Technical Background Information on Selecting / Designing a loudspeaker for STI Measurements.

Challenge Nr 1: The Sound Level Requirements from EN 60268-16

The STI standard 60268-16 requires the loudspeaker to be capable of delivering clean 60 dB(A) signal level of male speech spectra. That sounds easy, but is not, because of further details:

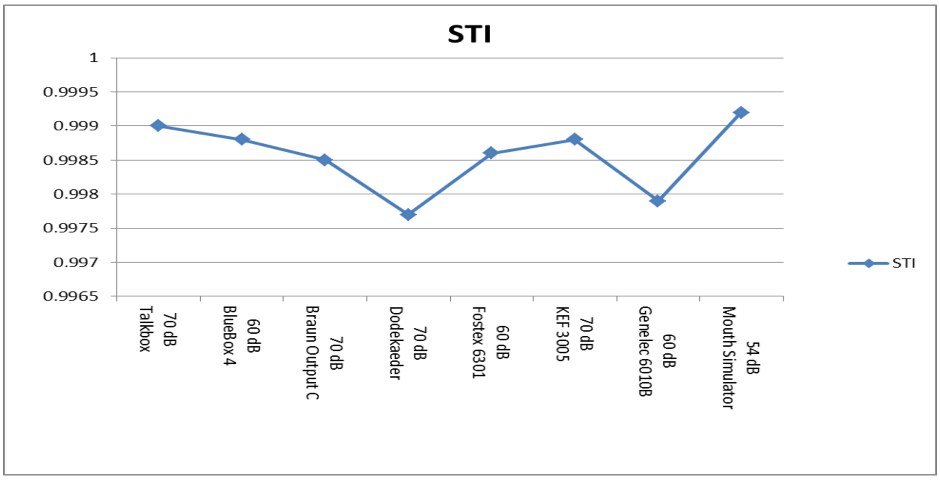
- 60 dB(A) with male on pink MLS spectra corresponds to a 62.9 dB(Z) signal level in the 125 Hz octave band
- As customer one may expect a STI speaker to be capable of delivering the Lombard level of signal output: 70 dB(A). So we are talking about 72.9 dB(Z) signal level in the 125 Hz band.
- The Direct Method of 60268-16 requires modulated pink noise as test signal - and it has about 13 – 14 dB crest factor. With IR systems using Pink random signals, the crest factor is often similar. In both cases, the 125 Hz signal level may achieve about 83 dB(Z).
- The loudspeaker gets driven in a housing far too small to can deliver its normal bass response. Using bass reflex design is an option to improve this aspect; just there is a price to pay: It reduces the MTF quality in the 125 Hz band slightly (but not to a relevant amount).

So it is not easy to design a reference speaker for STI measurements covering all these requirements, because this means that the loudspeaker chassis must be capable to drive a **large linear** membrane deflection. Practical experience shows, that a linear membrane movement capability of up to about 5 mm is required; something very few small chassis can do. Typical 10 cm or 12 cm (outside dimension, not size of the active "piston") chassis have an X-Max of about 2 mm. Only high end chassis like the Tangband W4-1879*** and a few other models have such X-Max qualities. Mostly the price is a good indicator for the level of chassis - such quality costs about € 150.- (single unit price). *** Author will maybe build a speaker based on this model.

With smaller chassis than 10 cm diameter, the Lombard level can mostly not get achieved - we are here still talking about using the standards male spectra without post processing tricks like the AE / B&K ECHO Box uses. The BlueBox, as mentioned in the stimulus build manual to the Excel stimulus build template, is not capable of delivering clean 70 dB(A); while 60 dB(A) is no problem.

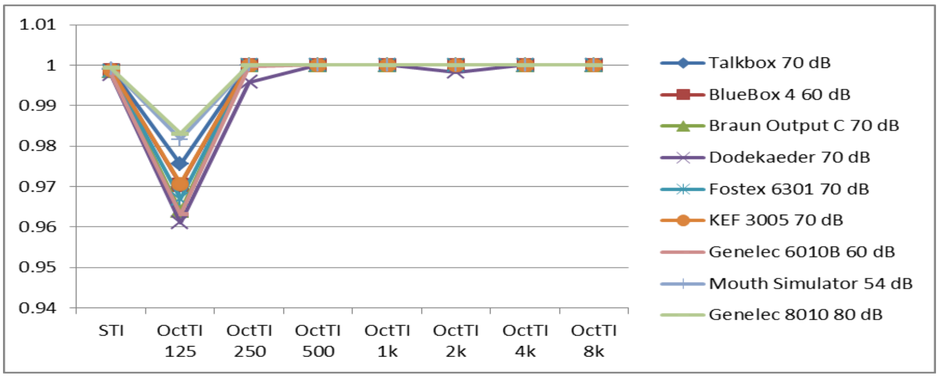
The author made a test with several speaker types, below you find some results. The measurements were made (in about 5 dB steps) at the highest still fully clean signal level the device could deliver - which varies a lot depending on the speaker and its design (all measurements were made at 0.5 m / with the signal level recalculated to 1m / in free field):

***Remark**: Both charts (next page) measurements were made some years ago using ARTA, which uses a multi-sine based stimulus with about 9 dB crest factor. Therefore be prepared that with other system up to 5 dB lower usable level max limits could result - e.g. with the Genelec 8010 with other systems 80 dB will not be possible. All other levels could / should work. Equalization was done with a 1/3 octave equalizer (except NTi's TalkBox, which has a build in DSP-based frequency response correction).*



What one can see out of the picture is, that all the speakers, when used at a level where they still work in a linear way, deliver a by far signal clean enough to not deteriorate the precision of the STI measurements - having no negative impact on the second digits behind the dot of STI result figures.

The following picture shows very well, how critical the 125 Hz band is regarding clean MTF results:



Challenge Nr 2: The Directivity Requirements from EN 60268-16

The STI standard requires speech directivity similar to a human mouth. Regarding shape, the loudspeakers should be close to a human head. The mouth opening should get emulated by a 20 mm opening or similar. The standard requires a mouth simulator or HATS (head and torso simulator) or equivalent. Just the standard has a back door, allowing "in absence of an artificial mouth" small speakers up to 100 mm cone diameter. The author finds this rather frustrating (and not helpful), since the directivity of 20 mm or 100mm sound openings are worlds apart and can have a relevant negative impact in critical situation. This while the standard does not allow 2-way speakers, although in room measurements no negative impacts appear. Further: The meaning of "in absence" is not clear (For end user only ?).

As result of above, *no STI test loudspeaker on the EU market the author knows about, complies to 60268-16 rules without using the standards back door*.

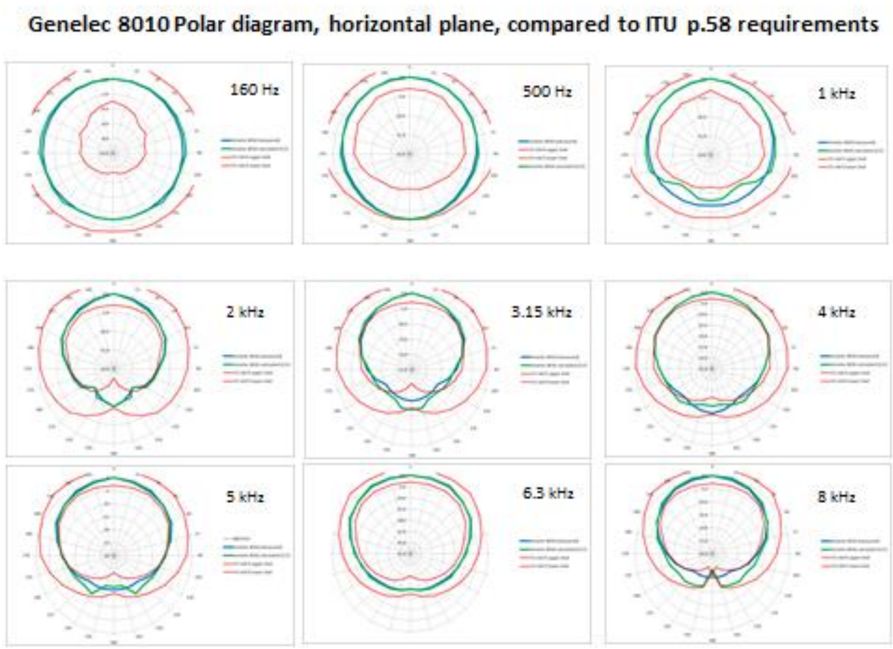
As consequence, the author of this article uses today 3 types of reference speakers - depending on which fits best to the expectations / analysis needs:

- A self-made (passive) product called **BlueBox** with an 8 cm chassis (Visaton B80; about 6.5 cm cone size), fully compatible to the standard, but not capable of delivering the Lombard 70 dB(A) signal level. (The Lombard level represents the voice level people in emergency situations use when talking into public address waring systems (the main market of STIPA)).

- The **Genelec 8010** is an active 2-way speaker. 2-way is not allowed according to the standard and may not be used for talking into command / push-button microphones, since 2 way speaker have uneven and undefined near field spectra. For STI measurements in rooms (conference, class, video conference, speech rooms) no disadvantages are to be expected - and none have been detected by the author so far. (HiFi speakers have also proven enough, that multi chassis products can deliver precise signals). The Genelec 8010 has some interesting qualities:

    - Nearly perfect horizontal directivity due to 19 mm tweeter and speaker size and shape like a human head. In vertical direction this is also fairly true - especially when the speaker is mounted upside down…
    - Capable of 70 dB(A) Lombard Level; 2 x 25 W power with speaker coil overheat protection.
    - Easy to use; it sits well balanced on a microphone stand and can easily get moved about in a room. The balanced signal input allows long cables without hum pickup.
    - If you don't equalize the speakers frequency response (e.g. due to no short term possibility, to go to a really good anechoic chamber), and just implement the male spectra in DIRAC's SFD, you are typically 1 dB outside the 60268-16 specs - leading to a moderate - if any - error. *But the speaker is not 60268-16 compliant.*

Picture below: Genelec 8010 horizontal directivity based on supplier data (6100).



- The **Fostex 6301B speaker** (now updated by Fostex to a Nx version). There are very few speakers on the market suitable for STI measurements, and this active loudspeaker gets proposed by several German acoustic companies for STI measurements. As broadband system, it is not in the same class of signal quality as the Genelec 8010; but with a carefully configured stimulus build, not only the standard can get fulfilled, but also the Lombard signal level should be possible. For details see the author's Manual regarding stimulus build.

Challenge Nr 3: The Directivity Requirements - Part II

The directivity aspect is again and again a topic in technical papers. A detailed analysis was done by Peter Mapp; presented at the 137. AES Convention in Los Angeles, October 2014 (AES Paper 9156). He came through computer simulation to the conclusion, that the speaker's cone diameter should be 50 – 60 mm or less; elsewise errors up to about 0.1 STI can occur.

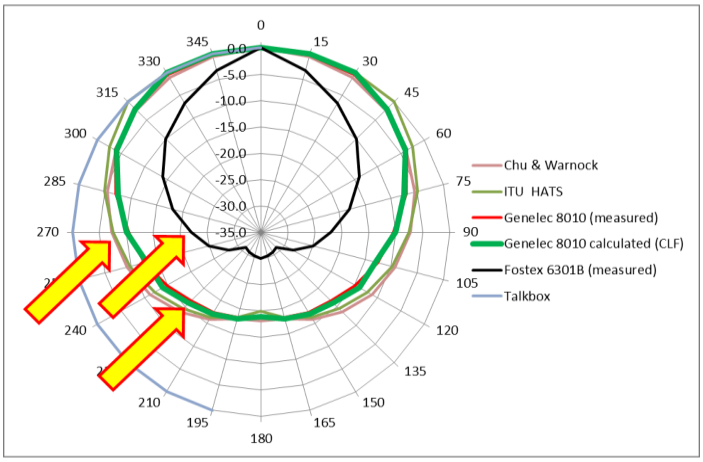
The author knows only one STI speaker on the market to fulfil the 2- 2.5 inch speaker size demand: The AE / B&K ECHO Box. *With speaker cones in this size, male RTA spectra is not possible; that must get done in post processing.* DIRAC has an extensive set so tools for all kinds of STI model conversions, not only for the ECHO speaker - any DIRAC user can apply them freely.

Using an Omni can lead in several situations to more precise results than using a 5 inch chassis based speaker - as to expected and also confirmed by some of Peter Mapp's investigations. Some STI specialists use the Omni for the teacher position in a class room, when the teacher writes on the black board and talks at the same time. The critical question is, what signal level shall be chosen, since the EN 3382-3 standard uses an other signal level (and spectra) than 60268-16 - different signal levels as directional speaker or not.

Below a picture of the directivity of the Fostex 6301 and the Genelec 8010 speakers; the Fostex based on own free field (flat roof corner) measurements, the Genelec based on supplier data.

This is not the place to go into details, since there is also a difference between just regarding the head of a speaker only, or including the impact of the rest of the body - what the HATS (head and torso simulator) does. Further the directivity should be measured in (at least) 2 planes.

The picture below shows directivity measurement / requirements curves from leading researchers, the ITU standard values, and results from measuring a Fostex 6301 speaker (own measurements). To measure the NTi Talkbox was planned, but did not get done (blue line).

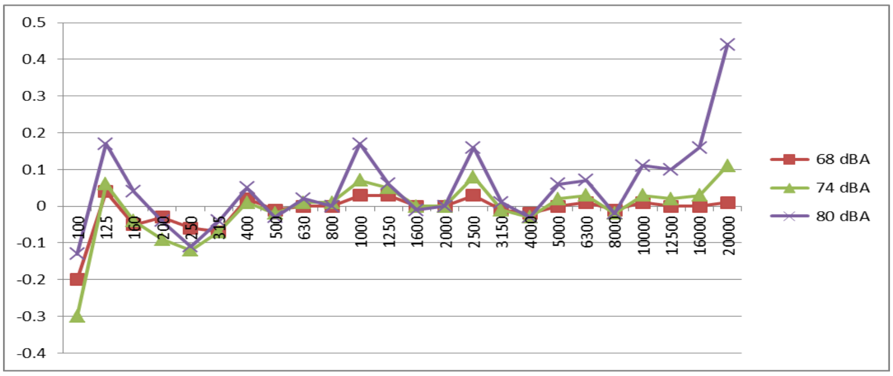


In the document mentioned further above, Peter Mapp addresses the possibility of a 2-way solution based on a coax designed loudspeaker. The author of this paper made once a test with a coax home cinema speaker - but was not convinced by the results and did **so far** no further investigations.

Challenge Nr 4: Measuring the speaker IR - at what level ?

Measuring the speakers IR is a critical process for the stimulus quality. The measured signal should be free of noise and distortions. Sounds fine, just there is an inherent conflict in this requirement: The higher the signal level, the lower the noise - but distortions get higher.

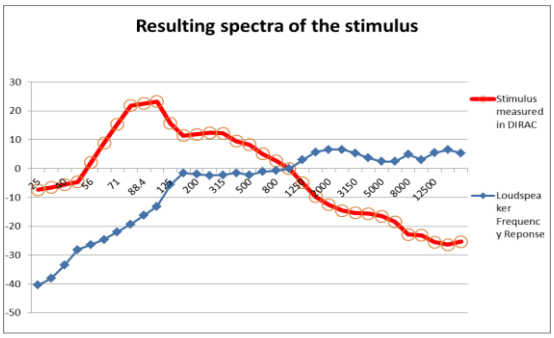
Here an example of a 2-way speaker, measured from 62 dB to 80 dB in 6 dB steps. *Signal used was PinkMLS;* the curves are relative to the 62 dBA measurement:



The picture shows how the relative signal level changes when going for higher levels - expression of non linearities. Taking the 62 dB level or lower would deliver the lowest distortions, but at a price of having noise in the frequency range below 100, especially below 50 Hz, **which must get taken into account when setting the low end cut** (-> template, manual)**.**

0.1 dB is about 1% of distortions (ITU-p.51: > 1% THD); the chart shows an increasement form 62 to 68 dB by less than 0.5%. So one may expect at 62 dB very low distortios - maybe 0.2% - since distortions develop (depending on there order) linearly of progressive with the signal level.

The stimulus should have a spectra something like below, to get best possible robustness for not tending to side bands in critical deconvolution situiations - eg. when measuring a room with very poor STI due to long distances, corners, etc, and the room is quite.



The experiences of the author showed good results with a stimulus level from 25 to 60 Hz on a level similar to the 1 kHz to 2 kHz bands. If the IR frequency response measurement of the test loudspeaker you selected as base for the SFD process, has a lot of low end noise, you need to tailor the low end cut; it must not only reflect the frequency response decay of the speaker, but also the overlapping noise.

The template has pale brown and violet data block groups:

Stimulus improvement Iterations — IR: BlueBox 3 FG 80 dBA.wav — Path: Average

Cut low-end — given by the standard

Freqency	STI Male	Loudspeaker Frequency Reponse	Stimulus measured in DIRAC	Resulting Error 1	Hz	2. Shape table data input	Stimulus measured in DIRAC	Resulting Error 2: Error of Stimulus V1	Hz	3. Shape table data input	Stimulus measured in DIRAC	Resulting Error 3
		-63.86	Offsets	-66.80		Offset of this block		-130.92		Offset of this block		-130.95
25	-50	-40.37	59.46	2.29	25	-50	-7.32	2.31	25	-50	-7.31	2.32
32	-45	-38.06	60.21	0.35	32	-45	-6.58	0.36	32	-45	-6.57	0.37
40	-40	-33.51	61.25	0.94	40	-40	-5.54	0.95	40	-40	-5.53	0.96
50	-35	-28.12	62.18	2.26	50	-35	-4.61	2.27	50	-35	-4.6	2.28
56	-30	-26.38	68.695	5.52	56	-30	2.14	5.76	56	-30	2.155	5.78
63	-20	-24.64	75.21	3.77	63	-20	8.89	4.25	63	-20	8.91	4.27
71	-10	-21.97	82.12	3.35	70	-10	15.375	3.41	70	-10	15.385	3.42
79	0	-19.30	89.03	2.93	80	0	21.86	2.56	80	0	21.86	2.56
88.4	10	-16.21	89.875	-3.13	88.4	10	22.62	-3.59	88.4	10	22.55	-3.66
100	10	-13.12	90.72	0.80	100	9.20	23.38	0.26	100	8.94	23.24	0.12
125	10	-5.65	83.58	1.13	125	8.87	15.84	0.19	125	8.68	15.68	0.03
160	10	-1.52	78.2	-0.12	160	10.12	11.34	-0.18	160	10.30	11.46	-0.06
200	10	-1.91	78.47	-0.24	200	10.24	11.9	-0.01	200	10.25	11.93	0.02
250	10	-2.38	79.02	-0.16	250	10.16	12.45	0.07	250	10.09	12.43	0.05
315	10	-2.25	78.57	-0.48	315	10.48	12.17	-0.08	315	10.56	12.23	-0.02
400	8	-1.49	76.16	-0.13	400	8.13	9.57	0.08	400	8.05	9.54	0.05
500	6	-2.26	74.89	-0.17	500	6.17	8.27	0.01	500	6.16	8.27	0.01
630	4	-1.11	71.73	-0.18	630	4.18	5.1	-0.01	630	4.19	5.12	0.01
800	2	-0.68	69.45	-0.03	800	2.03	2.72	0.04	800	1.99	2.71	0.03
1000	0	0.00	66.75	-0.05	1000	0.05	-0.01	-0.01	1000	0.06	0	0.00
1250	-2	2.95	61.97	0.12	1250	-2.12	-4.9	0.05	1250	-2.17	-4.93	0.02
1600	-4	5.73	57.16	0.09	1600	-4.09	-9.69	0.04	1600	-4.13	-9.71	0.02
2000	-6	6.57	54.01	-0.22	2000	-5.78	-12.6	-0.03	2000	-5.75	-12.57	0.00
2500	-8	6.56	51.97	-0.27	2500	-7.73	-14.55	0.01	2500	-7.74	-14.54	0.02
3150	-10	5.39	51.11	-0.30	3150	-9.70	-15.39	0.00	3150	-9.70	-15.38	0.01
4000	-12	3.66	50.91	-0.23	4000	-11.77	-15.63	0.03	4000	-11.80	-15.64	0.02
5000	-14	2.47	50.07	-0.26	5000	-13.74	-16.47	0.00	5000	-13.74	-16.46	0.01
6300	-16	2.45	48.14	-0.21	6300	-15.79	-18.46	-0.01	6300	-15.78	-18.45	0.00
8000	-18	4.88	44	0.08	8000	-18.08	-22.82	0.06	8000	-18.14	-22.85	0.03
10000	-20	3.11	43.85	0.16	10000	-20.16	-23.07	0.04	10000	-20.20	-23.07	0.04
12500	-20	5.42	41.44	0.06	12500	-20.06	-25.55	-0.13	12500	-19.93	-25.47	-0.05
16000	-20	6.47	41.4	1.07	16000	-21.07	-26.32	0.15	16000	-21.22	-26.4	0.07
20000	-20	5.36	42.25	0.81	20000	-20.81	-25.38	-0.02	20000	-20.79	-25.36	0.00

High end side band management — These fields are for stimulus generation based on calculation

BlueBox 3 2016-06-30 — These fields are for measurement in the anechoic chamber

Hz	4. Shape table data input	Stimulus measured in DIRAC	Resulting Error 4	Hz	5. Shape table data input	Result measured in the anechoic chamber	Resulting Error	Hz	6. Shape table data input	Result measured in the anechoic chamber	Final Resulting Error	1/3 octave STI male target value
	Offset of this block		-134.34		Offset of this block		-134.70		Offset of this block		-134.70	
25	-50	-134.34	-124.71	25	-50	-134.70	-84.7	25	-50	-134.70	-84.7	
32	-45	-134.34	-127.40	32	-45	-134.70	-89.7	32	-45	-134.70	-89.7	
40	-40	-134.34	-127.85	40	-40	-134.70	-94.7	40	-40	-134.70	-94.7	
50	-35	-134.34	-127.46	50	-35	-134.70	-99.7	50	-35	-134.70	-99.7	
56	-30	-134.34	-130.72	56	-30	-134.7	-104.7	56	-30	-134.7	-104.7	
63	-20	-134.34	-138.98	63	-20	-134.70	-114.7	63	-20	-134.70	-114.7	
70	-10	-134.34	-146.31	70	-10	-134.7	-124.7	70	-10	-134.7	-124.7	
80	0	-134.34	-153.64	80	0	-134.70	-134.7	80	0	-134.70	-134.7	
88.4	10	-134.34	-160.55	88.4	10	-134.7	-144.7	88.4	10	-134.7	-144.7	
100	8.82	-134.34	-157.46	100	166.28	-134.70	-144.7	100	310.98	-134.70	-144.7	58.4
125	8.65	-134.34	-149.99	125	158.64	-134.70	-144.7	125	303.34	-134.70	-144.7	58.4
160	10.36	-134.34	-145.86	160	156.22	-134.70	-144.7	160	300.92	-134.70	-144.7	58.4
200	10.23	-134.34	-146.25	200	156.48	-134.70	-144.7	200	301.18	-134.70	-144.7	58.4
250	10.04	-134.34	-146.72	250	156.76	-134.70	-144.7	250	301.46	-134.70	-144.7	58.4
315	10.58	-134.34	-146.59	315	157.17	-134.70	-144.7	315	301.87	-134.70	-144.7	58.4
400	8	-134.34	-143.83	400	151.83	-134.70	-142.7	400	294.53	-134.70	-142.7	56.4
500	6.15	-134.34	-142.60	500	148.75	-134.70	-140.7	500	289.45	-134.70	-140.7	54.4
630	4.18	-134.34	-139.45	630	143.63	-134.70	-138.7	630	282.33	-134.70	-138.7	52.4
800	1.96	-134.34	-137.02	800	138.98	-134.70	-136.7	800	275.68	-134.70	-136.7	50.4
1000	0.06	-134.34	-134.34	1000	134.4	-134.70	-134.7	1000	269.1	-134.70	-134.7	48.4
1250	-2.19	-134.34	-129.39	1250	127.2	-134.70	-132.7	1250	259.9	-134.70	-132.7	46.4
1600	-4.15	-134.34	-124.61	1600	120.46	-134.70	-130.7	1600	251.16	-134.70	-130.7	44.4
2000	-5.75	-134.34	-121.77	2000	116.02	-134.70	-128.7	2000	244.72	-134.70	-128.7	42.4
2500	-7.76	-134.34	-119.78	2500	112.02	-134.70	-126.7	2500	238.72	-134.70	-126.7	40.4
3150	-9.71	-134.34	-118.95	3150	109.24	-134.70	-124.7	3150	233.94	-134.70	-124.7	38.4
4000	-11.82	-134.34	-118.68	4000	106.86	-134.70	-122.7	4000	229.56	-134.70	-122.7	36.4
5000	-13.75	-134.34	-117.87	5000	104.12	-134.70	-120.7	5000	224.82	-134.70	-120.7	34.4
6300	-15.78	-134.34	-115.89	6300	100.11	-134.70	-118.7	6300	218.81	-134.70	-118.7	32.4
8000	-18.17	-134.34	-111.46	8000	93.29	-134.70	-116.7	8000	209.99	-134.70	-116.7	30.4
10000	-20.24	-134.34	-111.23	10000	90.99	-134.70	-114.7	10000	205.69	-134.70	-114.7	28.4
12500	-19.88	-134.34	-108.92	12500	89.04	-134.70	-114.7	12500	203.74	-134.70	-114.7	
16000	-21.29	-134.34	-107.87	16000	86.58	-134.70	-114.7	16000	201.28	-134.70	-114.7	
20000	-20.79	-134.34	-108.98		88.19	-134.70	-114.7			-134.70	-114.7	

In the brown iterations the impact of the frequency response of the speaker gets calculated - we are in a kind of simulation situation. The nonlinearities as above do not get corrected - only in the iteration using the violet fields, where the speaker's impact gets measured. The brown fields have the job to allow optimizing the low end cut and to check in the iteration process works well. Further it has the advantage of not requiring the use if an anechoic chamber, while in the later, violet iteration phases, using the speaker in the measurement loop, is a must.

With the speaker shown above, only moderate changes between the brown and violet measurements sub process will be obvious - the speaker does an fairly good job. But if the results look like this:

Her we are in a very different situation, also because the frequency response measurements was done using MLS (and not PinkMLS). ***Principally the frequency response IR for the SFD process should never get measured with MLS or lin sine sweep****; the low end noise will be for sure too disturbing.*

Using MLS (instead of PinkMLS), a 10 dB higher level (90 dBA) got used for a reasonable SNR, contributing much to the high end inconsistencies in the picture.

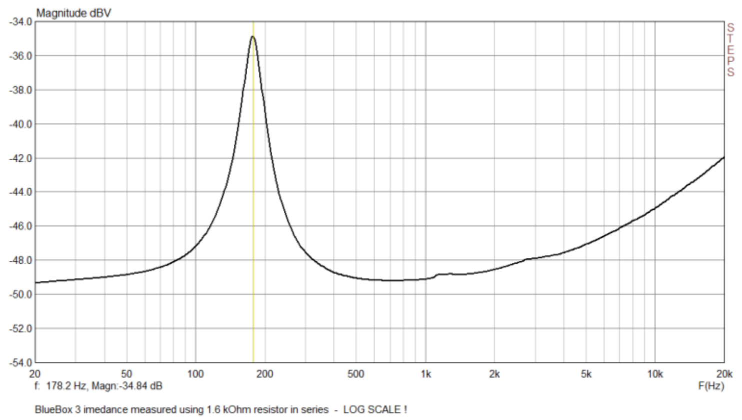
With a situation as just shown, the main part of the iteration work will - even when using PinkMLS - need to be done with the speaker in the loop (violet fields). *After doing all clarification work with the development template (mainly brown fields), use for the final work the productive template (all block with violet header).*

Preparing you to find the Right Decisions

Now we are going into deeper technical issues - the energy demand and the resonance frequency of the loudspeaker.

The resonance frequency is extremely important, because a loudspeaker can only efficiently radiate sound **above** their resonance frequency - this in contrast to measurement microphones, which have only a really flat frequency response **below** the membrane resonance frequency.

Taking the BlueBox speaker,..



Picture: At 178 Hz the speaker has about 5 fold of the DC impedance.

…we see that it has a resonance at about 178 Hz. But the test spectra goes down to 89 Hz for male measurements. This, in combination with the radiation physics for speakers, is the reason; why in the graph in the middle of page 2 shows a dramatic MTF drop at 125 Hz.

This is a simple test, and if you have a speaker already you are thinking about to use, you can make the test yourself. See at the end of this document. *The lower this resonance is, the better.*

But there is also an other point which plays a deciding role: The power going into the speaker:

The BlueBox as shown here, required about 150 mWrms signal power for emitting the 60 dBA signal. That’s not much, but the signal has a (measured) crest factor of nearly 13 dB for this speaker - meaning nearly 3 Watt peak signal level. If you want the speaker to emit the Lombard signal level of 70 dBA, then about 30 Wpeak are required - and much of the energy is in the frequency region, where a small speaker is not capable to emit its nominal power rating and has therefore a much reduced load capability.

So here you have the detailed explanation, why to reach 60 dBA is easy, and to reach 70 dBA of clean male spectra on pink MLS is a huge challenge with small speakers.

Challenge Nr 5: Finding the Right Decision

To do the right thing - here some aspects:

- Do you really need 70 dBA signal level capability ? Do you really need emitted male spectra ? Do you really need the combination of both ? That depends much on your applications.
- To select or develop a loudspeaker box for STI can be a time consuming job. If that happens as student, within a dissertation etc., where in depth learning is important - then cost will not play much of a role. But in a commercial company, mostly “time is money”. If the AE / B&K ECHO speaker covers your requirements, this will likely be by far the most economical solution. The limitations of ECHO are: No emitted male spectra, no automatic distance measurements, as available in DIRAC for 3382-3 measurements. The STI value for male spectra gets calculated in post processing - and not measured directly.
- On the other hand: If male spectra combined with Lombard signal level (70 dBA) is a must, then there is likely only one product that can be bought ready for use: The NTi TalkBox, which includes a DSP for frequency response linearization (In SFD therefore only the 60268-16 spectra must get generated - a small job). To make such a product with a chassis cone of max 100 mm is not easy - if you want to do such, it will cost time. The author himself has so far not had the opportunity to test the combination DIRAC with the TalkBox; but he rented it a few years go a few times, and got very good results (in those days in combination with Arta). Testing it with DIRAC before buying such a device is much adwised; don’t forget to test also the 3382-3 distance measurement capability; it could someday come in as being very helpful although you don’t need it now.
- 2-way loudspeakers are in a much better position as broadband devices to achieve precise and powerful results. The user must decide himself if he wants to override the standard when measuring in rooms. This based on the general opinion, that with distances above about one meter, 2-way systems can be used instead of broadband speakers.

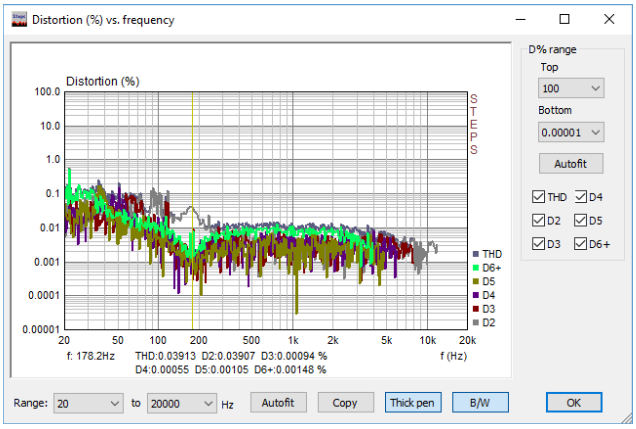
Feedback is welcome to the author: alastair.gurtner@sunrise.ch

Attachment: Measuring the Resonance Frequency of a Speaker

Its easy to do; you take the line or phone output of your sound converter, send the signal through a resistor of about 2 kOhm and measure the signal level over the speakers coils (terminals) with the line / microphone input. Using a large resistor value compared to the speakers impedance, the signal sent into the speaker works like a current source and the voltage is proportional to the impedance.

Principally you can use also DIRAC for the measurement; the author used a specialised software for electro acoustical measurements: a heterodyne style of tool: STEPS.

This has some advantages, on the one side it measures with a steady standing pure sine signal (no chirp - frequency steps), but also to be sure the output of the converter is not overloaded (“soft short cut”). After the measurement it is possible to switch to a distortions view, showing the output was not overloaded by the resistor value (here 1.6 kOhm, rather a low load for such outputs).:



In the picture below, between the speaker and the converter (blue panel) there is a small pale grey box; it holds the 1.6 kOhm resistor in series to the speaker. Output level is short to 1 Vrms. Black in the front (right side) you see the power amplifier for the BlueBox speaker, capable of clean 60 W into 8 Ohm - but not in use in this measurement here (could have been used if the converters output would have been overloaded by the resistor).