INFORMATICA MUSICALE

UNIVERSITA' DEGLI STUDI DI CATANIA
DIPARTIMENTO DI MATEMATICA E INFORMATICA
LAUREA TRIENNALE IN INFORMATICA
A.A. 2018/19
Prof. Filippo L.M. Milotta

PROJECT ID: 0A

PROJECT TITLE: SOUND RECOVER: NON-LINEAR FREQUENCY-COMPRESSION SCHEME

AUTHOR: Napoli Spatafora Maria Ausilia

Indice

1. Main Goals

- ***Show a medical application of the sound***

"O you men who think or say that I am malevolent, stubborn or misanthropic, how greatly do you wrong me, you do not know the secret causes of my seeming, from childhood my heart and mind were disposed to the gentle feelings of good will, I was even ever eager to accomplish great deeds, but reflect now that for six years I have been a hopeless case, aggravated by senseless physicians, cheated year after year in the hope of improvement, finally compelled to face the prospect of a lasting malady (whose cure will take years or, perhaps, be impossible), born with an ardent and lively temperament, even susceptible to the diversions of society, I was compelled early to isolate myself, to live in loneliness, when I at times tried to forget all this, O how harshly was I repulsed by he doubly sad experience of my bad hearing, and yet it was impossible for me to say to men speak louder, shout, for I am deaf. Ah how could I possibly admit such an infirmity in the one sense which should have been more perfect in me than in others, a sense which I once possessed in highest perfection, a perfection such as few surely in my profession enjoy or have enjoyed - O I cannot do it!"

Ludwig Van Beethoven, Heiligenstadt Testament (1802)

Social isolation — especially as we age — increases the risk of numerous mental and physical health (depression, heart disease, abnormal immune systems, dementia and Alzheimer's disease). According to several studies, social isolation is associated with a "reduction in lifespan similar to that caused by smoking 15 cigarettes a day." One big reason people become socially isolated is because of hearing loss. Often, as hearing becomes challenging, people avoid social, business situations where interaction is key — and instead choose to withdraw and isolate themselves.

- ***Outline background information from the field of audiology***

The frequencies (octave and semi-octave) checked during tonal audiometric test be part of range 125-8000 Hz and are classified in:

  - Low: 125 Hz, 250 Hz, 500 Hz;
  - Mid: 750 Hz, 1000 Hz, 2000 Hz;
  - High: 3000 Hz, 4000 Hz, 6000 Hz, 8000 Hz.

The loss of hearing could take place in one or more these macro-bands. Usually, high frequencies are damaged by hearing loss (figure 1).

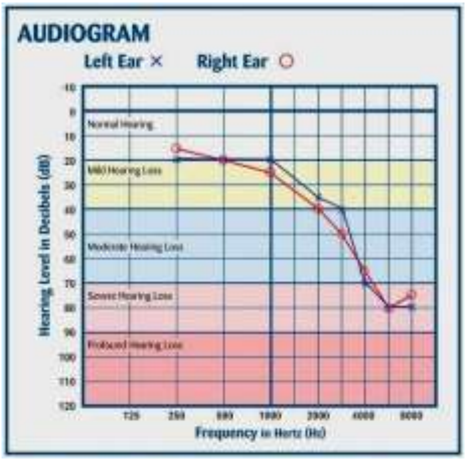

Figure 1 Audiogram with slope on high frequencies

- ***Learn the importance of perceptual bandwidth and how frequency compression extends it***

It is commonly accepted that the normal human auditory bandwidth encompasses the range of frequencies from 20 to 20,000 Hz but hearing instrument bandwidth is part of range 100-8,000/10,000 Hz. However, the audibility of a sound such as a pure tone depends not only on its frequency but also on its level. The hearing instrument should amplify the sound in order to balance the hearing loss, but the maximum gain couldn't be enough on the high frequencies and so people don't sense the sound. The compression shifts the high-frequency information to lower-frequency region that is less damaged.

- ***SoundRecover and SoundRecover 2: commercial non-linear frequency-compression algorithms of Phonak***

  Phonak led the way in modern frequency lowering technology with the introduction of SoundRecover in 2008. Since then, extensive worldwide field studies with adults and children have found increased detection, distinction and recognition of high frequency sounds, better speech understanding and significant improvement in intonation and overall voice quality for users. For those with more extreme severe-to-profound losses, however, including left corner audiograms and ski-slope losses, the benefits have been limited due to the restricted audible bandwidth in which frequency compression could be applied. The new SoundRecover2 algorithm aims to restore the audibility of relevant high frequency sounds while leaving intact the low frequency structures important for good sound quality.

2. References

[1] Phonak's Posters:

- 028-0952-02/V1.00/2013-06/cu Printed in XXXX © Phonak AG All rights reserved
- 028-1512-02/V1.00/2016-01/ © Phonak AG All rights reserved
- 028-1528-02/V1.00/2016-01/ © Phonak AG All rights reserved

[2] H. McDermott, D. Baldwin, M. Nyffeler: The importance of perceptual bandwidth and how frequency compression extends it. The Hearing Journal 2010, May

[3] Y. Mao, J. Yang, E. Hahn, L. Xu: Auditory perceptual efficacy of nonlinear frequency compression used in hearing aids: A review. Journal of Otology 2017, December

[4] U. Ambrosetti, F. Di Berardino, L. Del Bo. Audiologia Protesica. Edizioni Minerva Medica. Prima edizione, 2014

In order to write this paper, I use technical Phonak's information [1] to know every aspect of the two algorithms. I find out about hearing instrument parameters from references [2] and [4]. The paper [3] is needed for global look on clinical evidence. Finally, I add to this composition from my personal knowledge, past studies and experience.

3. Didactic Topics

A. Perceptual Importance of High Frequencies

For people with hearing loss, it is crucial that they can perceive and discriminate high-frequency sounds easily and accurately. There are three main areas where high-frequency sounds are important:

- **Speech intelligibility**: Many sounds or phonemes that contribute significantly to speech intelligibility include mainly high-frequency components (figure 2). Vowels are dominated by greater energy in the low frequencies whereas voiceless fricatives are dominated by greater energy in the high frequencies. For example, the phoneme /s/ is used in the English language to identify plurals and more; other phonemes are /sh/ and /f/. (They are discriminative and important for Italian language, too). Depending upon the age and gender of the speaker this phoneme has a spectral peak between 4 and 10 kHz. For men, the energy maximum is usually between 4 and 6 kHz and for women it is often between 7 and 10 kHz. In every language there are many such speech sounds, which can be distinguished only if the high-frequency components of the signal are audible. Moreover, the perception of this high-frequency range is particularly important for children in language acquisition, as it provides the potential to understand language and learn to reproduce it correctly.

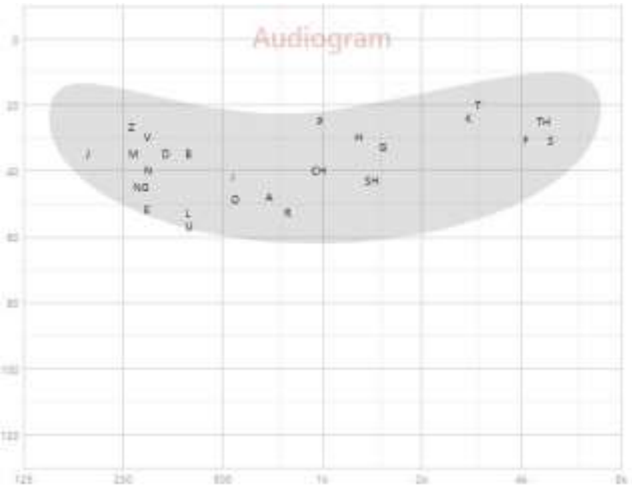

Figure 2 Audiogram with frequency content of the speech sound in English language.

- **Speech understanding in noise**: Speech understanding is particularly difficult when a listener is attempting to understand speech in a noisy environment. The high frequency part of the speech signal is particularly important because, unlike the low frequency part, it is less susceptible to being masked by the relatively intense low-frequency components of many common types of noise. Therefore, it is particularly significant in such acoustic environments that the high-frequency phonemes are audible and distinguishable.
- **Localization**: The perception of these high-frequency characteristics gives valuable information for the identification and localization of sound sources. It is significant that this high frequency information is needed from both ears.

B. Hearing Instrument Bandwidth

In general, bandwidth is defined in terms of the range of frequencies that can be carried by a communication channel. Widening the bandwidth means increasing the frequency range, thereby enabling more information to be delivered through that channel.

In the past, the high-frequency bandwidth limit of analog hearing aids usually resulted mainly from their electroacoustic performance. With high-powered instruments, it was often difficult to obtain adequate sound output levels at frequencies above about 4000 Hz. In recent years, however, receiver technology has improved to the extent that bandwidth limitations are imposed instead by other factors.

In all digital hearing instruments there is an absolute limit on bandwidth resulting directly from the sampling process. Sampling is required to convert the sound signals at the input of the hearing instrument (HI) into a stream of separate digital representations. The sampling rate must be high enough to ensure that the continuously varying acoustic signal is represented in the digital processor with adequate fidelity. The selection of sampling rate is based on a fundamental principle of digital signal processing that states that the highest frequency that can be represented adequately after sampling is slightly less than half the sampling rate. For normal-hearing listeners, the upper frequency limit is generally assumed to be 20,000 Hz, so the required sampling rate is more than 40,000 Hz.

Unfortunately, the use of relatively high sampling rates can have undesirable side-effects. The digital signal processor inside any modern hearing aid is programmed to modify the sound signals at a rate equal or proportional to the sampling rate. One practical effect of this relationship is that higher sampling rates require higher power consumption, which reduces battery life. Designers of digital HIs are faced with a difficult trade-off: widening the acoustic bandwidth of the device means shortening the battery lifetime. Consequently, it is common for the sampling rate in hearing instruments to be approximately 20,000 Hz. This choice means that the upper limit of the bandwidth in terms of sound produced by the HI must be about 10,000 Hz. In some devices, the sampling rate may be as low as 16,000 Hz, resulting in an acoustic bandwidth of less than 8000 Hz.

C. Perceptual Bandwidth

A more useful practical definition of bandwidth would specify the range of frequencies at which tones can easily be made comfortably loud. This is illustrated in figure 3, which shows the level in dB SPL (vertical axis) required to produce the same loudness for tones heard across a wide range of frequencies (solid curve).

In this graph, a tone at 1000 Hz is shown as having a level of 60 dB SPL, which would be comfortably loud for an average listener with normal hearing. To maintain the same loudness as the frequency changes, the level of the tone would need to be adjusted by less than about 10 dB across a frequency range from approximately 80 Hz up to nearly 20,000 Hz. At frequencies below 80 Hz, the level would need to be increased for the same perceived loudness. For example, a tone at 20 Hz would have to be presented at about 100 dB SPL to be heard as equal in loudness to the 1000-Hz tone at 60 dB SPL. This demonstrates that audible bandwidth depends strongly on the sound level, even for normal-hearing listeners. Generally, the effective perceptual bandwidth can be increased by raising the level of sounds. Figure 3 also shows equal-loudness data for older listeners who were assumed to have normal hearing (dashed curve). Although those listeners had no signs of ear disease, their average sensitivity to high-frequency tones was much poorer than that of younger listeners (solid curve). At 10,000 Hz, for instance, the level difference between these two groups was almost 20 dB for the same loudness. Even larger differences are evident at higher frequencies. In contrast, the listeners' age had no effect on the equal-loudness data for frequencies below about 2000 Hz. These measurements are consistent with findings from many research studies that "high-frequency hearing sensitivity tends to decline as a person ages, even in

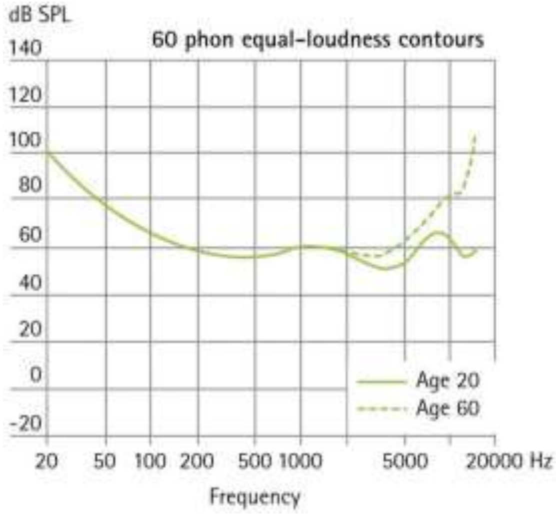

Figure 3 Equal loudness contours for young (solid curve) and older (dashed curve) listeners with normal hearing. The curves show a moderate level (vertical axis) that is perceived as equally loud across frequency (horizontal axis).

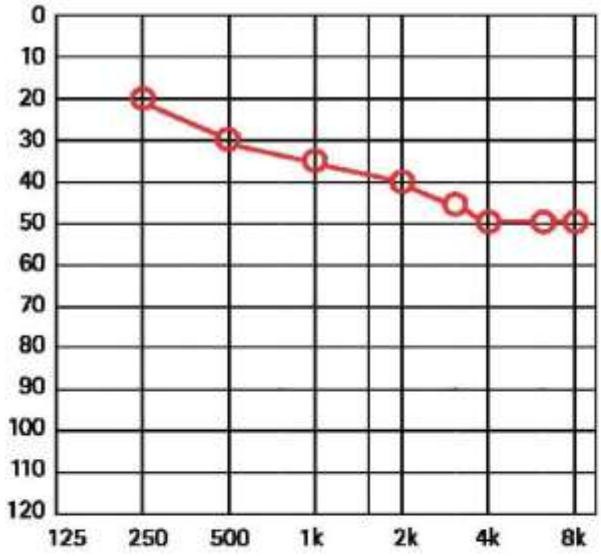

Figure 4 Audiogram for a typical sloping mild to moderate hearing loss.

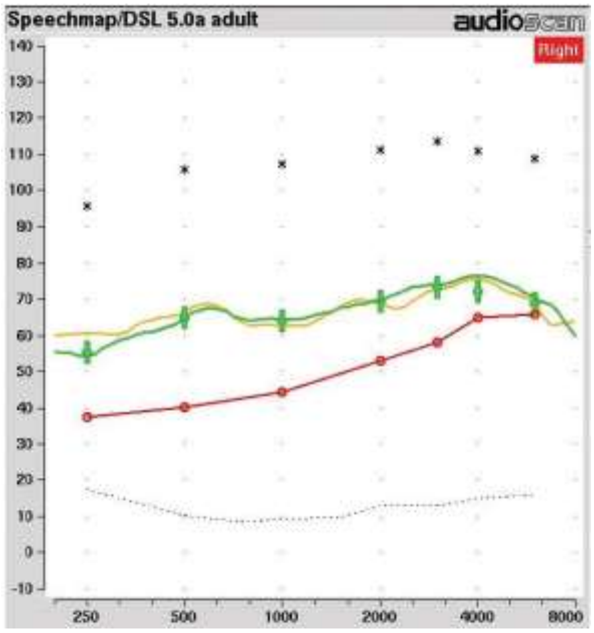

Figure 5 The results of fitting two HIs for the audiogram (red curve) shown in Figure 3. The Phonak HI (green curve) had SoundRecover disabled. The yellow curve shows comparable results from a different manufacturer's HI reported to provide extended bandwidth.

the absence of any specific pathology". Thus, when the bandwidth of hearing becomes narrower as a result of age-related hearing loss, the usual reason is a change in sensitivity at high rather than low frequencies. Furthermore, a similar type of bandwidth reduction can result from many common causes of hearing impairment, including exposure to excessive noise, various diseases, side-effects of ototoxic drugs, and other etiologies. How do these considerations apply to a person who uses a hearing instrument? The answer is complicated by the presence of two interacting factors. First, there is the particular configuration of each HI user's hearing impairment, as characterized by the audiogram. The second factor is the effective bandwidth of the HI, which depends on its gain and maximum output level, parameters that inevitably vary as a function of frequency. In addition, certain sound-processing techniques such as frequency lowering can affect the perceptual bandwidth.

A conventional audiogram records a person's threshold of hearing at a number of discrete frequencies. The lowest frequency is usually 125 or 250 Hz, while the highest frequency may be up to 8000 Hz.

Figure 4 shows a typical sloping hearing loss of mild to moderate severity, with thresholds of 50 dB HL at and above 4000 Hz. This audiogram is shown as the red curve in figure 5 that is the Speechmap-DSL plot[1]. Also shown in the latter figure is a fitting of a Phonak HI with wide bandwidth, with the proprietary frequency-shifting algorithm SoundRecover disabled (green curve).

As Figure 5 demonstrates, the Phonak HI without SoundRecover was able to provide useful audibility of the test signal (speech at an average level of 65 dB SPL) up to at least 6000 Hz. The yellow curve in the same figure shows, for comparison, results for another premium-level HI that claims extended bandwidth to 10,000 Hz. The measurements plotted in Figure 5 demonstrate clearly that these two instruments result in almost identical perceptual bandwidths when fitted to suit a common audiogram configuration. However, neither HI would provide useful audibility for frequencies higher than about 6000 Hz, even though the maximum available gain for those frequencies was selected in each device. It is noteworthy that this

[1] Speechmap is commercial fitting environment that allows to show which frequential region should be amplify and with which gain; It calibrates it through speech signal.
DSL is prescriptive algorithm.

restriction on audibility above 6000 Hz is present even for a mild to moderate hearing loss with thresholds in this region of
only 50 dB HL. This limitation on perceptual bandwidth is a consequence of particular characteristics of both the audiogram and the technical performance of the HIs when fitted for that audiogram.
What can be done to overcome this limitation?

D. Technical Description of SoundRecover and SoundRecover2

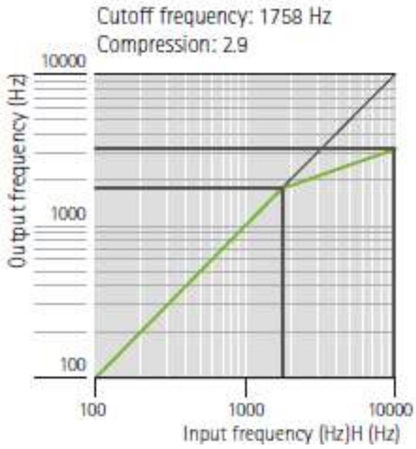

Figure 6 The response curve of nonlinear frequency compression, SoundRecover, for a cut-off frequency of 1758 Hz with a compression ratio of 2.9.

The objective of SoundRecover is to restore the audibility for high frequency inputs up to approximately 10 kHz. This unique approach is designed to compress the signal above a specified and adjustable cut-off frequency. The amount of compression applied to this frequency band is specified by the compression ratio. All frequencies below the defined cut off frequency starting point remain unchanged, preserving the quality of sounds delivered to the hearing aid user.
Frequencies which are further away from the cutoff frequency (higher frequency direction) are shifted to a greater extent to lower frequencies than a frequency which is closer to the cut off frequency. For example, for a given compression setting, the maximum energy of a female /s/ (usually around 9 kHz) will be shifted further towards the cut-off frequency than the maximum energy of a male /s/ (5 kHz) although the correct frequency order is still maintained.
The compression works without any delay, therefore there are no time constants required or artifacts heard. Figure 6 shows the response curve of the frequency transmission of a non-linear frequency compression. The cut-off frequency in this example is at 1758 Hz and the compression ratio at 2.9:1.
Thus, it is possible to calculate the maximal input frequency $f_{IN,max}$ which depends on the sampling rate of the maximal output frequency $f_{OUT,max}$ using the following formula:

$$f_{OUT,max} = f_{IN,max}^{\frac{1}{CR}} * f_{cutoff}^{1-\frac{1}{CR}}$$

This realization of SoundRecover in all Phonak hearing aids allows for an individual setting of the cut-off frequency dependent on the hearing loss between 1.5 kHz and 6 kHz. The compression ratio is automatically adjusted to a value between 1.5:1 and 4:1 according to the selected cut-off frequency.
The cut-off frequency and compression ratio values are combined within Phonak fitting software to define one SoundRecover setting which clinicians are able to adjust, in order to fine tune the settings.
The SoundRecover setting can be adjusted to strengthen or weaken how SoundRecover affects sounds. The lower the cut-off frequency and higher the compression ratio, the stronger SoundRecover affects the sounds. To weaken how SoundRecover settings, a higher cut-off frequency and a lower compression ratio should be set.

I. SoundRecover VS SoundRecover2

In order to widen the reach of SoundRecover in 2016, the new SoundRecover2 algorithm is designed to allow operation with lower cut-off frequencies and weaker compression ratios, thereby extending the benefits of frequency compression to a broader audience of children and adults.

The significant difference with SoundRecover2 is that the extent of frequency lowering, i.e. the area of protection and the starting point of compression, is not fixed but is instead set adaptively as a function of the input signal. This adaptive nature is realized by the use of two cut-off frequencies, of which only one is active at any moment in time. Based on the momentary energy distribution in the input signal, the system determines instantaneously which one of the two cut-off frequencies is applied. Thus, the functional principle of SoundRecover2 is similar to that in SoundRecover; with the increased sophistication that it now switches automatically between two possible starting points of compression, respectively between a "lower" and an "upper" cut-off frequency. As in SoundRecover, frequency lowering is always carried out with a predefined constant compression ratio regardless of which cut-off frequency is momentarily active.

Hence, SoundRecover2 instantaneously maps input components to the output depending on their energy content. This adaptive frequency lowering processing is accomplished by simply recognizing the different energy distributions of tonal and noise-like structures of the input signal. In case of more low frequency content, frequency compression takes place with the upper cut-off frequency in order to "protect" the low frequency sounds from being compressed. In case of more high frequency content, frequency compression takes place with the lower cut-off frequency to restore audibility of the high frequency sounds. When applied to speech signals, this strategy leaves vowels intact while allowing compression of important high frequency information in fricatives down to sufficiently low output frequencies.

II. Outcomes

Figure 7 presents the spectrograms of the sample sentence "my name is asa" (a) without frequency lowering, (b) with SoundRecover and (c) with SoundRecover2 processing.

Figure 7(a) shows pronounced formant structures up to 5.5 kHz at 0.2 seconds to 0.5 seconds and two high frequency /s/ phonemes at 1.2 seconds and 1.9 seconds.

In figure 7(b), showing SoundRecover with a cut-off frequency of 1500 Hz and compression ratio of 2.1, the audible bandwidth extends up to approximately 4000 Hz. The /s/ phonemes at 1.2 and 1.9 seconds are compressed down into a frequency area between 2.5 and 4 kHz. Note that the spectral structures above the cut-off frequency of 1500 Hz at the beginning of the sentence are not fully preserved at this maximum setting.

In figure 7(c), showing SoundRecover2 with a lower cut-off frequency of 1479 Hz, upper cut-off frequency of 3600 Hz, and compression ratio of 1.4, the audible bandwidth extends up to approximately 4000 Hz as well. Note the preservation of the spectral structures up to the upper cut-off of 3600 Hz at the beginning of the sentence at 0.2 to 0.5 seconds, and the remapping of the two significant high frequency /s/ phonemes at 1.2 and 1.9 seconds down into a frequency area as low as between 2000 and 3000 Hz.

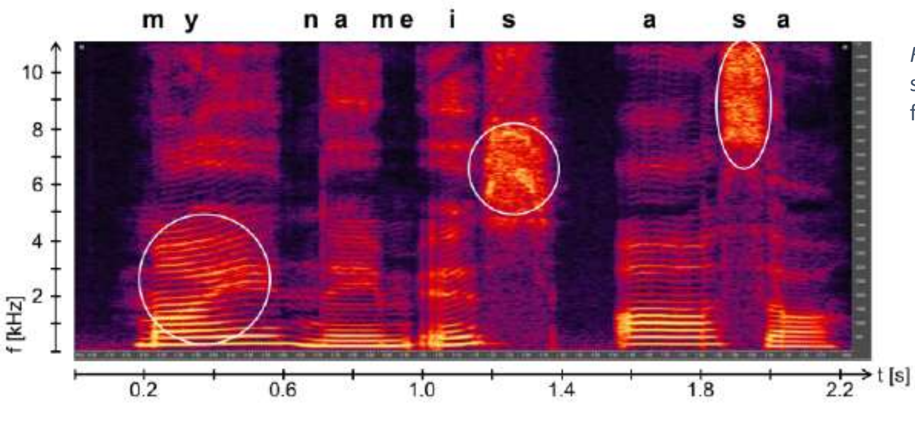

Figure 7(a) Spectrogram of the sample sentence "my name is asa" without frequency lowering.

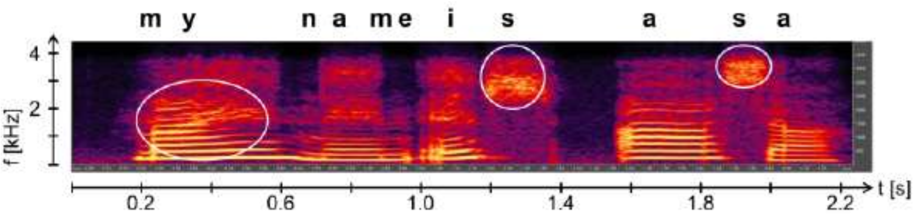

Figure 7(b) Spectrogram of the sample sentence “my name is asa” with SoundRecover (cut-off: 1500 Hz, compression ratio: 2.1).

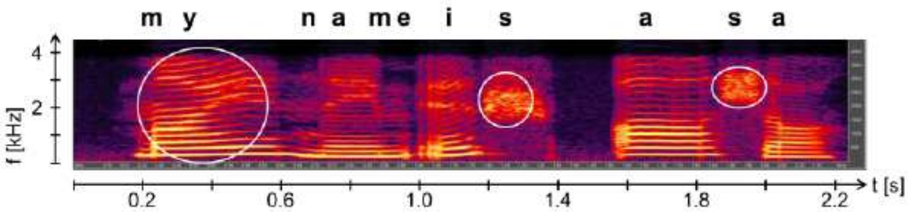

Figure 7(c) Spectrogram of the sample sentence “my name is asa” with SoundRecover2 (lower cut-off: 1479 Hz, upper cut-off: 3600 Hz, compression ratio: 1.4).

E. Clinical Evidence

The paper[3] is a review of the effects of NLFC (Non-Linear Frequency Compression) algorithm on speech and music perception and sound quality appraisal.

For vowel perception, it seems that the benefits provided by NLFC are limited, which are probably related to the parameter settings of the compression. For consonant perception, several studies have shown that NLFC provides improved perception of high frequency consonants such as /s/ and /z/. However, a few other studies have demonstrated negative results in consonant perception.

In terms of sentence recognition, persistent use of NLFC might provide improved performance. Compared to the conventional processing, NLFC does not alter the speech sound quality appraisal and music perception as long as the compression setting is not too aggressive.

The relevant factors tath regard to NLFC settings are time-course of acclimatization, listener characteristics, and perceptual tasks.

Although the literature shows mixed results on the perceptual efficacy of NLFC, this technique improved certain aspects of speech understanding in certain hearing-impaired listeners.

F. Example of fitting and fine tuning of SoundRecover2

Phonak Target 5.1 is the software for fitting of SoundRecover2. Figure 8(a) shows sample output curve for SoundRecover2 abled; whereas figure 8(b) shows the control panel of this algorithm.

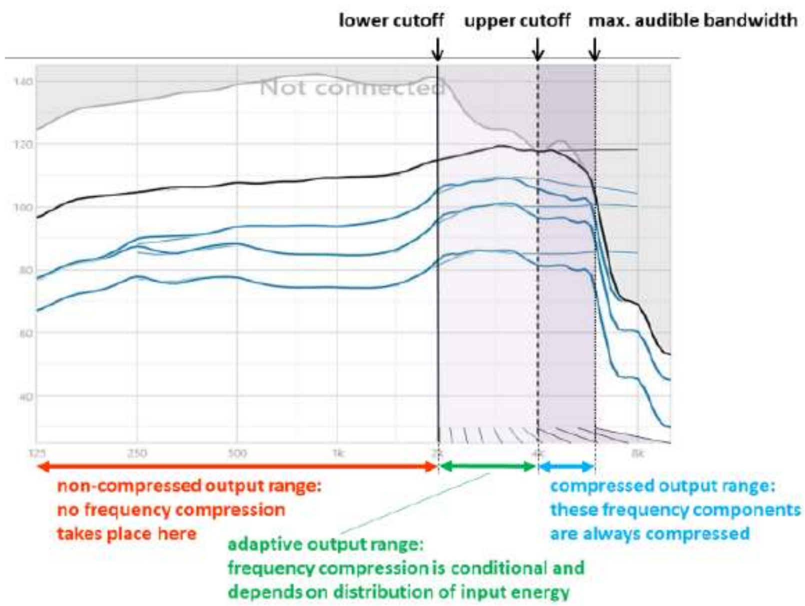

Figure 8(a) Sample output curve for SoundRecover2. Depending on the energy distribution of the input signal, the frequency compression starts either at the lower or at the upper cut-off frequency.

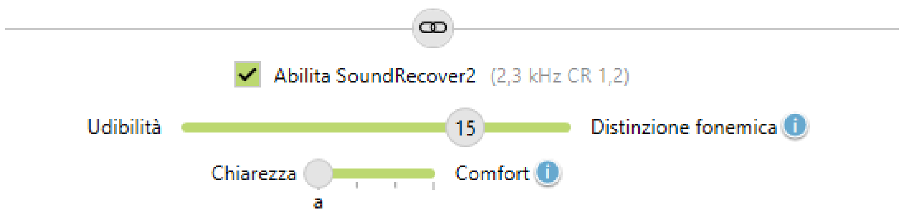

Figure 8(b) The control panel for SoundRecover2 in Phonak Target 5.1.