Loudness Using a Threshold Correction Factor

[1] Colin NOVAK; [2] Helen ULE; [1] Robert GAPSAR

[1] University of Windsor, Canada

[2] Akoustik Engineering Limited, Canada

ABSTRACT

Experimental studies have shown that for short gaps (2 to 5 ms) loudness and threshold are higher than for uninterrupted noise. Other studies have also shown that the present integration models for loudness do not adequately account for short duration phenomena. Studies have instead shown that the multiple look approach is the applicable method for these short-term circumstances. However, present technologies (i.e. FFT) are not adequate to deal with short duration sounds across the entire frequency spectra. A compromised approach is taken here to account for the threshold phenomena in the presence of gaps while using an integration model. This approach is referred to as a threshold correction factor.

Keywords: Loudness, Multiple Looks, Threshold Correction Factor I-INCE Classification of Subjects Number(s): 08.4, 63.1

1. INTRODUCTION

The ability to hear and discriminate sounds within our environment is a critical sensory mechanism that enables humans to communicate and to react to auditory stimuli. Communication plays a critical role in the maintaining of social relationships as well as the ability to hear, react and analyse sounds within the environment. As such, it can be said that hearing ability is paramount to an individual's ability to understand its surroundings and is a significant contribution to an overall quality of life.

For engineering applications, the goal is often to find the source mechanisms of a sound in the hope of either attenuating the noise or to improve its quality from a perceptional perspective. A fundamental psychoacoustic metric used to achieve this is loudness, a model for which many other metrics rely on for the basis of their calculation algorithms. Loudness is said to be a metric, which closely matches the perceived intensity of a sound.

Complex models are often needed for the estimation of loudness for real sounds. These models are divided into two fundamental types depending on the nature of the sound. These include loudness models for steady sounds, which do not change with time and more complex unsteady models used for the calculation of loudness of unsteady sounds. Several methods to calculate both categories of sounds can be found in the literature

The focus of this work is on the more complex determination of loudness for unsteady sounds. A generally accepted approach to the calculation of unsteady loudness use the method temporal or long term integration where the intensity of the unsteady signals are integrated over time. Psychoacoustic studies have found this method to be acceptable for sounds which do not change significantly over durations of approximately 100 ms or longer. However, it has been known for many years, as given by Exner (5), that the absolute thresholds of sounds are strongly dependent upon duration and frequency. Experiments have shown that the perceived absolute thresholds are increased for sounds, which are short in duration or in the presence of gaps. The use of temporal integration methods cannot account for the short duration data, and therefore, cannot always be considered to be good predictors of loudness for all signal types.

An alternative model of the human auditory system is called the multiple look approach. The premise of the multiple look theory is that the auditory system takes many samples or "looks" of a stimulus and stores them into memory for later processing. The specific processing performed is dependent on the makeup of the signal contained in the successive look. For the case where gaps or burst of information is present, the short term looks are processed immediately as an auditory perception. If the signal is steadier in nature, the looks are instead stored for a longer period and then processed as an integrated signal over time. This is very similar to the concept of a leaky integrator model.

A negative aspect of the implementation of the multiple look approach is that it cannot be adequately implemented using present day technology over the full auditory frequency range. This is due to limitations of present day digital signal processing techniques. That is, the DSP cannot adequately sample short duration signals with a low enough frequency resolution to cover the full auditory range.

The goal of this work was to develop an alternative model, which will support the documented experiments that support the multiple look theory for unsteady sounds, while at the same time is viable to implement for sounds across the human perceptible frequency range without the restrictions of present day signal processing capabilities. The proposed model is a hybrid multiple look approach which uses level correction factors in conjunction with a temporal integration method in order to adequately represent the perceived loudness levels of sounds that have gaps in the stimulus signal.

2. Background

2.1 Calculation of Unsteady Loudness

Loudness calculation methods restricted for the use on steady sounds are relatively simple and are easy to correlate to the results of auditory experiments. This is partially due to the fact that the designs of experiments which focus on steady sounds are fairly simple to implement, often using pure tones, and have good repeatability. The development of these steady models has given much insight to the present understanding of hearing perception and to the workings of the auditory system. Most real sounds encountered in daily life though are unsteady in nature. Examples include traffic noise, machinery noise on a factory floor or speech. For these sounds, an alternative loudness method is necessary to include the temporal, or time effects, of the human auditory system. These effects can be very complex and add a significant degree of complexity to the process of determining the loudness for these time varying sounds. The multiple look model developed in this work is a hybrid approach that uses and existing temporal integration loudness model; in this case the Glasberg and Moore time varying loudness (TVL) model (7). The ideas and approach for temporal integration and the Glasberg and Moore TVL model are described below.

2.1.1 Loudness Using Temporal Integration

Munsun originally poised the concept of a temporal integration period in 1947 to describe the perceived increase in loudness with increasing signal duration (11). It has been generally accepted that absolute thresholds of hearing are strongly dependant on the duration of the stimulus signal, at least for sounds lasting between 200 to 300 ms and that the auditory system is able to summate an internal representation of a signal over this period (12)(9). In fact, according to Zwislocki, it is usually taken that the sound intensity necessary for detection increases as the duration of the sound decreases (21).

Auditory temporal integration is described as a simple accumulation of acoustic stimuli over time, or energy integration, which is used for the detection or discrimination of sounds. According to Plomp & Bouman, this assumption has been based on observations that the absolute threshold for detecting sounds, usually described as a sound pressure level, decreases with increasing duration of the sound (18). This increase in performance is been modelled as a simple accumulation of intensity over time. Green described this behaviour in 1960 for the case of absolute threshold as the auditory system's energy integrator (8)(2). Penner argued against the theory of the auditory system integrating energy over time. He suggested that neural activity is instead combined over time as opposed to acoustic energy (17).

In addition to the general lack of consensus as to exactly what is combined over time, disagreement also exists as the how it is combined (10). Most agree that the auditory system does not in actuality integrate the acoustic stimuli in the same sense as a mathematical integration operation. Despite this, most existing time varying loudness models use what is referred to as a "leaky integrator" approach. This is synonymous to the accumulated weight of water being poured into a cup over time; only the cup has holes in it at various heights, which allow the water to leak out.

2.1.2 Glasberg and Moore's Time Varying Model

The publication of Glasberg and Moore's 2002 paper (7) described an unsteady loundess model, which was somewhat of an extension of their stationary loudness procedure. The output of the TVL model was both a both short-term and long-term loudness level. The authors described the usefulness of having both values using the example of speech as a noise source. They related short-term loudness as being useful for the measure of the intensity of a speech syllable. Long-term loudness on the other hand would be useful for the measure of the intensity of a much longer speech signal such as a sentence (7).

In order for the TVL model to accommodate the full audible frequency range and not lose resolution at higher frequencies for short duration signals, the model's use of six parallel FFTs to calculate spectral information over six bandwidths, calculated over decreasing lengths of time, for obtaining spectral information in increasing frequency ranges. The ranges of the bandwidths are 20 to 80 Hz, 80 to 500 Hz, 500 to 1250 Hz, 1250 to 2540 Hz, 2540 to 4050 Hz, and 4050 to 15000 Hz each having segment durations of 64, 32, 16, 8, 4, and 2 ms, respectively. The excitation pattern and instantaneous loudness levels are then calculated in the same fashion as their stationary model. The short term loudness is calculated by temporally averaging the instantaneous levels, thus providing a running average for the signal. The long term loudness is subsequently calculated by temporal averaging of the short-term loudness. While this model has shown to provide good correlation to the latest 2003 equal loudness contours (1), the use of temporal averaging has not shown to provide adequate prediction for noise bursts or sounds in the presence of gaps (20).

2.2 Research Supporting Multiple Look Approach for Loudness

Many listening experiments though have shown that the auditory system does not use a process that is wholly synonymous to temporal integration. For example, it is unlikely that the auditory system would integrate over time for a task as simple as the detection of a pure tone presented in quiet. It has been suggested (10) that it may be more appropriate to consider the auditory process as a combination of information from multiple independent "looks".

The concept of the multiple looks theory is that the auditory system takes sequential samples, or looks, of the sound information and either immediately processes the information as a perception or stores the information for future processing. The decision to process or store the information is dependent on the nature of the stimuli. If for example the sound has large sudden increases or decreases in amplitude or gaps in the flow of the stimuli, then the sounds are processed immediately as independent samples. This is applicable for sounds which change over short durations from 1 ms to approximately 5 to 10 ms. If, on the other hand the sound is more continuous over a much longer time period then the "looks" are instead thought to be integrated over time.

Viemeister and Wakefield published the first real evidence of the multiple look theory in 1991. This work demonstrated the validity of the theory through two very important experiments (20). Their first experiment measured the detectability in quiet of two very short pulse signals compared to a single pulse. The results averaged over all the tests subjects demonstrated the threshold of detectability of the two pulse pair with increasing separation distance compared to a single pulse stimulus. It was demonstrated that for a separation of 1 ms that the level of detectability was 4 dB lower than for a single pulse. In other words the single pulse sound would need to be 4 dB greater to have the same perceived loudness as the two equal amplitude pulse pair. For separations larger than about 5 ms, the detectabilities are averaged to have a level of approximately 1.6 dB lower than those for a single pulse. The significant point was that the detectibility levels increase with separations larger than 1 ms but then stop increasing once the separation has reach 5 ms. These results are inconsistent with those obtained using a long time constant leaking integrator.

Pedersen conducted several studies to investigate the temporal processing of the auditory system, which resulted in data that supports the multiple look theory. To study how listeners temporally integrate sounds to discriminate their loudness, Pedersen published several 2006 papers that focused on how listeners apply weighting to various temporal segments of a sound when judging loudness (13)(14)(15). The outcome was a temporal weighting curve showing the importance of different temporal locations of the sound. It was shown that listeners emphasize onsets and offsets in their temporal weighting of a sound, which showed that loudness integration is not a simple process as assumed in many loudness models. It was also demonstrated that listeners changed their pattern of temporal weighting if they are provided with feedback or a hint of the signal. This reinforces the work performed by Moore (10). Also, a change in the spectral content in the middle of a sound, demonstrating the onset of a new event, is shown to be weighted more heavily. Thus, it was shown that listeners pay attention to salient events within sounds, phenomena not possible with simple integration but only supported by a multiple look approach. Pedersen concluded that temporal variation is made available in the sensory system to allow for overall judgement of the properties of sound, such as loudness, and, “this information is weighted and analyzed in complex ways, which is not adequately described as a simple summation process,” but can be explained by the multiple look theory (16).

Much work has been done over the past 80 years or so in the development of loudness models. Progress has also been accomplished in the initial development of time varying loudness models using time integration techniques. While this approach has shown good results for some unsteady sounds, others have shown that integration is not the likely mechanism employed by the auditory system at all times. A more likely approach is some form of the multiple look approach.

3. Approach

Conventional approaches for calculating unsteady loudness using a multiple look model have been proposed. The caveat to these are that the procedure requires that a Fast Fourier Transform be applied to very short segments of the stimulus having lengths of approximately 1 or 2 ms. This is not possible given the limitation in frequency resolution that this would impose on the processed signal.

While the development of a true multiple look approach is desired, for this work, a calculation method that is alternative to the present loudness models, but still retains both the spirit and ability to account for auditory phenomenon, which the present models are incapable of was developed. This hybrid approach is one which samples the stimulus signals as 1 ms looks and processes the information to account for known auditory characteristics. It was further decided to focus on the specific characteristic of gap detection, as this is one phenomenon, which has been documented experimentally but has not been demonstrated to be included in any other loudness model. The following is a description of the methodology of the proposed model.

3.1 Proposed Model

The process begins with the input of a single channel of stimuli which represents a binaural diotic signal presented to the outer ear. The signal is sampled as a 16-bit resolution WAV file with a 32 kHz sampling rate. This will result in a file containing 32 samples for every 1 ms of stimulus data. The length for each look was chosen to be 1 ms. Studies have reported this to be the minimum length for audibility (6).

For the WAV file, each of the samples is given as a hexadecimal number. A calibration factor taken from the acquisition system is applied to each sample. The calibration factor scales the maximum value representing the full scale deflection of the acquisition file and fits this between the full scale deflection of the WAV file, or between the values of 32 768 and -32 768 for a 16 bit file. Each of the samples is next converted from a peak pressure value to a root mean square (RMS) value. This converts all samples to all positive hexadecimal values. Finally, in order to represent the 32 samples of sound level as a single 1 ms sound, a one millisecond equivalent sound level is calculated. This is an energy mean of the noise level averaged over the 1 ms measurement period.

Once the 1 ms sound levels have been calculated, intelligent processing of the noise information can be performed. Specifically, the signal is scanned for short duration gaps spanning in length from 1 ms to 5 ms. If a gap is found, a detectibility shift is applied with amplitude dependant on the length of the gap. While this can be user defined, a gap is taken to be when there is a 25 dB drop in level from one millisecond sample to the next. The 25 dB drop for recognition of a gap is taken from Shailer’s 1983 publication on ‘*Gap Detection as a Function of Frequency, Bandwidth and Level*’ (19).

Once a drop is found, length of the gap is determined and an amplitude adjustment is made based on the threshold shifts experimentally determined by Viemeister. If the gap is determined to be longer than 10 ms, no adjustment is applied and the gap is instead defined as a drop in level and the search parameter is reset.

Once the file has been entirely searched and all detectability shifts have been applied, the file WAV file is reconstructed into its original form and loudness is calculated. This is done using the Glasberg’s TVL model.

3.2 Test Procedure

In order to test the proposed model, a test procedure was established using several recorded sounds including stationary and time varying pure tones, white noise, warble tones as well "real life" speech and mechanical sounds. Some of the sounds were altered so as to insert gaps in the signals of known location and duration to test and debug operation of the multiple look gap correction computer code.

Once the test signals were recorded, and in some cases modified with reference gaps, they were processed into 16 bit WAV files suitable for input into the multiple look gap correction and subsequently loudness models. The time varying loudness model used to perform the loudness calculation was the TVL model. As stated earlier, some of the test signals were also stationary sounds. While the TVL loudness model is purported to accurately calculate loudness for stationary sounds, these sounds were also processed for loudness using a program that follows the DIN 45631 steady loudness standard (3). Differences in the results between the two models are expected to be minimal.

4. Results

As an initial test of the multiple look gap correction model, and its adaptation to the TVL model, pure sinusoidal tones were generated at 1000 Hz and tested using the various models. The obvious thing to note is that a sinusoidal wave is a continuous sound wave with no gaps. For this study, gaps were inserted into the wave in the centre of each 10 ms segment for the first 50 ms. The next 20 ms duration had no gaps inserted. The 70 ms signal was then repeated for a total signal length of 2000 ms.

The test results for the steady sinusoidal signals without the inserted gaps are given in Table 1. The test results for the steady sinusoidal signals with the inserted gaps into the signals are given in Table 2. Listed are the sound level for the tones, the steady loudness level calculated using the method specified by the DIN 45631 standard, the calculated loudness level using the time varying Cambridge model and the loudness level using the multiple look gap correction model.

Table 1 – Loudness level for 1000 Hz sinusoidal signals without inserted gaps calculated using DIN 45631, TVL model and with multiple look gap correction model

Signal Sound Pressure (dB)	Stationary Loudness Level (Phons) from DIN 45631	Time Varying Loudness Level (Phons) from TVL Model	Time Varying Loudness Level (Phons) using Multiple Look Gap
60	55.2	58.2	58.2
65	65.4	65.5	65.5
70	72.3	71.5	71.5
73	74.9	74.2	74.2
80	81.8	79.7	79.7
85	87.2	84.8	84.7
90	93.7	90.3	90.3
94	98.2	94.7	94.7

Table 2 – Loudness level for 1000 Hz sinusoidal signals with gaps inserted calculated using DIN 45631, TVL model and with multiple look gap correction model

Signal Sound	Stationary	Time Varying	Time Varying

Pressure (dB)	Loudness Level (Phons) from DIN 45631	Loudness Level (Phons) from TVL Model	Loudness Level (Phons) using Multiple Look Gap Adjustments
60	66.5	68.4	70.3
65	73.1	74.1	75.7
70	78.3	79.2	80.7
73	80.4	81.1	82.5
80	85.7	85.9	87.2
85	90.6	90.0	91.1
90	95.7	94.7	95.9
94	99.9	97.5	98.6

Inspection of Table 1 shows very little difference between the TVL and multiple gap models. This is expected given that this sinusoidal signal has no gaps. The DIN results showed varying differences, which prompted the use of the TVL model for all subsequent tests.

Inspection of Table 2 shows a marked change in loudness level for all models. This is not surprising given that the "gapped" model does sound significantly different than the original sinusoidal wave and thus should not be expected to have the same loudness level. The important observation is that the multiple look model has a consistent 1 to 2 phon increase over the TVL model. This is expected given the predictable gap duration and spacing that was applied. The conclusion that can be drawn here using a simple sinusoidal wave is that the resulting loudness level calculation follows intended adjustments set out by the development of the gap detection model. It can further be said that this was accomplished by intelligent decisions based on the content of the 1 ms looks.

The calculated loudness results for the steady mechanical sounds are given in Table 3. Listed is the measured sound level for the sounds at which they were recorded and subsequently analysed. Also given are the steady loudness levels calculated for each signal using the method specified by the TVL model and the loudness level using the multiple look gap correction model.

Table 3 – Loudness levels for steady mechanical sounds (white noise, warble and diesel) calculated using the TVL and multiple look gap correction models.

Signal Description	Time Varying Loudness Level (Phons) from TVL Model	Time Varying Loudness Level (Phons) using Multiple Look Gap Adjustments
White Noise without gaps	86.5	86.5
White Noise with gaps	85.0	85.6
Warble	79.0	79.1
Diesel Engine	70.7	70.6

As expected, the calculated loudness levels for the white noise signal containing no gaps was the same for both the TVL model alone and with the implementation of the multiple look gap adjustment model. At a minimum this is an indication that the multiple look model did not produce erroneous results. For the white noise signal with the inserted gaps, an increase of 0.6 dB is realized by implementation of the multiple look model over the application of the TVL model alone. While an immediate application of this result cannot be given for this artificial sound, the result does provide the predicted outcome, thus showing merit to the model.

As was for the case of the white noise with the gap inserted, an increase in loudness level is also given for the warble sound, albeit a much smaller increase. Unlike the white noise of sinusoidal signals with gaps, the time trace is relatively steady and full and more absent visually of numerous gaps. The one sound sample that showed an anomaly was the result for the diesel engine. Upon closer post inspection of the time signal, it became evident that the signal while rough does not have any found gaps as defined by the multiple look gap adjustment algorithm. The anomaly in the results was the fact that the multiple look loudness level results actually shows a decrease in loudness level by 0.1 phons. While not at all significant, a decrease is unexpected. It has been determined that an inaccuracy of up to 0.1 phons can occur during the regeneration of the modified file back into the 16 bit hexadecimal WAV format. This is due to the fact that the 32 samples within each look are treated as an average during the regeneration process.

Two time varying sounds were also analysed using the TVL model and the multiple look model. The two sounds evaluated were both spoken sentences. The evaluation of unsteady loudness for speech signals is a common for the application of speech recognition and intelligibility metrics. As such, they were included in this study. The first sentence was comprised of the phrase, “Suzie sold seashells by the seashore”. This sentence was chosen for its smooth cadence and expected lack of gaps in the recorded signal. The second sentence was comprised of the phrase, “Clickity clack, the train went down the track”. This sentence was chosen specifically for its much rougher cadence and greater chance to have gaps within the recorded sentence. The calculated loudness level results for the two time varying sounds are given in Table 4.

Table 4 – Loudness levels for time varying sinusoidal sweep and speech sounds calculated using the TVL model and multiple look gap correction model.

Signal Description	Time Varying Loudness Level (Phons) from TVL Model	Time Varying Loudness Level (Phons) using Multiple Look Gap Adjustments
Spoken Sentence “Suzie”	84.0	84.0
Spoken Sentence “Train”	90.9	94.1

As stated above, the “Suzie sold seashells by the seashore” sentence is very smooth with the syllables joined together with a great degree of sibilance. This is evident by the loudness level result with both the TVL model and the multiple look gap adjustment model producing the same result. Such an outcome can be applied to the application and understanding of alternative psychoacoustic metrics, particularly those concerned with speech transmission, intelligibility and recognition. All of which are metrics for which their outcomes are related to the presence, or lack of, sibilance and alternatively harshness.

The second sentence, “Clickity clack, the train went down the track”, resulted in a noticeable increase in loudness level with application of the multiple look gap adjustments. As with the first sentence, this result shows significant implication and usefulness to speech metrics. The result also follows the perceived difference in loudness for this harder sentence when compared to the former.

Given the presented results, it has been demonstrated that the multiple look gap adjustment program does have the ability to use the looks contained within a stimulus to identify the presence of gaps within the signal. Once found, an intelligent procedure is used to determine the length of the gap and apply the appropriate adjustment factor; one which follows the published empirical data.

5. Conclusions and Recommendations

Upon review of the results, as well as recalling the stated study objectives, these are the conclusions and recommendations that have been reached.

5.1 Conclusions

1. The objective of this work was to develop a hybrid multiple look approach which uses level correction factors in conjunction with temporal integration methods in order to adequately represent the perceived loudness levels in the presence of gaps in a stimulus signal. A program

was developed which divides the input signal into 1 ms looks, checks for the presence of gaps and makes the appropriate adjustments. The adjusted file is then converted to a state such that it can be applied to a loudness integration model.
2. It was intended that the developed multiple look with gap correction abilities model would be integrated into an existing loudness model using integration theories. The model developed and presented in this work was used in conjunction with the known TVL model for time varying loudness. It should be noted that the multiple look algorithm presented in this work can immediately be used with any time varying loudness model which accepts a WAV file as an input.
3. The focus of the multiple look model developed in this work was on the hearing phenomenon of gap detection. Other stimuli and resulting hearing sensations have been identified in the literature as not being adequately addressed by the present temporal integration models. Given that the fundamental aspect of this model included the division of the signal into short duration looks for intelligent decision making and processing, it can easily be adapted to include other phenomenon such as burst signal, something which is important to account for temporal pre-masking effects.
4. It was intended that any computer code developed in this study for the multiple look model would be open and be easily adaptable to allow for modifications to the programs parameters and correction values in order to accommodate any new empirical data in the future. The code used was a public domain Ruby language which is relatively simple to understand and edit with freely available editors.
5. Finally, it was intended that any method developed should be well suited for use in other psychoacoustic metrics. Many existing metrics such as sharpness, fluctuation strength and roughness begin with the calculation of loudness. Given that the multiple look model has shown to improve present loudness models for the case of gaps being present in the input signal, inclusion of it in these other metrics would be similarly beneficial.

5.2 Recommendations

1. The model and subsequent code developed using the multiple look theory was designed to integrate seamlessly with other loudness calculation software. As part of this, the program presented here was required to reconstruct the modified information contained within the individual looks back into a 16 bit WAV file for processing of loudness by the other calculation software. It was determined that during this reconstruction process that some temporal resolution of the 1 ms information can be lost. As a result, it was determined that in some circumstances an approximate 0.1 phon inaccuracy in loudness level can result in the final calculation. While this is not a significant value, improvements can be made and are being recommended to modify the treatment of the 32 hexadecimal format samples contained within each of the looks to eliminate this shortcoming in the software.
2. As was demonstrated in the results, the perception of speech can be dependent on the content of the signal, including the presence of gaps. One of the applications where the multiple look model demonstrated particular promise was in the ability to analyse speech information. The understanding and application of evaluation models for speech recognition are ever increasing. This is particularly true given the aging demographic and increased interest in the treatment of hearing loss. Another application of the recognition of speech within automated systems such as voice activated electronics within automobiles. It is recommended that the application of multiple looks be expanded into the specific area of the recognition and treatment of speech as a stimulus.
3. The multiple look approach presented in this dissertation was specific to the application of the detection and adjustment for gaps present in the input signal presented to the ear. It was demonstrated in the literature review section that gap detection, while important, is not the only shortcoming associated with the present day loudness calculation models. This is especially true for those that rely on long term integration techniques for treatment of the temporal component of the sound. It is recommended that the model be expanded to include other distinct sound components. An example of this would be the inclusion of burst noise, an area which is important to the phenomenon of temporal pre-masking and one which is ignored by both the TVL model and the time varying method adapted by DIN as 45631-A1 (4).

ACKNOWLEDGEMENTS

The authors wish to acknowledge the technical support provided by Bruel & Kjaer, which was necessary to see the research to a successful fruition. Also acknowledged is the financial support provided by the Ontario Graduate Scholarship program.

REFERENCES

1. Charbonneau, J., Novak, C. J., & Ule, H. J. (2009). Loudness Prediction Model Comparison Using the Equal Loudness Contours. *Acoustics Week in Canada. 37(3)*, pp. 64-65. Niagara-on-the-Lake: Canadian Acoustical Association.
2. Dallos, P. J., & Olsen, W. O. (1964). Integration of Energy at Threshold with Gradual Rise-Fall Tone Pips. *Journal of the Acoustical Society of America , 36*, 743-751.
3. Deutsches Institut fur Normung. (1991). DIN 45631 - Procedure for Calculating Loudness Level and Loudness. Berlin, Germany: DIN.
4. Deutsches Institut fur Normung. (2007). DIN 45631/A1 - Calculation of Loudness Level and Loudness from the Sound Spectrum - Zwicker Method - Amendment 1: Calculation of the Loudness of Time-Variant Sound. Berlin, Germany: Deutsches Institut fur Normung.
5. Exner, S. (1876). Zur Luhre von den Gehorsempfindungen. *Pflugers Archiv , 13*, 228-253.
6. Fitzgibbons, P. J. (1983). Temporal Gap Detection in Noise as a Function of Frequency, Bandwidth, and Level. *The Journal of the Acoustical Society of America , 74* (Number 1), 67-72.
7. Glasberg, B. R., & Moore, B. C. (2002). A Model of Loudness Applicable to Time-Varying Sounds. *The Journal of the Audio Engineering Society , 50* (5), 331-342.
8. Green, D. M. (1960). Auditory Detection of a Noise Signal. *Journal of the Acoustical Society of America , 32*, 121-131.
9. Moore, B. C. (2004). *An Introduction to the Psychology of Hearing* (Fifth ed.). London: Elsevier.
10. Moore, B. C. (2003). Temporal Integration and Context Effects in Hearing. *Journal of Phonetics , 31*, 563-574.
11. Munson, W. A. (1947). The Growth of Auditory Sensaton. *Journal of the Acoustical Society of America , 19*, 584-591.
12. Oxenham, A. J., & Moore, B. C. (1994). Modeling the Additivity of Nonsimultaneous Masking. *Hearing Research , 80*, 105-118.
13. Pedersen, B. (2006). *Auditory Temporal Resolution and Integration.* Aalborg: Aalborg University.
14. Pedersen, B. (2006). Discrimination of Temporal Patterns on the Basis of Envelope and Fine-Structure Cuses. *Auditory Temproal Resolution and Integration: Stages of Analyzing Time-Varying Sounds*, (pp. 85-96). Aalborg.
15. Pedersen, B. (2006). Temporal Masking in the Auditory Identification of Envelope Patterns. *Auditory Temporal Resolution and Integration: Stages of Analyzing Timer-Varying Sounds*, (pp. 67-82). Aalborg.
16. Pedersen, B., & Ellermeir, W. (2005). Temporal Weighting in Loudness Judgements of Level-Fluctuating Sounds. *149th Meeting of the Acoustical Society of America* (pp. 13-23). Vancouver: Acoustical Society of America.
17. Penner, M. J. (1972). Neural or Energy Summation in a Poisson Counting Model. *Journal of Mathematical Psychology , 9*, 286-293.
18. Plomp, R., & Bouman, M. A. (1959). Relation Between Hearing Threshold and Duration for Tone Pulses. *Journal of the Acoustical Society of America , 31*, 749-758.
19. Shailer, M. J., & Moore, B. C. (1983). Gap Detection as a Function of Frequency, Bandwidth, and Level. *The Journal of the Acoustical Society of America , 74* (Number 2), 467-473.
20. Viemeister, N. F., & Wakefield, G. H. (1991). Temporal Integration and Multiple Looks. *The Journal of Acoustical Society of America , 90* (2), 858-865.
21. Zwislocki, J. J. (1960). Theory of Temporal Auditory Summation. *Journal of the Acoustical Society of America , 32*, 1046-1060.