2 Acoustic Stimuli Used in Diagnostic Audiology

Samuel R. Atcherson, M. Bryson Howard, and Jason W. Johnson

Abstract

This chapter provides an overview of the acoustic stimuli commonly used in auditory and vestibular assessments. In audiology clinics, these stimuli include sounds with relatively simple and static acoustic properties, such as pure tones, spectrally and temporally dynamic sounds such as frequency- and amplitude-modulated tones, synthesized speech sounds that mimic the complex acoustic characteristics of consonants and vowels in natural spoken language, and recorded or live human speech. These and other types of acoustic stimuli serve a wide range of diagnostic purposes in audiology. Clinical applications include basic behavioral assessments, such as the familiar pure tone detection threshold, elicitation of otoacoustic emissions to evaluate cochlear function, evocation of myogenic responses to assess vestibular function, and verification of aided audibility using speech-evoked cortical auditory potentials, among others. The outcomes of diagnostic tests in audiology can be influenced by various acoustic properties of sound stimuli, including intensity, frequency, presentation rate, duration, polarity, modulation depth, and envelope characteristics. Therefore, clinicians can benefit from a basic understanding of acoustic stimuli to effectively apply them in a variety of audiological contexts.

Keywords: frequency, intensity, spectral, temporal, pure tone, noise, speech, click, toneburst, chirp

2.1 Introduction

The assessment of the auditory and vestibular systems falls within the scope of practice for audiologists. To carry out these assessments, audiologists have access to a variety of stimuli. Although both acoustic and non-acoustic stimuli can be used, this chapter focuses exclusively on acoustic stimuli—that is, sounds presented specifically to the ear, primarily through air conduction. These stimuli possess multiple acoustic properties that can influence the measured response, whether behavioral or physiological. Some properties, such as the frequency or intensity of a pure tone, are relatively easy to conceptualize in terms of perceived pitch or loudness. Other properties, such as the polarity or bandwidth used to evoke an auditory brainstem response (ABR), are less straightforward to define in terms of perceptual qualities but can still significantly affect the characteristics of the measured response. It is crucial for audiologists to understand the sounds they use in the clinic not only from the perspective of their own listening experiences but also through a foundational understanding of their acoustic properties. This knowledge allows for more informed and effective clinical practice. While this chapter focuses on acoustic stimuli applicable across various diagnostic tests, readers are encouraged to consult Chapter 3 for information on basic instrumentation and the calibration of stimuli.

2.1.1 Listening Conditions for Acoustic Stimuli

Audiologists use a variety of diagnostic instrumentation presenting acoustic stimuli to one or both ears. Acoustic stimuli are most often presented unilaterally (or monaurally or monotically) to one ear at a time to obtain ear-specific information. The type of transducer (e.g., headphone, earphone, soundfield, probe) used and the relative hearing sensitivity differences between ears will dictate whether or not the stimulus is truly ear-specific. For example, if there is a large difference in hearing sensitivity between ears, presenting a stimulus at too high an intensity level in the poorer test ear may acoustically leak and/or may be transmitted via skull/bone vibration, either of which results in crossover to the other ear. Whenever ear-specific information is required, the audiologist may need to mask the nontest ear with a noise (i.e., masker) to ensure that the nontest ear is not contributing to the results (see, e.g., Chapter 10). In addition to assessing hearing sensitivity, audiologists have several speech tests, including speech-in-noise (SIN) tests, at their disposal, which can also be unilaterally presented to gain a sense of functional speech understanding abilities in each ear or bilaterally presented to examine overall speech understanding abilities (see, e.g., Chapter 11).

Special Consideration

When stimuli presented to one ear is detected by the opposite ear, this is known as *crossover hearing*. This is an undesirable effect during frequency-specific air- and bone-conduction hearing assessments, which calls for masking the nontest ear with noise. However, there are benefits to crossover hearing via bone conduction in the case of bilateral atresia or single-sided (unilateral) deafness.

Some tests, however, involve presenting stimuli to both ears in a bilateral (or binaural) manner. Bilateral stimuli containing the same information presented to both ears constitute *diotic* stimulation, whereas bilateral stimuli with different information presented to each ear constitute *dichotic* stimulation. Diotic stimulation can occur through simultaneous presentation of the same stimuli via independent transducers in each ear, by presenting the stimuli in the soundfield through loudspeakers, or by using a bone-conduction oscillator that vibrationally transmits information to both cochleas. While dichotic stimulation is rare in real-world settings, presenting different stimuli to each ear can provide valuable insights into the central auditory system and higher-order auditory processing abilities. One well-known example of dichotic stimulation is the Dichotic Digit (DD) test,[1] where pairs of single-syllable numbers are presented separately to each ear, such as “one five” to the right ear and “nine four” to the left ear. The patient verbally responds by recognizing and reporting all four numbers, if possible. The DD test is an example of a *binaural integration* test. Another type of

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

dichotic stimulation is the Competing Sentences (CS) test.[2] In the CS test, different sentences are presented to each ear, and the patient is instructed to attend to one ear and repeat what they heard, all while ignoring the other ear. The CS test is an example of a *binaural separation* test.

While most diagnostic audiology tests are presented unilaterally, the specific anatomy and physiological correlates of progressively higher levels of the auditory, vestibular, and related systems may reveal information about one side of the head versus the other. Stimuli presented to one ear may provide clues about the function of structures on the same side as the stimulated ear, referred to as an *ipsilateral* condition. Conversely, when stimuli provide information about the function of structures on the opposite side of the head relative to the stimulated ear, this is referred to as a *contralateral* condition. For instance, the acoustic reflex test can assess the ipsilateral and contralateral pathways of the acoustic reflex arc, which is particularly useful for identifying possible lesions beyond the cochlea (see, e.g., Chapter 16). The ABR, a common evoked potential test, can also be used to assess ipsilateral and contralateral central auditory system pathways up to the level of the lateral lemniscus and inferior colliculus using a unilateral stimulus. These pathways are recorded with strategically placed electrodes that measure neurophysiological dipole activity. Another evoked potential, the vestibular evoked myogenic potential (VEMP) test, has two types—cervical (cVEMP) and ocular (oVEMP)—each with distinct ipsilateral and contralateral characteristics. For example, using air-conduction stimulation, the cVEMP is primarily an ipsilateral response (e.g., right ear and right sternocleidomastoid muscle), whereas the oVEMP is primarily a contralateral response (e.g., right ear and left inferior oblique muscle). Exceptions do exist; for instance, although the oVEMP is primarily a contralateral response, a small ipsilateral response may also occur.[3] More detailed information on VEMP tests is provided in Chapter 24.

Related to the stimulation of one or both ears, audiology equipment and acoustic stimulus materials are often organized into *channels* of stimulation. For example, most audiometers are two-channel systems, with one channel dedicated to the right ear and the other to the left. While stimuli can be presented to one ear using a single channel (monotic), clinicians may also need to independently control stimulus intensity between channels or present stimuli in diotic or dichotic configurations. Acoustic stimuli can be presented as *mono* (one channel) or *stereo* (two channels). Examples of how two independent channels may be used include the following:

- Testing one ear in one channel while masking the contralateral (nontest) ear with the other channel.
- Conducting SIN testing by presenting speech to one ear via one channel while controlling competing noise in the same ear using the other channel to vary the signal-to-noise ratio (SNR).
- Presenting dichotic stimuli with different speech inputs to each ear using two channels.

In summary, readers should consider the diverse tests administered by audiologists, focusing on the available stimulus options, whether stimulation is unilateral or bilateral, and whether the stimuli share the same or different information (e.g., diotic or dichotic). A summary of these listening conditions is provided in ▶ Table 2.1.

Table 2.1 Listening conditions

Listening condition	Description
Unilateral	Refers to something occurring in or involving only one ear or one side of the body. In audiology, it typically describes hearing conditions or tests that affect or involve one ear. Also known as monaural or monotic.
Bilateral	Refers to something occurring in or involving both ears or both sides of the body. In audiology, it commonly refers to hearing conditions, tests, or measurements that involve both ears. Also known as binaural.
Diotic	Refers to the presentation of identical auditory stimuli (e.g., pure tones, speech) to both ears simultaneously. Diotic stimuli are commonly used in hearing threshold tests and speech recognition assessments. The same signal is delivered to both ears at the same intensity.
Dichotic	Refers to the simultaneous presentation of different auditory stimuli to each ear. In dichotic listening tests, two different sounds, such as speech or tones, are presented at the same time—one to each ear. Can have binaural summation or binaural separation types.
Ipsilateral	Refers to something occurring or located on the same side of the body or the same ear.
Contralateral	Refers to something occurring or located on the opposite side of the body or ear.
Mono	Refers generally to presentation of stimuli to one ear via one channel in a monotic manner.
Stereo	Refers generally to presentation of stimuli to both ear via two channels, though the stimuli may be diotic or dichotic.

Special Consideration

Many clinicians continue to use compact discs (CDs) to present speech materials through a CD player routed to the audiometer. Unless otherwise specified, the digitized material on CDs are split into right and left ear channels, which are then set up and independently controlled using the audiometer. The clinicians must then ensure that they have set up the channel inputs and outputs correctly. It is also worth noting that identical stimuli may be on both the left and right channels (for either mono or stereo presentation), or different stimuli in the left and right channels (also for either mono or stereo presentation).

Special Consideration

As testing techniques continue to evolve and advance along with new applications in instrumentation, we may come to see combinations of listening conditions used. For example, modern-day auditory steady-state response (ASSR) testing may permit the audiologist to present four of the same frequency-specific stimuli (e.g., 500, 1,000, 2,000, and 4,000 Hz) in both ears (i.e., binaural and diotic) and yet they are presented at different rates between the ears (e.g., binaural and dichotic) so that neurons of the central auditory system can differentially phase-lock to them. See Section 2.2.1 and Chapter 21 to learn more.

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

2.1.2 Visualizing the Properties of Acoustic Stimuli

Acoustic stimuli can be visualized (and characterized) using a variety of tools. Readers can find more detailed information about these tools elsewhere, but for this chapter, a basic understanding of the various properties of acoustic stimuli can help us understand how they interact with (and are influenced by) the anatomy and physiology to arrive at the results we obtain. In general, sound can be characterized by physical characteristics such as its frequency content, intensity, and overall duration. Frequencies correspond to the number of cycles per second of a sound wave, measured in *hertz* (Hz). Acoustic stimuli can comprise one or more frequencies. The intensity of sound waves is measured in *decibels* (dB), usually in *sound pressure level* (SPL) or converted to *hearing level* (HL). The duration of sound is typically measured in milliseconds (ms) or seconds (s) and, in the case of very short-duration click stimuli, microseconds (µs).

Pearl

Although audiometric pure tones are physically measured in decibels in sound pressure level (dB SPL) using calibration tools such as a sound level meter, dB SPL is converted to decibels in hearing level (dB HL) on an audiogram. Because thresholds in dB SPL are not equal across frequencies, conversion to dB HL forces all thresholds to a zero dB HL line. This greatly aids in patient comprehension during the explanation of audiogram test results by the clinician.

Visualizing acoustic stimuli in the *temporal (or time) domain* is useful for observing how the stimulus waveform changes as a function of time. Here, we can look at factors such as rise time, plateau time, fall time, and repetition rate. An oscilloscope is especially useful to see what the acoustic stimulus looks like. We can visualize the simplicity or complexity of frequency content using a spectrum analyzer, which attempts to deconstruct stimuli into its individual or predominant frequencies in the *frequency domain*. Both the time and frequency domains can also provide measured intensity (or amplitude) levels for the acoustic stimuli of interest. The *envelope* is another important temporal property of sound stimuli for clinical testing. The envelope of a sound characterizes its slow variations in peak amplitude over time. However, abrupt *onsets* or *offsets* of sounds at their beginning and end may produce audible distortion or clicks from the transducer. It is common, then, to "ramp up" and "ramp off" the amplitude of sounds at their onsets and offsets (i.e., rise and fall times) over several milliseconds to reduce *spectral splatter*. This aspect of tonal stimuli can impact the accuracy of a behavioral hearing test; thus, it is one of the performance aspects measured during annual equipment calibrations. A term related to stimulus envelopes is *windowing*. Different types of window functions can be applied to stimuli (e.g., Linear, Hamming, Hanning, Blackman) that differ in their spectral onset and offset effects. Last but not least, and related to the movement of a transducer diaphragm (e.g., insert earphone or supra-aural headphone), the physical polarity of the stimulus may influence test results. When the transducer diaphragm moves toward the ear canal, it is of *condensation* polarity (also called positive polarity), whereas diaphragm movement away from the ear canal is of *rarefaction* polarity (also called negative polarity). These terms are based on the movement of air molecules within the ear canal and, as a consequence, the physiological movements of the eardrum, ossicles, and hydromechanical movements in the fluid-filled inner ear. As there can be both physiological and stimulus transducer effects (e.g., electromagnetic) by the stimulus, sometimes an *alternating* polarity pattern (i.e., positive/negative polarity) will be required.

Special Consideration

Spectral splatter is often described as the spread of energy as a consequence of rapid (or instantaneous) onsets or offsets of a stimulus. To a listener, any rapid change in the stimulus yields an audible "pop" or "click." Physiologically, these rapid onsets result in broadband stimulation of the cochlea. This is a desirable stimulus quality for some audiological tests, such as the ABR. Other times, this splatter is unwanted with efforts to minimize the splatter by windowing the onsets and offsets.

Special Consideration

Clinicians cannot discount the importance of stimulus calibration, especially to ensure that the intensity and the purity of signals presented to patients meet expected standards. Regular calibration ensures that equipment and related accessories are in good working order and that the results obtained clinically are deemed reliable regardless of where the patient was tested. In other words, calibration helps with consistency across clinics.

To aid readers in visualizing acoustic stimuli throughout the chapter, we will be using two free software to create images to feature acoustic stimuli in the time and frequency domains or to show how stimuli interact with the cochlea and early neural structures. One software, Audacity, is a free, open source, cross-platform software for recording and editing sounds (http://www.audacityteam.org). With this software, various stimuli can be (1) generated or imported, (2) edited, (3) analyzed, and (4) recorded and played back. ▶ Fig. 2.1 shows an example of how Audacity can be used to create a pure tone and analyze its frequency spectrum. In addition, there are some nice export functions to save audio files and produce data to create graphs. The other software, CochSim, is an interactive Windows PC simulator program for demonstrating how the cochlea analyzes sounds (https://www.phon.ucl.ac.uk/resource/cochsim/). CochSim is not public domain software, but it may be used and copied without charge as long as the program remains unmodified and continues to carry its copyright notice. ▶ Fig. 2.2 illustrates some of the features of the CochSim to assist in our understanding of how acoustic stimulation interacts with the oval window, the basilar membrane, hair cells, and early neural structures.

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

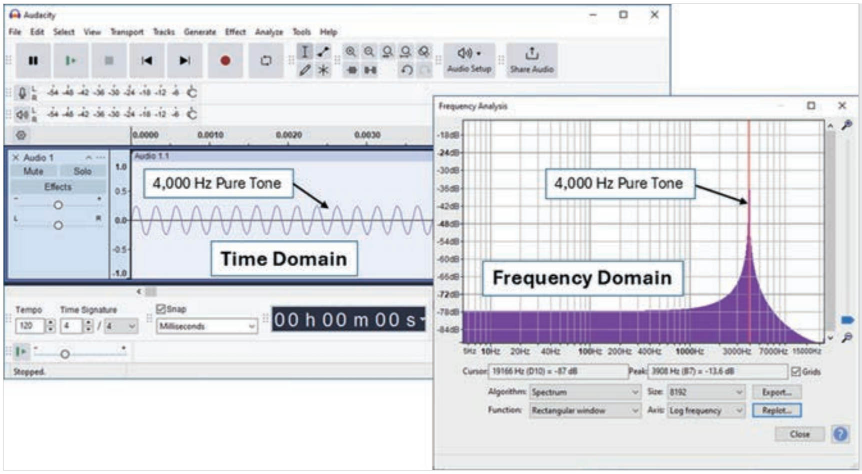

Fig. 2.1 Screenshots of the Audacity software showing a 4,000 Hz pure tone in the time domain (top left) and frequency domain (bottom right).

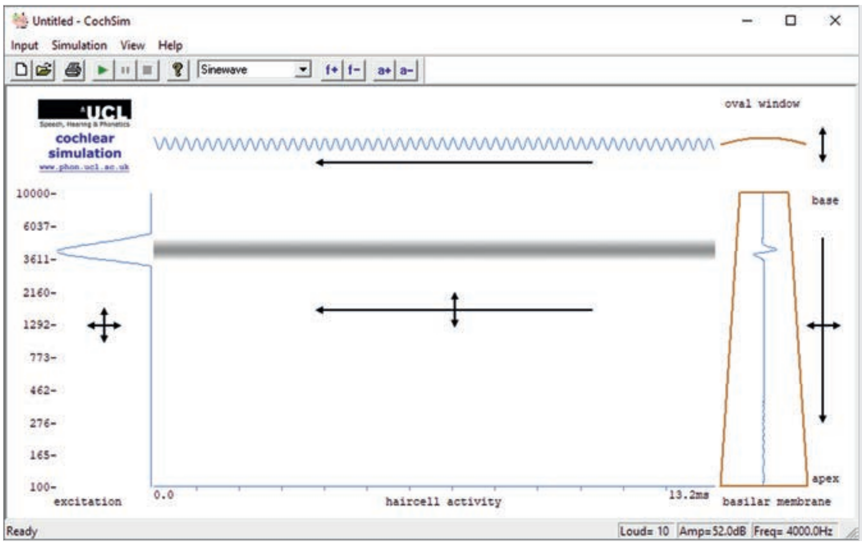

Fig. 2.2 Screenshot of the CochSim software. At the top of the software, there is a 4,000 Hz sine wave. At the top right corner, the oval window moves up and down in tandem with the sine wave. On the right-hand side of the screen outlined as a brown trapezoidal shaped box is the basilar membrane arranged from high frequencies (top) to low frequencies (bottom). Inside of the brown box shows the basilar membrane as a traveling wave in blue. In the large spacious area in the center of the screen, hair cell activity is represented based on the movements of the traveling wave. Finally, on the left side of the screen, summed neural activity is shown along with the frequencies of stimulation. The *arrows* are placed by the author to show directions the various simulation software features go.

2.2 Types of Stimuli Used in Audiology

2.2.1 Tonal Stimuli

In this section, a variety of tonal stimuli is described. Generally, these tones are used for frequency-specific assessments, but they may be modified or combined with other tones for unique assessment purposes. ▶ Table 2.2 lists the various clinical tests that may employ the use of tonal stimuli.

Pure Tone

A *pure tone* is a periodic sine wave that oscillates at a single, fixed frequency, meaning it has no other harmonics or overtones. It is highly effective for frequency-specific purposes, especially for the assessment of hearing whether subjective and behavioral or objective and physiological. Previously, ▶ Fig. 2.1 showed an example of a 4,000 Hz pure tone in the time and frequency domains. There are other testing possibilities using pure tones that may be found in audiology. For example, using more than one transducer, it is also possible to have two or more pure tones presented that mix in the ear canal to one that becomes a complex stimulus at the eardrum, yet functionally separates again at the level of the cochlea and beyond. ▶ Fig. 2.3 shows an example of a two-tone stimulus using the CochSim software.

Pearl

Tuning forks can be helpful sources of pure tone stimulation for bedside examinations. The tuning fork is struck and the tines oscillate in a repeating, periodic manner. Depending on the specific tuning fork test, the stimulus can be delivered via air or bone conduction.

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

Modulated Tones

Several applications in diagnostic audiology implement pure tones with constant frequency and peak amplitude. However, there are applications in diagnostic audiology that incorporate variations or *modulation* to pure tone frequency (FM) or amplitude (AM) for a specific purpose. Modulation in frequency can be visualized in the time domain as variation in the period of a sinusoid over time, while AM is evident as periodic variation in peak amplitude over time (compare FM and AM in ▶ Fig. 2.4). It is common to use frequency-modulated *warble tones* when establishing soundfield thresholds because these stimuli are less sensitive to standing waves in the test booth than are standard pure tones. Other clinical applications for modulated tones are more related to attention than acoustics. Particularly in pediatric or developmentally delayed populations, clinicians often discover that switching tone types (e.g., using frequency-modulated tones) can add novelty and better hold the attention of distracted testers than utilizing only one type of tonal presentation (see also narrowband and FRESH noises in Section 2.2.2 as alternative to tonal stimulation).

Special Consideration

Pulsing a tone (pulsed tones) refers to changing the presentation of the tone from continuous to adding variations in the amplitude to the tone. This gives the tone a pulsing (rhythmic) quality that has been shown to be more easily detected, especially with tinnitus patients.[6]

2.2.2 Noise

In this section, a variety of noise stimuli is described. Generally, noise is used to mask the nontest ear or to assess perceptual SIN functions, but they may also be used as stimuli for hearing sensitivity thresholds. ▶ Table 2.3 lists the various clinical tests that may employ the use of noise.

Table 2.2 Tonal stimuli used in audiology

Test	Tonal stimuli used
Audiometry	Pure tone audiometry (125, 250, 500, 1,000, 1,500, 2,000, 3,000, 4,000, 6,000, and 8,000 Hz pure tones, which may be pulsed or warble) Tone decay (500, 1,000, 2,000, or 4,000 Hz pure tones) High-frequency audiometry (e.g., 10,000 through 20,000 Hz pure tones)
Immittance	Tympanometry (e.g., 226 and 1,000 Hz pure tones) Acoustic reflex thresholds (500, 1,000, 2,000, and 4,000 Hz pure tones) Acoustic reflex decay (500 and 1,000 Hz pure tones)
Otoacoustic emissions	Distortion-product otoacoustic emissions (e.g., f1 and f2 tone pairs at a fixed ratio [e.g., 1.22] such as f2 = 2,000 Hz and f1 = 1,640 Hz)
Evoked potentials	Auditory steady-state response (e.g., amplitude modulation and/or frequency modulation tones presented at specific rates to cause neural phase-locking in the auditory brainstem and cortex; compare to narrowband chirps in Section 2.2.4 For other evoked potentials, brief tones (i.e., tonebursts) are better described as transient stimuli (see also Section 2.2.4)
Other	Frequency (Pitch) Pattern Test (FPT)[4] (e.g., 880 and 1,100 Hz presented in triads) Duration Pattern Test (DPT)[5] (e.g., 1,000 Hz presented in short [e.g., 250 ms] and long [e.g., 500 ms] durations)

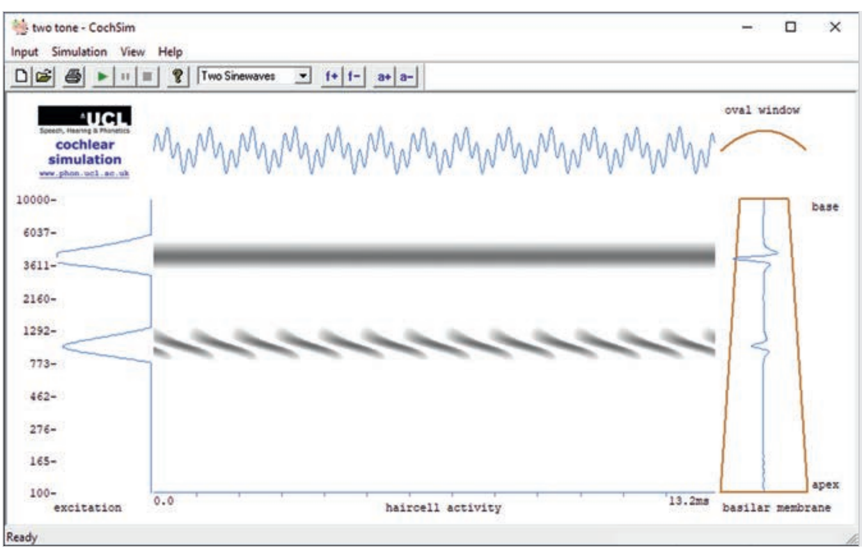

Fig. 2.3 Screenshot of a two-tone stimulation using the CochSim software. A two-tone stimulus of 1,000- and 4,000 Hz pure tones was custom generated in Audacity and imported into the CochSim software. Note distinct cochlear excitation at two different frequencies.

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

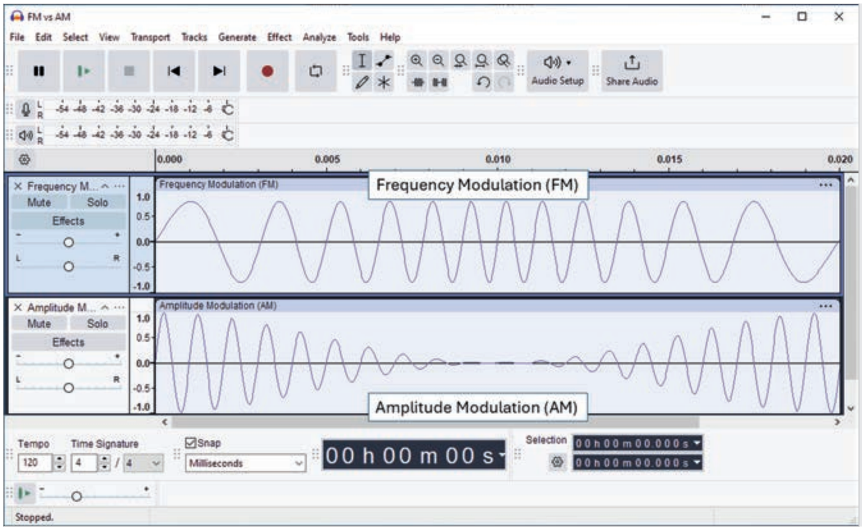

Fig. 2.4 Screenshots of the Audacity software showing stimulus examples of frequency modulation (FM; top) stimuli and amplitude modulation (AM; bottom). Note how the time domain presentations of the FM stimulus changes in frequency, but not amplitude, while the AM stimulus changes in amplitude but not frequency.

Table 2.3 Noise stimuli used in audiology

Test	Noise stimuli used
Audiometry	FRESH noise as an alternative to pure tones, warble tones, or pulsed tones Masking nontest ear for pure tone audiometry or speech testing (e.g., narrowband noise, speech-shaped white noise)
Immittance	Acoustic reflex thresholds (e.g., broadband noise of 300–3,000 Hz wide)
Otoacoustic emissions	Broadband and narrowband noises have been used to explore ipsilateral, contralateral, and bilateral suppression
Evoked potentials	Though rare, noise maskers may be used for nontest ears, including exploration of contralateral masking (e.g., auditory steady-state response)
Other	Gaps-in-Noise (GIN)[7] test (e.g., gaps of various durations are interspersed throughout the 6-s-long bursts of broadband noise)

White Noise

White noise consists of a full range of audible frequencies each played at the same average intensity (▶ Fig. 2.5). Perceptually, it is a constant, static-like sound that masks other sounds, including tinnitus (i.e., minimum masking level). In some diagnostic applications, white noise may be bandlimited, yet maintain their effectiveness. For example, noise-based acoustic reflex threshold (ART) stimuli can be comprised of shaped white noise. That is, the low band, high band, and broadband could be bandpass filtered at 125–1,600, 1,600–4,000, and 125–4,000 Hz, respectively. The filter roll-off rate for these bandpass filter cutoffs is typically set to 12 dB per octave (or greater).

Narrowband

Narrowband noise (NBN) comprises a collection of frequencies around a central focus frequency typically with one-third or one-half octave bandwidths around the specified center frequency. It is often used to mask the nontest ear during pure tone audiometry. NBNs have also been used with infants and young children during soundfield testing to keep their attention. Pure tones are generally uninteresting to infants and young children and produce unwanted standing waves in the soundfield test booth. Unfortunately, NBNs often underestimate hearing thresholds because broad stimulus energy can fall into regions of better hearing. An alternative NBN for threshold testing, while retaining its attention-getting attributes, is the FRESH (FREquency-Specific Hearing assessment) noise.[8] This new NBN has much steeper filter slopes to limit the energy to desired frequency regions in the cochlea. ▶ Fig. 2.6 shows spectral and simulated physiological differences for the NBN and FRESH noise.

Speech-Shaped Noise

In the case of speech recognition testing or SIN testing, neither white noise nor NBN is effective enough to challenge the listener in a manner that better emulates real-world listening. Thus, a speech-shaped (or speech-weighted) white noise would be a better option. This is a type of steady-state noise specifically designed to mimic the frequency spectrum of human speech.[9] ▶ Fig. 2.7 illustrates how a speech-shaped noise looks like a white noise, but the spectrum shows characteristic peaks expected of averaged human speech. Specifically, vowels tend to be more intense and in the lower frequencies (~ < 2,000 Hz), whereas consonants tend to be less intense and in the higher frequencies (~ > 2,000 Hz).

2.2.3 Speech Stimuli

In this section, a variety of speech stimuli is described. Generally, speech is used for speech understanding in quiet or presented with background noise, but it may also be used as stimuli for hearing sensitivity estimates. ▶ Table 2.4 lists the various clinical tests that may employ the use of speech.

Speech

In diagnostic audiology, speech stimuli are used to assess how well a patient hears, understands, and ultimately processes speech. Speech stimuli will vary in complexity and

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

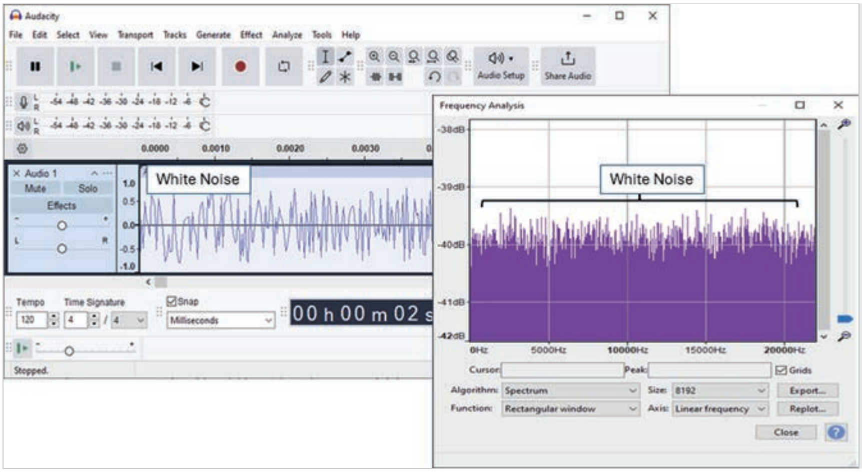

Fig. 2.5 Screenshots of the Audacity software showing a white noise stimulus in the time domain (top left) and frequency domain (bottom right). Note the averaged equal energy across frequencies that is characteristic of white noise.

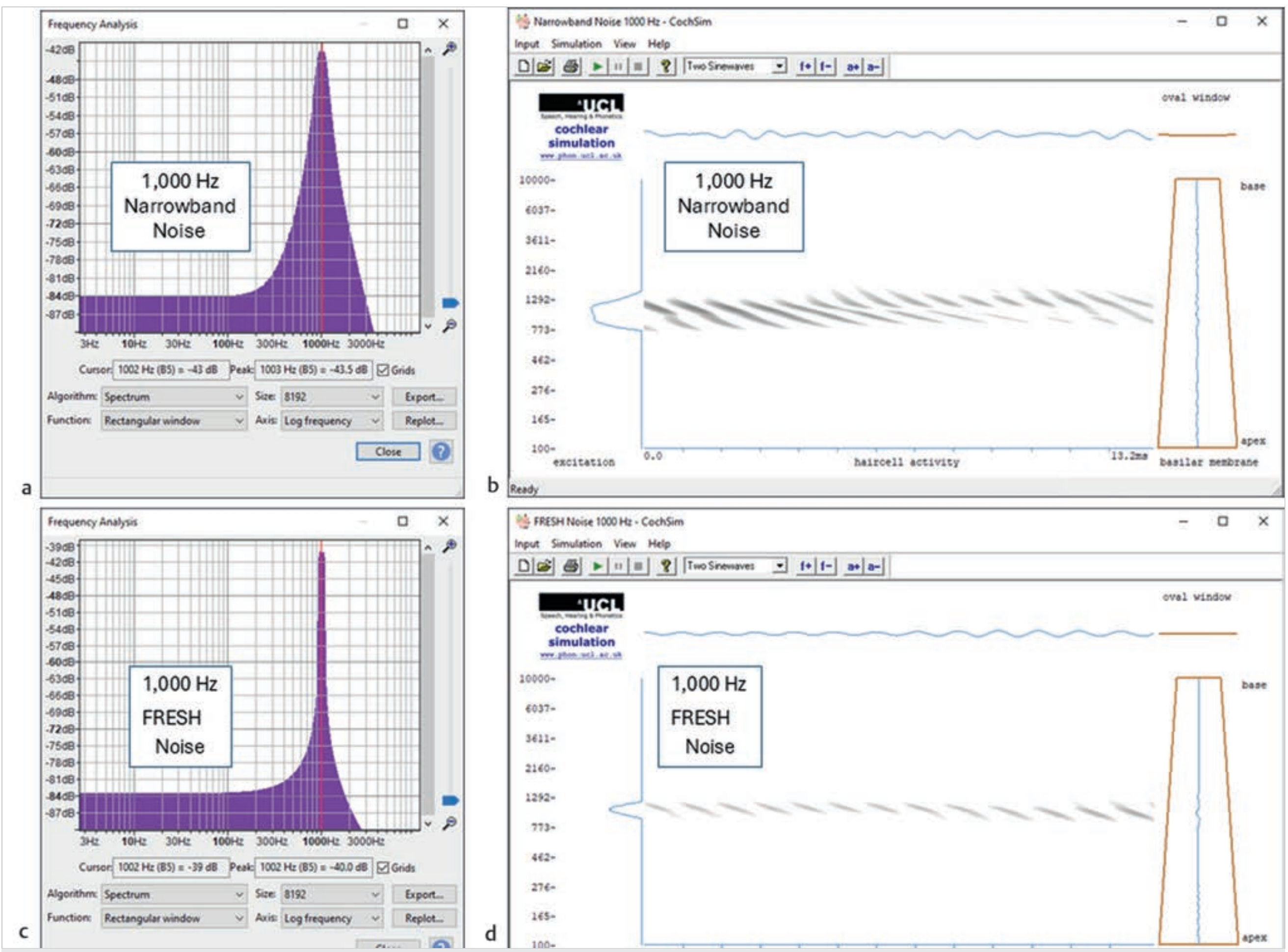

Fig. 2.6 Screenshots of narrowband **(a, b)** versus FRESH **(c, d)** noise using the Audacity and CochSim software. Although the center frequency remains at 1,000 Hz for both stimuli, the narrowband noise has greater spread of energy compared with the FRESH noise as shown in the spectra (**a** vs. **c**) and the physiological simulation (**b** vs. **d**).

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

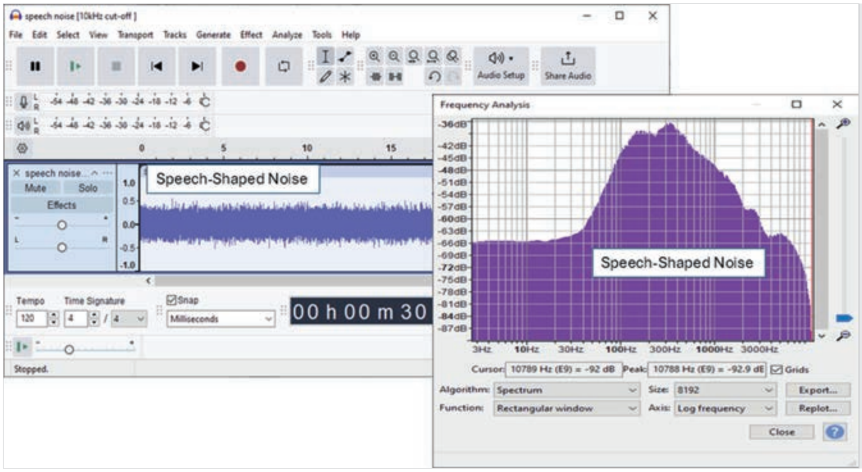

Fig. 2.7 Screenshots of the Audacity software showing a speech-shaped noise stimulus in the time domain (top left) and frequency domain (bottom right). Note the spectrum of the speech-shaped carrying typical long-term average peaks associated with vowels (generally < 2,000 Hz) versus consonants (generally > 2,000 Hz). Note also how different speech-shaped noise spectra is from white noise spectra.

Table 2.4 Speech stimuli used in audiology

Test	Speech stimuli used
Audiometry	Speech awareness threshold (e.g., varies) Speech recognition threshold (e.g., spondee words) Word recognition (e.g., monosyllabic words)
Immittance	None
Otoacoustic emissions	None
Evoked potentials	Speech (or complex) auditory brainstem response[10] (e.g., synthesized 40-ms consonant-vowel /da/ stimulus that evokes an initial transient auditory brainstem response and frequency-following response
Other	Competing Sentences test[11] (e.g., dichotic complete sentences presented to each other, but patient ignores one ear and responds with what they heard in the test ear) Connected Speech Test (CST)[12] (e.g., female talker with average speech intelligibility presented against an adjustable signal-to-noise ratio for the multitalker babble)

structure and include both periodic and aperiodic components (▶ Fig. 2.8). Voiced sounds, such as *vowels*, are periodic because they are produced by the regular vibration of the vocal cords, creating a waveform with a predictable pattern. These sounds contain distinct frequency bands, called *formants*, which shape the unique qualities of each vowel and help listeners identify different phonemes. In contrast, unvoiced *consonant* sounds, such as sibilants (e.g., /s/ and /ʃ/), are aperiodic; they do not have a regular waveform and are characterized by noise-like properties. Speech sounds also vary in pitch, intensity, and duration, contributing to spoken language's rich and dynamic nature. Understanding these properties is essential in audiology for evaluating and managing hearing and speech perception.

Speech stimuli can be presented through recordings or live voice. Speech audiometry adds a dimension to the audiological assessment that accounts for the efficacy of the performance of a person's entire auditory system, both peripheral and central. Different types of speech stimuli are used in audiology and include monosyllabic words, sentences, spondees (i.e., two-syllable words with equal stress on both syllables), nonsense syllables, and words designed to assess detection or recognition at varying intensity levels. Recorded speech provides a more standardized and consistent testing procedure. This method typically includes a calibration tone (usually a sustained 1,000 Hz pure tone), which, when used correctly, helps minimize variability in presentation levels. The calibration tone will have the same root mean square (RMS) level as the speech signal (▶ Fig. 2.9). When live voice is used, it is referred to as monitored live voice (MLV). The term "monitored" indicates that the clinician actively observes the volume unit meter on the audiometer to ensure that their voice presentation remains steady around 0 dB without significant variations in intensity. Proper use of this method requires practice and attention to factors such as presentation level, presentation rate or speed, regional accents, pronunciation, and other potential variables that could skew test results by introducing unintended cues or complications to the task. Due to the variability associated with MLV speech presentation, a recorded speech is preferred.[13] Readers are encouraged to consult Chapter 11 for a more in-depth discussion of speech materials and practices used in audiometry.

Synthesized Speech Sounds

Although not currently widespread in use at this time, some synthesized speech stimuli (/da/, /ba/, /ga/, or other customized speech-like sounds) have been used for audiological assessment. These stimuli contain several features that mimic natural speech but are used whenever there is a need for more acoustic control of certain parameters, such as frequency, duration, amplitude, and formant transitions. One popularized synthetic speech sound used clinically is the /da/ stimulus used to evoke the ABR and the frequency-following response.[10]

Speech-in-Noise Testing

SIN is used to simulate real-world conditions. It is often used to create a challenging listening environment to assess how well a

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

Fig. 2.8 Screenshot of the Audacity software showing a recording of the phrase "Say the word ball." We can easily see amplitude variations over time (top) and the spectrogram (below) shows spectral changes over time.

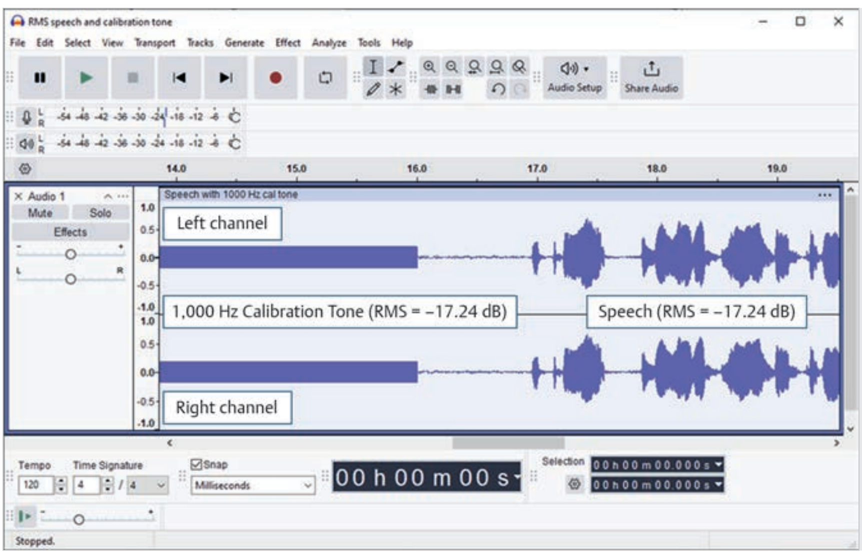

Fig. 2.9 Screenshot of the Audacity software showing a two-channel audio recording of speech test material preceded by a 1,000 Hz calibration tone, each having the same root mean square levels. Note also that the audio track has two channels with the left channel at the top and the right channel at the bottom. The calibration tone is used to adjust the volume unit meter on the audiometer prior to testing.

patient might perform in background noise. SIN materials will include targeted speech words or sentences while a competing background noise is presented. Example SIN tests include the QuickSIN (Quick SIN test),[14] the Hearing in Noise Test (HINT),[15] and the BKB-SIN (Bamford–Kowal–Bench SIN test).[16] Some SIN tests use speech-shaped white noise (see Section 2.2.2), such as HINT. Other SIN tests use speech babble or multitalker babble, which is a speech noise that contains either nonintelligible speech or the energy of many people talking at the same time. Performance on SIN tests can be used also to gauge SNR needs that can influence and optimize a patient's hearing aid or cochlear implant selection and fitting needs. An interesting nonintelligible speech noise is the International Speech Test Signal (ISTS),[17] which consists of segments of natural speech from native speakers of six different languages. This unique characteristic of the ISTS essentially makes it independent of language.

Pearl

SNR is an important concept in SIN testing. When the speech signal is at the same level as the noise, the SNR is 0 dB. When the speech signal is louder than the noise, the magnitude of difference in dB is reflected in a positive manner, such as +10 dB SNR. Conversely, when the speech signal is softer than the noise, the magnitude of difference is reflected in a negative manner, such as −10 dB SNR. During speech audiometry, the speech signal and noise may be independently controlled by two channels or the SNR is already prerecorded into the same channel.

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

Pearl

When it comes to masking speech sounds, two terms come to mind when using noise to mask speech sounds. When steady-state noise, such as white, narrowband, or speech-shaped noise is used, it is considered a form of *energetic masking*. Conversely, when the noise contains elements of speech and is used to mask other speech sounds, it is considered a form of *informational masking*. Interestingly, a single talker babble would result in greater informational masking than would multitalker babble. That is, the more talkers merged into the multitalker babble, the more it takes on an energetic masking form. However, multitalker babble would still result in greater informational and energetic masking than a speech-shaped noise.

Complex Speech Stimuli for Auditory Processing Disorders Testing

Multiple tests typically are included in accessing auditory processing, and many of these tests task the listener with comprehension of complex presentations of speech. These presentations might include digitally altering the speech signal, which increases comprehension difficulty. Examples of this type of stimuli might include filtered speech (using a low, high, or bandpass filter) or time-compressed speech (digitally altered playback of a recorded spoken passage making the presentation unnaturally fast). Another type of listening task to assess auditory processing skills includes tests that challenge the binaural integration or separation of the listener. These tests present independent information to left and right channels and task the listener with comprehension of one side, while ignoring the other or comprehension of both sides at the same time. Along with the DD test described in Section 2.1.1, the Staggered Spondaic Word (SSW) test is a binaural integration dichotic test that presents different spondees to each ear.[18] The SSW is unique as the two spondees are time staggered between the left and right channels such that the second syllable of one spondee overlaps with the first syllable of the second spondee, and the patient needs to repeat both spondees for scoring. As described earlier in Section 2.1.1, CS test is an example of binaural separation, where two different sentences are presented to the two channels and the patient is instructed to ignore one ear, while repeating the sentence they heard in the attended ear.[2,11]

2.2.4 Transient Stimuli Used in Audiology

In this section, a variety of *transient* stimuli are described. Generally, transient stimuli have rapid (or instantaneous) onsets and overall brief durations. They are often used in objective, physiological tests, but in some cases they may be used for some temporal processing tests (e.g., temporal gap detection tests). ▶ Table 2.5 lists the various clinical tests that may employ the use of transient stimuli.

Table 2.5 Transient stimuli used in audiology

Test	Transient stimuli used
Audiometry	None
Immittance	Wideband acoustic immittance[19] (e.g., clicks or chirps)
Otoacoustic emissions	Transient-evoked otoacoustic emissions (e.g., 100-µs click or chirp; see Dau et al[20])
Evoked potentials	Broad uses (i.e., click, chirp, noise bursts, gaps in noise, and abrupt any speech stimuli) Short duration tonebursts (e.g., 500, 1,000, 2,000, and 4,000 Hz) for electrocochleography, auditory brainstem response, middle latency response, and late latency responses (including mismatch negativity and P300) Auditory steady-state response (e.g., chirps presented at specific rates to cause neural phase-locking in the auditory brainstem and cortex)
Other	Random Gap Detection Test (RGDT)[21] (e.g., click or toneburst pairs randomly separated gaps between 0 and 40 ms)

Clicks

The shortest duration stimulus commonly used in diagnostic audiology is the *click* stimulus. To produce a click stimulus, a positive or negative square wave electrical impulse (100-µs duration) is delivered to the transducer, which in turn produces a pressure condensation or rarefaction. Click stimuli have a very broad spectrum. In the following sections, compare how a click differs from tonebursts and chirps. Temporal, spectral, and physiological simulations of the broadband click (and CE-Chirp) are shown in ▶ Fig. 2.10 and ▶ Fig. 2.11.

Tonebursts

A very brief pure tone consisting of only a few cycles is referred to as a *toneburst*. A typical toneburst will have a duration of less than 10 ms. Tonebursts typically have brief but gradual onsets and offsets to reduce spectral splatter. There is an interesting tradeoff between the duration of the toneburst and its frequency specificity. The shorter the toneburst, the less frequency specificity it has around its central frequency (i.e., greater sideband energy). Although tonebursts have less frequency specificity than a standard pure tone, they are brief enough in duration that they can be used to elicit evoked potentials, such as for ABR threshold estimation and VEMPs. ▶ Fig. 2.12 shows four common tonebursts and their frequencies.

Chirps

A *chirp* stimulus consists of a brief pure tone sweep where the frequency of the signal shifts continuously from cycle to cycle, typically from low to high frequency. Chirp stimuli are often constructed using various models of cochlear delay with the goal of rapidly stimulating the cochlea progressively from the apical to basal ends of the cochlea. In effect, chirp stimuli will temporally excite the cochlea broadly to promote greater neural synchrony.[22,23,24] Compared with clicks, chirp stimuli can elicit larger evoked potential responses. Readers are directed again to ▶ Fig. 2.10 and ▶ Fig. 2.11 to see how the temporal, spectral, and physiological simulations of the CE-Chirp compare with the 100-µs click.

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

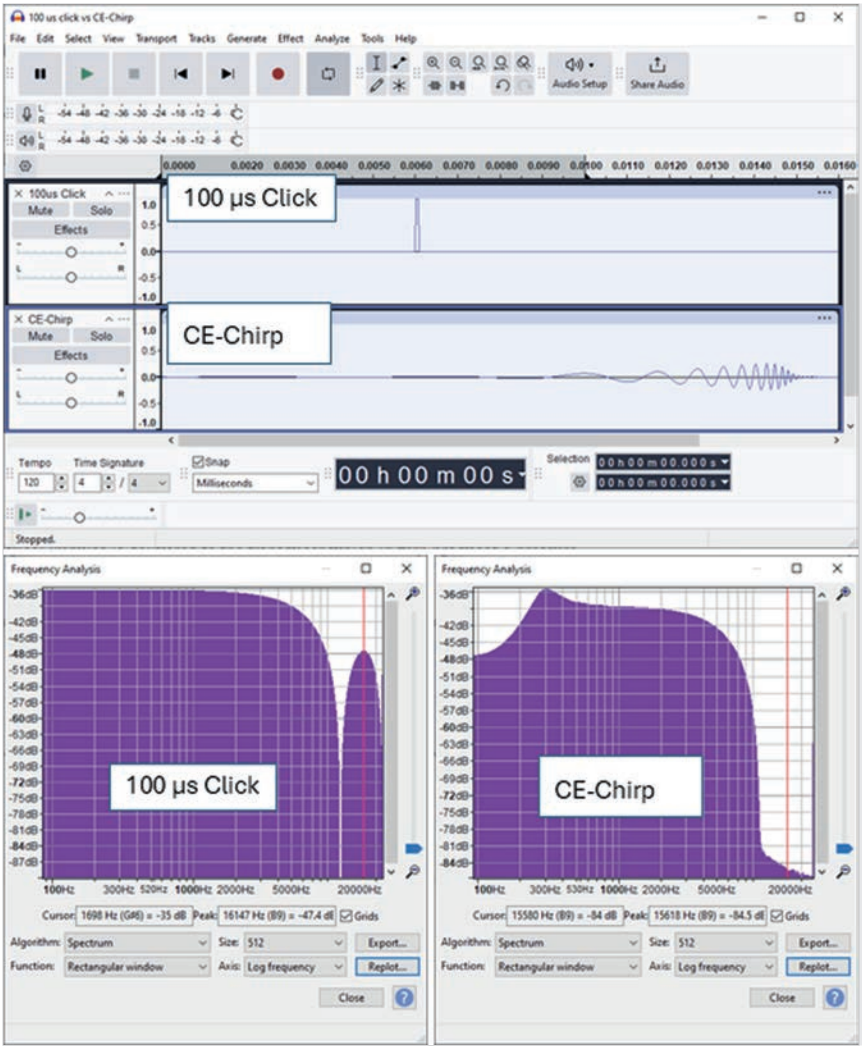

Fig. 2.10 Screenshots of the Audacity software showing a 100-µs click and a broadband CE-Chirp stimulus in the time domain (top) and frequency domain (bottom). Note the click is a brief square pulse and the CE-Chirp progresses from low to high frequency, which results in broadband spectrum quite similar to the click.

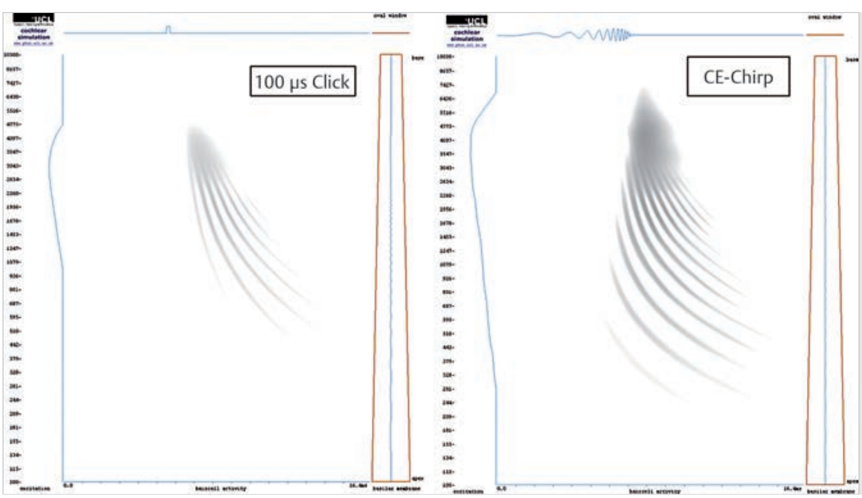

Fig. 2.11 Screenshots of the CochSim software comparing the 100-µs click and a broadband CE-Chirp stimulus. Note the greater spread of energy in the cochlea making the CE-Chirp more effective than the click.

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

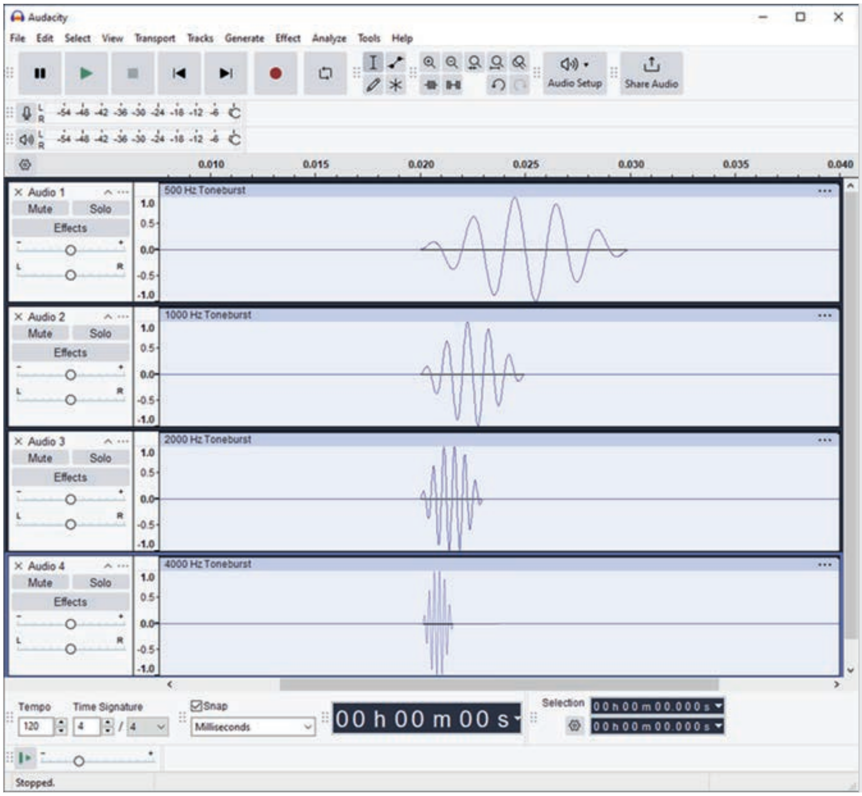

Fig. 2.12 Screenshot of the Audacity software showing four tonebursts (500, 1,000, 2,000, and 4,000 Hz, top to bottom). Tonebursts are intended to be brief (no more than a few cycles in duration) and often use brief onsets and offsets (e.g., two cycles).

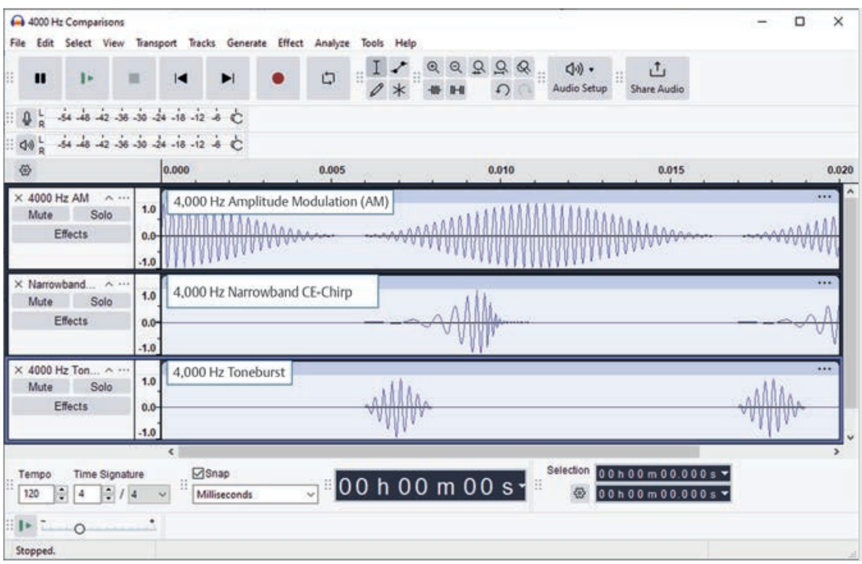

Fig. 2.13 Screenshot of the Audacity software showing a 4,000 Hz carrier tone amplitude modulated at 90 Hz (top), a 4,000 Hz narrowband CE-Chirp stimuli presented at a rate of 90 Hz (i.e., one stimulus approximately every 11 ms), and, for temporal comparison, a 4,000 Hz toneburst (bottom).

From the broadband models, narrowband CE-Chirp have also been developed that offer frequency-specific threshold applications for ABR and ASSR. Narrowband CE-Chirp stimuli for threshold estimation evoked potential tests help clinicians arrive at thresholds closer to behavioral thresholds compared with tonebursts (for ABR) and appear to be producing more robust responses compared with AM stimuli (for ASSR). ▶ Fig. 2.13 illustrates the temporal differences between an AM stimulus, narrowband CE-Chirp, and toneburst with target stimulus frequency of 4,000 Hz.

2.3 Other Stimuli Used in Diagnostic Audiology

As this chapter aptly illustrates, there are many types of acoustic stimuli used in audiology diagnostics; however, not all audiology tests employ acoustic stimuli presented through air conduction. Acoustic stimuli can also be presented through bone-conduction devices such as a bone oscillator and tuning forks. Bone conduction devices deliver stimulation by vibrating

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved

the skull in a manner that excites both inner ears or one of the inner ears in the case of single-sided deafness or contralateral masking. When hearing loss is severe enough, patients may note that they feel high-intensity, low-frequency acoustic stimuli rather, resulting in a vibrotactile (or tactile) response. In the case of a third window disorder (e.g., superior semicircular canal dehiscence), a patient may hear that same vibratory stimulus when the bone oscillator is placed on their wrist, elbow, or ankle! With the exception of VEMPs, where the saccule and utricle can be sensitive to both air- and bone-conducted stimuli, the assessment of vestibular disorders generally does not use acoustic stimuli. Rather, the different structures of the vestibular system are sensitive to the influence of gravitational, linear acceleration and rotary effects applied to the head and/or body as we gauge the health of vestibular, visual, and proprioceptive systems. One interesting stimulus that is generally unexpected is a bithermal, warm and cold stimulus (air or water) to isolate and assess the function of the horizontal semicircular canals. The use of bone conduction will be found in various chapters of this book, but interested readers are referred to Chapter 24 for coverage of the different types of vestibular assessments.

2.4 Conclusion

Understanding and familiarizing oneself with the stimuli utilized in diagnostic audiometry is an essential component of clinical competence. Audiologists work with a diverse range of acoustic stimuli, including pure tones, filtered forms of white noise, clicks, chirps, speech stimuli, altered speech stimuli, synthetic speech stimuli, and combinations of these forms. Each type of stimulus serves specific diagnostic purposes, ranging from basic behavioral assessments to complex evaluations of cochlear, central auditory, and vestibular functions. A thorough understanding of the various acoustic properties of these stimuli, such as intensity, frequency, duration, modulation, polarity, and envelope characteristics, can sometimes be helpful in the interpretation of test outcomes. For example, subtle variations in the acoustic properties of a stimulus can affect the elicitation of otoacoustic emissions, the outcomes of SIN testing, or the amplitude and latency of auditory evoked potentials. Clinicians who grasp these nuances are better equipped to select and modify stimuli to suit the unique needs of their patients and to address specific diagnostic questions. In addition, the manner in which stimuli are presented (e.g., unilaterally, bilaterally, diotically, or dichotically) can provide valuable insights into auditory processing and the integrity of central and peripheral auditory pathways. Finally, incorporating this knowledge into clinical practice not only enhances the accuracy and reliability of diagnostic testing but also contributes to improved patient outcomes. By understanding both the technical and practical aspects of acoustic stimuli, audiologists can make informed decisions about their application across a variety of diagnostic scenarios. This understanding ensures that the audiologist can meet the demands of diverse clinical cases while advancing their own expertise.

References

[1] Musiek FE. Assessment of central auditory dysfunction: the dichotic digit test revisited. Ear Hear. 1983; 4(2):79–83

[2] Shivashankar N, Willeford JA. Competing Sentence test: a test for central auditory dysfunction. NIMHANS J. 1990; 8(1):43–46

[3] Govender S, Rosengren SM, Colebatch JG. The effect of gaze direction on the ocular vestibular evoked myogenic potential produced by air-conducted sound. Clin Neurophysiol. 2009; 120(7):1386–1391

[4] Musiek FE, Pinheiro ML. Frequency patterns in cochlear, brainstem, and cerebral lesions. Audiology. 1987; 26(2):79–88

[5] Musiek FE, Baran JA, Pinheiro ML. Duration pattern recognition in normal subjects and patients with cerebral and cochlear lesions. Audiology. 1990; 29 (6):304–313

[6] Lentz JJ, Walker MA, Short CE, Skinner KG. Audiometric testing with pulsed, steady, and warble tones in listeners with tinnitus and hearing loss. Am J Audiol. 2017; 26(3):328–337

[7] Musiek FE, Shinn JB, Jirsa R, Bamiou DE, Baran JA, Zaida E. GIN (Gaps-In-Noise) test performance in subjects with confirmed central auditory nervous system involvement. [published correction appears in Ear Hear. 2006 Jun;27(3):228]. Ear Hear. 2005; 26(6):608–618

[8] Lantz J. FRESH Noise (White Paper). 2003. Accessed November 12, 2024 at: https://natus.bynder.com/web/18c3fb5314896a5d/resource-library-madsen-astera-/?mediaId=D022E6ED-3D21–4701-AF4AC83F4915B38F

[9] Phatak SA, Allen JB. Consonant and vowel confusions in speech-weighted noise. J Acoust Soc Am. 2007; 121(4):2312–2326

[10] Skoe E, Kraus N. Auditory brain stem response to complex sounds: a tutorial. Ear Hear. 2010; 31(3):302–324

[11] Willeford JA, Burleigh JM. Sentence procedures in central testing. In: Katz J, ed. Handbook of Clinical Audiology. 4th ed. Baltimore, MD: Williams & Wilkins; 1994;256–268

[12] Cox RM, Alexander GC, Gilmore C. Development of the Connected Speech Test (CST). Ear Hear. 1987; 8(5) Suppl:119S–126S

[13] Billings CJ, Olsen TM, Charney L, Madsen BM, Holmes CE. Speech-in-noise testing: an introduction for audiologists. Semin Hear. 2023; 45(1):55–82

[14] Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2004; 116 (4 Pt 1):2395–2405

[15] Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am. 1994; 95(2):1085–1099

[16] Bench J, Kowal A, Bamford J. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. Br J Audiol. 1979; 13(3):108–112

[17] Holube I, Fredelake S, Vlaming M, Kollmeier B. Development and analysis of an International Speech Test Signal (ISTS). Int J Audiol. 2010; 49(12): 891–903

[18] Katz J. The use of staggered spondaic words for assessing integrity of the central auditory nervous system. J Aud Res. 1962; 2:327–337

[19] Shore SE, Nuttall AL. High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone bursts. J Acoust Soc Am. 1985; 78(4): 1286–1295

[20] Dau T, Wegner O, Mellert V, Kollmeier B. Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. J Acoust Soc Am. 2000; 107(3):1530–1540

[21] Elberling C, Don M, Cebulla M, Stürzebecher E. Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay. J Acoust Soc Am. 2007; 122(5):2772–2785

[22] AlMakadma H, Kei J, Yeager D, Feeney MP. Fundamental concepts for assessment and interpretation of wideband acoustic immittance measurements. Semin Hear. 2023; 44(1):17–28

[23] Keefe DH, Feeney MP, Hunter LL, Fitzpatrick DF. Comparisons of transient evoked otoacoustic emissions using chirp and click stimuli. J Acoust Soc Am. 2016; 140(3):1949

[24] Keith RW. Random Gap Detection Test. Auditec of St Louis Ltd.; 2000. Accessed November 12, 2024 at: www.auditec.com.

from: Atcherson, Audiology Diagnosis (9781684201440)
© 2026. Thieme. All rights reserved