OLE WOLFF

2-way intelligent audio interface with devices via IoT & AI

WP06

Feb. 2025

Table of Contents

Introduction

With AI (Artificial Intelligence) entering the scene and IoT (Internet of Things) being reality the recent years, some brand-new possibilities for communicating with a new generation of devices arise.

Imagine if you could have an intelligent 2-way audio communication with your numerous electrical devices you have in your position today?
That goes for house holding machines, stereo system, car control, fitness equipment etc.
A conversation with your future washing machine could be you asking:
"I have a wool sweater with a stain of unknown cause. Could you give me your best advice of washing this sweater clean?"
And the washing machine could answer:

"I will recommend you soak your sweater in lukewarm water with added wool soap for about 2 hours. After that, I'll wash it for you on program 4".

This White Paper deals with the issue of tuning an IoT loudspeaker to give the very optimal speech intelligibility performance.

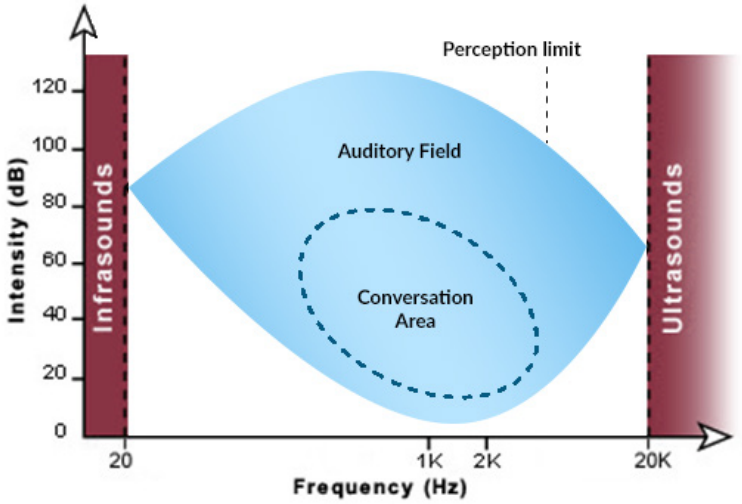
Range of human perception of sound

The human perception of sound corresponds to a specific frequency band and a specific range of sound intensities, all perceived by our ear.
The human ear (full functional hearing capability of an average young person) perceives frequencies from 20Hz to 20kHz. Acoustic vibrations outside this frequency band are considered to be "vibrations" and as such not a sound.

During aging, the human perception of sound degrades, especially in the higher frequency region either caused by normal aging of the body functions or by exposure to very high sound pressure levels over lifetime.

The sound intensity range the human ear can handle (dynamic range) goes from the threshold of hearing (0dB SPL) peaking at 1-2kHz to a maximum perception limit at around 120dB at 800Hz. (see *Fig. 1*).

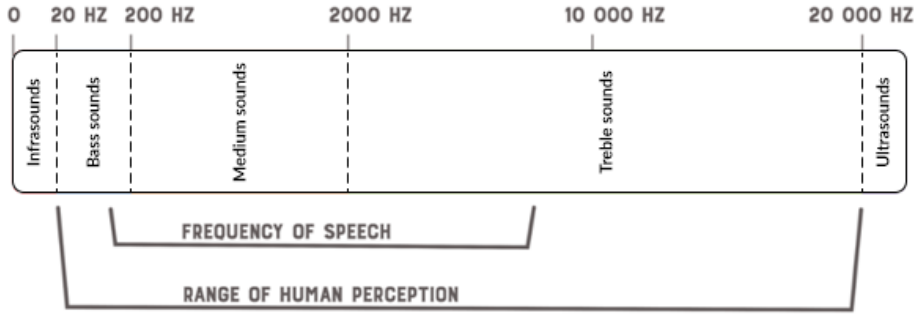
Fig. 1



Frequency range of human speech

The frequency range of human speech is very limited compared to the total frequency band of the human range of hearing.
Speech frequency range is a span from 300Hz to around 4kHz. (see *Fig. 2*)

Fig. 2



This fact is actually the reason why the telephone network is designed to handle just this frequency range and these telephone standards were established a century ago.

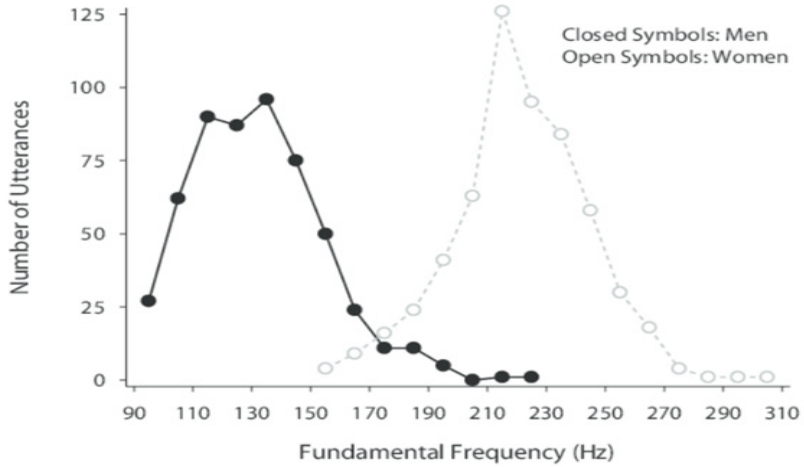
Formant frequencies (male & female) determining human speech recognition

Formant frequencies are those frequencies in the speech spectrum, which are dominant for males and females respectively.

In other words, the frequency content, which makes us able to distinguish if the voice is coming from a male or a female and contain the information that we actual know the person who is talking. - This is also defined as “human speech recognition”.

These formant frequencies are mainly distributed in the low to mid-range of frequencies in the voice spectrum. Below is seen the fundamental frequency span of the formant frequencies for male and female respectively (see *Fig. 3*)

Fig. 3



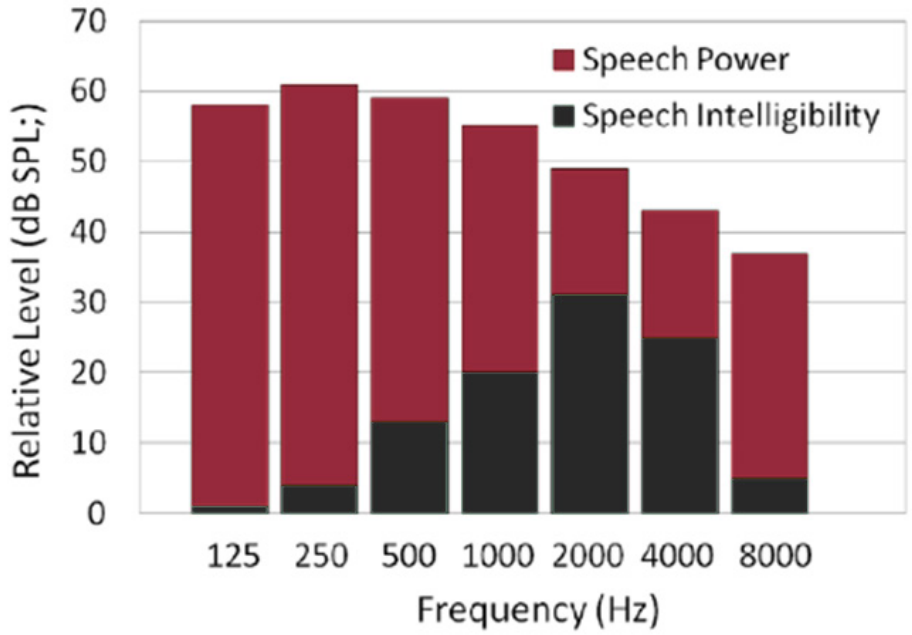
Human speech intelligibility

For an audio device reproducing speech, it is not necessary to reproduce the formant frequencies in speech except if there is a requirement for the voice to have a male or female character.

Rather, focus should be on emphasizing the frequency range in speech where most of the speech intelligibility information is present.

This intelligibility frequency range in human speech can be seen in *Fig. 4* showing the total speech power in red and the speech intelligibility part in black.

Fig. 4



Loudspeaker designs optimized for machine generated speech

With *Fig. 4* as a starting point, loudspeakers for IoT devices can be optimized to have most of their sensitivity positioned in the green marked frequency range.

It is possible to use DSP (Digital Signal Processing) for achieving this emphasized sensitivity response, but as always, it is good audio engineering practice to tune to preferred response in the loudspeaker design first and then maybe fine tune the rest by use of DSP.

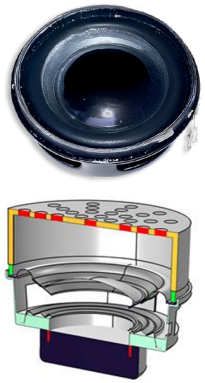
The reason is the retention of optimal dynamic range and signal to noise ration.
As examples for this, two different loudspeaker sizes and form factors are chosen.

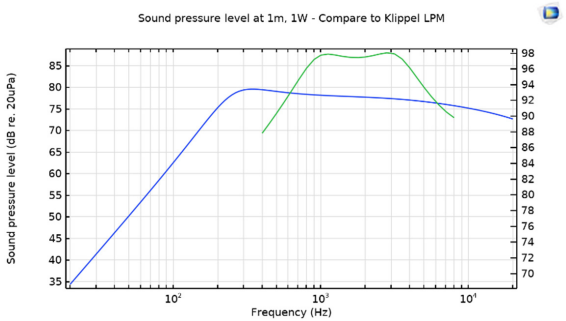
A circular loudspeaker for larger IoT devices such as washing machines, dish washers etc. and a rectangular micro loudspeaker suited for Smart Glasses or similar very small IoT driven devices.
The circular Ø31mm loudspeaker with its preferred tuned versus untuned frequency response is shown in *Fig. 5*. (green curve is after tuning for best speech intelligibility, blue curve is before).

Parameters tuned in the driver and built-in are:

- Moving mass reduced by 45%
- Mechanical compliance reduced by a factor of 5
- Front resonator established
- Closed back cavity of 10ml

Fig. 5



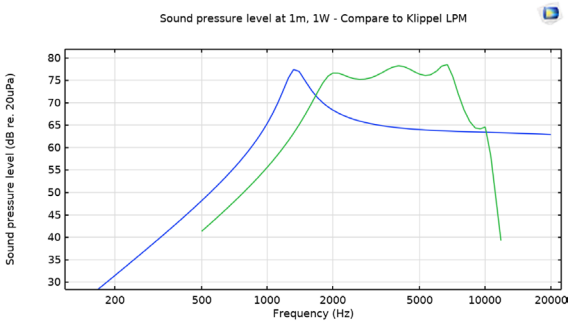


Correspondingly, the rod-shaped 6mmx12mm loudspeaker with its tuned versus untuned frequency response is shown in *Fig. 6*. (green curve is after tuning for best speech intelligibility, blue curve is before).

Parameters tuned in the driver and built-in are:

- Moving mass reduced by 10%
- Mechanical compliance reduced by 50%
- Front resonator established at the short side of loudspeaker (extension of length)
- Closed back cavity of 1ml

Fig. 6



Conclusion

Iot combined with AI opens for a lot of advantages making life on daily basis easier and in general giving improved quality of life.

Not only in 2-way audio communication with the numerous devices we have in our homes for convenience, but also for more serious purposes, for example in medicine, where e.g. blood pressure monitors in smart watches in real-time could be genuine lifesavers.

In this white paper it was shown, that a loudspeaker system can be optimized for a specific application. By focusing on the performance on the frequency range most crucial to speech intelligibility, the sensitivity of the system could be increased.

In case of questions or a wish for a dialog, the following e-mail address and mobile phone number can be used:

Morten Kjeldsen, B.Sc.E.E.
Specialist, Micro Acoustics
Advanced Development, Ole Wolff A/S

e-mail: mk@owolff.com
Mobile: +45 26 32 28 79