Application Note

Predicting the Speech Transmission Index (STI) in Bose Modeler 6.11

August 14, 2019

Summary:

From its introduction on the Apple Macintosh in the 1980's to recent times, the Bose Modeler sound system software has always aimed at providing fast and sophisticated tools to predict the perhaps most important attribute of sound systems: speech intelligibility.

Two core components of the Modeler program are essential to accomplish this task: First, the Hybrid Energy Decay Curve (HEDC^{TM}) yields a computationally efficient yet accurate representation of the system's impulse response and then second, the most accurate method to predict the intelligibility of a sound system, the Speech Transmission Index (STI), is applied to post-process this data in order to predict a single-number score for the expected intelligibility. The prediction of the STI follows a standardized procedure as defined in the IEC standard IEC 60268-16. This standard has been revised several times over the years since its introduction in 1988 and Modeler uses the year of publication as the identifier of the respective version of the underlying algorithm.

So far, Modeler has offered the following STI versions:

- 1988 (uni-gender),
- 2003 (Male and Female),
- 2003 (STIPA Male).
- 1998 (Male and Female),
- 2011 (Male and Female).

With the introduction of the 6.11 release, Modeler adds the following STI variant:

- 2019 (Male).

This application note is intended to give an overview on the available options for the prediction of the STI and to give recommendations which version should be used for a given application. This application note is *not* intended to give an introduction to the STI method itself or to the underlying Modulation Transfer Function (MTF) and its interpretation. It is assumed that the reader has a basic understanding of the program, its settings and how to run predictions.

The information in this application note is organized into two primary sections:

Section 1: Description of available STI versions
Section 2: Which STI version to use for a given application?

Associated Reading:

Modeler 6.11 Software Help File.

For Bose technical papers on intelligibility, please see:
https://pro.bose.com/en_us/products/software/acoustical_design/modeler_software.html#ProductTabs_tab1

Section 1 – Description of available STI versions

Selecting the STI version

In order to perform STI predictions in Modeler, you need to have at least a model of the space you want to analyze and also at least a single loudspeaker. You may then analyze STI either in form of a coverage map or for individually located listeners. For each sample of the map and/or for each listener, an HEDC-based MTF is predicted and post-processed to yield the single-number STI score.

In order to adjust the STI version:

1. Select ***Design – D²RASTIc – Speech (STI)***. Alternatively, you may use the keyboard shortcut ***F8*** or enable the STI and D²RASTIc buttons in the **Map Toolbar**.
2. Select the ***Simulation*** tab in the ***Detail View*** (keyboard shortcut: ***Ctrl-1***).
3. Use the ***Algorithm*** dropdown menu to select the desired version (publication date).
4. Use the ***Source*** dropdown menu to select the desired gender.

NOTE: The *Source* dropdown menu is only available for the STI versions 1998, 2003 and 2011. For the 1988 version, only an uni-gender spectrum is available and for both the STIPA (2003) and the 2019 version, only a Male spectrum is available.

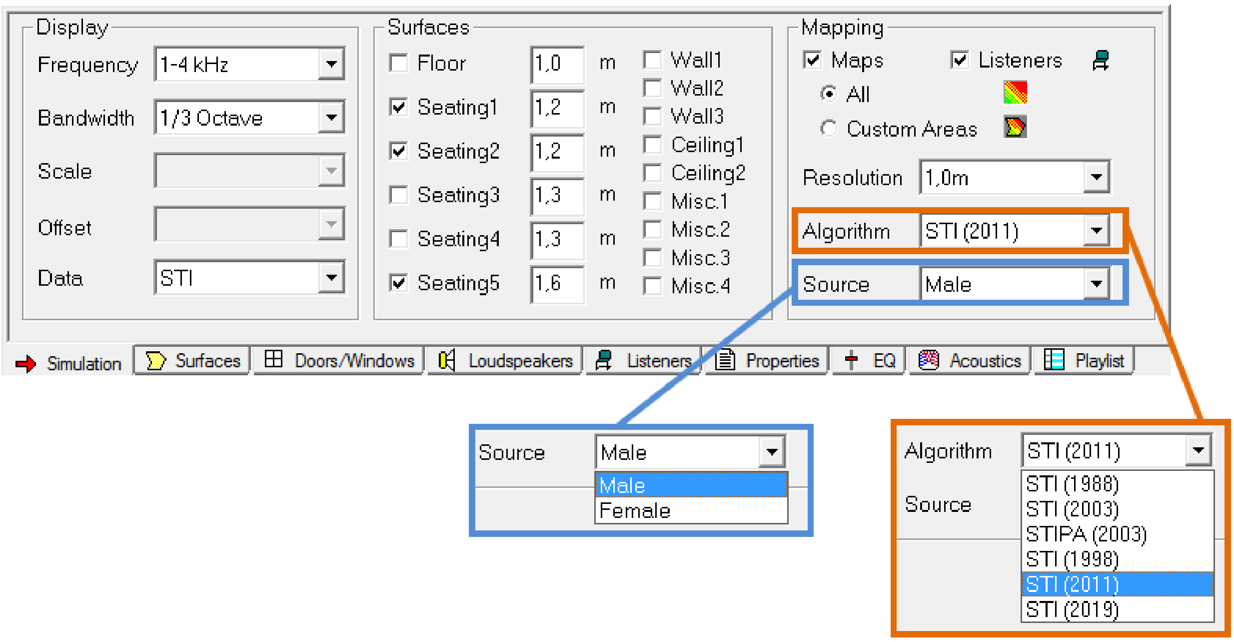

Figure 1 – Selecting the STI version in the Simulation Tab of the Detail View

Evolution of the various STI versions

The first version of the STI standard IEC 60268-16 was published in 1988. Subsequently, it has been revised in the years 1998, 2003 and 2011. Throughout the year 2019, the latest edition (‘Ed. 5’) will be published by the IEC and other standard organizations.

The changes that were applied to the STI algorithm over the years can be categorized into three main buckets:

1. Signal spectrum.
2. Octave-band weighting factors.
3. Level-dependent masking effects.

In the first version of the STI standard from 1988, a gender-independent speech signal spectrum and a simple frequency weighting scheme was used.

With the first revision in 1998, gender specific test signals were introduced for male and female talkers. Consequently, also specific sets of weighting factors, relating to each gender, were created. In addition, the weighting scheme was extended by introducing so-called redundancy factors (β).

With the second revision in 2003, a further refinement was included in the algorithm: Level-dependent masking effects. This addition is particularly important for the designer of sound systems: With the 2003 version, STI will deteriorate at high sound pressure levels, a situation commonly encountered in sound systems design.

As a side note, the 2003 version also introduced a new STI derivative named STIPA, which was co-developed by Bose Professional, TNO in the Netherlands and Gold Line. STIPA was specially developed as a fast measurement method that allows determining the speech transmission index of PA systems using handheld meters. Today, STIPA is available from a number of measurement systems and the de-facto standard for measuring intelligibility in the Pro Sound and Fire Alarm industries.

The third revision of STI standard was published in 2011. While the major aim of this revision was “to provide a more comprehensive, complete and unambiguous standardization of the STI methodology”, another important change was also incorporated: A new, continuous function for the prediction of auditory masking effects.

The latest revision of STI standard is scheduled to be published at the end of year 2019. Most notably, Female STI has been discontinued and thus, this gender option is not available anymore for STI 2019. For the Male version of the STI, the 2019 version introduces a modified source spectrum for male talkers where the octave band levels in the 125 and 250 Hz octave bands have been reduced.

The following subsections attempt to illustrate the various changes in more detail.

Speech spectra in the various STI versions

As mentioned before, the standardized spectra for speech have been revised with the 1998 version of the STI standard. In 1988, only a single, gender-independent, speech signal spectrum was available while in 1998 and in later versions, a distinction is made between male and female talkers. For the 2019 revision and Male STI, the levels in the 125 and 250 Hz octave bands has been *reduced* by about -6 dB and -3 dB, respectively.

All the various spectra employ some kind of low-pass response with the highest energy in the 125 and 250 Hz octave bands and a slope towards higher frequencies of roughly -6dB /octave. Thus, compared to a signal with a pink spectrum, the level of speech is significantly attenuated in the higher frequency bands. At 8 kHz, this difference may get as large as 25 dB for male speech and it is crucial that this aspect is remembered when predicted levels in Modeler are compared to disturbing background noise. A second important attribute only affects the Female spectrum: It does not contain any level in the 125 Hz band.

NOTE: Since Modeler's SPL predictions (and displays) are always based on a pink input spectrum, no specific display exists in Modeler that allows inspection of the actual speech spectra. Nonetheless, for STI predictions, Modeler internally considers the proper speech spectrum for the selected STI version and gender.

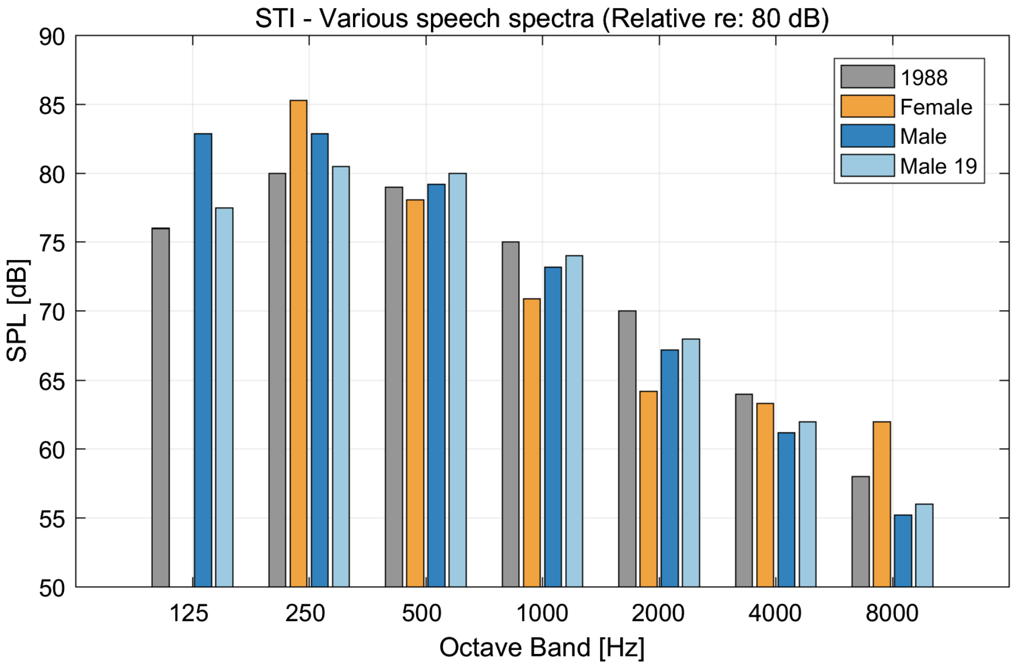

Figure 2 – STI speech spectra: 1988 (uni-gender), Male, Female

Figure 2 shows the three different speech signal spectra relative to a constant level of 80 dB SPL in each band. So for a sample or a listener location that has a flat frequency response from 125 Hz up to 8 kHz and octave band levels of 80 dB = const., the bars in Fig. 2 represent the speech signal levels that would have actually been utilized within Modeler's implementation of the STI algorithm.

NOTE: In case where you would like to get (numerical) access to the *actual* speech levels at certain listener locations, including the transfer function of the sound system in the room acoustical environment under consideration, please use the Right-Click-Menu of the MTF tab and select 'Export Listeners'.

Frequency weightings in the various STI versions

Together with the changes to the speech signal spectra in version 1998, a new scheme for weighting and averaging the contributions from the various octave bands was introduced. While only a simple weighting scheme was used in 1988, the 1998 and the following versions use so-called alpha and beta factors, where the latter are used to describe the redundancy of information in two adjacent frequency bands.

Figure 3 shows the various weighting factors for the 1988 version as well as for the Male and Female gender options of the 1998, 2003, 2011 and 2019 versions. Note that the beta redundancy-factors are depicted in between octave bands in order to illustrate the interaction between the two respective bands.

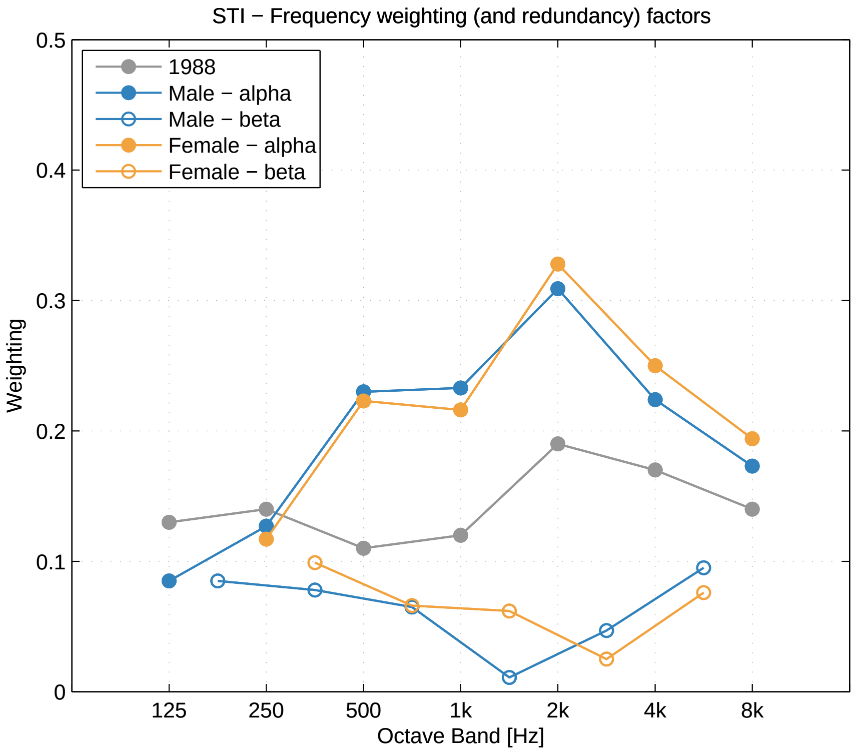

Figure 3 – STI Frequency weighting factors: 1988 (uni-gender), Male, Female

Level-dependent masking effects in the various STI versions

As already described above, one aspect of STI evolution is of particular importance for the designer of sound systems, especially if the system is supposed to run at high SPL, e.g. in a sports venue or in a voice alarm or emergency sound system. This component of the STI algorithm is termed 'Level-dependent masking' and it causes the STI to decrease for sound pressure levels above roughly 85 to 90 dB of total SPL. Level-dependent masking was introduced in the 2003 version and refined in the 2011 version. Due to a coarse SPL resolution, the 2003 version yielded a step-like response of STI vs. SPL. In contrast, the 2011 and later versions use a continuous masking function that yields a smooth response of STI vs. SPL. Figure 4 shows the implementation of the various masking schemes for Male speech in Modeler 6.11. The sound pressure level is increased in increments of 1 dB and A-weighted before summing for total SPL.

Please note that in the original revision of this application note, the SPL was shown as un-weighted SPL, so be careful when you compare figures.

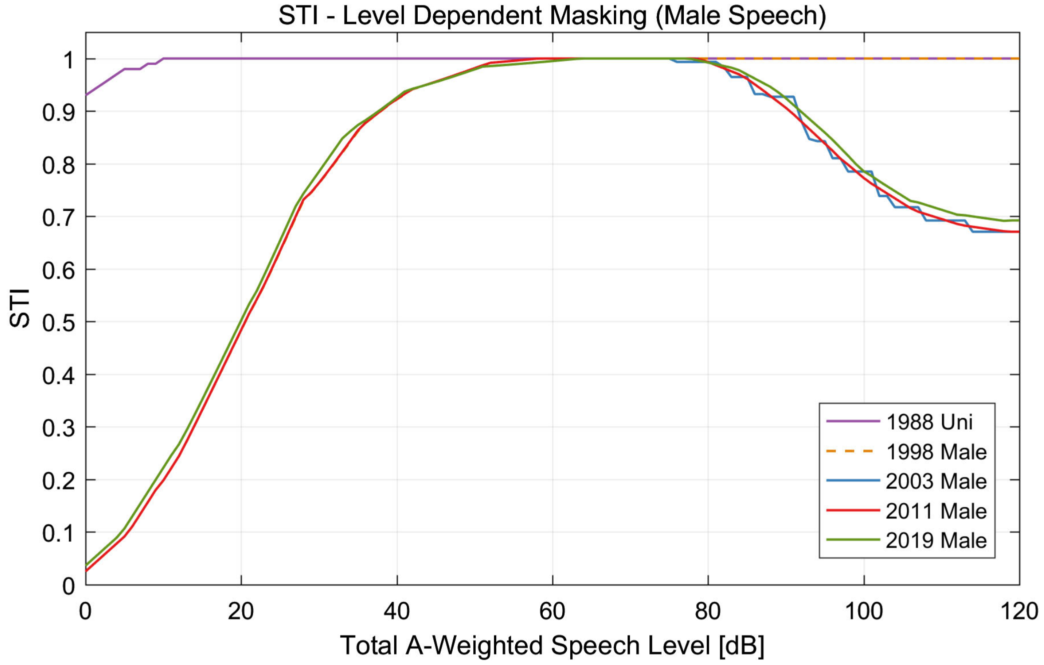

Figure 4 – Level-Dependent Masking: All (Male) STI versions as included in Modeler 6.11

NOTE: It is important to understand that the shown reduction in STI cannot be interpreted as a (constant) *relative* reduction for a given STI at a given (total) sound pressure level, i.e. the graph shall *not* be read as a reduction of STI in percent. For the above graph, a perfect MTF (consisting of Ones only) was assumed. In case the MTF is deteriorated by reverberation, noise or echoes, the relative reduction due to masking will be smaller than depicted above.

The next figure shows again the level dependency of the STI, but this time for the Female STI versions (1998-2011).

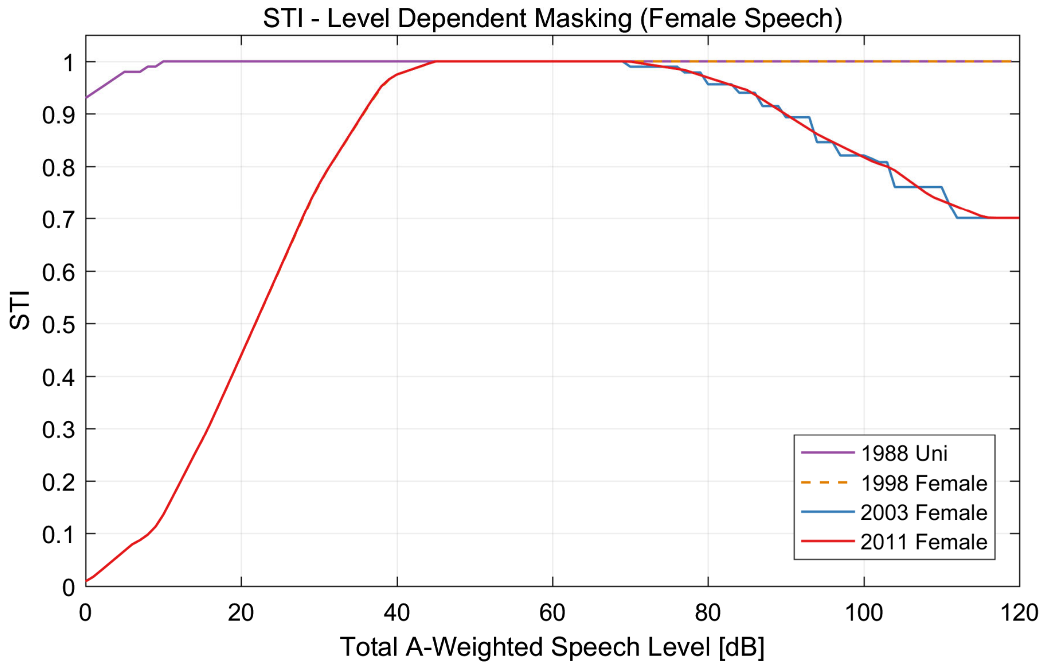

Figure 5 – Level-Dependent Masking: All (Female) STI versions as included in Modeler 6.11

Figures 4 and 5 do not only show a decrease of SPL towards high SPL, the STI is also reduced significantly towards very low sound pressure levels. This effect is caused by a part of the STI-algorithm that applies the so-called 'Auditory Reception Threshold' which is mimicking the absolute threshold of hearing of the human auditory system. While Modeler's implementation of the STI versions 1998 to 2019 is exactly according to the standard(s), Modeler's 1988 version uses a much lower Auditory Reception Threshold, resulting in STI's that don't decrease much at low SPL. This imposes no practical problem for the prediction of sound systems which are typically run at much higher levels. Nonetheless, should one desire to predict STI for unamplified sources, like natural speech, it is recommended to use any of the other versions (see also Section 2). For completeness, note that the effect of the Auditory Reception Threshold is identical for the 1998 to 2019 versions.

Section 2 – Which STI version to use for a given application?

Some generic advice

You may ask yourself: Why have so many different STI versions been implemented in Modeler? What version should I use for a certain modelling or design application? It turns out that each STI version in Modeler has its own justification.

As just one example, various versions of the STI standard are referenced in a variety of national and international standards that cover emergency sound systems or voice alarm systems and thus, it is of greatest importance for the system designer to perform predictions exactly according to the referenced standard versions. In this context, note that you will need to check carefully whether the STI standard is cited as dated reference, e.g. like 'IEC 60268-16:1998', or as an undated reference, where the year of publication is omitted. In the latter case, the latest published version of the IEC 60268-16 shall apply. At the time of writing (August 2019), this is the STI version from 2011 but from end of 2019 onwards, the 2019 version will be the most recent one.

Specifics for each STI version

The following paragraphs aim at providing some practical advice on the application of the various versions of the STI. Note that in sound system design, it has become common practice to use the Male gender version of the STI. Also, the STIPA meters for measuring STI are only validated for Male STI and therefore apply the corresponding parameters. Thus, the following list makes no specific distinction between Male and Female versions. Nonetheless, depending on SPL, the background noise spectrum and lastly, room acoustics, female voices may yield a slightly better STI score.

1988 (ed.1). The sole reason why Modeler offers this old STI version is to allow comparisons to predictions that have been formerly performed with Mac Modeler or very early versions of Modeler 6. We recommend to *not* use STI 1988 for your daily design work, no matter whether it is in room- or in electro-acoustics.

1998 (ed.2). The 1998 version of the STI is up-to-date in terms of frequency weightings and signal spectra for the female speech but doesn't employ auditory masking effects at high SPL. Thus, we recommend using STI 1998 for the purpose of optimizing your sound system design from a pure electro-acoustical perspective. Excluding the auditory effects allows the designer to focus on optimizing essential system attributes, like the direct-to reverberant ratio. You may also use STI 1998 to predict the intelligibility of un-amplified speech or to evaluate room acoustical scenarios but the recently added ed. 5 may be more suitable for scenarios that include background noise.

2003 (ed.3). The most important change in STI 2003 is certainly the addition of level-dependant masking effects, which make the STI result depend on sound pressure level (see Figs. 4 and 5). In addition, this third edition of the STI-standard introduced the STIPA method which is featured in all current STIPA meters and many other measurement systems. Most of them provide options to select the desired STI version and some also allow to turn on/off masking effects, independent of the chosen edition. Use the 2003 version if you need to predict according to exactly this revision. In most other cases, it is advised to use edition 4 from 2011. You may also use STI 2003 if you need to run predictions on how to trade-off an increased signal-to-noise ratio against auditory masking effects, but unless you need to predict exactly according to this standard revision, the 2011 version is much better suited for this task, see below for details. For a typical analysis of room-acoustical scenarios with un-amplified talkers, the results from STI 2003 will be very close, if not identical, to those from STI 1998.

2011 (ed.4). Currently, the 2011 version is the most commonly applied version of the STI method. Many international and national standards for emergency sound or voice alarm systems, like ISO 7240-19, EN 50849, CEN TS 54-32, BS 5839-8 or DIN VDE 0833-4 include an undated reference to the STI standard and as it was mentioned above, at the time of writing, the 2011 version of the STI standard is the currently applicable revision. Except for a few cases at high SPL, STI 2011 will yield almost identical results compared to STI 2003. Nonetheless, the updated masking scheme will avoid some undesired effects that occurred with the 2003 version, where very small changes in level could cause noticeable differences in STI (see Figs. 4 and 5). We therefore recommend using the 2011 (or later) versions of the STI in cases where you need to perform very precise predictions of STI as a function of SPL and background noise. Again, for a typical analysis of un-amplified talkers and low SPL, the results from STI 2011 will be very close, if not identical, to those from STI 1998 and STI 2003.

2019 (ed.5). There are two main changes that have been applied in the 2019 version of the STI standard. Female STI has been removed, thus it is not available as a gender option if STI 2019 has been selected. Second, as already mentioned, the Male speech spectrum has been adjusted in order to achieve better alignment to other research resources on human speech spectra. The reduction of speech level has two effects on STI: first, if significant background noise is present in the 125 and 250 Hz octave bands, the reduced speech level will lead to a reduced signal-to-noise ratio and in turn to a lower STI value. On the other hand, the upward masking effects will be reduced, leading to a small increase in STI. Depending on the speaker selection for the sound system, the altered speech spectrum may also significantly reduce the power that needs to be transmitted by the low frequency drivers. If those drivers are the limiting factor for the maximum achievable SPL of the sound system, then the 2019 spectrum will typically lead to slightly higher maximum sound pressure levels for speech transmission. Note that for the calculation of power consumption by the loudspeaker drivers, Modeler assumes a pink input spectrum and if maximum achievable SPL with other signal spectra is required, special procedures need to be followed. Please see Figs. 4 to evaluate the changes in level-dependant masking due to the altered speech spectrum.

We recommend using the 2019 version of the STI once it has been published in its final version by the IEC. In any case, STI 2019 represents the ‘state-of-the-art’ in terms of modelling system performance by means of the STI method.

Conclusion

We certainly hope the concepts and recommendations discussed in this application note are helpful to you. We hope you will enjoy working with Bose Modeler Sound System Software. While we do think that Modeler 6.11 provides the best combination of speed of prediction and versatility of STI versions being applied to sound systems, we encourage you to send us your suggestions about how to improve the tools or the workflow for predicting speech intelligibility in Modeler.