Combination tone: Absent but audible component

Kaoru Ashihara
National Institute of Advanced Industrial Science and Technology, AIST, 1–1–1 Higashi, Tsukuba, 305–8566 Japan

Abstract: When two or more tones are presented simultaneously, a listener can sometimes hear other tones that are not present. These other tones called ‘combination tones’ are thought to be induced by nonlinear activities in the inner ear. It is difficult to demonstrate this phenomenon because a listener cannot easily distinguish combination tones from primary tones. This paper introduces a unique method called the ‘sweep tone method’ by which combination tones can be perceptually distinguished from primary tones relatively easily. The importance of the non-linear characteristics of the intact auditory system is described.

Keywords: Combination tone, Outer hair cell, Sensory hearing loss, Sweep tone

PACS number: 43.64.Kc [doi:10.1250/ast.27.332]

This article contains the supplementary media files. Underlined file names in the article correspond to the supplementary files. For more information, see http://www.asj.gr.jp/2006/data/ast2706.html.

Supplementary files
tone_a.wav, tone_b.wav, c_demo.wav, pat_1.wav, pat_2.wav

1. INTRODUCTION

When two tones whose frequencies are close to each other are presented simultaneously, sometimes a listener with an intact auditory system hears a third tone that is not present. A fourth tone may also be heard. In those cases, the presented tones are called ‘primary tones’ and the third and fourth tones are called ‘combination tones.’ When the frequencies of primary tones are \( f_1 \) and \( f_2 \) \((f_1 < f_2)\), components at the frequencies \( 2f_1 - f_2 \) and \( f_2 - f_1 \) are usually perceived. More combination tones are heard for \( f_2/f_1 < 1.5 \) [1]. In daily environments, however, a combination tone is not always recognized by the listener. To recognize a combination tone, the listener has to distinguish it from primary tones, which is not very easy. This paper shows that a combination tone can be distinguished relatively easily from primary tones by the sweep tone method. How a combination tone is related to the dynamic range of the inner ear will be described.

2. LISTENING TO COMBINATION TONE

Prepare a stereophonic audio reproduction system. It is recommended to use two loudspeakers in order to prevent the system from producing intermodulation distortions. Place two loudspeakers close to each other in front of a listener, as can be seen in Fig. 1, because it is preferred that primary tones come from the same direction.

First, reproduce the included file named tone_a.wav. A pure tone at 2.05 kHz will be presented by the left loudspeaker. Adjust its level to be about 30 dB above the threshold. Then reproduce the file named tone_b.wav at the same level. An ascending tone whose frequency sweeps from 2.1 kHz to 2.6 kHz will be presented by the right loudspeaker. Now, reproduce the file named c_demo.wav. C_demo.wav is a combination of the two sounds previously heard, tone_a.wav and tone_b.wav. A pure tone and an ascending sweep will be presented by the left and right loudspeakers, respectively. You can hear primary tones A and B shown in Fig. 2. Can you detect anything else? Some listeners may hear a descending sweep, which is a combination tone.

It is known that when the frequencies of the primary tones are \( f_1 \) and \( f_2 \), components at the frequencies \( 2f_1 - f_2 \) and \( f_2 - f_1 \) will be heard. When the tones shown in Fig. 2 are presented, \( 2f_1 - f_2 \) sweeps from 2 kHz to 1.5 kHz so that the descending sweep can be heard. This is shown in Fig. 3. The broken lines in the figure represent the combination tone of \( 2f_1 - f_2 \). The audibility of the combination tone varies considerably among listeners [1,2]. If only small effects are observed, adjust the level and try again.

Note that only a constant pure tone and an ascending tone are presented by the loudspeakers. Descending sweeps cannot be recorded with a microphone placed near the
If the channel separation and the linearity of the sound reproduction system are very poor, descending sweeps may be produced by the system as intermodulation distortions.

When the presentation level is increased, some listeners may detect another ascending sweep in a lower-frequency region, shown as dotted lines in Fig. 3. This component whose frequency is $f_2 - f_1$ is also a combination tone and is sometimes called a ‘difference tone.’

Listen to the files pat1.wav and pat2.wav. Their time-frequency patterns are shown in Fig. 4. Although these examples contain only ascending sweeps, some listeners will detect descending sweeps. In the case of
The 2\(f_1 - f_2\) component is a descending sweep. In the case of pat_2.wav, the \(f_2 - f_1\) component is a descending sweep.

It is thought that a combination tone is the production of nonlinear activity in the intact inner ear. Listeners with sensory hearing deficiency cannot hear a combination tone.

### 3. MEASUREMENT OF COMBINATION TONE

Until the 1980s, researchers tried to investigate a combination tone mostly in psychoacoustic experiments. To measure a combination tone in psychoacoustic paradigms, it has to be perceptually distinguished from primary tones by the listener or its effects have to be observed objectively. Researchers have proposed several methods: the canceling method [2–5], loudness-balancing method [3], and forward-masking method [6,7].

When a combination tone is recognized by a listener, it can be canceled out by the addition of a cancellation tone with the same frequency as and a phase opposite to the combination tone. In the canceling method, the listener has to adjust the amplitude and phase of the cancellation tone to cancel the combination tone. When the combination tone is canceled, the cancellation tone is supposed to be approximately equal in amplitude and opposite in phase to the combination tone. This method requires skilled listeners.

In the loudness-balancing method, a comparison tone is presented shortly after the primary tones have been presented. The frequency of this comparison tone is the same as that of the combination tone. A listener has to adjust the amplitude of the comparison tone to balance its loudness with that of the combination tone. In this method, the combination tone has to be perceptually distinguished from the primary tones by the listener, but this is not easy.

When a low-level probe signal at the frequency of the combination tone is presented immediately after the primary tones, the probe signal is subjected to the forward-masking caused by the combination tone. The loudness of the combination tone can be estimated by measuring the masking effects caused by the combination tone. This is the forward-masking method.

The loudness-balancing and the forward masking methods can be used to estimate the loudness of the combination tone. They are not preferred, however, for the measurement of the detection threshold for the combination tone. The detection threshold for the combination tone can be directly measured by the sweep tone method because the listener can easily distinguish a combination tone from a primary tones using this method [8]. If the listener is able to discriminate between ascending and descending tones, no other special skill is required in this method.

Recently, distortion product otoacoustic emissions (DPOAEs) have often been used to study the functions of the peripheral auditory system. Otoacoustic emissions (OAEs) are sounds emitted from the ear and can be recorded with a probe microphone inserted in the outer ear [9]. When primary tones at \(f_1\) and \(f_2\) are presented simultaneously, acoustic emissions at the frequency \(2f_1 - f_2\) can be observed. This is a DPOAE. It is believed that a DPOAEs are produced when active motions of outer hair cells amplify vibrations on the basilar membrane.

### 4. ORIGINS OF COMBINATION TONE

Although the mechanisms responsible for combination tones are not fully understood, it is widely accepted that combination tones are originally produced in the inner ear. Although the outer ear and middle ear are assumed to be linear transducers up to about 100 dB SPL [10–13], the combination tone at \(2f_1 - f_2\) can be induced by primary tones at an input level of about 15 dB above the threshold. The perception of the combination tone strongly depends on the frequency ratio of primary tones \((f_2/f_1)\) suggesting that the combination tone originates somewhere in the cochlear where frequency analysis takes place. Because DPOAEs can be observed only in living animals, some active mechanisms are thought to be involved [14].

### 5. NONLINEAR CHARACTERISTICS AND DYNAMIC RANGE OF INNER EAR

The wide dynamic range of an intact auditory system is thought to be due to the active amplification in the inner ear. It seems that weak sounds are selectively amplified by active mechanisms in the inner ear. The input-output function of the inner ear, therefore, shows strong non-linear characteristics. The input-output function of the inner ear is schematically shown in Fig. 5. The broken lines represent the characteristics of the intact inner ear. Because of the selective amplification, this line crosses the threshold at the input level of about 0 dB and, therefore, the dynamic range of the intact inner ear is as wide as about 120 dB. This selective amplification seems to be achieved by active movements of the outer hair cells.

The solid line that crosses the threshold at the input level of about 40 dB represents the characteristics of a damaged inner ear. When outer hair cells are not functioning, low-level sounds are not amplified. This is why the dynamic range in sensory hearing loss is limited.

The active amplification in the inner ear seems to have strong frequency selectivity. The sharp dip in the tuning curve of the normal auditory system disappears after an animal’s death [14]. One of the main reasons why the compensation of sensory hearing loss is difficult is that conventional hearing aids can amplify low-level sounds but cannot mimic the strong frequency selectivity of the amplification in the inner ear.
Nonlinear characteristics in the auditory system are closely related to the intact function of the inner ear. The study of the properties of combination tones may give us more knowledge about the functions of the inner ear.

REFERENCES