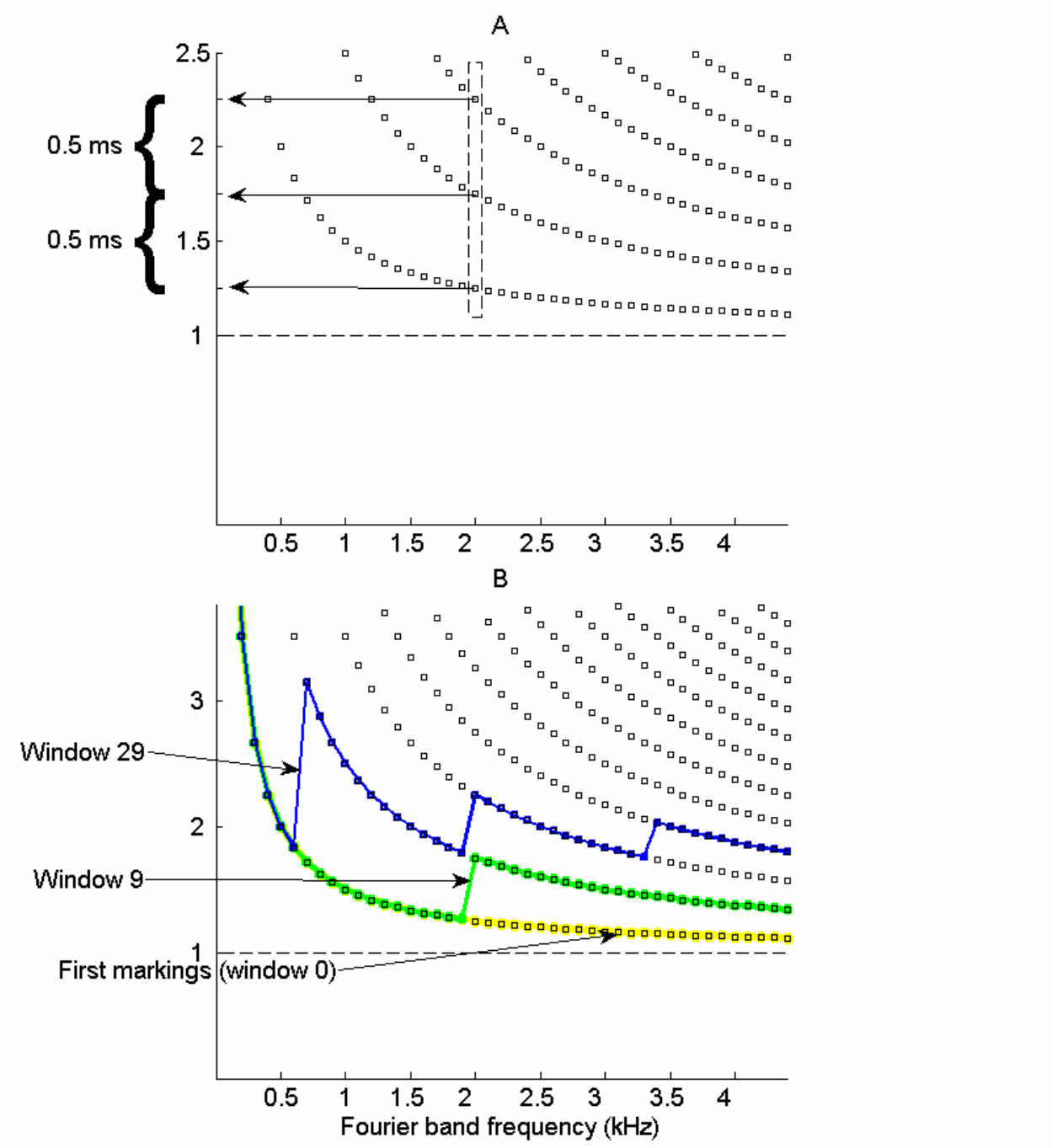


Figure 50. Point separation at a given frequency and the evolution of latency by frequency analysis to the rarefaction click. A. Markings at 2 kHz are distributed at points separated by 1/2 kHz (seconds) = 0.5 ms. B. The evolution of latency by frequency analysis showing jumps at 2 kHz. The first windows that has a new latency at 2 kHz are indicated by yellow (first markings; window 0), green (window 9) and blue (window 29). The markings in window 9 (green) show a 0.5-ms jump at 2 kHz. The markings in window 29 (blue) show

jumps at 0.7, 2 and 3.4 kHz of 1.43, 0.5 and 0.29 ms, respectively. Straight dash line indicates the additional 1 ms

The three individual windows that form the "jumps" at 2 kHz are shown in Figure 50B to illustrate the evolution of latency by frequency analysis. In windows 9 and 29 (indicated by green and blue), the marking at 2 kHz jumps to a new latency that is 0.5 ms away from the previous one. Also, the markings in window 29 (blue) also show jumps at 0.7 and 3.4 kHz by $\frac{1}{0.7\,\text{kHz}}\,\text{seconds} = 1.43\,\text{ms}$ and $\frac{1}{3.4\,\text{kHz}}\,\text{seconds} = 0.29\,\text{ms}$, respectively, indicating that 0.7 kHz corresponds to 1.43 ms and 3.4 kHz corresponds to 0.29 ms. Appendix C demonstrates further the calculations and the synthesized Fourier components before and after the jumps to investigate why the jumps may occur.

4.2.2 Condensation click

The Fourier pattern in each of the first 100 consecutive analyses (plus 1 ms) to a condensation click is calculated and shown in Figure 51. The predicted latencies in each analysis also frequently present a saw-tooth pattern.

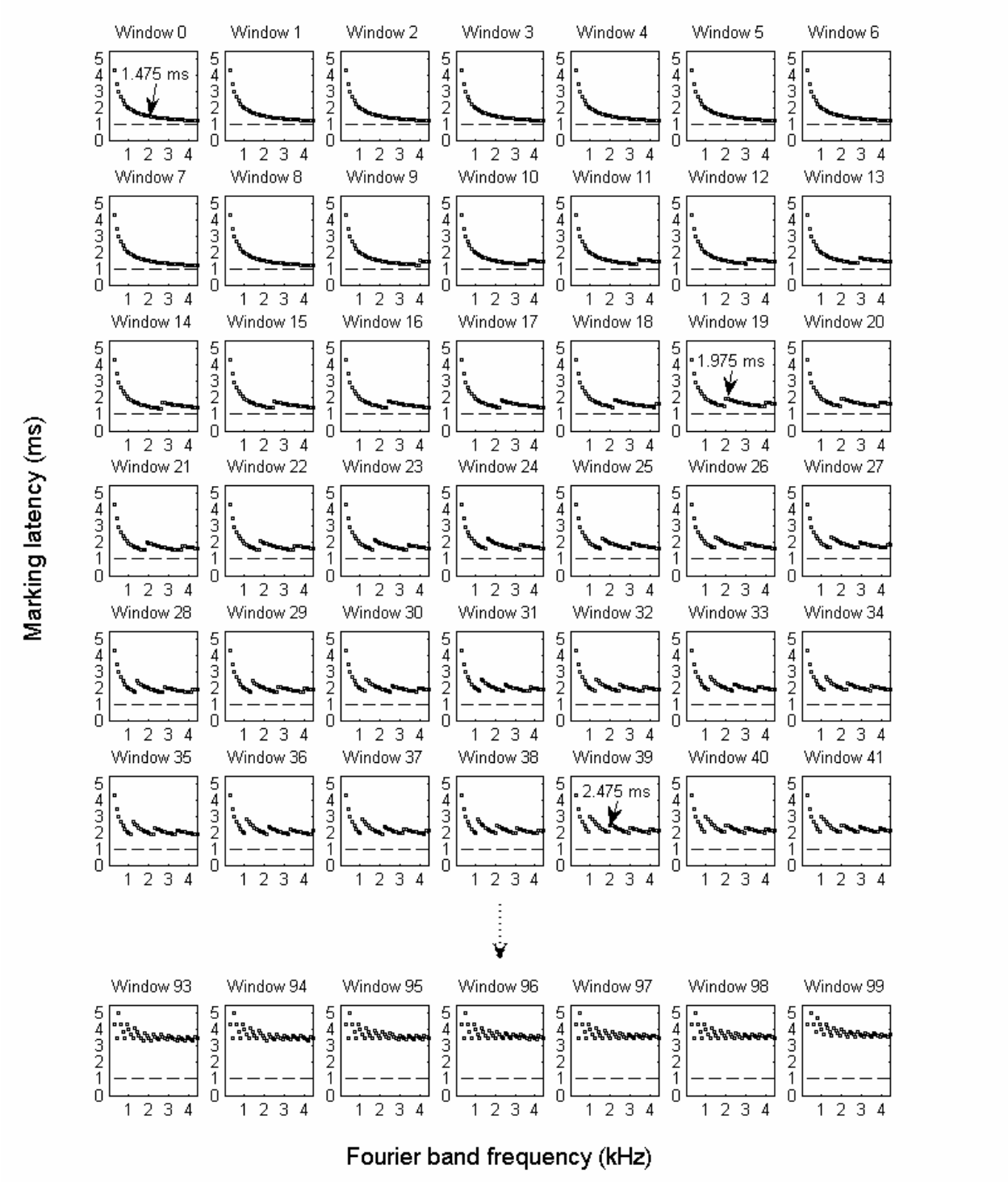


Figure 51. Predicted latency pattern in each of the consecutive windows to the condensation click. Markings beyond the time and frequency ranges are not shown. The markings usually present a saw-tooth pattern. Arrows indicate the new values (latency 'jumps') at the frequency of 2 kHz in windows 0, 19 and 39.

The condensation click is generated as the signal: x(n) = 100, when n = 0; x(n) = 0, when n ≠ 0. Sampling rate: 40 kHz. Window size: 400 points.

Similarly, at a given Fourier band frequency, the predicted latency to the condensation click is usually the same for several consecutive analyses then jumps to a new latency that is separated from the previous one by the inverse of the Fourier band frequency. The marking stays at the new timing for several analyses before it changes again. For instance, tracing the markings at the same frequency of 2 kHz, as indicated by the arrows in Figure 51, the latency is 1.475 ms for the first 19 analyses (windows 0 to 18), 1.975 ms for the next 20 analyses (windows 19 to 38), etc. This indicates that, at a given frequency, the markings of all consecutive analyses to the condensation click are precisely (same numbers in Matlab) distributed at points that are separated by the inverse of that Fourier band frequency. In this example, because windows 0, 19, and 39 are the windows that have a new latency at 2 kHz, these windows are taken for examples later to illustrate the evolution of latency by frequency analysis.

The combined Fourier pattern (plus 1 ms) of all the 400 analyses is demonstrated in Figure 52 by plotting all the markings together. The saw-tooth pattern in each analysis is now embedded in the combined pattern. It again shows the above-mentioned characteristic that, to the condensation click, the markings at a given frequency are precisely distributed at points that are separated by the inverse of that Fourier band frequency. For instance at 2 kHz, the 400 markings are distributed at 21 points that are separated by $\frac{1}{2\,\text{kHz}}\,\text{seconds} = 0.5\,\text{ms}$ (Figure 53A).

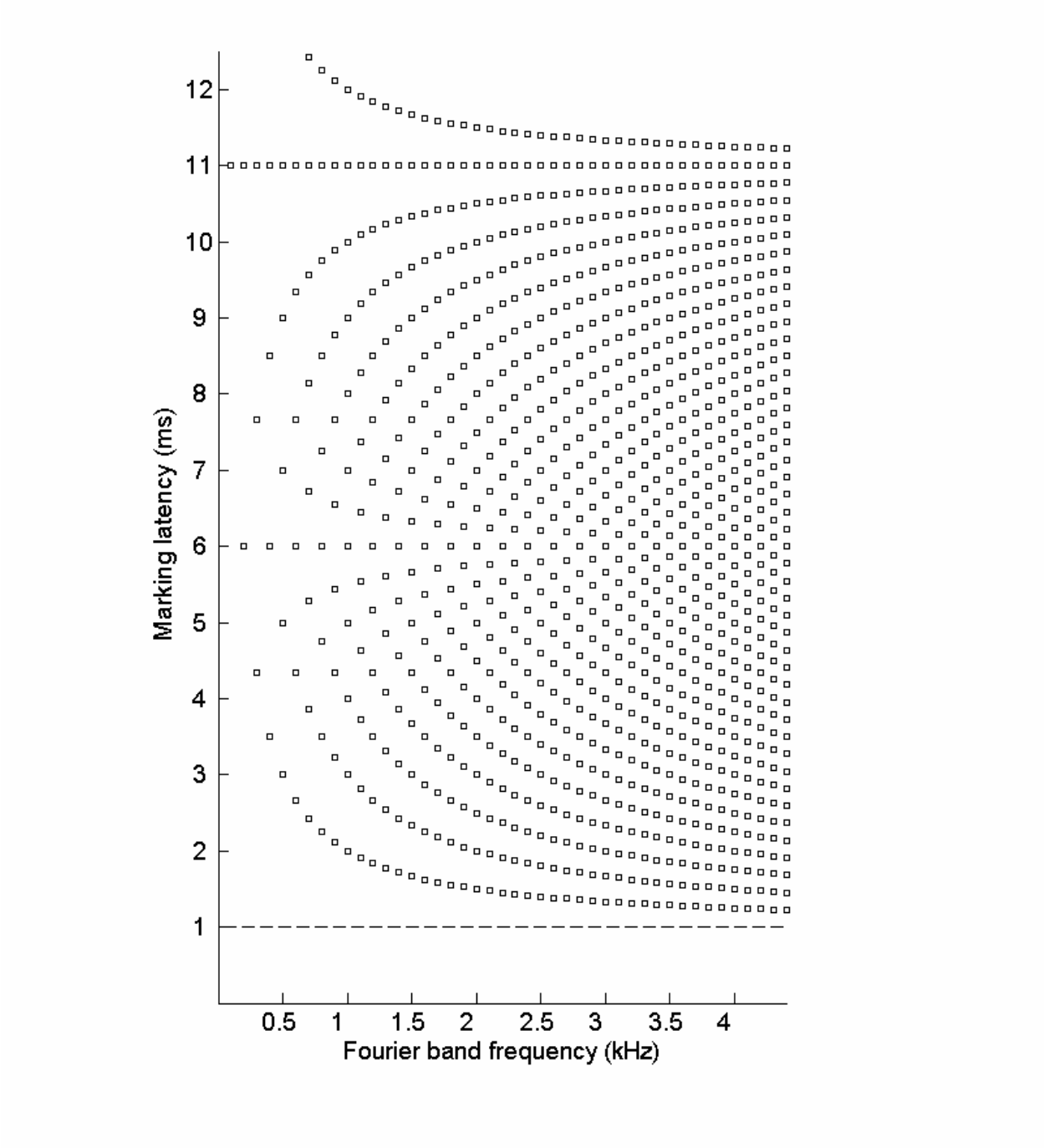


Figure 52. The combined Fourier pattern (plus 1 ms) of all the 400 consecutive analyses to the condensation click. Straight dash line indicates the additional 1 ms.

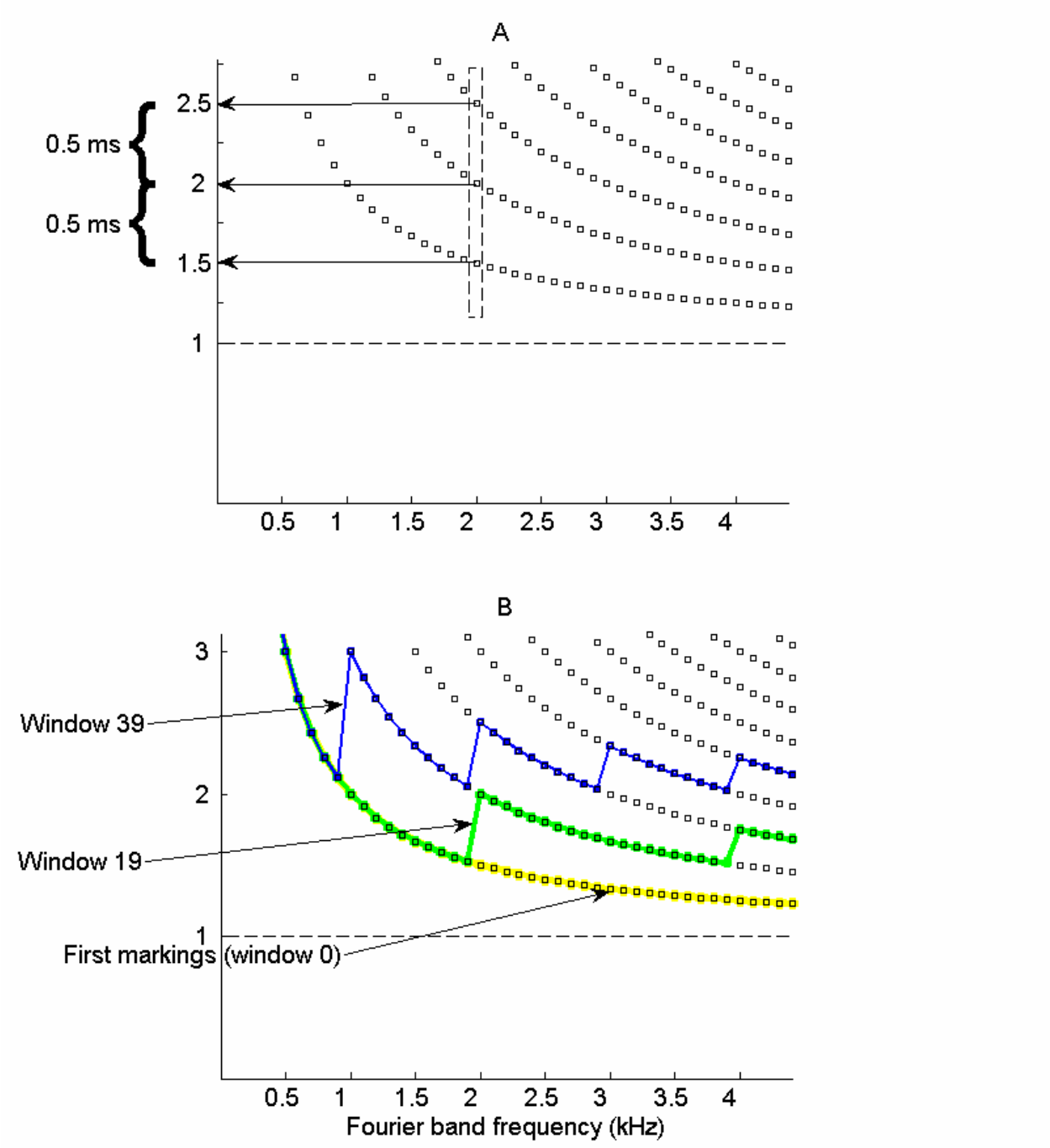


Figure 53. Point separation at a given frequency and the evolution of latency by frequency analysis to the condensation click. A. Markings at 2 kHz are distributed at points separated by 1/2 kHz (seconds) = 0.5 ms. B. The evolution of latency by frequency analysis showing jumps at 2 kHz. The first windows that has a new latency at 2 kHz are indicated by yellow (first markings; window 0), green (window 19) and blue (window 39). The markings in window 19 (green) show jumps at 2 and 4 kHz of 0.5 and 0.25 ms, respectively.

The markings in window 29 (blue) show jumps at 1, 2, 3 and 4 kHz of 1, 0.5, 0.33 and 0.25 ms, respectively. Straight dash line indicates the additional 1 ms.

The three individual windows that form the "jumps" at 2 kHz are shown in Figure 53B to illustrate the evolution of latency by frequency analysis. In windows 19 and 39 (indicated by green and blue), the marking at 2 kHz jumps to a new latency that is 0.5 ms away from the previous one. In addition to 2 kHz, the markings in window 39 (blue) show jumps at 1, 3 and 4 kHz of 1, 0.33 and 0.25 ms, respectively, indicating that 1 kHz corresponds to 1 ms, etc. See Appendix C for the calculations and the synthesized Fourier components before and after the jumps to for why the jumps may occur.

4.2.3 Combined click pattern

The combined Fourier pattern plus 1 ms to both click polarities is shown in Figure 54B. In the classical view, in a peristimulus-time histogram, the click responses in an auditory nerve fiber show peaks separated by the inverse of the characteristic frequency, and the peaks from rarefaction and condensation clicks neatly interleave (Figure 54A). Inspecting the exact numbers in Matlab, the results show that the predicted latency pattern shared the same characteristics.

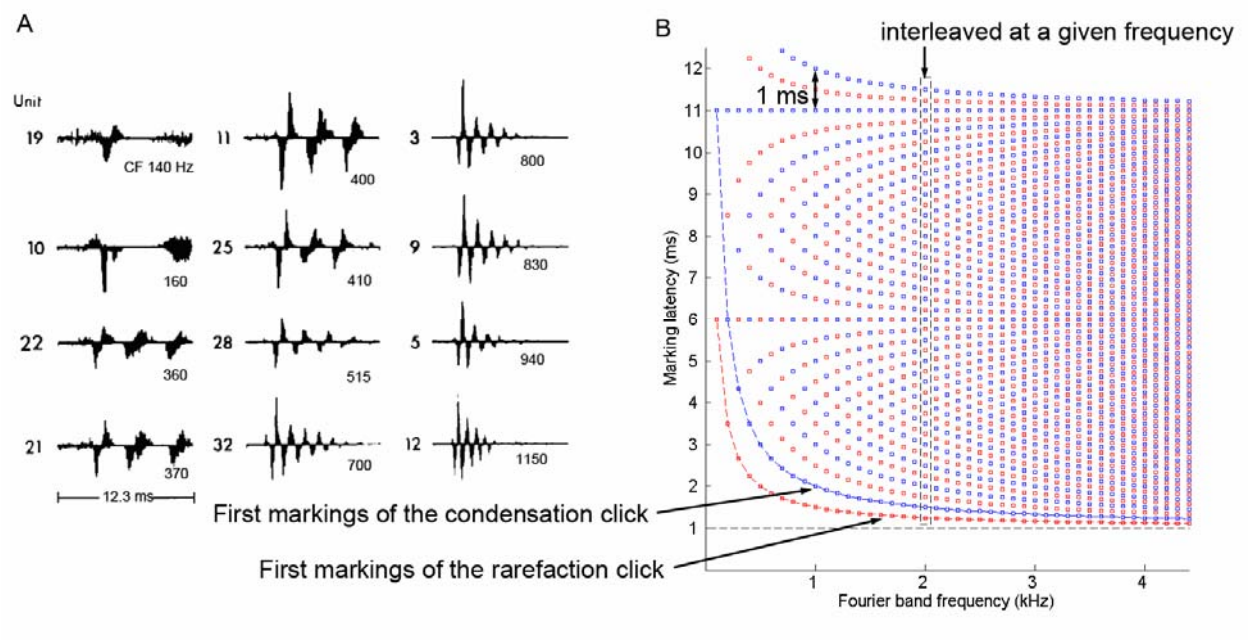


Figure 54. Animal vs. predicted latency pattern to rarefaction and condensation clicks. A. Compound peristimulus-time histograms of classical click responses in single auditory nerve fibers of cats. Upper: rarefaction. Reprinted with permission from Fig. 4, Pfeiffer, R. R. and D. O. Kim (1972). "Response patterns of single cochlear nerve fibers to click stimuli: descriptions for cat." J Acoust Soc Am 52(6): 1669-1677. Copyright 1972, Acoustic Society of America. B. The predicted latency pattern shares the same characteristics in animal studies. Red: rarefaction. Blue: condensation. Markings to the same click are distributed at points separated by 1/frequency (e.g., the denoted interval of 1 ms at 1 kHz). The points from the rarefaction and condensation clicks neatly interleave. Straight dash line indicates the additional 1 ms.

4.2.4 Various tone bursts

The combined Fourier patterns (plus 1 ms) to single tone bursts of 0.5, 1, 2 and 4 kHz are calculated and shown in Figure 55. The results show that the markings are better synchronized (distributed close to some specific points with a smaller variation) at or around the stimulus frequency. The markings have larger variation at frequencies away from the stimulus frequency.

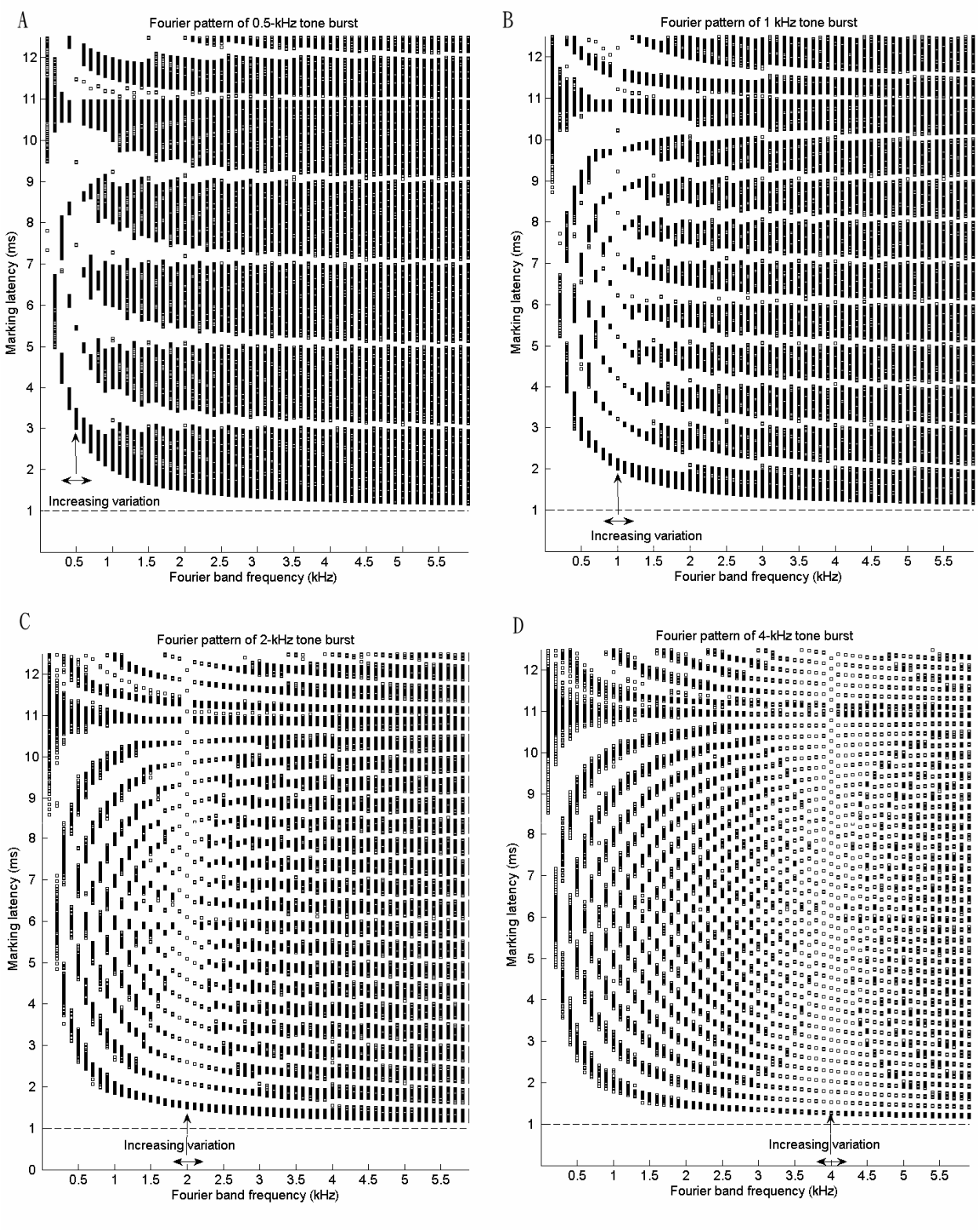


Figure 55. Combined Fourier patterns (plus 1 ms) to single tone bursts of 0.5, 1, 2 and 4 kHz. The tone bursts are generated by x(n) = sin(2 π · f · n/40000), where f = 0.5, 1, 2 or 4 kHz in panels A to D,

respectively, 1≦n≦400. Sampling rate = 40 kHz. Window size = 400 points. Arrows indicate the stimulus frequency and the best synchronized band.

The markings around the stimulus frequency are synchronized to specific points that are separated by the inverse of the stimulus frequency. For example in Figure 55C, in which the stimulus frequency is 2 kHz, the markings around 2 kHz are synchronized to points that are approximately separated by 1/ 2 kHz (seconds) = 0.5 ms.

Previous animal studies show that, in a single auditory nerve fiber, the dispersion of spikes around the mean becomes larger as the frequency of the stimulating tone decreases (see 2.2.3). The results reveal the same characteristic that, at a given frequency, the variation becomes larger as the stimulus frequency decreases (Figure 56). For example at 2 kHz, the variation is the smallest when the stimulus is 2 kHz, is median when the stimulus is 1 kHz, and is the largest when the tone burst is 0.5 kHz.

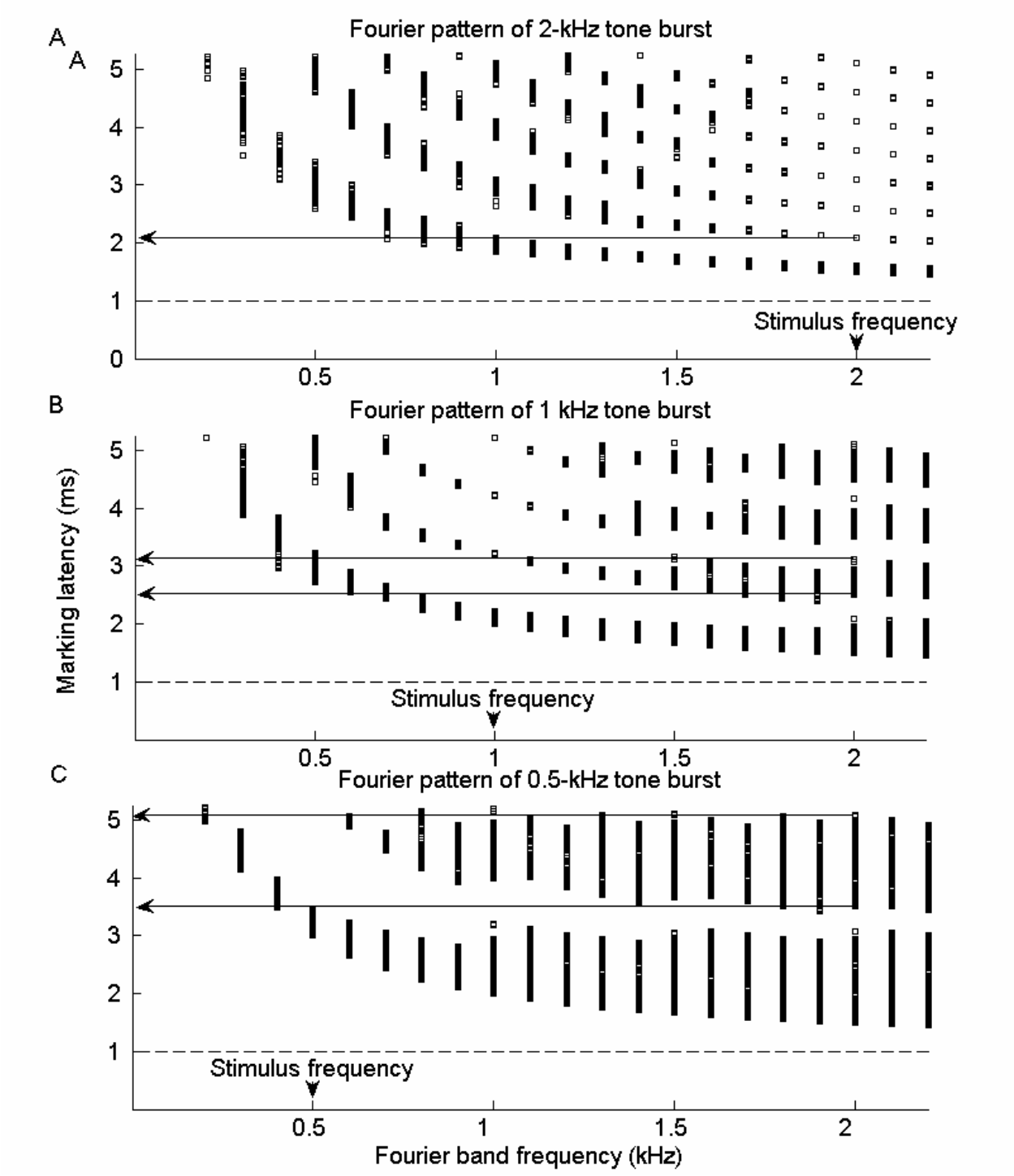


Figure 56. At a given frequency, the variation becomes larger as the stimulus frequency decreases. A. Smallest variation at 2 kHz when the stimulus is 2 kHz. B. Median variation at 2 kHz when the stimulus is 1 kHz. C. Largest variation at 2 kHz when the stimulus is 0.5 kHz. Arrows indicate the increasing variations. The Fourier patterns are enlarged from the previous figure.

5.0 DISCUSSION

For each of the stimuli considered, the latency at each frequency predicted by the Fourier pattern is highly correlated with the latency in the auditory nerve fiber that has the same characteristic frequency. In other words, the predicted latencies by Fourier pattern are significantly related to neural latencies across stimulus (a rarefaction click, a condensation click and various tone bursts), characteristic frequency (about 0.1 to 10 kHz), time sections (first spikes and after first spikes), and species (chinchillas, barn owls and to the extent they represent). Several features of these findings are now considered and the results correspondingly interpreted according to key issues that originally motivated this work.

5.1 PHASE ENCODING AMONG AUDITORY NERVE FIBERS

The results show that the Fourier pattern varies among different signals. The rarefaction click, the condensation click, and the tone bursts have different Fourier patterns. While endeavoring to control all other variables (sampling rate, window size and calculating method), the results indicate the significant effect of stimulus on Fourier pattern among the rarefaction click, the condensation click, and the tone bursts.

All correlations relevant to Hypothesis 1 under examination proved to be positive, namely of high value and thus statistically significant, supporting the hypothesis. Specifically,

the predicted latency pattern by the Fourier pattern is significantly related to the neural latency pattern in the auditory nerve fibers across the tested range of characteristic frequency among the tested stimuli.

Because the neural latency patterns of the selected two species to the stimuli tested are representative to a considerable extent of other animal species [22], the high correlation held in both species suggest that the predicted latency patterns by Fourier pattern are highly related to the neural latency patterns to various stimuli across different species, to a considerable extent.

In a correlational study, the square of the correlation coefficient shows the proportion of variation in one of the variables that is accounted for by the other. The results reveal that the Fourier pattern, a mathematical mechanism in Fourier analysis, accounts for a high proportion of variation in the latency pattern in the auditory nerve fibers. Specifically, to a rarefaction click, r^2 equals 0.95 for regression and 0.68 for raw data. To a condensation click, r^2 equals 0.99 for regression and 0. 82 for raw data. To tone bursts, r^2 equals 0.99 for chinchillas' regression and 0. 99 for chinchillas' raw data, and 0.99 for barn owls' regression and 0. 93 for barn owls' raw data. Thus, the results indicate that the Fourier pattern, and hence phase encoding, accounts for a high proportion of the variance in auditory neural responses under scrutiny.

In the conventional view, different positions of the cochlear partition are accessed via the delay line imposed by the propagation of a traveling wave in mammals [1]. Interestingly, results from animal studies have shown direct evidence that an identical auditory neuron responds with different latencies to different stimuli such as rarefaction and condensation clicks (e.g., [85, 88]). For example, *the same auditory unit* # 187 (from their raw data in [88]) with the characteristic frequency of 189 Hz responded at 2.36 ms to a rarefaction click (Figure 57A), but responded at 5.08 ms to a condensation click (Figure 57D). Nevertheless, auditory neurons with *the same*

characteristic frequency of 189 Hz in different species respond similarly with the latency of about 3 ms to a rarefaction click (Figure 22, the 1-ms neural and synaptic transmission was subtracted in the figure).

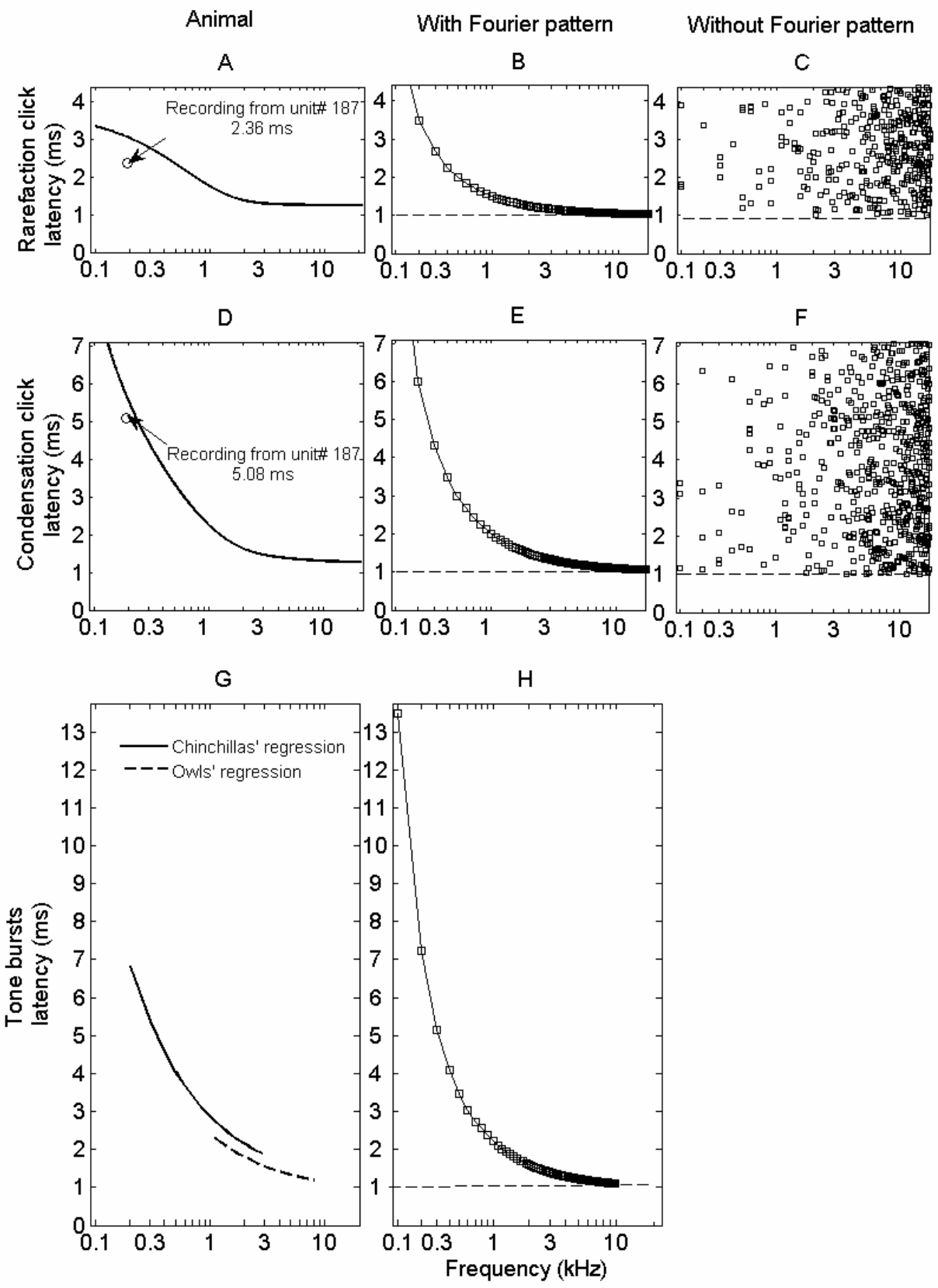

Animal
With Fourier pattern
Without Fourier pattern
A
B
C
D
E
F
G
H
Rarefaction click latency (ms)
Condensation click latency (ms)
Tone bursts latency (ms)
Recording from unit# 187
2.36 ms
Recording from unit# 187
5.08 ms
Chinchillas' regression
Owls' regression
Frequency (kHz)

Figure 57. The changeable delay line in animal recordings vs. the corresponding Fourier pattern. An identical auditory neuron responds with different latencies to different stimuli (A,D and G show response regressions to a rarefaction click, a condensation click and tone bursts, respectively). Arrows indicate the same auditory unit # 187 (from Temchin et al., 2005) with the characteristic frequency of 189 Hz responded at 2.36 ms to a rarefaction click (A), but responded at 5.08 ms to a condensation click (D). The predicted latency patterns in B, E and H are all significantly related to the neural latency patterns in A, D and G, respectively. The latencies without the Fourier pattern (C and F, which are calculated by randomly marking at an average rate of 60 Hz for 10 repetitions) are not related to the neural latencies.

As mentioned above (2.1.4), the latency pattern in auditory nerve fibers to a stimulus is mainly accounted for by the anatomical and physiological interactions in the cochlea. For example, basal neurons are usually discharged sooner because they are anatomically closer to the middle ear. For another example, at a given cochlear position, a rarefaction click usually evokes with a shorter latency than a condensation click in mammals, because it pulls out the eardrum and displaces the basilar membrane to scala vestibuli [122, 215].

The above two examples suggest that the delay line imposed by the propagation of a traveling wave in mammals is *species-independent* to an identical stimulus, but is *stimulus-specific* for an identical neuron. In other words, the delay line imposed by the anatomical and physiological interactions in the hearing organs can change. There seems to be a theoretical delay line that changes only when the stimulus differs but does not change among various species while the stimulus is the same.

The results of the present study suggest that these anatomical and physiological interactions in the hearing organs across different species actually bespeak a common mathematical-like, latency-pattern analyzer among auditory nerve fibers when stimulated by an identical stimulus. Such a mechanism serves to encode phase information of the stimulus.

Consequently, in contrast to the several designs and dimensions of hearing organs across phyla and classes of species, Nature has been conservative, namely in its endeavor to maintain an essentially universal phase encoder. Different priorities seem to have driven the evolution of hearing-organ mechanics, culminating with the mammalian system, which at the first level (at least) constitutes an elaborate broadband impedance matching device subsequently enhanced to address still other priorities (e.g. keen sensitivity). From this perspective, phase encoding is concluded to be forethought, not an afterthought, of the design of hearing organs across species. Indeed, from this perspective, the challenge to Nature is to have made evolutionary modifications without degrading phase encoding, given the delay lines inherent to the most recent designs that serve mammalians.

5.2 PHASE ENCODING WITHIN AN AUDITORY NERVE FIBER

The specific spike distribution in a single auditory nerve fiber is hypothesized (see 2.4.2) to be accounted for by the Fourier pattern created by, in effect, heavily overlapped Fourier analyses (a process of moving-window Fourier analysis wherein the window is moved forward in a point-wise manner). The specific spike distribution in a single auditory nerve fiber to a stimulus is usually viewed as the temporal code to that stimulus.

5.2.1 Temporal code to single rarefaction or condensation click

To a single click, the results show that the markings at a given frequency have a specific distribution. The markings locate at some specific points. The number of these points, especially

at low frequencies, is much smaller than the number of total Fourier analyses (e.g., at 2 kHz, there are 21 points in Figure 49, which are from 400 analyses). The markings are precisely overlapped (having the same numbers in Matlab) at these points. Furthermore, these points are separated by the inverse of the Fourier band frequency. These findings are strongly consistent with the classical click responses in auditory nerve fibers.

The classical click responses in auditory nerve fibers, however, are usually reported at low frequencies (< 4 kHz) (see 2.2.5). The results in the present study thus are inevitably in accord with the classical neural responses, yet show patterns also at high frequencies not yet reported in the literature (to the author's knowledge). This difference can and should be explored by several investigations in the future.

First, interpreting the difference observed at high frequencies requires considerations of intensity information. A theoretically ideal click, as used in the present study, has homogeneous energy across all frequencies. The amplitudes (intensity information) of sinusoid components across all frequencies are the same. Because markings are distributed at more points at high frequencies, the mean intensity reflected by one of these points is weaker. For example in Figure 49, there are 41 points at 4 kHz but only 6 points at 0.5 kHz. Therefore, if the intensity represented at high frequencies is close to or less than the neuron's threshold, there might not be a peak in the peristimulus-time histogram.

The difference observed at high frequencies also may be due to inadequate temporal resolution in previous studies. For example, a minimum resolution of about 50 μs may be required to identify the possible existed pattern at 4 kHz. And the demand of temporal resolution increases as the characteristic frequency increases.

Also, future studies may be required to investigate the latency pattern at high frequencies in various species with different cochlear mechanics. Different animal species may synchronize auditory neural activity with different levels of precision at high frequencies. For example, the tone-burst responses in barn owls are featured with significant phase locking at high frequencies up to 10 kHz [25]--uncommon in other species, which may be related to a greater need for precise localization for catching a prey in total darkness. The pattern shown at high frequencies thus may be observed only in some species.

In summary, at a given frequency, the markings from consecutive Fourier analyses to a single click are distributed at specific points. Inspecting the exact numbers in Matlab, these points are separated by the inverse of the Fourier band frequency. Also, these points to rarefaction and condensation clicks neatly interleave. The results clearly demonstrate the characteristics of the classical click responses in animal studies and thus strongly support Hypothesis 2. The results suggest that the temporal code in a single auditory nerve fiber to a click is accounted for by phase encoding in a mechanism that mimics the process of consecutive Fourier analyses.

5.2.2 Temporal code to a single tone burst

At least for low-characteristic-frequency fibers around the stimulus frequency, in the peristimulus-time histogram in animal studies, the neural spikes to a single tone burst are usually synchronized to specific points. Specifically, an auditory neuron tends to fire at a particular phase of a stimulating tone, and the inter-spike intervals tend to occur at integer multiples of the inverse of the stimulus frequency. In addition, the dispersion around the mean becomes larger as the stimulus frequency decreases. In other words, the dispersion around the mean is larger in

auditory nerve fibers whose characteristic frequencies are away from the stimulus frequency, and smaller at or around the stimulus frequency (see 2.2.3). All such characteristics form a specific spike distribution in a given auditory nerve fiber.

The results of the present study show that, around the stimulus frequency, the markings have a specific distribution. The markings are synchronized to points that are separated by the inverse of the stimulus frequency. At a given frequency, the markings' dispersion around the mean increases as the stimulus frequency decreases. These results clearly are in accord with the findings in previous animal studies.

Similarly as discussed above, the intensity information is probably important when interpreting the predicted pattern to a tone burst. The intensity (amplitude) of any two sinusoid components of the tone burst can be different in a single Fourier analysis. The amplitude of the sinusoid components also may vary at a given frequency in consecutive Fourier analyses. In other words, the intensity information between any two of the sinusoid components to the tone burst can be different. Therefore, the intensity differences need to be considered when interpreting the predicted latency pattern to a tone burst. Again, if the intensity represented by a predicted point is close to or less than the neuron’s threshold, there might not be a peak in the peristimulus-time histogram.

In conventional view, the specific distribution of the spikes in a single auditory neuron to a tone burst represents the temporal code of frequency encoding (ultimately, pitch perception), and also the phase-lock phenomenon (see 2.2.3). Specifically, the temporal code of a tone burst in animal studies can be characterized as the specific points synchronized and the level of spike dispersion.

The results of the present study demonstrate that, around the stimulus frequency, the markings are synchronized to points that are separated by the inverse of the stimulus frequency. Also, the markings' dispersion around the mean is larger at frequencies away from the stimulus frequency, and smaller at or around the stimulus frequency. From the intensity considerations and the observed pattern of latency groupings, the results are consistent with the observed neural latency pattern and, again, support Hypothesis 2. The findings suggest that the temporal code of a tone burst (ultimately, pitch perception) and also the phase-lock phenomenon are accounted for by a phase encoding process in the hearing organ that mimics the process of consecutive Fourier analyses.

6.0 CONCLUSION AND IMPLICATIONS

Among various frequencies, each marking accounts for a large percentage of a corresponding peak latency in the peristimulus-time histogram. For each of the stimuli considered (a rarefaction click, a condensation click, and various tone bursts), the latency predicted by the Fourier pattern is highly correlated with the latency observed in the auditory nerve fiber across frequencies and species, at least to a considerable extent. The results suggest that the hearing organ analyzes not only amplitude spectrum but also phase information in Fourier analysis.

At a given frequency, the results show that the markings are synchronized to stimulus-specific points. To single clicks, these points are separated by the inverse of the Fourier band frequency. Also, these points to rarefaction and condensation clicks neatly interleave. To single tone bursts, these points are separated by the inverse of the stimulus frequency. Also, the markings' dispersion around the mean is larger at frequencies away from the stimulus frequency, and smaller at or around the stimulus frequency. These results support that the specific spike distribution (i.e., temporal code) in a single auditory nerve fiber is accounted for by phase encoding in a mechanism that mimics the process of consecutive Fourier analyses.

This phase-encoding mechanism that mimics the process of consecutive Fourier analyses is proposed to be the common mechanism that, in the face of species differences in peripheral auditory hardware, accounts for the considerable similarities across species in their latency-by-frequency functions, in turn assuring optimal phase encoding in the hearing.

7.0 FUTURE STUDIES

In addition to phase and frequency, the third kind of fundamental information of a sound is intensity. Future studies are required to investigate how the hearing organ encodes intensity in outputs that reflect synchronization of auditory neural activity consistent with the Fourier pattern.

Another required study at the theoretical level is, in the cochlear-implant-like envelope detection processing, whether rearranging sinusoid components with the Fourier pattern improves correlation of temporal fine structure. This can be tested mathematically and is essential because that, as Wilson and Dorman indicated [6], the loss of temporal fine structure information is responsible for a major portion of the difficulty and bad outcome of implant users.

Implementing the Fourier pattern in cochlear implants should replicate the Fourier pattern in auditory nerve fibers, at least to a considerable extent. If rearranging sinusoid components with the Fourier pattern significantly improves correlation of temporal fine structure, implementing the Fourier pattern may convey accurate temporal fine structure information and improve associated auditory perceptions. This can be verified in actual implant users.

A possible approach is to manipulate the stimulating pattern with and without the Fourier pattern to investigate the effect on various perceptions, such as speech recognition in noise, music perception, lexical tone recognition, gender identification and talker recognition.

In addition to the research about biological representations, theoretical validations and clinical implementation, there is an area in basic hearing science that requires more

investigations. The concept of Fourier pattern opens a new avenue in understanding auditory perceptions derived from phase. For example, since each tone burst (which contributes to a pitch perception) has a specific Fourier pattern, now it is possible to unify the amplitudes of all the decomposed sinusoids of the tone burst and reconstruct two signals from the same set of sinusoids with and without the specific Fourier pattern to investigate the question whether there is a (Fourier) pattern-only pitch perception. This will be a "pattern" pitch perception, because the unified amplitude of each sinusoid theoretically provides no frequency information contributing to the pitch perception.

Phase, as one of the three kinds of fundamental information of a sound, affects the relationship among various auditory variables and broadly impacts auditory functions. The proposed mechanism of phase encoding suggests new relationships and makes new predictions among these auditory variables and functions (Figure 58). It may guide future studies to investigate the suggested relationships and evaluate the new predictions.

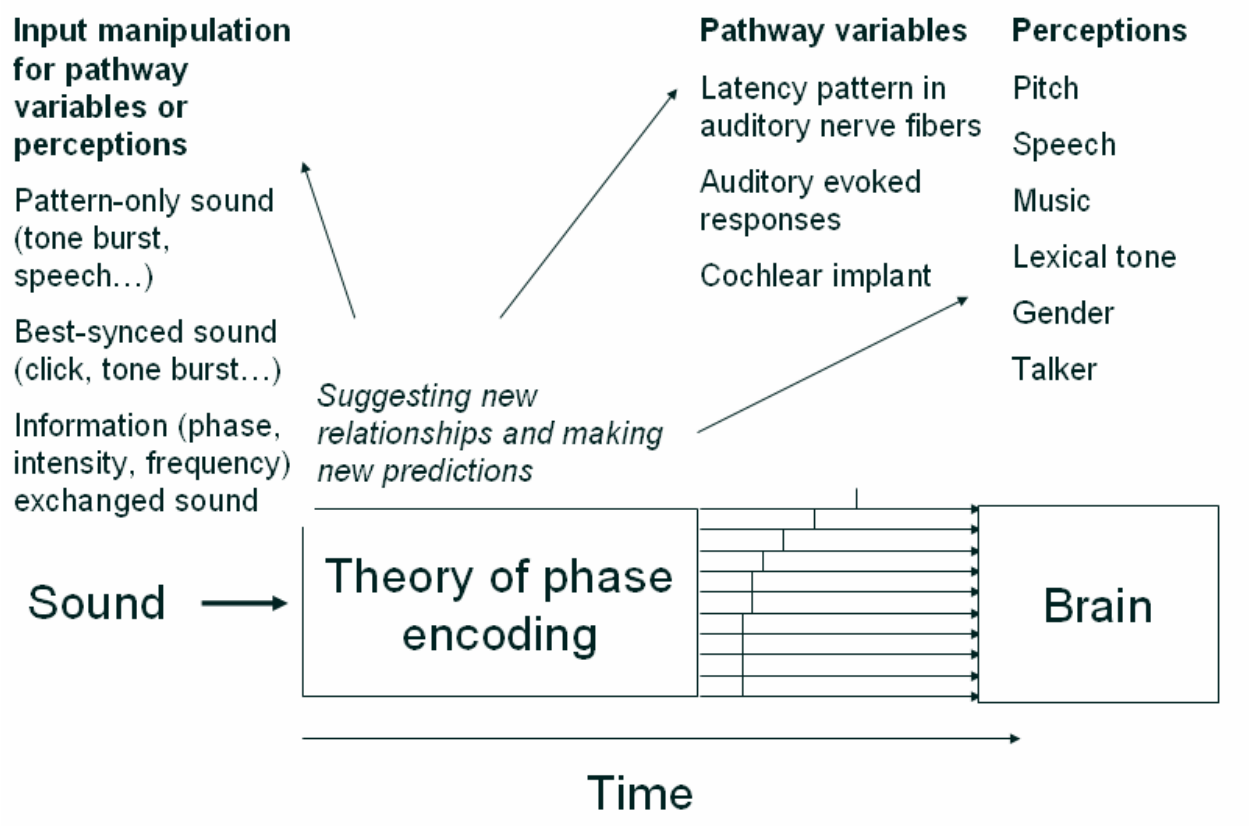


Figure 58. The theory of phase encoding suggests new relationships and makes new predictions among various auditory variables and functions.

In summary, future studies in interdisciplinary areas among mathematics, auditory physiology, bioengineering and otolaryngology are required to find out adequate pieces of the puzzle of the similar phase encoding in hearing organs across species and individuals, and hopefully to put it together. A likely picture is proposed in which hearing organs of animals across species, which survive long-term biological evolutions, encode a sound by synchronizing auditory neural activity with a common pattern, probably the Fourier pattern, by which the center auditory system can reconstruct and comprehend the original sound accurately. Better hearing hence helps the animals in competitions such as preventing risks and finding food to win the natural selections.

APPENDIX A

A.1 CONNECTION BETWEEN FOURIER PHASE AND COSINE PHASE

Taking the same computational steps to calculate a marking latency, as in 3.2.1, it can be further demonstrated how Fourier phase (the angle of a frequency term, a complex number, to the real axis in its polar expression) relates to cosine phase (the starting phase of a waveform generated by cosine function), assuming a sampling rate of f_s Hz for N points to yield the input signal $x(n)$ in an N-point analysis window:

A.1.1 $\boxed{x(n) \qquad 0 \le n \le N-1}$

Figure 59A and E show a rarefaction click and a tone burst (that have been used in Figure 33 and Figure 10 respectively), with f_s = 40000 Hz and N = 400 (chosen arbitrarily for demonstration), analyzed as follows:

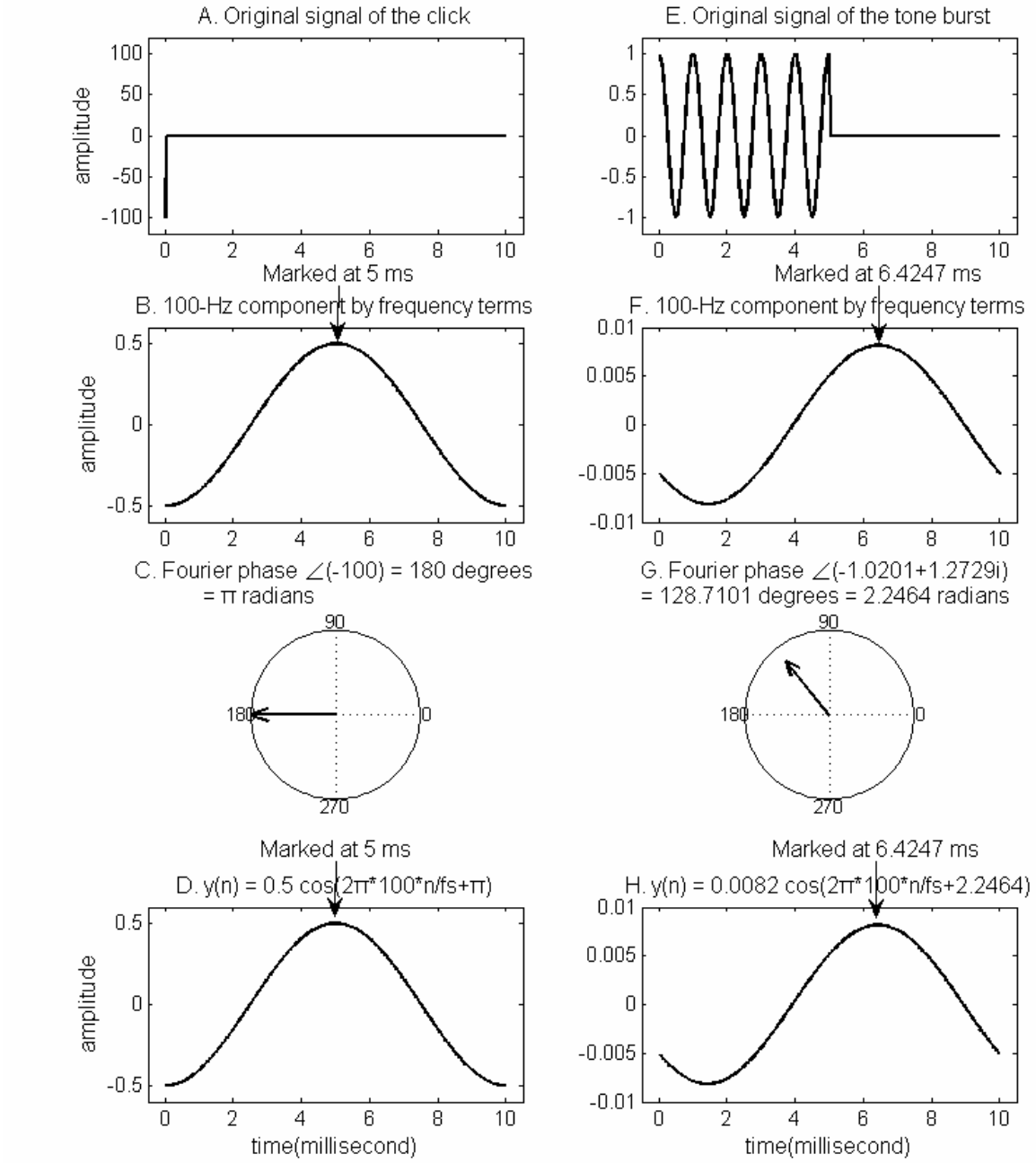


Figure 59. Fourier phase vs. cosine phase of the sinusoid that is synthesized from that Fourier component. The 100-Hz component can be synthesized either from a positive and a negative frequency term (in B and F) or by the cosine function: y(n)=Acos(2πfn/fs+θ), where 0 ≦ n <400, A=2|X(c)|/N, θ = ∠X(c), f =

100 and fs (sampling rate) = 40000, in D and H. X(c) is a Fourier frequency term. Window size (N): 400 points. A. Original signal of the rarefaction click. B. The 100-Hz component of A, synthesized from frequency terms. C. The Fourier phase of X(1) of the click. D. A cosine function plotted by the above equation of y(n)=Acos(2πfn/fs+θ), where A=2|X(1)|/400 = 0.5 and θ equals the Fourier phase of X(1) = π. E. Original signal of the 1-kHz tone burst. F. The 100-Hz component of E, synthesized from frequency terms. G. The Fourier phase of X(1) of the tone burst. H. A cosine function plotted by the above equation of y(n)=Acos(2πfn/fs+θ), where A=2|X(1)|/400 = 0.0082 and θ equals the Fourier phase of X(1) = 2.2462 radians.

An N -point FFT is then computed for the input signal $x(n)$:

A.1.2 $$x(n) \xrightarrow{\text{N-point FFT}} X(k) \qquad 0 \le k \le N-1$$

According to the definition of the inverse discrete Fourier transform (IDFT),

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} \mathrm{X}(k) \cdot \mathrm{e}^{i\frac{2\pi \mathrm{nk}}{N}}$$

synthesize a temporal waveform $x_c(n)$ as per the equation below, where c indicates the "component", from a positive and a negative frequency term. Because the DC component (c = 0) is a straight line, thus bearing no phase information, it is ignored:

A.1.3 $$x_c(n) = \frac{1}{N}(\mathrm{X}(c) + \mathrm{X}(N-c)) \cdot \mathrm{e}^{i\frac{2\pi \mathrm{mk}}{N}} \qquad 1 \le \mathrm{c} < \frac{\mathrm{N}}{2}$$

For instance, Figure 59B and F show $x_1(n)$ of the click and tone burst, respectively. The 100-Hz component is chosen arbitrarily for demonstration.

To investigate the relation between the Fourier phase and the cosine phase, assume a *new cosine function*:

A.1.4 $$y(n) = A \cdot \cos(2\pi f \cdot \frac{n}{f_s} + \theta) \qquad \mathrm{A} \ge 0,\ 0 \le \theta < 2\pi$$

where A is the magnitude and θ is the starting phase (cosine phase). Letting $x_c(n) = y(n)$, the equation becomes:

$$\frac{1}{N}(\mathrm{X}(c) + X(N-c)) \cdot \mathrm{e}^{i\frac{2\pi \mathrm{mk}}{N}} = A \cdot \cos(2\pi f \cdot \frac{n}{f_s} + \theta).$$

At the time $n = 0$, the value of the temporal waveform, $x_c(0)$, is determined by the phase, and the first amplitude maximum of $x_c(n)$ is delayed accordingly. When $n = 0$, the value of the synthesized waveform is:

$$x_c(0) = \frac{1}{N}(X(c) + X(N-c)) \cdot \mathrm{e}^0 = \frac{X(c) + X(N-c)}{N}$$

which equals $\frac{2|X(c)|\cos(\angle X(c))}{N}$, because $X(c)$ and $X(N-c)$ are complex conjugates (e.g., $X(1)$ = -1.0201+1.2729i and $X(399)$ = -1.0201-1.2729i for the tone burst in Figure 59E, and

$$\frac{X(c)+X(N-c)}{N} = \frac{-1.0201+1.2729i-1.0201-1.2729i}{40000} = \frac{-2.0402}{40000} = \frac{2\times 1.6312\times\cos(2.2464)}{40000}$$

$$= \frac{2|X(c)|\cos(\angle X(c))}{N}).$$

When $n=0$, the value of the synthesize waveform from the new cosine function is:

$$y(0) = A\cdot\cos(2\pi f\cdot\frac{0}{f_s}+\theta) = A\cdot\cos(\theta).$$

Because $x_c(n) = y(n)$, i.e., $x_c(0) = y(0)$, the equation becomes

$$\frac{2|X(c)|\cos(\angle X(c))}{N} = A\cdot\cos(\theta).$$

One condition that satisfies $\frac{2|X(c)|\cos(\angle X(c))}{N} = A\cdot\cos(\theta)$ is when the Fourier phase, $\angle X(c)$, equals the cosine phase of the synthesized sinusoid, θ, and the magnitude A equals $\frac{2|X(c)|}{N}$. For example, submitting the Fourier phase (π) and $\frac{2|X(c)|}{N} = \frac{2|X(1)|}{400} = \frac{2|-100|}{400} = 0.5$ to θ and A to the equation in A.1.4, respectively, the waveform from the new cosine function (Figure 59D) is the same as that from the frequency terms (Figure 59B). Another example, submitting the Fourier phase (128.7 degrees or 2.2464 radians) and $\frac{2|X(c)|}{N} = \frac{2|X(1)|}{400} = \frac{2|\text{-}1.0201+1.2729\text{i}|}{400} = 0.0082$ to θ and A, respectively, the waveform from the new cosine function (Figure 59H) is the same as that from the frequency terms (Figure 59F).

A.2 SELECTION OF THE HYPOTHESIZED REFERENCE

The hypothesized reference is set to mark the first amplitude maximum in a sinusoidal signal.

The sinusoid signal $y(n) = A \cdot \cos(2\pi f \cdot \frac{n}{f_s} + \theta)$ has multiple amplitude maxima over time when $\cos(2\pi f \cdot \frac{n}{f_s} + \theta) = 1$. That is, a point is an amplitude-maximum when $2\pi f \cdot \frac{n}{f_s} + \theta$ equals multiples of 2π (e.g., 0, $\pm 2\pi$, and $\pm 4\pi$). Letting

$$2\pi f \cdot \frac{n}{f_s} + \theta = 2\pi l$$

where l is an integer and one of the angles, $2\pi l$, is *the hypothesized reference*. The equation becomes

$$\frac{n}{f_s} = \frac{2\pi l - \theta}{2\pi f}$$

which indicates the timing for each amplitude-maximum.

The first amplitude maximum occurs between time 0 and the period of the sinusoid. In other words, the range of the first amplitude maximum is

$$0 < \frac{2\pi l - \theta}{2\pi} \cdot \frac{1}{f} \le \frac{1}{f}$$

Therefore, the range of $2\pi l - \theta$ is

$$0 < 2\pi l - \theta \le 2\pi$$

Because the range of θ is defined in A.1.4 as $0 \le \theta < 2\pi$, then $2\pi l - \theta$ must be smaller than or equal to $2\pi l - 0$, or $2\pi l$, and $2\pi l - \theta$ must be bigger than $2\pi l - 2\pi$. The range of $2\pi l - \theta$ becomes

$$0 = 2\pi l - 2\pi < 2\pi l - \theta \le 2\pi l - 0 = 2\pi$$

where l equals 1.

In other words, only when l = 1 or $2\pi l$ equals 2π, the timing of the maxima $\frac{2\pi l - \theta}{2\pi f}$ is within the range of $0 < \frac{2\pi l - \theta}{2\pi} \cdot \frac{1}{f} \leq \frac{1}{f}$. When $l \leq 0$, the timing is less than 0. When $l \geq 2$, the timing is larger than the period, $\frac{1}{f}$. Therefore, l is selected as 1 and the hypothesized reference is set to 2π.

A.3 LATENCY OF THE FIRST AMPLITUDE MAXIMUM OF

$$y(n) = A \cdot \cos(2\pi f \cdot \frac{n}{f_s} + \theta)$$

When the hypothesized reference is set to 2π, the latency of the first amplitude maximum of the waveform from the new cosine function in A.1.4 can be calculated by the timing of $\frac{2\pi - \theta}{2\pi f}$, as illustrated in A.2.

For example in Figure 59D, the first amplitude maximum is at $\frac{2\pi - \theta}{2\pi f} = \frac{2\pi - \pi}{2\pi \cdot 100} = 0.005$ seconds (or 5 ms). For the other example in Figure 59H, the first amplitude maximum is at $\frac{2\pi - \theta}{2\pi f} = \frac{2\pi - 2.2464}{2\pi \cdot 100} = 0.0064247$ seconds (or 6.4247 ms).

A.4 SUMMARY

The latency of the first amplitude maximum of the signal from the new cosine function (as previously defined in A.1.4) is the same as that of the signal synthesized from Fourier frequency terms, when these two signals represent each other. When the Fourier phase, $\angle X(c)$, is submitted to replace the cosine phase, θ, the marking latency ($\frac{2\pi-\theta}{2\pi f}$) calculated from the cosine phase of the new cosine function equals the marking latency ($\frac{2\pi-\angle X(c)}{2\pi f}$) calculated by time-shift property of Fourier transform (see 3.2.1), according to the phase difference between the hypothesized reference 2π and the Fourier phase $\angle X(c)$. The results from these two computational methods are the same but the latter (calculation via the time-shift property of the Fourier transform) skips unnecessary conversion and simplifies the process.

APPENDIX B

CALCULATING FOURIER PATTERN OF A 1-KHZ TONE BURST

The tone burst, as shown in Figure 40, is generated by the equation $x(n) = \sin(2\pi f \cdot \frac{n}{f_s})$, where f = 1 kHz, $1 \le n \le 400$, f_s (sampling rate) = 40,000. Window size is set to 400 points.

The tone burst is then sampled with a moving window process wherein the window is moved forward in a point-wise manner, i.e., sampling the stimulus repeatedly one point at a time while moving forward in time from the onset of the stimulus (also see 3.2.2). There will be totally 799 windows (i.e., 400×2-1) that contain at least one non-zero point of the 400-point stimulus.

To demonstrate in an affordable space, processing in the first 100 windows will be detailed. This part of the illustration includes 100 individual inputs for the Fourier analyses, 100 individual Fourier patterns, and the figure when plotting all the 100 individual Fourier patterns together.

Figure 60 shows the first 100 individual inputs sampled from the moving window process. A fast Fourier transform is then calculated in each window for each 400-point input. Each window contain at least one non-zero point of the stimulus.

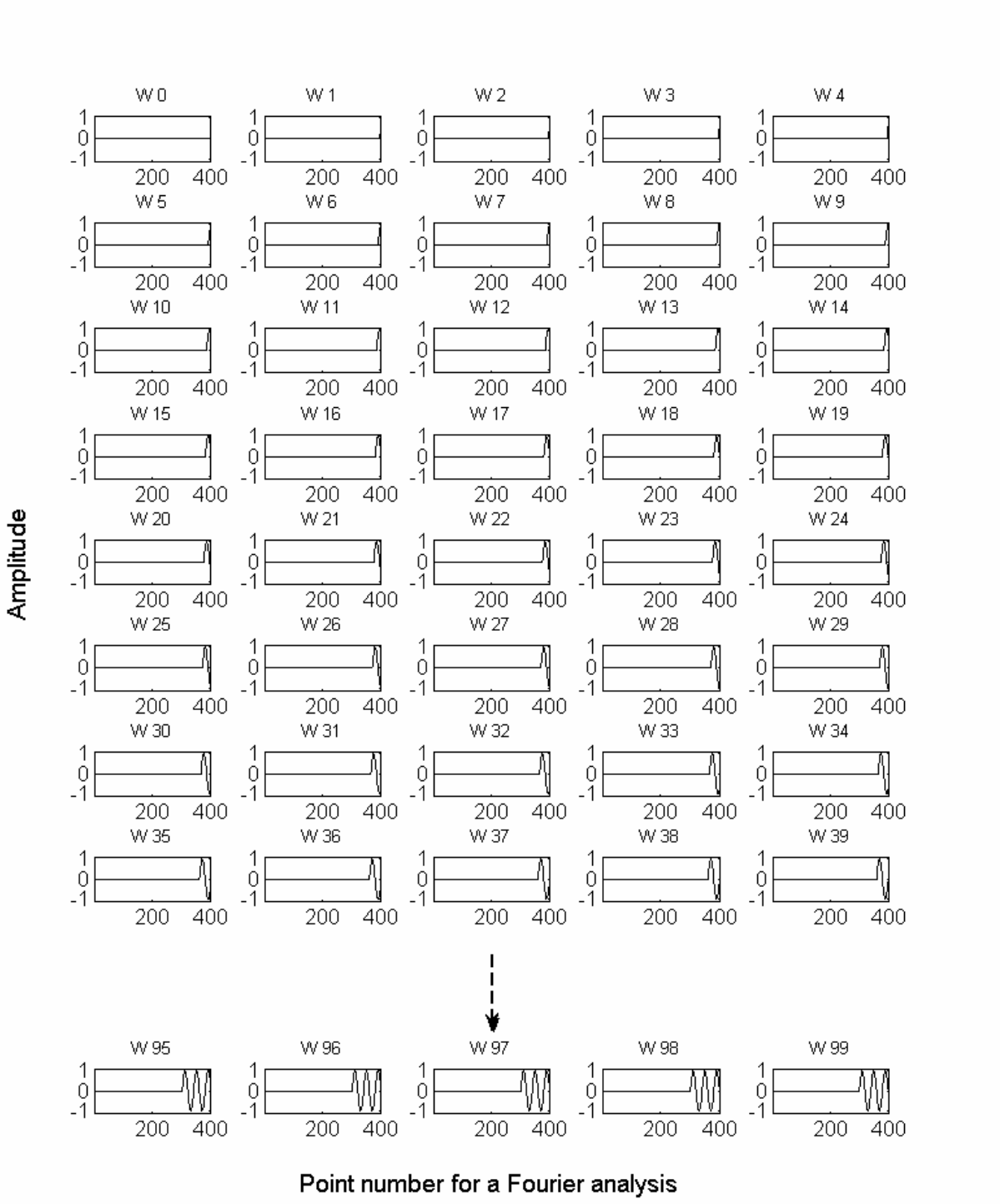


Figure 60. The first 100 individual inputs sampled from the 1-kHz tone burst. Each window contains at least one non-zero point of the stimulus.

The Fourier pattern of each window is then calculated as described in 3.2. For instance in window 99, the steps 1, 2 and 3 for calculating the marking latency at 1 kHz are illustrated in Figure 61. The first amplitude maximum of the 1-kHz sinusoid component is marked at 0.725 ms. In step 4, the time difference between window 99 and window 0, which equals $\frac{99-0}{40000} = 0.002475$ seconds (or 2.475ms), is then added to the marking latency, 0.725 ms, in this single analysis. Therefore, the marking latency at 1 kHz plus 1 ms (the synaptic and neural transmission; see 2.2.7) in window 99 is 0.725 + 2.475 + 1 = 4.2 ms.

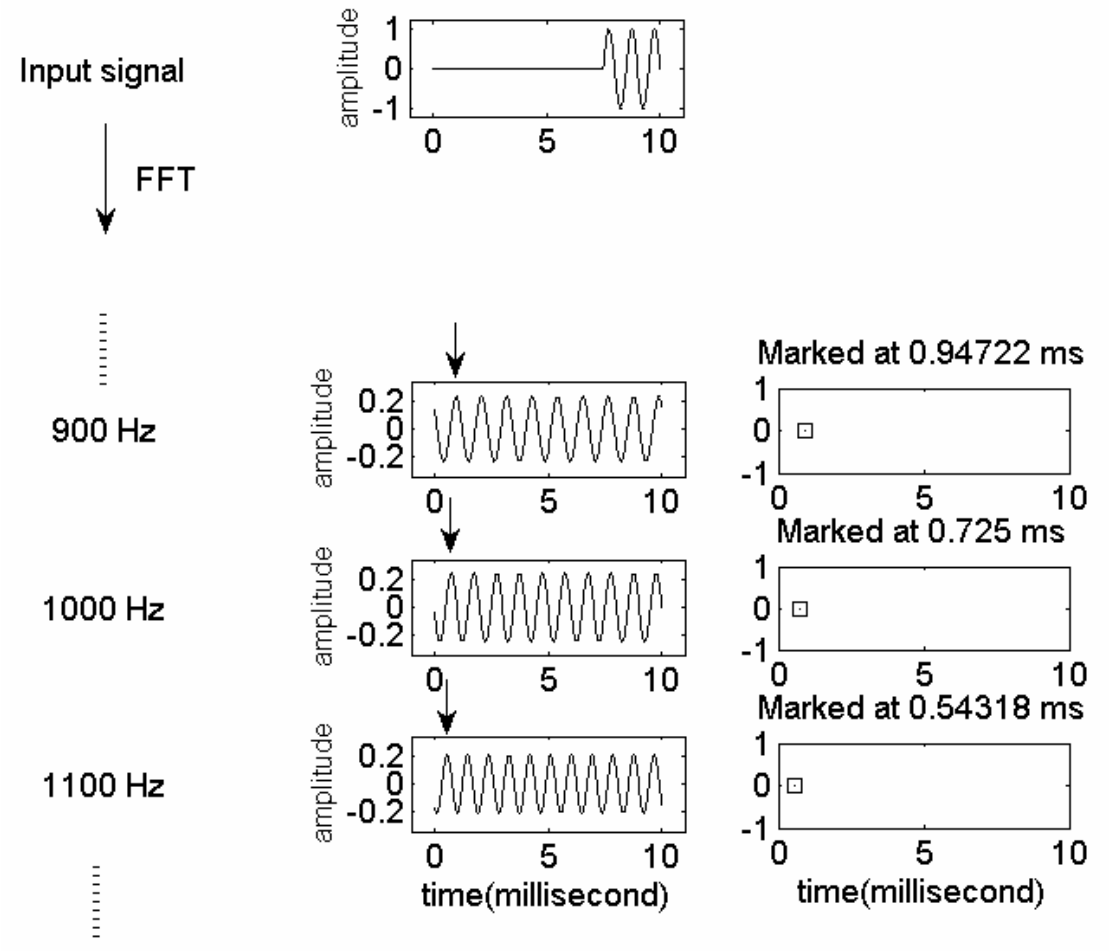


Figure 61. Fourier analysis in window 99 of the 1-kHz tone burst. Sampling rate: 40 kHz. Window size: 400 points. Only three sinusoid components around 1 kHz are shown. Arrows indicate the marking times.

To improve clarity when plotting them together later, the markings are colored, every five windows, in the order of yellow, cyan, red, green and blue, respectively. The first 100 individual Fourier patterns are shown in Figure 62, Figure 63, Figure 64 and Figure 65. The marking latency at 1 kHz plus 1 ms in window 99, which equals 4.2 ms, is indicated by the arrow in Figure 65.

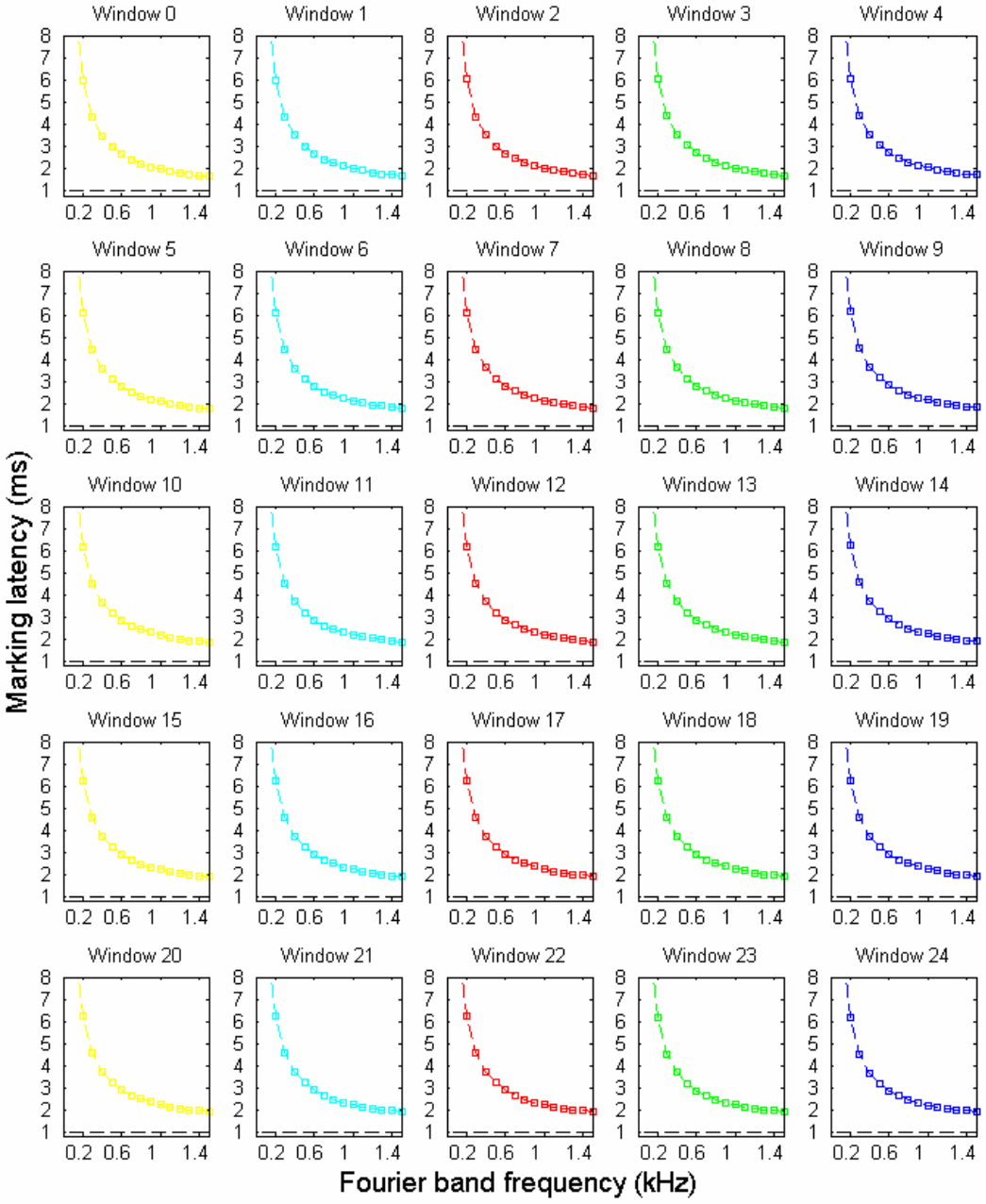


Figure 62. Windows 0 to 24 of the first 100 individual Fourier patterns of the 1-kHz tone burst. Markings in an identical window are in the same color for clarity when plotting them together later.

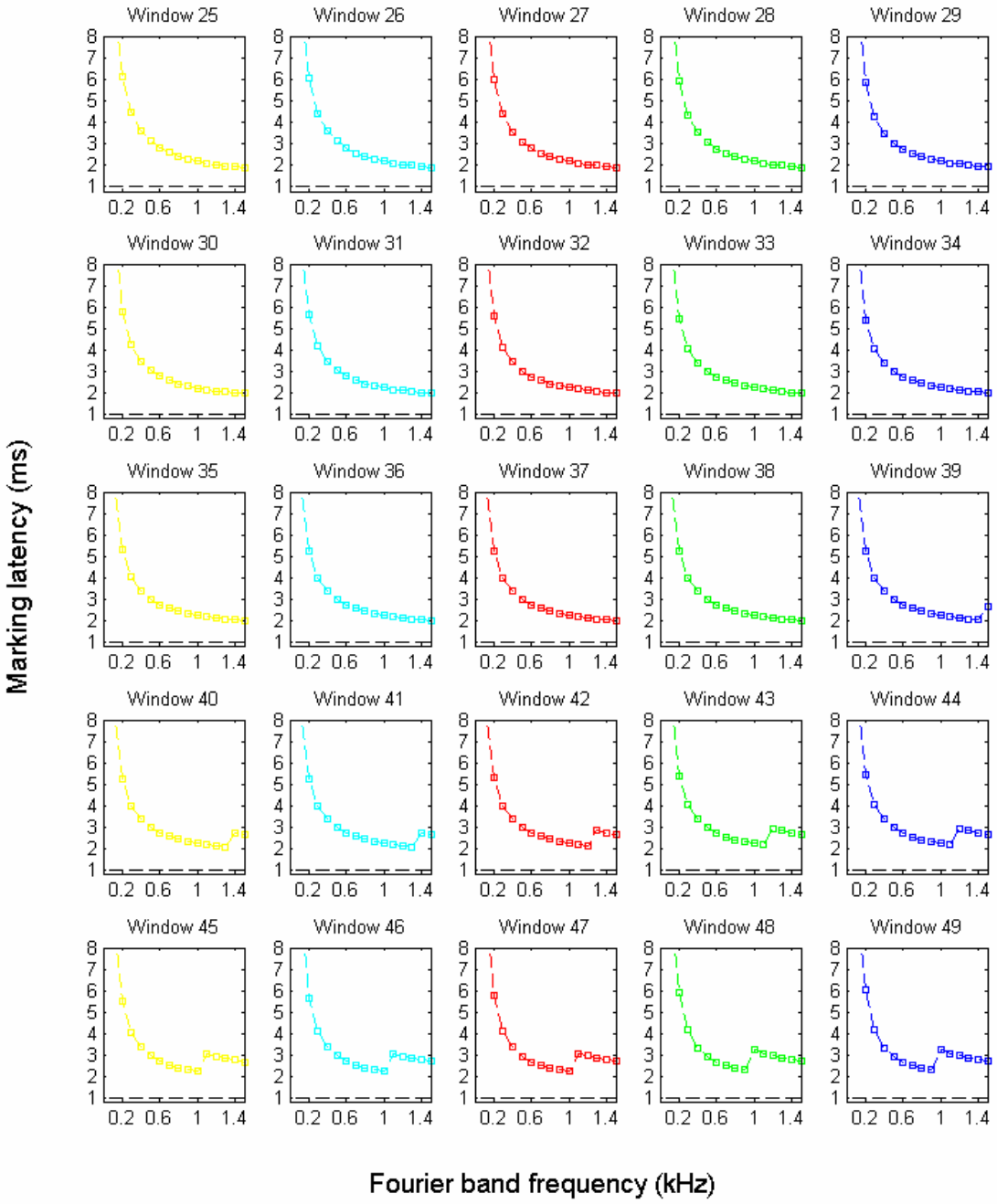


Figure 63. Windows 25 to 49 of the first 100 individual Fourier patterns of the 1-kHz tone burst. Markings in an identical window are in the same color for clarity when plotting them together later.

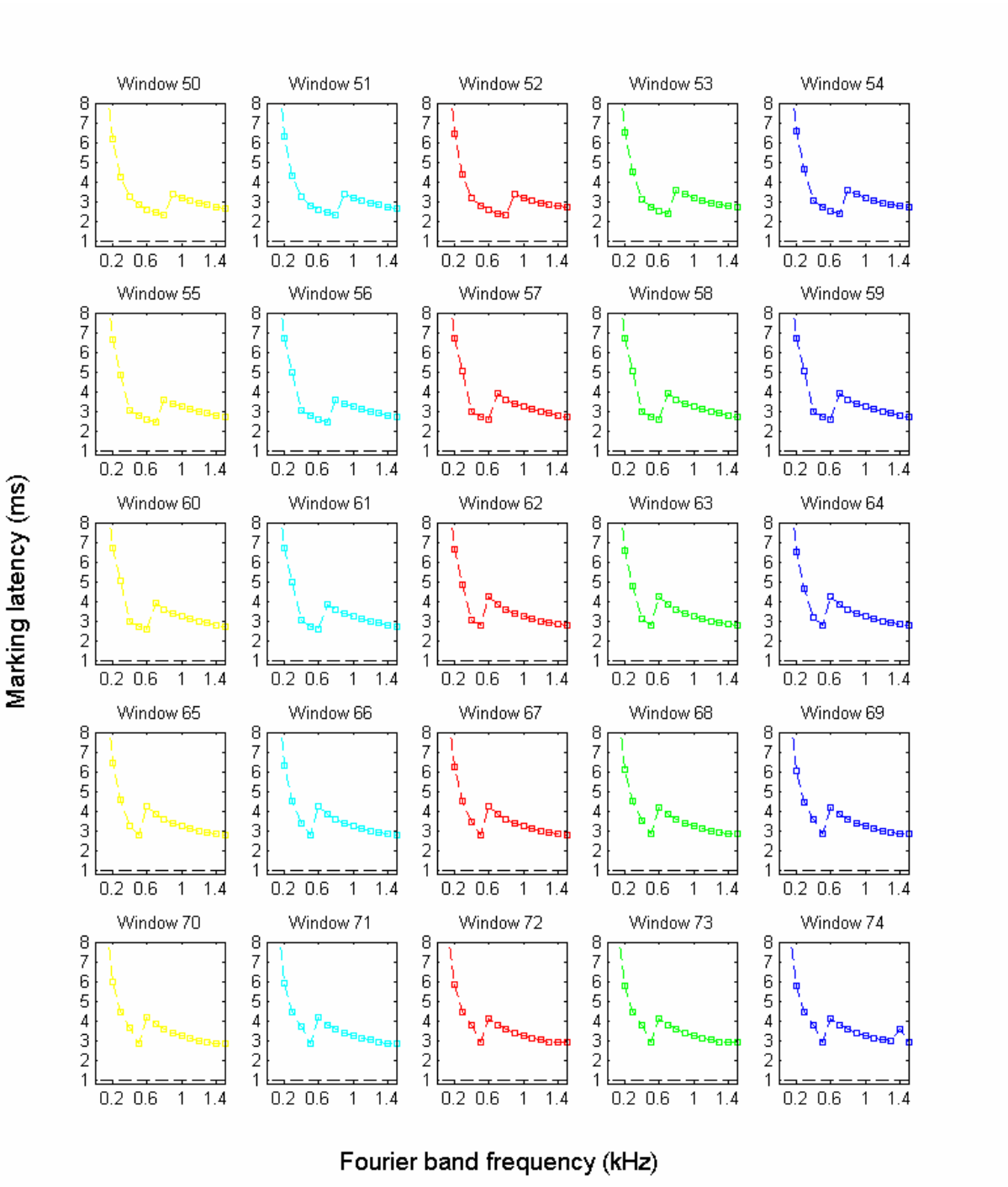


Figure 64. Windows 50 to 74 of the first 100 individual Fourier patterns of the 1-kHz tone burst. Markings in an identical window are in the same color for clarity when plotting them together later.

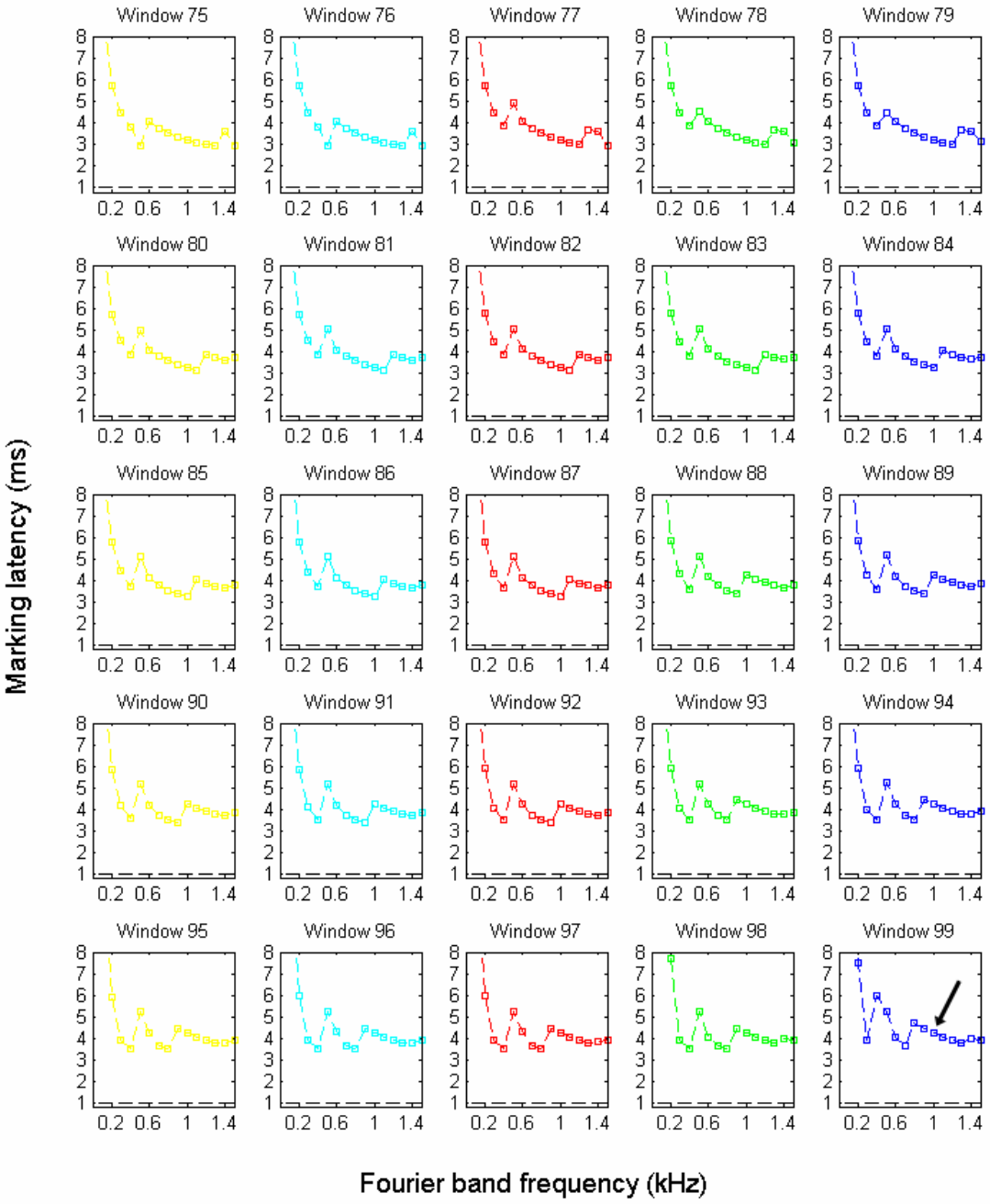


Figure 65. Windows 75 to 99 of the first 100 individual Fourier patterns of the 1-kHz tone burst. Markings in an identical window are in the same color for clarity when plotting them together later. Arrow indicates the predicted latency at 1 kHz in window 99.

Plotting all of the first 100 individual Fourier patterns together, Figure 66B shows the combined Fourier pattern from the first 100 windows. The dash line between two markings is removed in Figure 66A. The color coding remains the same as in the individual Fourier patterns.

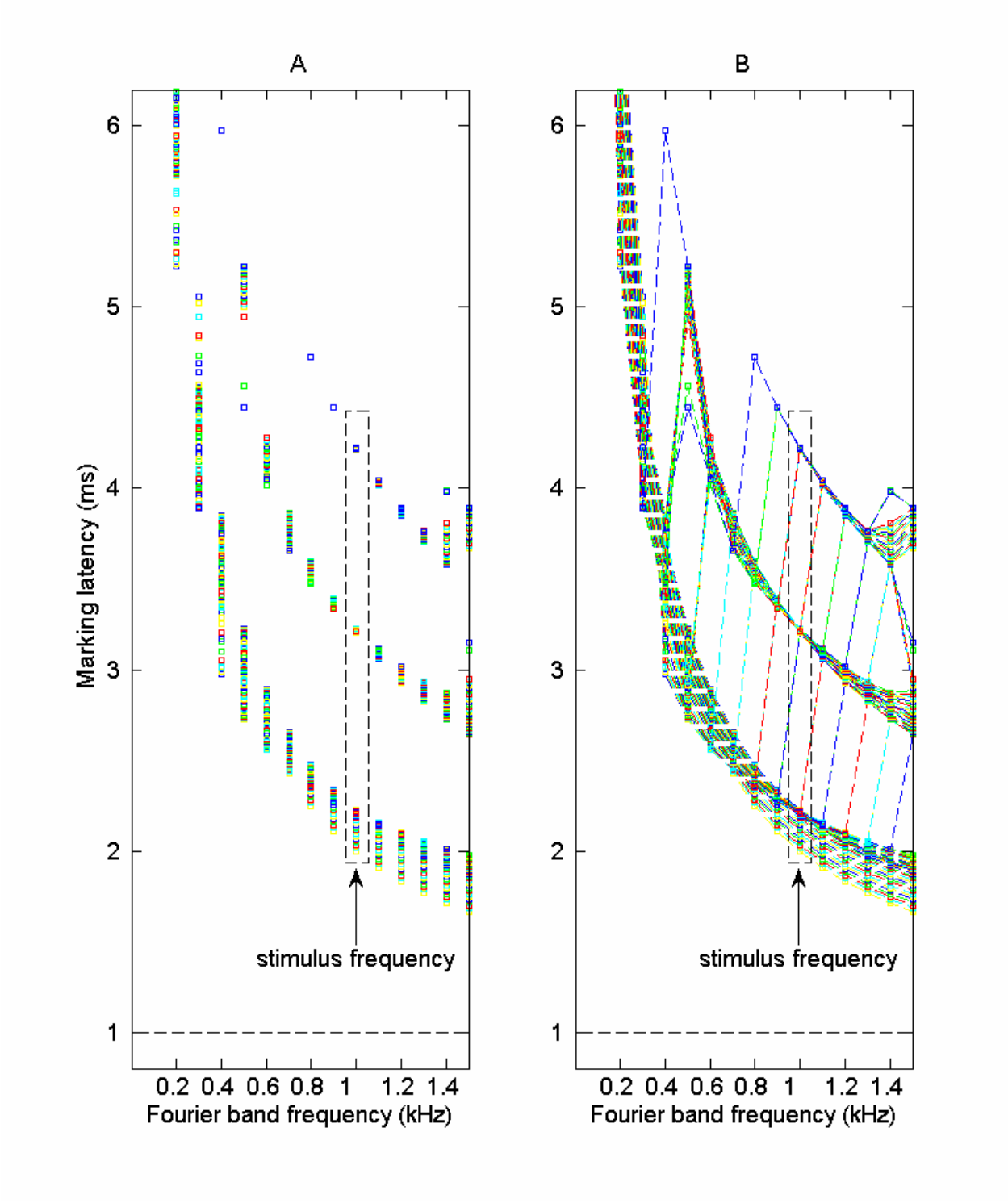


Figure 66. The combined Fourier pattern of the 1-kHz tone burst from the first 100 windows. A. Dash lines are removed for clarity. B. The markings in an identical window are connected with the same-color dash line. Colors repeat every 5 windows

In the same manner, the rest 699 windows are processed. Figure 67 shows the final pattern when plotting all the 799 windows together. For clarity, not all the markings are shown. The color coding for the first 100 windows remains the same and the markings in the rest windows are black. Because the plotting procedure goes from the first to the last, a marking can be overlapped and invisible by another in a later window.

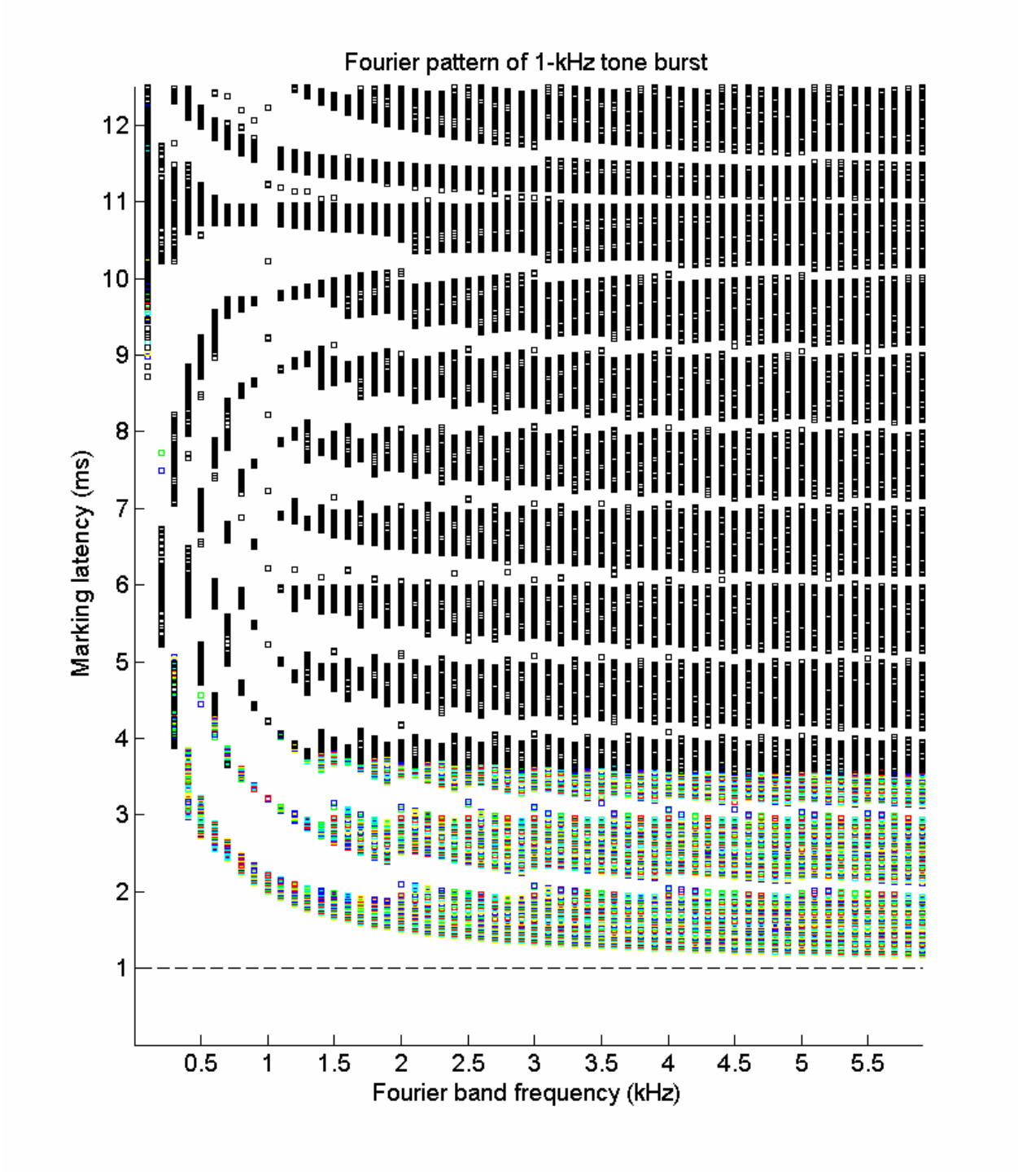


Figure 67. The combined Fourier pattern of the 1-kHz tone burst. The color coding for the first 100 windows remains the same and the markings in the rest windows are black. Because the plotting procedure goes from the first to the last, a marking can be overlapped and invisible by another in a later window.

This is how the final Fourier pattern is formed by combining all the individual analyses of the 1-kHz tone burst.

By comparing a marking at a given frequency in Figure 62, Figure 63, Figure 64 or Figure 65 with the same latency-frequency location in Figure 66, the origin of a marking can be traced to elucidate where the variation comes from. Another way to investigate the variation is to compare two same-frequency markings in any two windows. Again, one window generates a single marking at a given frequency.

APPENDIX C

THE JUMPS IN PREDICTED LATENCIES TO CLICKS

At a given frequency, the predicted latency to a click is usually the same for several consecutive analyses then jumps to a new latency. The marking stays at the new timing for several analyses before it changes again. For instance, at 2 kHz, the predicted latency to the rarefaction click ($x(n) = -100$, when $n = 0$; $x(n) = 0$, when $n \neq 0$) is 1.225 ms for the first 9 analyses (windows 0 to 8), 1.725 ms for the next 20 analyses (windows 9 to 28), and then jumps to 2.225 ms (window 29). The following two sections provide two examples at different frequencies (2 vs 1 kHz) with different click polarities (rarefaction vs. condensation). These examples will show the calculations and illustrate the synthesized Fourier components before and after the jumps to investigate why the jumps may occur.

C.1 JUMPS AT 2 KHZ TO THE RAREFACTION CLICK

Taking the first two jumps in predicted latencies at 2 kHz to the rarefaction click for an example, because these jumps occur between windows 8 and 9 and between 28 and 29, the

following figure shows calculations for windows 0, 8, 9, 28 and 29. Window 0 is also selected because it is the first and the reference window.

The top panels in Figure 68 show the inputs of these five windows. The predicted latency at 2 kHz for each single window is calculated according to the steps 1 to 4 described in 3.2, with additional 1 ms (the synaptic and neural transmission). For example, the predicted latency in windows 0 is $1000\times D(20,0)+1$ (ms) = $1000(\frac{2\pi-3.4558}{2\pi\cdot 2000}+\frac{0}{40000})$ + 1 = 1.225 ms. In window 8, it equals $1000\times D(20,8)+1$ (ms) = $1000(\frac{2\pi-5.969}{2\pi\cdot 2000}+\frac{8}{40000})$ + 1 = 1.225 ms. Because they share the same value, they superimpose at the same point, 1.225 ms.

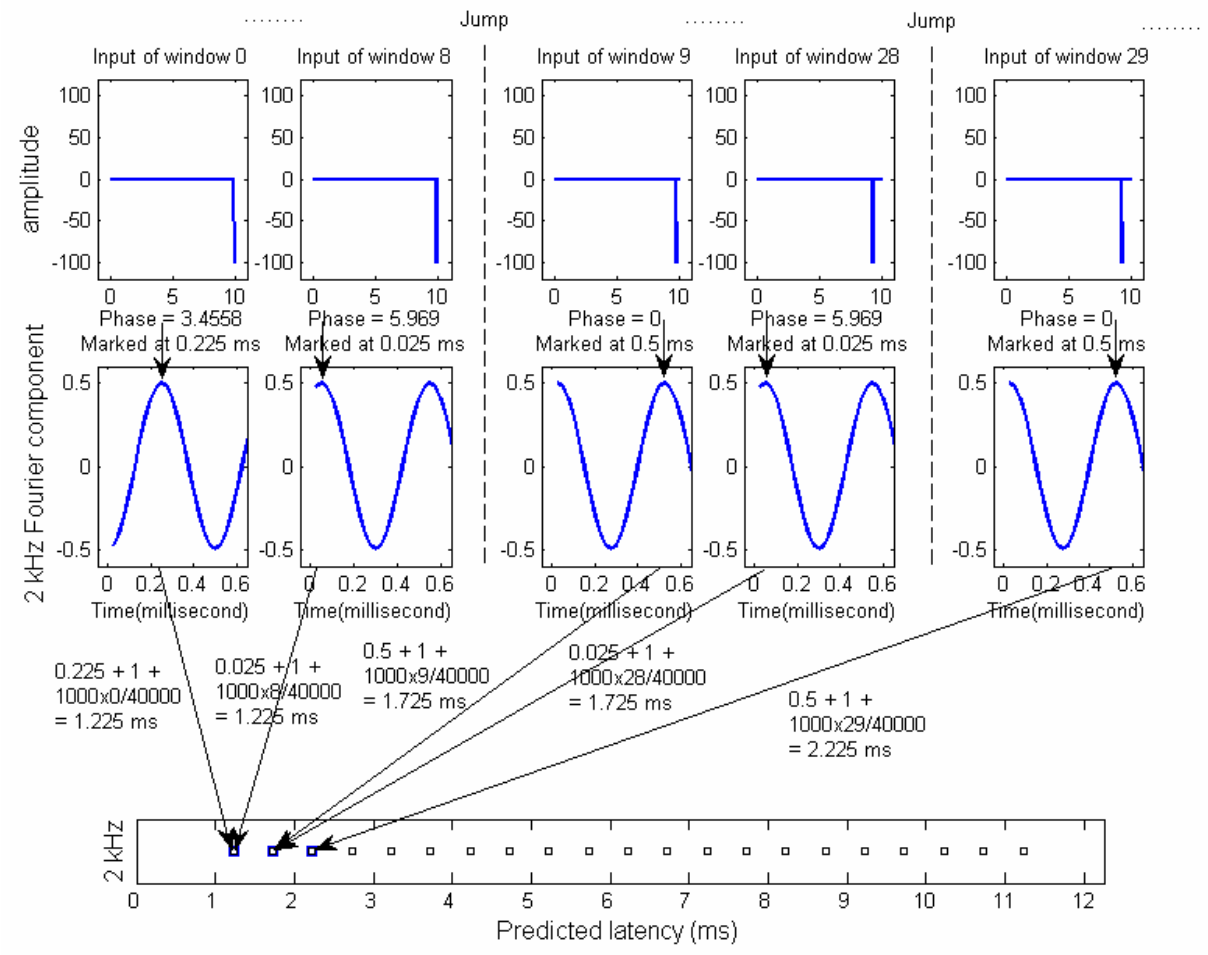


Figure 68. The first two jumps in predicted latencies at 2 kHz to the rarefaction click. Calculations in windows 0, 8, 9, 28 and 29 are demonstrated because the first two jumps occur between windows 8 and 9 and between 28 and 29. The top panels show the inputs of these five windows. The middle panels show the synthesized Fourier components at 2 kHz. Fourier phases and marking latencies of the single window are indicated above the middle panels. The bottom panel shows the predicted latencies after adding 1 ms (of the synaptic and neural transmission) and the latency correction.

The predicted latency in window 9 is $1000\times D(20,9)+1$ (ms) = $1000(\frac{2\pi-0}{2\pi\cdot 2000}+\frac{9}{40000})$ + 1 = 1.725 ms. Because it is 0.5 ms larger than 1.225 ms, there shows a jump in the predicted marking latencies (bottom panel).

The predicted latency in window 28 equals $1000 \times D(20,28)$ + 1 (ms) = $1000(\frac{2\pi-5.969}{2\pi\cdot 2000}+\frac{28}{40000})$ + 1 = 1.725 ms. Because it has the same value as that in window 9, they superimpose at the same point, 1.725 ms.

The predicted latency in window 29 is $1000\times D(20,29)+1$ (ms) = $1000(\frac{2\pi-0}{2\pi\cdot 2000}+\frac{29}{40000})$ + 1 = 2.225 ms. Because it is 0.5 ms larger than 1.725 ms, the second jump occurs.

Looking at the middle panels, the latency of the first amplitude maximum decreases (from 0.225 to 0.025 ms) as the time relationship between the stimulus and the window changes from window 0 to window 8. In the mean while, the Fourier phase increases from 3.4558 to 5.969 radians. When the Fourier phase increases to 2π (that is set to 0 according to the defined range 0 $\leqq$ Fourier phase < 2π), a jump occurs.

This could also be observed at the second jump. The Fourier phase increases from 0 to 5.969 radians when it goes from window 9 to 28. When the Fourier phase increases to 2π in window 29, a jump occurs again.

C.2 JUMP AT 1 KHZ TO THE CONDENSATION CLICK

Taking the first jump in predicted latencies at 1 kHz to the condensation click (x(n) = 100, when n = 0; x(n) = 0, when n $\neq$ 0) for another example, because the jump occur between windows 38 and 39, calculations in these two windows are detailed. Windows 0, 1, and 2 are also selected to show consecutive analyses.

Looking at the middle panels in Figure 69, the Fourier phase of the 1-kHz component increases in the consecutive windows 0, 1 and 2 from 0.15708, 0.31416 to 0.47124 radians. In window 38, it increases to 6.1261 radians, which is close to 2π. Then a jump occurs when the Fourier phase continues to increase and reaches 2π in window 39.

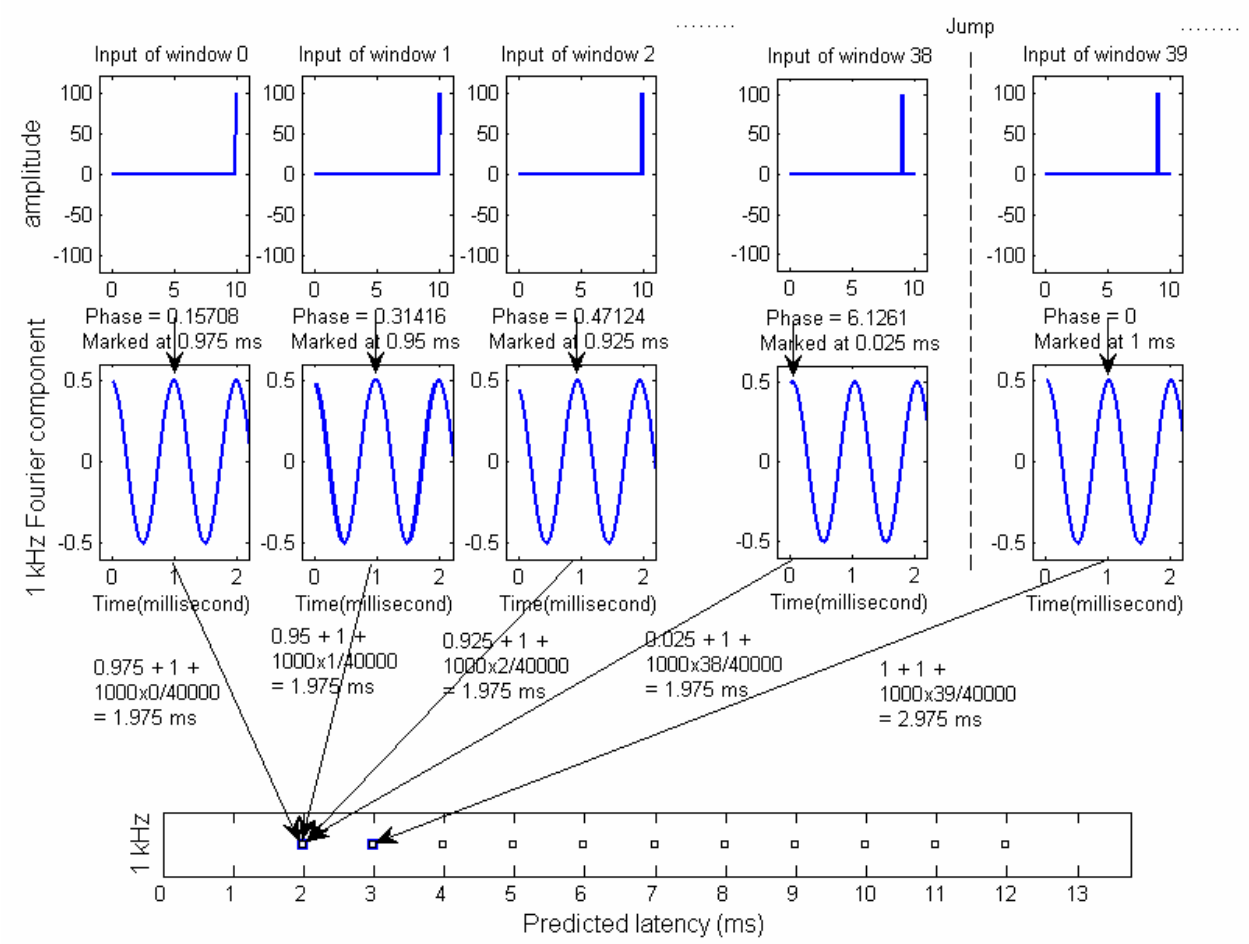


Figure 69. The first jump in predicted latencies at 1 kHz to the condensation click. Calculations in windows 0, 1, 2, 38 and 39 are selected to demonstrate calculations in consecutive windows and the jump that is between windows 38 and 39. The top panels show the inputs of these five windows. The middle panels show the synthesized Fourier components at 1 kHz. Fourier phases and marking latencies of the single window are indicated above the middle panels. The bottom panel shows the predicted latencies after adding 1 ms (of the synaptic and neural transmission) and the latency correction. The jump occurs when the Fourier phase crosses zero.

This example consists with the example at 2 kHz to the rarefaction click – the jumps occur if the Fourier phase crosses zero while the stimulus passes through different windows.

C.3 SUMMARY

To summarize the observations, as the window is moved toward the click in a point-wise manner, the Fourier phase at a given frequency continues to increase. When the Fourier phase crosses zero (or multiples of 2π), there is a jump of the predicted latency at that frequency. This suggests that jumps are due to phase unwrapping and correspond to shift in phase of 2π.

BIBLIOGRAPHY

1. Durrant, J.D. and J.H. Lovrinic, *Bases of hearing science*. 3rd ed. 1995, Baltimore: Williams & Wilkins. 333 p.
2. Van Cleve, J.V. and Gallaudet College., *Gallaudet encyclopedia of deaf people and deafness*. 1987, New York: McGraw-Hill.
3. Flint, P.W. and C.W. Cummings, *Cummings otolaryngology head & neck surgery*. 5th ed. 2010, Philadelphia, PA: Mosby/Elsevier.
4. Bailey, B.J., J.T. Johnson, and S.D. Newlands, *Head & neck surgery--otolaryngology*. 4th ed. 2006, Philadelphia, PA: Lippincott Williams & Wilkins.
5. Clark, G.M., et al., *A multiple electrode cochlear implant.* J Laryngol Otol, 1977. **91**(11): p. 935-45.
6. Wilson, B.S. and M.F. Dorman, *Cochlear implants: a remarkable past and a brilliant future.* Hear Res, 2008. **242**(1-2): p. 3-21.
7. Loeb, G.E., *Are cochlear implant patients suffering from perceptual dissonance?* Ear Hear, 2005. **26**(5): p. 435-50.
8. Shannon, R.V., *The relative importance of amplitude, temporal, and spectral cues for cochlear implant processor design.* Am J Audiol, 2002. **11**(2): p. 124-7.
9. Smith, Z.M., B. Delgutte, and A.J. Oxenham, *Chimaeric sounds reveal dichotomies in auditory perception.* Nature, 2002. **416**(6876): p. 87-90.
10. Wilson, B.S., et al., *Better speech recognition with cochlear implants.* Nature, 1991. **352**(6332): p. 236-8.
11. White, M.W., M.M. Merzenich, and J.N. Gardi, *Multichannel cochlear implants. Channel interactions and processor design.* Arch Otolaryngol, 1984. **110**(8): p. 493-501.
12. Kayser, C., N.K. Logothetis, and S. Panzeri, *Millisecond encoding precision of auditory cortex neurons.* Proc Natl Acad Sci U S A, 2010. **107**(39): p. 16976-81.
13. Kong, Y.Y., G.S. Stickney, and F.G. Zeng, *Speech and melody recognition in binaurally combined acoustic and electric hearing.* J Acoust Soc Am, 2005. **117**(3 Pt 1): p. 1351-61.
14. Kong, Y.Y., et al., *Music perception with temporal cues in acoustic and electric hearing.* Ear Hear, 2004. **25**(2): p. 173-85.
15. Gfeller, K., et al., *Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise.* Ear Hear, 2007. **28**(3): p. 412-23.
16. Dallos, P., et al., *Cochlear inner and outer hair cells: functional differences.* Science, 1972. **177**(46): p. 356-8.
17. Dallos, P. and J.D. Durrant, *On the derivative relationship between stapes movement and cochlear microphonic.* J Acoust Soc Am, 1972. **52**(4): p. 1263-5.
18. von Bekesy, G., *Experiments in hearing*. 1960, New York: McGraw-Hill Book Co.
19. Durrant, J.D. and L.L. Feth, *Hearing Science--The Bases*. in press, Boston: Pearson.

20. Ruggero, M.A. and N.C. Rich, *Timing of spike initiation in cochlear afferents: dependence on site of innervation.* J Neurophysiol, 1987. **58**(2): p. 379-403.
21. Manley, G.A., *Cochlear mechanisms from a phylogenetic viewpoint.* Proc Natl Acad Sci U S A, 2000. **97**(22): p. 11736-43.
22. Ruggero, M.A. and A.N. Temchin, *Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods.* J Assoc Res Otolaryngol, 2007. **8**(2): p. 153-66.
23. Anderson, D.J., et al., *Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: frequency and intensity effects.* J Acoust Soc Am, 1971. **49**(4): p. Suppl 2:1131+.
24. Hillery, C.M. and P.M. Narins, *Neurophysiological evidence for a traveling wave in the amphibian inner ear.* Science, 1984. **225**(4666): p. 1037-9.
25. Koppl, C., *Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba.* J Neurosci, 1997. **17**(9): p. 3312-21.
26. Manley, G.A., et al., *Peripheral auditory processing in the bobtail lizard Tiliqua rugosa IV. Phase locking of auditory-nerve fibres.* J Comp Physiol A, 1990. **167**: p. 129-138.
27. Shera, C.A., *Frequency glides in click responses of the basilar membrane and auditory nerve: their scaling behavior and origin in traveling-wave dispersion.* J Acoust Soc Am, 2001. **109**(5 Pt 1): p. 2023-34.
28. Manley, G.A., G.K. Yates, and C. Koppl, *Auditory peripheral tuning: evidence for a simple resonance phenomenon in the lizard Tiliqua.* Hear Res, 1988. **33**(2): p. 181-9.
29. Helmholtz, H.v. and A.J. Ellis, *On the sensations of tone as a physiological basis for the theory of music*. 2d English ed. 1954, New York,: Dover Publications. xix, 576 p.
30. Geisler, C.D., *From sound to synapse : physiology of the mammalian ear*. 1998, New York: Oxford University Press. xiv, 381 p.
31. Fourier, J.B.J., *La theorie analytique de la chaleur (The Analytic Theory of Heat).* Mem. Acad. R. Sci., 1822. **8**: p. 581-2.
32. Oppenheim, A.V. and R.W. Schafer, *Discrete-time signal processing*. 3rd ed. Prentice Hall signal processing series. 2010, Upper Saddle River [N.J.]: Prentice Hall : imprint of Pearson. xxviii, 1108 p.
33. Taft, D.A., D.B. Grayden, and A.N. Burkitt, *Across-frequency delays based on the cochlear traveling wave: enhanced speech presentation for cochlear implants.* IEEE Trans Biomed Eng, 2010. **57**(3): p. 596-606.
34. Chang, Y.T., et al., *Tone discrimination and speech perception benefit in Mandarin-speaking children fit with HiRes fidelity 120 sound processing.* Otol Neurotol, 2009. **30**(6): p. 750-7.
35. Musiek, F.E. and G.D. Chermak, *Handbook of (central) auditory processing disorder*. Vol. 1. 2006, San Diego: Plural Publishing INC.
36. Oppenheim, A.V. and J.S. Lim, *The importance of phase in signals.* Proc. IEEE, 1981. **69**: p. 529-541.
37. Moore, B.C., *The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people.* J Assoc Res Otolaryngol, 2008. **9**(4): p. 399-406.
38. Ghitza, O., *On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception.* J Acoust Soc Am, 2001. **110**(3 Pt 1): p. 1628-40.
39. Zeng, F.G., et al., *On the dichotomy in auditory perception between temporal envelope and fine structure cues.* J Acoust Soc Am, 2004. **116**(3): p. 1351-4.

40. Hopkins, K., B.C. Moore, and M.A. Stone, *Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech.* J Acoust Soc Am, 2008. **123**(2): p. 1140-53.
41. Lorenzi, C., et al., *Speech perception problems of the hearing impaired reflect inability to use temporal fine structure.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18866-9.
42. Drennan, W.R., S. Gatehouse, and C. Lever, *Perceptual segregation of competing speech sounds: the role of spatial location.* J Acoust Soc Am, 2003. **114**(4 Pt 1): p. 2178-89.
43. Xu, L. and B.E. Pfingst, *Relative importance of temporal envelope and fine structure in lexical-tone perception.* J Acoust Soc Am, 2003. **114**(6 Pt 1): p. 3024-7.
44. Batteau, D.W., *The role of the pinna in human localization.* Proc R Soc Lond B Biol Sci, 1967. **168**(11): p. 158-80.
45. Raykar, V.C., R. Duraiswami, and B. Yegnanarayana, *Extracting the frequencies of the pinna spectral notches in measured head related impulse responses.* J Acoust Soc Am, 2005. **118**(1): p. 364-74.
46. Overstreet, E.H., 3rd and M.A. Ruggero, *Development of wide-band middle ear transmission in the Mongolian gerbil.* J Acoust Soc Am, 2002. **111**(1 Pt 1): p. 261-70.
47. Rosowski, J.J., et al., *Measurements of middle-ear function in the Mongolian gerbil, a specialized mammalian ear.* Audiol Neurootol, 1999. **4**(3-4): p. 129-36.
48. Xue, S., D.C. Mountain, and A.E. Hubbard, *Electrically evoked basilar membrane motion.* J Acoust Soc Am, 1995. **97**(5 Pt 1): p. 3030-41.
49. Cohen, Y.E., et al., *Middle-ear development. V: Development of umbo sensitivity in the gerbil.* Am J Otolaryngol, 1993. **14**(3): p. 191-8.
50. Olson, E.S., *Observing middle and inner ear mechanics with novel intracochlear pressure sensors.* J Acoust Soc Am, 1998. **103**(6): p. 3445-63.
51. Puria, S. and J.B. Allen, *Measurements and model of the cat middle ear: evidence of tympanic membrane acoustic delay.* J Acoust Soc Am, 1998. **104**(6): p. 3463-81.
52. Guinan, J.J., Jr. and W.T. Peake, *Middle-ear characteristics of anesthetized cats.* J Acoust Soc Am, 1967. **41**(5): p. 1237-61.
53. Olson, E.S., *Direct measurement of intra-cochlear pressure waves.* Nature, 1999. **402**(6761): p. 526-9.
54. Raftenberg, M.N., *Flow of endolymph in the inner spiral sulcus and the subtectorial space.* J Acoust Soc Am, 1990. **87**(6): p. 2606-20.
55. Karavitaki, K.D. and D.C. Mountain, *Evidence for oscillatory fluid flow in the tunnel of Corti.* Biophys J, 2007.
56. Zwislocki, J.J., *Tectorial membrane: a possible sharpening effect on the frequency analysis in the cochlea.* Acta Otolaryngol, 1979. **87**(3-4): p. 267-9.
57. Zwislocki, J.J. and E.J. Kletsky, *Tectorial membrane: a possible effect on frequency analysis in the cochlea.* Science, 1979. **204**(4393): p. 639-41.
58. de Boer, E., *Can shape deformations of the organ of Corti influence the travelling wave in the cochlea?* Hear Res, 1990. **44**(1): p. 83-92.
59. Zwislocki, J.J. and W.G. Sokolich, *Velocity and displacement responses in auditory-nerve fibers.* Science, 1973. **182**(107): p. 64-6.
60. Zwislocki, J.J., *Phase opposition between inner and outer hair cells and auditory sound analysis.* Audiology, 1975. **14**(5-6): p. 443-55.
61. Woolf, N.K., A.F. Ryan, and R.C. Bone, *Neural phase-locking properties in the absence of cochlear outer hair cells.* Hear Res, 1981. **4**(3-4): p. 335-46.

62. Brundin, L., A. Flock, and B. Canlon, *Sound-induced motility of isolated cochlear outer hair cells is frequency-specific.* Nature, 1989. **342**(6251): p. 814-6.
63. Frank, G., W. Hemmert, and A.W. Gummer, *Limiting dynamics of high-frequency electromechanical transduction of outer hair cells.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4420-5.
64. Zhao, H.B. and J. Santos-Sacchi, *Auditory collusion and a coupled couple of outer hair cells.* Nature, 1999. **399**(6734): p. 359-62.
65. Salt, A.N. and T.E. Hullar, *Responses of the ear to low frequency sounds, infrasound and wind turbines.* Hear Res, 2010. **268**(1-2): p. 12-21.
66. Deol, M.S. and S. Gluecksohn-Waelsch, *The role of inner hair cells in hearing.* Nature, 1979. **278**(5701): p. 250-2.
67. Cheatham, M.A. and P. Dallos, *Response phase: a view from the inner hair cell.* J Acoust Soc Am, 1999. **105**(2 Pt 1): p. 799-810.
68. Zerlin, S., *Traveling-wave velocity in the human cochlea.* J Acoust Soc Am, 1969. **46**(4): p. 1015.
69. Gould, H.J. and O.A. Sobhy, *Using the derived auditory brain stem response to estimate traveling wave velocity.* Ear Hear, 1992. **13**(2): p. 96-101.
70. Donaldson, G.S. and R.A. Ruth, *Derived band auditory brain-stem response estimates of traveling wave velocity in humans. I: Normal-hearing subjects.* J Acoust Soc Am, 1993. **93**(2): p. 940-51.
71. Don, M., et al., *Auditory brainstem response (ABR) peak amplitude variability reflects individual differences in cochlear response times.* J Acoust Soc Am, 1994. **96**(6): p. 3476-91.
72. Kim, Y., M. Aoyagi, and Y. Koike, *Measurement of cochlear basilar membrane traveling wave velocity by derived ABR.* Acta Otolaryngol Suppl, 1994. **511**: p. 71-6.
73. Don, M., et al., *Successful detection of small acoustic tumors using the stacked derived-band auditory brain stem response amplitude.* Am J Otol, 1997. **18**(5): p. 608-21; discussion 682-5.
74. Don, M., et al., *The stacked ABR: a sensitive and specific screening tool for detecting small acoustic tumors.* Audiol Neurootol, 2005. **10**(5): p. 274-90.
75. Thornton, A.R., et al., *Verification of a new test of endolymphatic hydrops.* J Laryngol Otol, 1989. **103**(12): p. 1136-9.
76. Thornton, A.R., G. Farrell, and N.P. Haacke, *A non-invasive, objective test of endolymphatic hydrops.* Acta Otolaryngol Suppl, 1991. **479**: p. 35-43.
77. Donaldson, G.S. and R.A. Ruth, *Derived-band auditory brain-stem response estimates of traveling wave velocity in humans: II. Subjects with noise-induced hearing loss and Meniere's disease.* J Speech Hear Res, 1996. **39**(3): p. 534-45.
78. Murray, J.G., et al., *Tone burst auditory brain stem response latency estimates of cochlear travel time in Meniere's disease, cochlear hearing loss, and normal ears.* Am J Otol, 1998. **19**(6): p. 854-9.
79. Don, M., B. Kwong, and C. Tanaka, *A diagnostic test for Meniere's Disease and Cochlear Hydrops: impaired high-pass noise masking of auditory brainstem responses.* Otol Neurotol, 2005. **26**(4): p. 711-22.
80. Neely, S.T., et al., *Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli.* J Acoust Soc Am, 1988. **83**(2): p. 652-6.

81. Shera, C.A., J.J. Guinan, Jr., and A.J. Oxenham, *Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements.* Proc Natl Acad Sci U S A, 2002. **99**(5): p. 3318-23.
82. Siegel, J.H., et al., *Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering.* J Acoust Soc Am, 2005. **118**(4): p. 2434-43.
83. Mackersie, C., A.C. Neuman, and H. Levitt, *Response time and word recognition using a Modified-Rhyme monitoring task: list equivalency and time-order effects.* Ear Hear, 1999. **20**(6): p. 515-20.
84. Elberling, C., et al., *Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay.* J Acoust Soc Am, 2007. **122**(5): p. 2772-85.
85. Lin, T. and J.J. Guinan, Jr., *Auditory-nerve-fiber responses to high-level clicks: interference patterns indicate that excitation is due to the combination of multiple drives.* J Acoust Soc Am, 2000. **107**(5 Pt 1): p. 2615-30.
86. Lin, T. and J.J. Guinan, Jr., *Time-frequency analysis of auditory-nerve-fiber and basilar-membrane click responses reveal glide irregularities and non-characteristic-frequency skirts.* J Acoust Soc Am, 2004. **116**(1): p. 405-16.
87. Guinan, J.J., Jr. and N.P. Cooper, *Medial olivocochlear efferent inhibition of basilar-membrane responses to clicks: evidence for two modes of cochlear mechanical excitation.* J Acoust Soc Am, 2008. **124**(2): p. 1080-92.
88. Temchin, A.N., et al., *Wiener kernels of chinchilla auditory-nerve fibers: verification using responses to tones, clicks, and noise and comparison with basilar-membrane vibrations.* J Neurophysiol, 2005. **93**(6): p. 3635-48.
89. van der Heijden, M. and P.X. Joris, *Cochlear phase and amplitude retrieved from the auditory nerve at arbitrary frequencies.* J Neurosci, 2003. **23**(27): p. 9194-8.
90. Smolders, J.W. and R. Klinke, *Synchronized responses of primary auditory fibre-populations in Caiman crocodilus (L.) to single tones and clicks.* Hear Res, 1986. **24**(2): p. 89-103.
91. Hillery, C.M. and P.M. Narins, *Frequency and time domain comparison of low-frequency auditory fiber responses in two anuran amphibians.* Hear Res, 1987. **25**(2-3): p. 233-48.
92. Shamma, S.A., *Speech processing in the auditory system. I: The representation of speech sounds in the responses of the auditory nerve.* J Acoust Soc Am, 1985. **78**(5): p. 1612-21.
93. Brugge, J.F., et al., *Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds.* J Neurophysiol, 1969. **32**(3): p. 386-401.
94. Carney, L.H. and T.C. Yin, *Temporal coding of resonances by low-frequency auditory nerve fibers: single-fiber responses and a population model.* J Neurophysiol, 1988. **60**(5): p. 1653-77.
95. Krishna, B.S., *A unified mechanism for spontaneous-rate and first-spike timing in the auditory nerve.* J Comput Neurosci, 2002. **13**(2): p. 71-91.
96. Rose, J.E., et al., *Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey.* J Neurophysiol, 1971. **34**(4): p. 685-99.
97. Hind, J.E., et al., *Coding of information pertaining to paired low-frequency tones in single auditory nerve fibers of the squirrel monkey.* J Neurophysiol, 1967. **30**(4): p. 794-816.

98. Rose, J.E., et al., *Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey.* J Neurophysiol, 1967. **30**(4): p. 769-93.
99. Geisler, C.D., W.S. Rhode, and D.T. Kennedy, *Responses to tonal stimuli of single auditory nerve fibers and their relationship to basilar membrane motion in the squirrel monkey.* J Neurophysiol, 1974. **37**(6): p. 1156-72.
100. Horst, J.W., E. Javel, and G.R. Farley, *Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms.* J Acoust Soc Am, 1986. **79**(2): p. 398-416.
101. Shofner, W.P. and M.B. Sachs, *Representation of a low-frequency tone in the discharge rate of populations of auditory nerve fibers.* Hear Res, 1986. **21**(1): p. 91-5.
102. Anderson, D.J., *Quantitative model for the effects of stimulus frequency upon synchronization of auditory nerve discharges.* J Acoust Soc Am, 1973. **54**(2): p. 361-4.
103. Johnson, D.H., *The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones.* J Acoust Soc Am, 1980. **68**(4): p. 1115-22.
104. Joris, P.X., et al., *Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency.* J Neurophysiol, 1994. **71**(3): p. 1022-36.
105. Tollin, D.J. and T.C. Yin, *Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive.* J Neurosci, 2005. **25**(46): p. 10648-57.
106. Dreyer, A. and B. Delgutte, *Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: implications for sound localization.* J Neurophysiol, 2006. **96**(5): p. 2327-41.
107. Cariani, P.A. and B. Delgutte, *Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch.* J Neurophysiol, 1996. **76**(3): p. 1717-34.
108. Klump, G.M., et al., *AM representation in green treefrog auditory nerve fibers: neuroethological implications for pattern recognition and sound localization.* J Comp Physiol A Neuroethol Sens Neural Behav Physiol, 2004. **190**(12): p. 1011-21.
109. Tolnai, S., et al., *Dynamic coupling of excitatory and inhibitory responses in the medial nucleus of the trapezoid body.* Eur J Neurosci, 2008. **27**(12): p. 3191-204.
110. Rhode, W.S., G.L. Roth, and A. Recio-Spinoso, *Response properties of cochlear nucleus neurons in monkeys.* Hear Res, 2010. **259**(1-2): p. 1-15.
111. Simmons, A.M., J.J. Schwartz, and M. Ferragamo, *Auditory nerve representation of a complex communication sound in background noise.* J Acoust Soc Am, 1992. **91**(5): p. 2831-44.
112. Joris, P.X. and P.H. Smith, *The volley theory and the spherical cell puzzle.* Neuroscience, 2008. **154**(1): p. 65-76.
113. Temchin, A.N. and M.A. Ruggero, *Phase-locked responses to tones of chinchilla auditory nerve fibers: implications for apical cochlear mechanics.* J Assoc Res Otolaryngol, 2010. **11**(2): p. 297-318.
114. Blackburn, C.C. and M.B. Sachs, *Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis.* J Neurophysiol, 1989. **62**(6): p. 1303-29.
115. Taberner, A.M. and M.C. Liberman, *Response properties of single auditory nerve fibers in the mouse.* J Neurophysiol, 2005. **93**(1): p. 557-69.

116. Rhode, W.S., *Response patterns to sound associated with labeled globular/bushy cells in cat.* Neuroscience, 2008. **154**(1): p. 87-98.
117. FitzGerald, J.V., et al., *Delay analysis in the auditory brainstem of the rat: comparison with click latency.* Hear Res, 2001. **159**(1-2): p. 85-100.
118. Goblick, T.J., Jr. and R.R. Pfeiffer, *Time-domain measurements of cochlear nonlinearities using combination click stimuli.* J Acoust Soc Am, 1969. **46**(4): p. 924-38.
119. Guinan, J.J., Jr., T. Lin, and H. Cheng, *Medial-olivocochlear-efferent inhibition of the first peak of auditory-nerve responses: evidence for a new motion within the cochlea.* J Acoust Soc Am, 2005. **118**(4): p. 2421-33.
120. Kiang, N.Y., *Stimulus Coding in the Auditory Nerve and Cochlear Nucleus.* Acta Oto-laryngologica, 1965. **59**(2-6): p. 186-200.
121. Pfeiffer, R.R. and D.O. Kim, *Response patterns of single cochlear nerve fibers to click stimuli: descriptions for cat.* J Acoust Soc Am, 1972. **52**(6): p. 1669-77.
122. Recio, A. and W.S. Rhode, *Basilar membrane responses to broadband stimuli.* J Acoust Soc Am, 2000. **108**(5 Pt 1): p. 2281-98.
123. Shera, C.A., *Intensity-invariance of fine time structure in basilar-membrane click responses: implications for cochlear mechanics.* J Acoust Soc Am, 2001. **110**(1): p. 332-48.
124. Wagner, H., et al., *Microsecond precision of phase delay in the auditory system of the barn owl.* J Neurophysiol, 2005. **94**(2): p. 1655-8.
125. Spoendlin, H., *Innervation patterns in the organ of corti of the cat.* Acta Otolaryngol, 1969. **67**(2): p. 239-54.
126. Spoendlin, H., *Innervation densities of the cochlea.* Acta Otolaryngol, 1972. **73**(2): p. 235-48.
127. Spoendlin, H. and A. Schrott, *The spiral ganglion and the innervation of the human organ of Corti.* Acta Otolaryngol, 1988. **105**(5-6): p. 403-10.
128. Wittig, J.H., Jr. and T.D. Parsons, *Synaptic ribbon enables temporal precision of hair cell afferent synapse by increasing the number of readily releasable vesicles: a modeling study.* J Neurophysiol, 2008. **100**(4): p. 1724-39.
129. Buran, B.N., et al., *Onset coding is degraded in auditory nerve fibers from mutant mice lacking synaptic ribbons.* J Neurosci, 2010. **30**(22): p. 7587-97.
130. Liberman, M.C. and M.E. Oliver, *Morphometry of intracellularly labeled neurons of the auditory nerve: correlations with functional properties.* J Comp Neurol, 1984. **223**(2): p. 163-76.
131. Arbuthnott, E.R., I.A. Boyd, and K.U. Kalu, *Ultrastructural dimensions of myelinated peripheral nerve fibres in the cat and their relation to conduction velocity.* J Physiol, 1980. **308**: p. 125-57.
132. Alving, B.M. and W.M. Cowan, *Some quantitative observations on the cochlear division of the eighth nerve in the squirrel monkey (Saimiri sciureus).* Brain Res, 1971. **25**(2): p. 229-39.
133. Arnesen, A.R. and K.K. Osen, *The cochlear nerve in the cat: topography, cochleotopy, and fiber spectrum.* J Comp Neurol, 1978. **178**(4): p. 661-78.
134. Moser, T., A. Neef, and D. Khimich, *Mechanisms underlying the temporal precision of sound coding at the inner hair cell ribbon synapse.* J Physiol, 2006. **576**(Pt 1): p. 55-62.
135. Moser, T. and D. Beutner, *Kinetics of exocytosis and endocytosis at the cochlear inner hair cell afferent synapse of the mouse.* Proc Natl Acad Sci U S A, 2000. **97**(2): p. 883-8.

136. Goutman, J.D. and E. Glowatzki, *Time course and calcium dependence of transmitter release at a single ribbon synapse.* Proc Natl Acad Sci U S A, 2007. **104**(41): p. 16341-6.
137. Johnson, S.L., et al., *Tonotopic variation in the calcium dependence of neurotransmitter release and vesicle pool replenishment at mammalian auditory ribbon synapses.* J Neurosci, 2008. **28**(30): p. 7670-8.
138. Palmer, A.R. and I.J. Russell, *Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells.* Hear Res, 1986. **24**(1): p. 1-15.
139. Brown, M.C., *Morphology of labeled afferent fibers in the guinea pig cochlea.* J Comp Neurol, 1987. **260**(4): p. 591-604.
140. van den Honert, C. and P.H. Stypulkowski, *Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings.* Hear Res, 1984. **14**(3): p. 225-43.
141. DeWeese, M.R., M. Wehr, and A.M. Zador, *Binary spiking in auditory cortex.* J Neurosci, 2003. **23**(21): p. 7940-9.
142. Furukawa, S., L. Xu, and J.C. Middlebrooks, *Coding of sound-source location by ensembles of cortical neurons.* J Neurosci, 2000. **20**(3): p. 1216-28.
143. Heil, P. and D.R. Irvine, *First-spike timing of auditory-nerve fibers and comparison with auditory cortex.* J Neurophysiol, 1997. **78**(5): p. 2438-54.
144. Yang, Y., et al., *Millisecond-scale differences in neural activity in auditory cortex can drive decisions.* Nat Neurosci, 2008. **11**(11): p. 1262-3.
145. Liang, D.H., H.S. Lusted, and R.L. White, *The nerve-electrode interface of the cochlear implant: current spread.* IEEE Trans Biomed Eng, 1999. **46**(1): p. 35-43.
146. Hochmair-Desoyer, I. and E.S. Hochmair, *An eight channel scala tympani electrode for auditory prostheses.* IEEE Trans Biomed Eng, 1980. **27**(1): p. 44-50.
147. Brummer, S.B. and M.J. Turner, *Electrochemical considerations for safe electrical stimulation of the nervous system with platinum electrodes.* IEEE Trans Biomed Eng, 1977. **24**(1): p. 59-63.
148. Parkins, C.W., *Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys.* Hear Res, 1989. **41**(2-3): p. 137-68.
149. Kiang, N.Y. and E.C. Moxon, *Physiological considerations in artificial stimulation of the inner ear.* Ann Otol Rhinol Laryngol, 1972. **81**(5): p. 714-30.
150. Miller, C.A., et al., *Electrically evoked single-fiber action potentials from cat: responses to monopolar, monophasic stimulation.* Hear Res, 1999. **130**(1-2): p. 197-218.
151. Miller, C.A., et al., *Changes across time in the temporal responses of auditory nerve fibers stimulated by electric pulse trains.* J Assoc Res Otolaryngol, 2008. **9**(1): p. 122-37.
152. Shepherd, R.K. and E. Javel, *Electrical stimulation of the auditory nerve: II. Effect of stimulus waveshape on single fibre response properties.* Hear Res, 1999. **130**(1-2): p. 171-88.
153. Javel, E. and R.K. Shepherd, *Electrical stimulation of the auditory nerve. III. Response initiation sites and temporal fine structure.* Hear Res, 2000. **140**(1-2): p. 45-76.
154. Litvak, L., B. Delgutte, and D. Eddington, *Improved neural representation of vowels in electric stimulation using desynchronizing pulse trains.* J Acoust Soc Am, 2003. **114**(4 Pt 1): p. 2099-111.
155. Litvak, L.M., B. Delgutte, and D.K. Eddington, *Improved temporal coding of sinusoids in electric stimulation of the auditory nerve using desynchronizing pulse trains.* J Acoust Soc Am, 2003. **114**(4 Pt 1): p. 2079-98.

156. Rubinstein, J.T., et al., *Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation.* Hear Res, 1999. **127**(1-2): p. 108-18.
157. Middlebrooks, J.C. and R.L. Snyder, *Selective electrical stimulation of the auditory nerve activates a pathway specialized for high temporal acuity.* J Neurosci, 2010. **30**(5): p. 1937-46.
158. Eddington, D.K., *Speech discrimination in deaf subjects with cochlear implants.* J Acoust Soc Am, 1980. **68**(3): p. 885-91.
159. Fu, Q.J. and R.V. Shannon, *Effects of electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant.* Ear Hear, 1999. **20**(4): p. 332-44.
160. Fu, Q.J. and R.V. Shannon, *Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing.* J Acoust Soc Am, 1999. **105**(3): p. 1889-900.
161. Li, T. and Q.J. Fu, *Effects of Spectral Shifting on Speech Perception in Noise.* Hear Res, 2010.
162. Shannon, R.V., F.G. Zeng, and J. Wygonski, *Speech recognition with altered spectral distribution of envelope cues.* J Acoust Soc Am, 1998. **104**(4): p. 2467-76.
163. Fu, Q.J., et al., *Voice gender identification by cochlear implant users: the role of spectral and temporal resolution.* J Acoust Soc Am, 2005. **118**(3 Pt 1): p. 1711-8.
164. Kovacic, D. and E. Balaban, *Voice gender perception by cochlear implantees.* J Acoust Soc Am, 2009. **126**(2): p. 762-75.
165. Kovacic, D. and E. Balaban, *Hearing History Influences Voice Gender Perceptual Performance in Cochlear Implant Users.* Ear Hear, 2010.
166. Meister, H., et al., *The perception of prosody and speaker gender in normal-hearing listeners and cochlear implant recipients.* Int J Audiol, 2009. **48**(1): p. 38-48.
167. Cleary, M. and D.B. Pisoni, *Talker discrimination by prelingually deaf children with cochlear implants: preliminary results.* Ann Otol Rhinol Laryngol Suppl, 2002. **189**: p. 113-8.
168. Green, T., et al., *Talker intelligibility differences in cochlear implant listeners.* J Acoust Soc Am, 2007. **121**(6): p. EL223-9.
169. Vongpaisal, T., et al., *Children With Cochlear implants recognize their mother's voice.* Ear Hear, 2010. **31**(4): p. 555-66.
170. Vongphoe, M. and F.G. Zeng, *Speaker recognition with temporal cues in acoustic and electric hearing.* J Acoust Soc Am, 2005. **118**(2): p. 1055-61.
171. Binns, C. and J.F. Culling, *The role of fundamental frequency contours in the perception of speech against interfering speech.* J Acoust Soc Am, 2007. **122**(3): p. 1765.
172. Krishnan, A., et al., *Human frequency-following response: representation of pitch contours in Chinese tones.* Hear Res, 2004. **189**(1-2): p. 1-12.
173. Lane, H., et al., *Changes in sound pressure and fundamental frequency contours following changes in hearing status.* J Acoust Soc Am, 1997. **101**(4): p. 2244-52.
174. Langereis, M.C., et al., *Effect of cochlear implantation on voice fundamental frequency in post-lingually deafened adults.* Audiology, 1998. **37**(4): p. 219-30.
175. Laures, J.S. and K. Bunton, *Perceptual effects of a flattened fundamental frequency at the sentence level under different listening conditions.* J Commun Disord, 2003. **36**(6): p. 449-64.
176. Laures, J.S. and G. Weismer, *The effects of a flattened fundamental frequency on intelligibility at the sentence level.* J Speech Lang Hear Res, 1999. **42**(5): p. 1148-56.

177. Liu, S. and A.G. Samuel, *Perception of Mandarin lexical tones when F0 information is neutralized.* Lang Speech, 2004. **47**(Pt 2): p. 109-38.
178. Nardo, W.D., et al., *Differences between electrode-assigned frequencies and cochlear implant recipient pitch perception.* Acta Otolaryngol, 2007. **127**(4): p. 370-7.
179. Xu, Y., et al., *Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences.* J Acoust Soc Am, 2004. **116**(2): p. 1168-78.
180. Zhou, N. and L. Xu, *Development and evaluation of methods for assessing tone production skills in Mandarin-speaking children with cochlear implants.* J Acoust Soc Am, 2008. **123**(3): p. 1653-64.
181. Xin, L., Q.J. Fu, and J.J. Galvin, 3rd, *Vocal emotion recognition by normal-hearing listeners and cochlear implant users.* Trends Amplif, 2007. **11**(4): p. 301-15.
182. Balkany, T., et al., *Nucleus Freedom North American clinical trial.* Otolaryngol Head Neck Surg, 2007. **136**(5): p. 757-62.
183. Nie, K., A. Barco, and F.G. Zeng, *Spectral and temporal cues in cochlear implant speech perception.* Ear Hear, 2006. **27**(2): p. 208-17.
184. Roberts, B., B.R. Glasberg, and B.C. Moore, *Primitive stream segregation of tone sequences without differences in fundamental frequency or passband.* J Acoust Soc Am, 2002. **112**(5 Pt 1): p. 2074-85.
185. Hong, R.S. and C.W. Turner, *Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients.* J Acoust Soc Am, 2006. **120**(1): p. 360-74.
186. Brockmeier, S.J., et al., *Music perception in electric acoustic stimulation users as assessed by the Mu.S.I.C. test.* Adv Otorhinolaryngol, 2010. **67**: p. 70-80.
187. Chen, J.K., et al., *Music training improves pitch perception in prelingually deafened children with cochlear implants.* Pediatrics, 2010. **125**(4): p. e793-800.
188. Galvin, J.J., 3rd, Q.J. Fu, and G. Nogaki, *Melodic contour identification by cochlear implant listeners.* Ear Hear, 2007. **28**(3): p. 302-19.
189. Nimmons, G.L., et al., *Clinical assessment of music perception in cochlear implant listeners.* Otol Neurotol, 2008. **29**(2): p. 149-55.
190. Singh, S., Y.Y. Kong, and F.G. Zeng, *Cochlear implant melody recognition as a function of melody frequency range, harmonicity, and number of electrodes.* Ear Hear, 2009. **30**(2): p. 160-8.
191. Vongpaisal, T., S.E. Trehub, and E.G. Schellenberg, *Song recognition by children and adolescents with cochlear implants.* J Speech Lang Hear Res, 2006. **49**(5): p. 1091-103.
192. Galvin, J.J., 3rd, Q.J. Fu, and S. Oba, *Effect of instrument timbre on melodic contour identification by cochlear implant users.* J Acoust Soc Am, 2008. **124**(4): p. EL189-95.
193. Gfeller, K., et al., *Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients.* J Am Acad Audiol, 2002. **13**(3): p. 132-45.
194. Gfeller, K., et al., *Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal.* Ann Otol Rhinol Laryngol, 2002. **111**(4): p. 349-56.
195. Leal, M.C., et al., *Music perception in adult cochlear implant recipients.* Acta Otolaryngol, 2003. **123**(7): p. 826-35.
196. Brockmeier, S.J., et al., *Comparison of musical activities of cochlear implant users with different speech-coding strategies.* Ear Hear, 2007. **28**(2 Suppl): p. 49S-51S.

197. Wong, A.O. and L.L. Wong, *Tone perception of Cantonese-speaking prelingually hearing-impaired children with cochlear implants.* Otolaryngol Head Neck Surg, 2004. **130**(6): p. 751-8.
198. Wei, C., et al., *Psychophysical performance and Mandarin tone recognition in noise by cochlear implant users.* Ear Hear, 2007. **28**(2 Suppl): p. 62S-65S.
199. Wei, C.G., K. Cao, and F.G. Zeng, *Mandarin tone recognition in cochlear-implant subjects.* Hear Res, 2004. **197**(1-2): p. 87-95.
200. Schatzer, R., et al., *Temporal fine structure in cochlear implants: preliminary speech perception results in Cantonese-speaking implant users.* Acta Otolaryngol, 2010. **130**(9): p. 1031-9.
201. Han, D., et al., *Lexical tone perception with HiResolution and HiResolution 120 sound-processing strategies in pediatric Mandarin-speaking cochlear implant users.* Ear Hear, 2009. **30**(2): p. 169-77.
202. Wilson, B.S. and M.F. Dorman, *Cochlear implants: current designs and future possibilities.* J Rehabil Res Dev, 2008. **45**(5): p. 695-730.
203. Arnoldner, C., et al., *Speech and music perception with the new fine structure speech coding strategy: preliminary results.* Acta Otolaryngol, 2007. **127**(12): p. 1298-303.
204. Riss, D., et al., *1-year results using the Opus speech processor with the fine structure speech coding strategy.* Acta Otolaryngol, 2008: p. 1-4.
205. Riss, D., et al., *A new fine structure speech coding strategy: speech perception at a reduced number of channels.* Otol Neurotol, 2008. **29**(6): p. 784-8.
206. Donaldson, G.S., H.A. Kreft, and L. Litvak, *Place-pitch discrimination of single- versus dual-electrode stimuli by cochlear implant users (L).* J Acoust Soc Am, 2005. **118**(2): p. 623-6.
207. Firszt, J.B., et al., *Current steering creates additional pitch percepts in adult cochlear implant recipients.* Otol Neurotol, 2007. **28**(5): p. 629-36.
208. Koch, D.B., et al., *Using current steering to increase spectral resolution in CII and HiRes 90K users.* Ear Hear, 2007. **28**(2 Suppl): p. 38S-41S.
209. Chen, J., et al., *Simulated phase-locking stimulation: an improved speech processing strategy for cochlear implants.* ORL J Otorhinolaryngol Relat Spec, 2009. **71**(4): p. 221-7.
210. Nie, K., G. Stickney, and F.G. Zeng, *Encoding frequency modulation to improve cochlear implant performance in noise.* IEEE Trans Biomed Eng, 2005. **52**(1): p. 64-73.
211. Sit, J.J., et al., *A low-power asynchronous interleaved sampling algorithm for cochlear implants that encodes envelope and phase information.* IEEE Trans Biomed Eng, 2007. **54**(1): p. 138-49.
212. Glass, G.V., *Primary, secondary, and meta-analysis of research.* Educational Researcher, 1976. **5**: p. 3-8.
213. Christensen, L.B., *Experimental methodology*. 10th ed. 2007, Boston: Pearson/Allyn & Bacon. xvi, 528 p.
214. Liberman, M.C., *Auditory-nerve response from cats raised in a low-noise chamber.* J Acoust Soc Am, 1978. **63**(2): p. 442-55.
215. Hall, J.W., 3rd, *New handbook of auditory evoked responses*. 2007, Boston: Pearson Education, Inc. 750.