Echo Cancellation

Definition

Wireless phones are increasingly being regarded as essential communications tools, dramatically impacting how people approach day-to-day personal and business communications. As new network infrastructures are implemented and competition between wireless carriers increases, digital wireless subscribers are becoming ever more critical of the service and voice quality they receive from network providers. A key technology to provide near-wireline voice quality across a wireless carrier's network is echo cancellation.

Overview

Subscribers use speech quality as the benchmark for assessing the overall quality of a network. Regardless of whether or not this is a subjective judgment, it is the key to maintaining subscriber loyalty. For this reason, the effective removal of hybrid and acoustic echo inherent within the digital cellular infrastructure is the key to maintaining and improving perceived voice quality on a call. This has led to intensive research into the area of echo cancellation, with the aim of providing solutions that can reduce background noise and remove hybrid and acoustic echo before any transcoder processing. By employing this technology, the overall efficiency of the coding can be enhanced, significantly improving the quality of speech. This tutorial discusses the nature of echo and how echo cancellation is helpful in making mobile calls meet acceptable quality standards.

Topics

1. History of Echo Cancellation
2. Types of Echo
3. Causes of Echo
4. The Combined Problem on Digital Cellular Networks
5. The Process of Echo Cancellation
6. Controlling Acoustic Echo
7. Controlling Complex Echo in a Wireless Digital Network
8. Room for Improvement in the Handset

1. History of Echo Cancellation

The late 1950s marked the birth of echo control in the telecommunications industry with the development of the first echo-suppression devices. These systems, first employed to manage echo generated primarily in satellite circuits, were essentially voice-activated switches that transmitted a voice path and then turned off to block any echo signal. Although echo suppressers reduced echo caused by transmission problems in the network, they also resulted in choppy first syllables and artificial volume adjustment. In addition, they eliminated double-talk capabilities, greatly reducing the ability to achieve natural conversations.

Echo-cancellation theory was developed in the early 1960s by AT&T Bell Labs, followed by the introduction of the first echo-cancellation system in the late 1960s by COMSAT TeleSystems (previously a division of COMSAT Laboratories). COMSAT designed the first analog echo canceller systems to demonstrate the feasibility and performance of satellite communications networks. Based on analog processes, these early echo-cancellation systems were implemented across satellite communications networks to demonstrate the network's performance for long-distance, cross-continental telephony. These systems were not commercially viable, however, because of their size and manufacturing costs.

In the late 1970s, COMSAT TeleSystems developed and sold the first commercial analog echo cancellers, which were mainly digital devices with an analog interface to the network. The semiconductor revolution of the early 1980s marked the switch from analog to digital telecommunications networks. More sophisticated digital interface, multichannel echo-canceller systems were also developed to address new echo problems associated with long-distance digital telephony systems. Based on application-specific integrated circuit (ASIC) technology, these new echo cancellers utilized high-speed digital signal-processing techniques to model and subtract the echo from the echo return path. The result was a new digital echo-cancellation technique that outperformed existing suppression-based techniques, creating improved network performance.

The 1990s have witnessed explosive growth in the wireless telecommunications industry, resulting from deregulation that has brought to market new analog and digital wireless handsets, numerous network carriers, and new digital network infrastructures such as TDMA, CDMA, and GSM. According to the Cellular Telecommunications Industry Association (CTIA), new subscribers are driving the growth of the wireless market at an annual rate of 40 percent. With wireless telephony being widely implemented and competition increasing as new wireless carriers enter the market, superior voice transmission quality and customer service have now become key determining factors for subscribers evaluating a carrier's network. Understanding and overcoming the inherent echo problems associated with digital cellular networks will enable network operators and telcos to offer subscribers the network performance and voice quality they are demanding today.

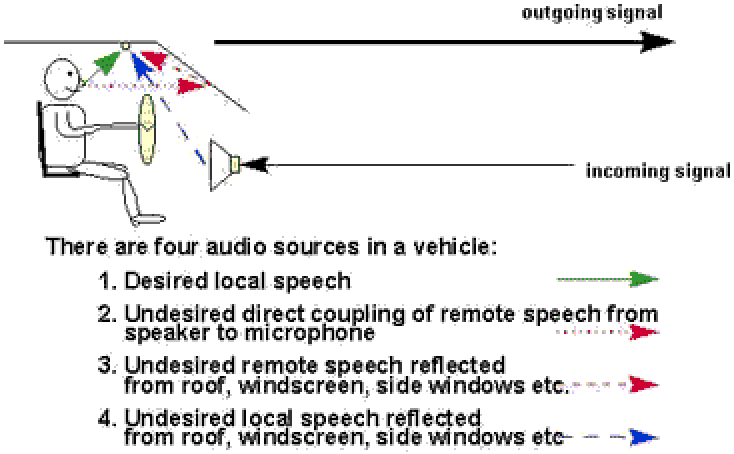
2. Types of Echo

Acoustic Echo

Acoustic echo is generated with analog and digital handsets, with the degree of echo related to the type and quality of equipment used. This form of echo is produced by poor voice coupling between the earpiece and microphone in handsets and hands-free devices. Further voice degradation is caused as voice-compressing encoding/decoding devices (vocoders) process the voice paths within the handsets and in wireless networks. This results in returned echo signals with highly variable properties. When compounded with inherent digital transmission delays, call quality is greatly diminished for the wireline caller.

Acoustic echo was first encountered with the early video/audioconferencing studios and—as *Figure 1* shows—now also occurs in typical mobile situations, such as when people are driving their cars. In this situation, sound from a loudspeaker is heard by a listener, as intended. However, this same sound also is picked up by the microphone, both directly and indirectly, after bouncing off the roof, windows, and seats of the car. The result of this reflection is the creation of multipath echo and multiple harmonics of echo, which, unless eliminated, are transmitted back to the distant end and are heard by the talker as echo. Predominant use of handsfree telephones in the office has exacerbated the acoustic echo problem.

Figure 1. Acoustic Echo on Digital Cellular

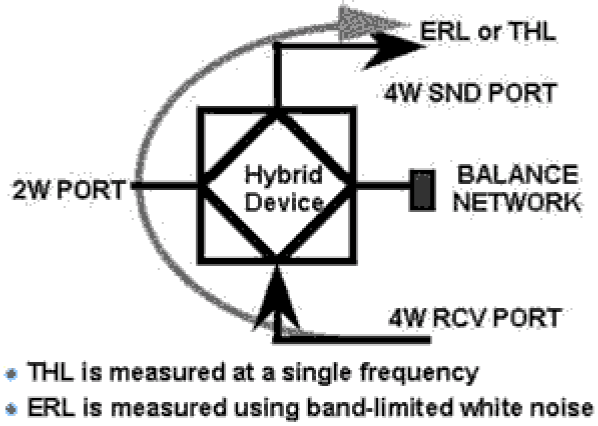


Hybrid Echo

Hybrid echo is the primary source of echo generated from the public-switched telephone network (PSTN). This electrically generated echo is created as voice signals are transmitted across the network via the hybrid connection at the two-wire/four-wire PSTN conversion points, reflecting electrical energy back to the speaker from the four-wire circuit.

Hybrid echo has been around almost since the advent of the telephone itself. The signal path between two telephones, involving a call other than a local one, requires amplification using a four-wire circuit. Although not a factor in itself on digital cellular networks, hybrid echo becomes a problem in PSTN–originated calls. The cost and cabling required rules out the idea of running a four-wire circuit out to the subscriber's premise from the local exchange. For this reason, an alternative solution had to be found. Hence, the four-wire trunk circuits were converted to two-wire local cabling, using a device called a "hybrid" (see *Figure 2*).

Figure 2. Hybrid Echo



Unfortunately, the hybrid is by nature a leaky device. As voice signals pass from the four-wire to the two-wire portion of the network, the energy in the four-wire section is reflected back on itself, creating the echoed speech. Provided that the total round-trip delay occurs within just a few milliseconds (i.e., within 28 ms), it generates a sense that the call is live by adding sidetone, which makes a positive contribution to the quality of the call.

In cases where the total network delay exceeds 36 ms, however, the positive benefits disappear, and intrusive echo results. The actual amount of signal that is reflected back depends on how well the balance circuit of the hybrid matches the two-wire line. In the vast majority of cases, the match is poor, resulting in a considerable level of signal reflecting back. This is measured as echo return loss (ERL). The higher the ERL, the lower the reflected signal back to the talker, and vice versa.

3. Causes of Echo

Acoustic echo apart, background noise is generated through the network when analog and digital phones are operated in hands-free mode. As additional sounds are directly and indirectly picked up by the microphone, multipath audio is created and transmitted back to the talker. The surrounding noise, whether in an automobile or in a crowded, public environment, passes through the digital cellular vocoder, causing distorted speech for the wireline caller.

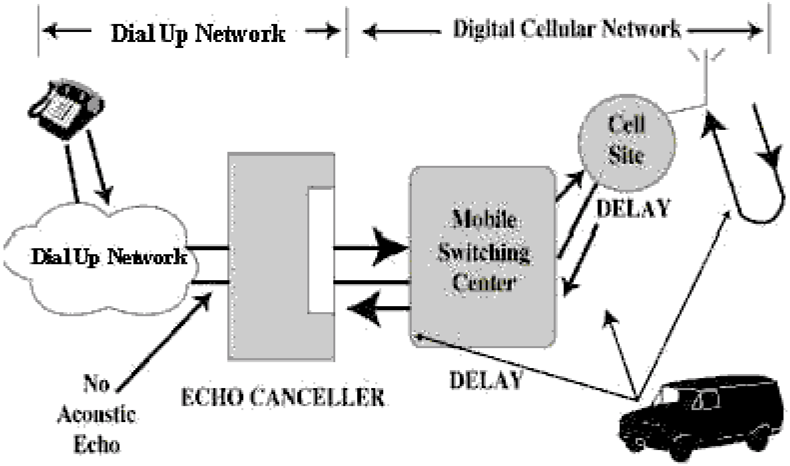
Digital processing delays and speech-compression techniques further contribute to echo generation and degraded voice quality in wireless networks. Delays are encountered as signals are processed through various routes within the networks, including copper wire, fiber optic lines, microwave connections, international gateways, and satellite transmission. This is especially true with mixed technology digital networks, where calls are processed across numerous network infrastructures.

Echo-control systems are required in all networks that produce one-way time delays greater than 16 ms. In today's digital wireless networks, voice paths are processed at two points in the network within the mobile handset and at the radio frequency (RF) interface of the network. As calls are processed through vocoders in the network, speech processing delays ranging from 80 ms to 100 ms are introduced, resulting in an unacceptable total end-to-end delay of 160 ms to 200 ms. As a result, echo cancellation devices are required within the wireless network to eliminate the hybrid and acoustic echoes in a digital wireless call.

4. The Combined Problem on Digital Cellular Networks

To deal with hybrid echo created by vocoder processing delays, it is mandatory for digital cellular mobile calls to have a group echo canceller installed—even for local calls. As a result, all calls on to the PSTN must pass through an echo canceller to remove what would otherwise be a noticeable and annoying echo, as shown in *Figure 3*.

Figure 3. Digital Cellular Network

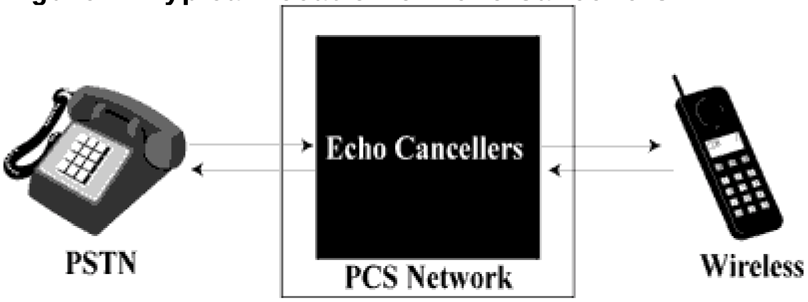


For example, consider a digital cellular mobile user who makes a call to the PSTN without an echo canceller in place. The user would hear his or her own speech being echoed back 180 ms or more later, even if the called person is in the same locality. The mobile user will either be using a hands-free system installed in his or her vehicle or a hand portable. In either case, these units will involve the occurrence of direct and indirect coupling between the microphone and the speaker, creating acoustic echo. In this situation, however, it is the PSTN user who suffers by experiencing poor speech quality. Hence, the echo canceller installed in the digital cellular network must be capable of handling both sources of echoes.

5. The Process of Echo Cancellation

In modern telephone networks, echo cancellers are typically positioned in the digital circuit, as shown in *Figure 4*. The process of canceling echo involves two steps. First, as the call is set up, the echo canceller employs a digital adaptive filter to set up a model or characterization of the voice signal and echo passing through the echo canceller. As a voice path passes back through the cancellation system, the echo canceller compares the signal and the model to cancel existing echo dynamically. This process removes more than 80 to 90 percent of the echo across the network. The second process utilizes a non-linear processor (NLP) to eliminate the remaining residual echo by attenuating the signal below the noise floor.

Figure 4. Typical Location of Echo Cancellers

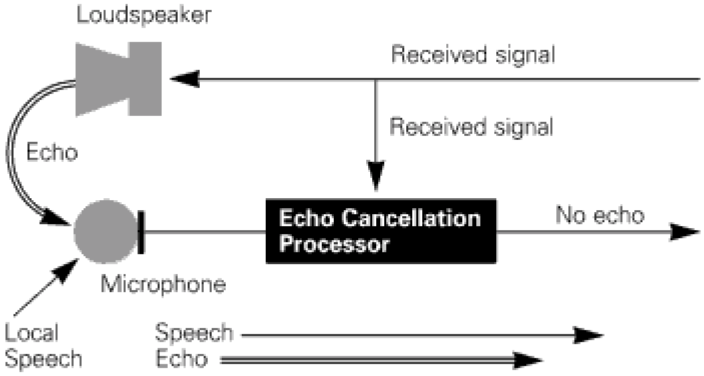


Today's digital cellular network technologies, namely TDMA, CDMA, and GSM, require significantly more processing power to transmit signal paths through the channels. As these technologies become even more sophisticated, echo control will be more complex. Echo cancellers designed with standard digital signal processors (DSPs), which share processing time in a circuit within a channel or across channels, provide a maximum of only 128 ms of cancellation and are unable to cancel acoustic echo. With network delays occurring in excess of 160 ms in today's mixed-signal network infrastructures, a more powerful, application-specific echo-cancellation technology is required to control echo across wireless networks effectively.

6. Controlling Acoustic Echo

In echo cancellation, complex algorithmic procedures are used to compute speech models. This involves generating the sum from reflected echoes of the original speech, then subtracting this from any signal the microphone picks up. The result is the purified speech of the person talking. The format of this echo prediction must be learned by the echo canceller in a process known as adaptation. It might be said that the parameters learned from the adaptation process generate the prediction of the echo signal, which then forms an audio picture of the room in which the microphone is located. *Figure 5* shows the basic operation of an echo canceller in a conference room type of situation.

Figure 5. Operation of an Acoustic Echo Canceller



During the conversation period, this audio picture constantly alters, and, in turn, the canceller must adapt continually. The time required for the echo canceller to fully learn the acoustic picture of the room is called the convergence time. The best convergence time recorded is 50 ms, and any increase in this number results in syllables of echo being detected.

Other important performance criteria involve the acoustic echo canceller's ability to handle acoustic tail circuit delay. This is the time span of the acoustic picture and roughly represents the delay in time for the last significant echo to arrive at the microphone. The optimum requirement for this is currently set at 270 ms—any time below this could result in echoes being received by the microphone outside the ability of the echo canceller to remove them, and hence in participants hearing the echoes.

Another important factor is acoustic echo return loss enhancement (AERLE). This is the amount of attenuation which is applied to the echo signal in the process of echo cancellation—i.e., if no attenuation is applied, full echo will be heard. A value of 65 dB is the minimum requirement with the non-linear processor enabled, based on an input level of -10 dBm white noise electrical and 6 dB of echo return loss (ERL).

The canceller's performance also relies heavily on the efficiency of a device called the center clipper, or non-linear processor. This needs to be adaptive and has a direct bearing on the level of AERLE that can be achieved.

7. Controlling Complex Echo in a Wireless Digital Network

Although acoustic echo is present in every hands-free mobile call, the amount of echo depends on the particular handset design and model that the mobile user has. On the market are a few excellent handsets that limit the echo present, but, due to strong price pressures, most handsets do not control the echo very well at all—in fact, some phones on the market have been determined to have a terminal compiling loss of 24 dB. Echo becomes a problem when the processing inherent to the digital wireless network adds an additional delay (typically in excess of 180 ms round-trip). This combination makes for totally unacceptable call quality for the fixed network customer, as shown in *Figure 6*.

Figure 6. Acoustic Echo in a Mobile Environment

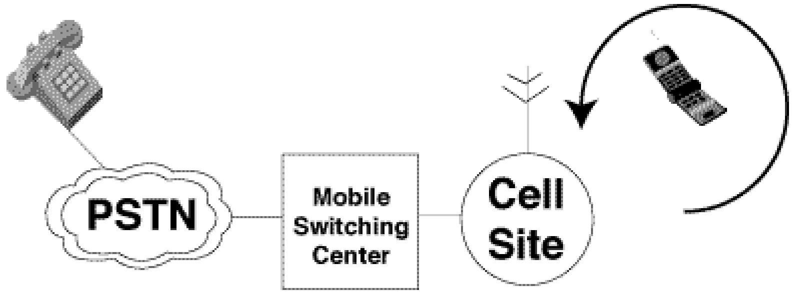
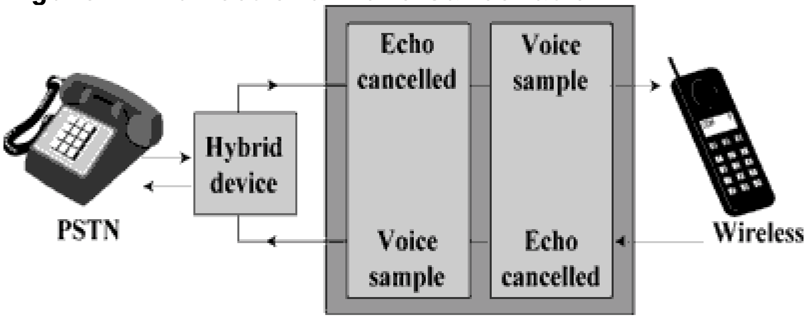


Figure 7. Bidirectional Echo Cancellation



Echo Canceller Board

This back-to-back configuration ensures a high audio quality for both PSTN and mobile customers. In addition, the echo canceller's software configuration is designed to provide a detailed analysis of background noises, including acoustic echo from the mobile user's end. Some echo cancellers incorporate a user-settable network delay, which enables network operators to fine-tune the echo control to suit their parameters via a menu option on the canceller's hand-held terminal or on the network management system (NMS).

8. Room for Improvement in the Handset

Applying effective echo control via the echo-cancellation platform is one way of improving the overall call clarity on digital cellular networks. Another derives from improvements that must be made within the handset or terminal itself. There also is considerable room to enhance the network itself, focusing principally on vocoder development.

Recent headlines have charted the ongoing commercial battles regarding which digital technologies will eventually emerge as the winners, as equipment manufacturers fight it out. However, this public battle will soon be overshadowed by another battle concerning handsets. At present, there are four major players in the digital cordless market. Europe has cordless telephony (CT2) and digital European cordless telephony (DECT), while Japan has the personal handyphone system (PHS) and the United States has personal communications services (PCS).

Connecting directly into the plain old telephone system, CT2 was one of the first digital technologies to provide low-cost mobile phones. Although the technology worked well, it had a fundamental problem: it could not handle cell handovers. DECT and GSM have overcome this problem and will eventually dominate European cellular services.

During the development of early cordless telephony, attention was paid to basic and enhanced functions and interworking with different network architectures. While the early generation of handsets looked very elegant and aesthetically pleasing, very little attention was paid to designing the handset with echo suppression/cancellation in mind. The result was that they looked good but were extremely poor at reducing acoustic echo.

In the setting of standards for GSM and PCS, handset design and the impact of different design approaches on call quality was researched. As a result, recommendations stated a range of parameters, including sidetone tolerance and echo return loss performance. With the resultant advent of new recommendations with much tighter requirements for handsets, there is a call for greatly improved designs to be implemented. This, complemented by ongoing improvements in network technology and echo cancellation techniques, will bring digital wireless telephony much closer to matching wireline quality.

9. The Future of Echo-Cancellation Technology

New digital cellular networks and network scenarios reflect a significant change taking place in the operation of echo cancellers. Instead of being a means of simple echo control, echo cancellers have now become highly sophisticated transmission equipment at the center of highly complex networks. Network operators and telcos implementing echo cancellation across their networks will hold the key to improved call quality, directly impacting their ability to provide enhanced network performance, maintain customer loyalty, increase talk time revenues, and reduce subscriber churn.

Self-Test

1. Echo-suppression devices, developed in the late 1950s, were first employed to manage echo generated primarily in satellite circuits.

    a. true

    b. false

2. The two distinct areas of echo cancellation are acoustic and hybrid echo cancellation.

    a. true

    b. false

3. Hybrid echo results from converting four-wire trunk circuits to two-wire local cabling using a device called a "hybrid."

    a. true

    b. false

4. Today's digital cellular network technologies, namely TDMA, CDMA, and GSM/PCS, require significantly less processing power to transmit signal paths through channels.

    a. true

    b. false

5. When a mobile user calls a fixed network customer, unacceptable call quality can result at the fixed end of the link. In addition, certain digital handsets add an offset to the transmit signal, causing degradation of digital cellular's speech compression performance.

    a. true

    b. false

6. The explosive growth in the wireless telecommunications industry in this decade has been a result of which of the following?

    a. new analog and digital wireless handsets in the market

    b. new digital network infrastructures such as TDMA, CDMA, and GSM

    c. numerous network carriers
    d. all of the above

7. Which of the following is the term for the amount of time that the canceller has to hold the model of the echo in order to recognize and cancel it?
    a. echo model
    b. tail circuit requirement
    c. echo prediction period
    d. convergence time

8. Although not a factor in itself on digital cellular networks, which of the following becomes a problem in PSTN–originated calls?
    a. attenuation
    b. echo return loss
    c. hybrid echo
    d. comfort noise

9. Digital processing delays are encountered as signals are processed through various network routes, such as which of the following?
    a. copper wire
    b. fiber optic lines
    c. microwave connections
    d. all of the above

10. Which factor is expected to bring digital wireless telephony much closer to matching wireline quality?
    a. new recommendations with much tighter requirements for handsets
    b. a call for the implementation of greatly improved handset designs
    c. ongoing improvements in network technology and echo cancellation techniques
    d. all of the above

Correct Answers

1. Echo suppression devices, developed in the late 1950s, were first employed to manage echo generated primarily in satellite circuits.

   a. true

   b. false

   See Topic 1.

2. The two distinct areas of echo cancellation are acoustic and hybrid echo cancellation.

   a. true

   b. false

   See Topic 2.

3. Hybrid echo results from converting four-wire trunk circuits to two-wire local cabling using a device called a "hybrid."

   a. true

   b. false

   See Topic 2.

4. Today's digital cellular network technologies, namely TDMA, CDMA, and GSM/PCS, require significantly less processing power to transmit signal paths through channels.

   a. true

   b. false

   See Topic 5.

5. When a mobile user calls a fixed network customer, unacceptable call quality can result at the fixed end of the link. In addition, certain digital handsets add an offset to the transmit signal, causing degradation of digital cellular's speech compression performance.

   a. true

   b. false

See Topic 7.

6. The explosive growth in the wireless telecommunications industry in this decade has been a result of which of the following?

    a. new analog and digital wireless handsets in the market

    b. new digital network infrastructures such as TDMA, CDMA, and GSM

    c. numerous network carriers

    d. all of the above

    See Topic 1.

7. Which of the following is the term for the amount of time that the canceller has to hold the model of the echo in order to recognize and cancel it?

    a. echo model

    b. tail circuit requirement

    c. echo prediction period

    d. convergence time

    See Topic 2.

8. Although not a factor in itself on digital cellular networks, which of the following becomes a problem in PSTN–originated calls?

    a. attenuation

    b. echo return loss

    c. hybrid echo

    d. comfort noise

    See Topic 2.

9. Digital processing delays are encountered as signals are processed through various network routes, such as which of the following?

    a. copper wire

    b. fiber optic lines

c. microwave connections

d. all of the above

See Topic 3.

10. Which factor is expected to bring digital wireless telephony much closer to matching wireline quality?

a. new recommendations with much tighter requirements for handsets

b. a call for the implementation of greatly improved handset designs

c. ongoing improvements in network technology and echo cancellation techniques

d. all of the above

See Topic 8.

Acronym Guide

acoustic echo
echo that occurs from sounds reflecting off surfaces in the surrounding environment

acoustic echo return loss (AERL)
the amount of signal level that the loudspeaker signal and its echo lose as they travel from the loudspeaker to the microphone

acoustic echo return loss enhancement (AERLE)
the amount of attenuation applied to the echo signal in the process of echo cancellation

acoustic tail circuit delay
the time span of the acoustic picture; it roughly represents the delay in time for the last significant echo to arrive at the microphone

ambient noise
background noise such as that found in a conference room (e.g., air conditioning and paper shuffling), it is often the result of using a car phone hands-free

attenuation
the decrease in the power of a signal

bandwidth
the range of frequencies that a device can handle

convergence time
the time required for the echo canceller to fully learn the acoustic picture of a room

decibel (dB)
the unit of measurement to measure sound pressure levels; it characterizes the way human beings actually perceive a sound

full duplex
simultaneous, interactive two-way communications; transmission in two directions simultaneously; this means that if parties at either end of the conversation talk at the same time they will both be heard

frequency
the rate at which an electrical current alternates, typically measured in Hertz (Hz)

half duplex
transmit in one direction at one time. This means that if parties at either end of the conversation talk at the same time, only one party will be heard

hertz (Hz)
a measurement of frequency in cycles per second; one cycle per second is 1 Hz

hybrid echo
echo that occurs where two-wire-to-four-wire conversion takes place

ISO 9001
an international quality standard for a company's procedures, covering areas such as product development, engineering, technical support, and manufacturing

microphone gating
technique used in speakerphones that results in half duplex performance, limiting the sound output in one of the two directions of transmission to minimize feedback

reverberation time
a criterion used to judge the acoustic characteristics of a room; the amount of time that sound continues to be heard after the source of the sound has stopped