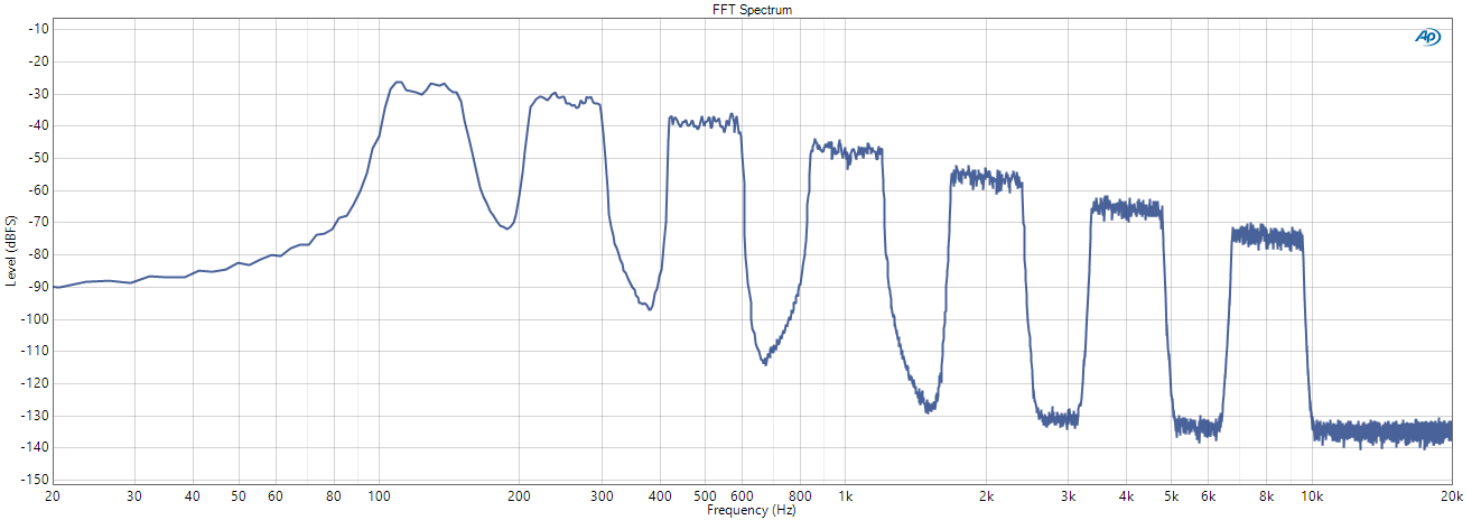
DATA SHEET

STI: SPEECH TRANSMISSION INDEX

An APx software option for the objective measurement of speech intelligibility



FFT of STI (STIPA) stimulus waveform

APPLICATIONS

- Public Address (PA) systems
- Voice Announce (VA) and emergency sound systems
- Communications channels and systems
- Meeting rooms and auditoria
- Direct (unamplified) communication in rooms, including vehicles
- Assistive hearing systems

HIGHLIGHTS

- Works with any APx analyzer hardware.
- Requires APx500 v4.2.1 or later measurement software.
- Download the installer, and purchase a software option key to enable the measurement.
- For in-depth information about APx and STI, download the Audio Precision Technote 128.

STANDARDS and REFERENCES

- IEC 60268-16
- NFPA 72, NFPA 1981:2013
- EUROCAE ED-112A
- UK-CAA Specification No. 15
- TRCA DO-214A

Overview

Perceptual audio measurements began as audio evaluation by groups of people, who rated what they heard for quality or intelligibility. This kind of testing is expensive and time consuming, so alternative methods have been developed to provide similar results using hardware and software tools.

For objective measurement of speech intelligibility, Audio Precision offers two software options for our APx analyzers: STI and ABC-MRT. This data sheet focuses on STI.

STI

STI (Speech Transmission Index) is an objective measure first introduced in the 1970s, used to predict the intelligibility of speech transmitted from a talker to a listener through a “transmission channel.”

Speech intelligibility is defined as the degree to which a normal person can understand speech, whereas speech quality is defined as the difference between idealization (the theoretical capabilities of a communication channel) and realization (the performance of the channel in practical use). The two are related, but they are not the same.

Conceptually, STI is based on the empirical finding that the normal fluctuations in speech signals carry the most relevant information related to speech intelligibility. These fluctuations (or modulations) result from the acoustic separation of sentences, words, and phonemes (the fundamental elements of speech). For clear speech, modulation rates extend from 0.5 Hz to 16 Hz, with maximum modulation at about 3 Hz.

Fluctuations in speech can be quantified in terms of modulation depth and modulation rate (frequency), which together form a modulation spectrum. STI uses the concept of the modulation transfer function (MTF), which is the ratio of modulation depth of the received signal to the modulation depth of the transmitted signal as a function of frequency. Any reduction of the MTF by the transmission channel is considered to result in some deterioration of speech intelligibility. Distortions in the transmission channel which can reduce the modulation depth (and hence intelligibility) include noise, reverberation, echoes, non-linear distortion, spectral distortion, and digital codecs.

APx STI SPEECH INTELLIGIBILITY SOFTWARE OPTION

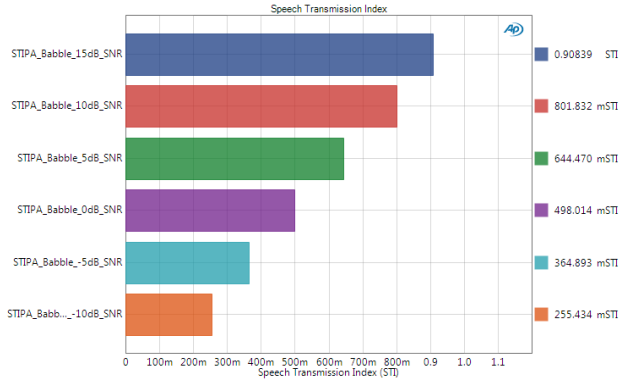
The spectral content of speech itself is also important. The long-term average speech spectrum includes the octave bands from 125 Hz through 8 kHz, but these frequencies do not contribute equally to intelligibility. Although low speech frequencies (which make up vowel sounds) contribute most to the power of a speech signal, it is the higher speech frequencies (consonants) which contribute the most to speech intelligibility. This is reflected in the weighting factors applied when calculating the STI.

STI should not be used for transmission channels containing vocoders (codecs which operate specifically on speech elements), but it may be used for digital codecs that operate on the entire signal. It also cannot be used on systems with aggressive noise-suppression algorithms, because the STI signal itself is modulated noise that is likely to be suppressed by the algorithm.

The APx STI option features an STI measurement and a Speech Level (IEC 60268-16) measurement.

The STIPA measurement

The APx STI implementation uses the STIPA method, a simplification of full STI. The STIPA measurement requires a special signal containing shaped noise filtered in ½-octave bands and modulated with two modulation frequencies per band. The signal has a unique FFT spectrum, as seen in the APx Signal Analyzer FFT display shown on the first page of this data sheet.



STIPA signal mixed with noise at various signal-to-noise ratios.

STIPA Analysis Details

STIPA Analysis Details

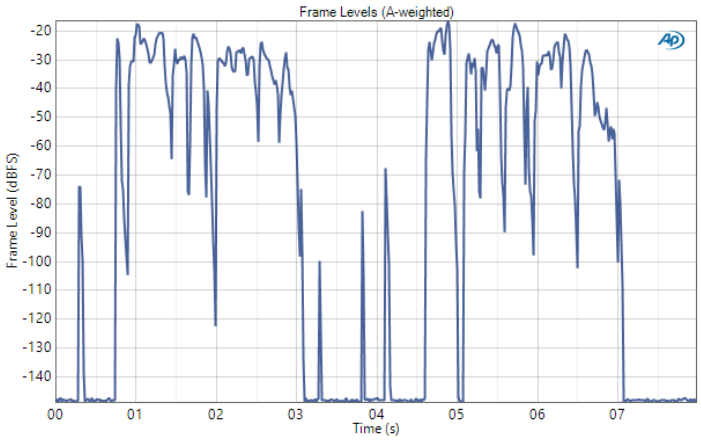
		125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
1	Ch1: STI = 0.991							
2	Leq (dB)	75.2	74.9	71.0	65.1	59.2	52.9	47.4
3	Level Ratio (dB)	-0.4	-0.1	0.1	0.0	-0.1	0.1	-0.3
4	f-mod1 (Hz)	1.60	1.00	0.63	2.00	1.25	0.80	2.50
5	Raw MTR1	1.170	0.969	1.062	0.996	1.024	0.987	0.995
6	Corrected MTR1	1.169	0.962	1.047	0.982	1.020	0.986	0.994
7	TI1	1.000	0.970	1.000	1.000	1.000	1.000	1.000
8	f-mod2 (Hz)	8.00	5.00	3.15	10.00	6.25	4.00	12.50
9	Raw MTR2	1.053	0.972	1.060	0.962	1.013	0.996	0.985
10	Corrected MTR2	1.051	0.966	1.045	0.948	1.009	0.995	0.984
11	TI2	1.000	0.986	1.000	0.921	1.000	1.000	1.000
12	Octave MTI	1.000	0.978	1.000	0.961	1.000	1.000	1.000

Tabular display of STIPA Analysis Details

The Speech Level measurement

The STI plug-in also adds a measurement named Speech Level (IEC 60268-16J) which determines the “speech level” as defined in Annex J of that standard. It applies A-weighting to the speech signal and removes the silent parts between words before calculating the level. The resulting speech level is a better estimate of the true signal level that can be used to select an appropriate level for the STIPA signal. The intent is similar to that of the Active Speech Level concept used in Recommendation ITU-T P.56.5

The Speech Level measurement has four results: Acquired Waveform, Frame Levels (A-weighted), Speech Level (A-weighted) and RMS Level (A-weighted). After A-weighting the signal, the measurement divides the signal into frames of 10 ms to 20 ms in length and calculates the rms level of each frame. The Frame Levels result is a plot of these levels versus time.



Frame Levels versus time

Perceptual Audio Software Options

In addition to the STI option discussed in this data sheet, Audio Precision also offers the following APx software options:

- ABC-MRT (speech intelligibility)
  ABC-MRT is an objective measurement of speech intelligibility using recorded speech and time-frequency patterns.
- PESQ (speech quality)
  PESQ is an established method for evaluating speech quality in low-bandwidth devices such as telephones, smart phones and hands-free devices.
- POLQA (speech quality)
  POLQA is the successor to PESQ, extending perceptual audio test to encompass the wideband measurement of speech quality, acoustic interfaces and more.

XVIII0330113734

Audio Precision
5750 SW Arctic Drive
Beaverton, OR 97005 USA

ap.com

(800) 231-7350

sales@ap.com