4pPP10. Modeling loudness for impaired ears and applications to fitting hearing aids

Brian C. Moore*

*Corresponding author's address: Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB, Cambridgeshire, United Kingdom, bjcml@cam.ac.uk

Models of loudness for impaired ears are based on the assumption that cochlear hearing loss can be partitioned into a component due to outer hair cell dysfunction, resulting in reduced frequency selectivity and a more rapid growth of neural response with increasing level (reduced compression), and a component due to inner hair cell dysfunction, which reduces the neural response at all levels. In the first two models that were developed in Cambridge, the filtering and compression that take place on the basilar membrane were modeled as sequential processes, which is not physiologically realistic. Nevertheless, the models were able to account for many aspects of loudness, and were used to develop methods of fitting multi-channel compression hearing aids that have proven to be effective. More recently, a model of loudness has been developed in which the filtering and compression are modeled using a physiologically plausible nonlinear filter bank. This has also been applied to the fitting of hearing aids. Factors not included in the models include central plasticity resulting from altered auditory input, possible consequences of the operation of the efferent system, and the influence of cognitive factors such as perceived distance of the sound source and perceived vocal effort.

Published by the Acoustical Society of America through the American Institute of Physics
LOUDNESS MODELS FOR NORMAL HEARING

The models of loudness for normal hearing proposed by Moore and co-workers (Moore and Glasberg, 1996; Moore et al., 1997) have the following stages: (1) Simulation of the transfer of sound from the source to the eardrum; (2) Simulation of the transfer of sound through the middle ear, which is most efficient for frequencies in the range 500 to 5000 Hz; (3) Calculation of an excitation pattern for the sound reaching the cochlea. The excitation pattern is defined as the magnitude of the output of the auditory filters plotted as a function of center frequency (CF, Moore and Glasberg, 1983; Glasberg and Moore, 1990); (4) Transformation from excitation to specific loudness, which is a kind of “loudness density”, representing the loudness per ERBN, where ERBN is the average value of the equivalent rectangular bandwidth of the auditory filter at moderate sound levels for listeners with normal hearing (Glasberg and Moore, 1990). This transformation involves a compressive nonlinearity, reflecting the compression that occurs on the basilar membrane (BM) (Robles and Ruggero, 2001). The specific loudness plotted as a function of CF is called the specific loudness pattern. The CF is transformed to an ERBN-number scale (Glasberg and Moore, 1990; Moore, 2012), with units of Cams, which corresponds approximately to a scale of distance along the BM (Moore, 1986). The overall loudness of a sound, in sones, is assumed to be equal to the total area under the specific loudness pattern. The timbre of the sound is thought to be determined partly by the shape of the specific loudness pattern (Bladon and Lindblom, 1981; Moore, 2012), although dynamic aspects of the sound also play a role in timbre perception.

LOUDNESS MODELS FOR IMPAIRED HEARING

The perception of loudness may be affected by at least five changes that occur with cochlear hearing loss: (1) Elevation in absolute threshold, which may be caused by loss of function of outer hair cells (OHCs), loss of function of inner hair cells (IHCs), neural degeneration, or a combination of all of these (Schuknecht, 1993). Reduced functioning of the stria vascularis may also be involved (Schuknecht, 1993; Schmiedt, 1996), but it is assumed that this can be modeled indirectly as reduced functioning of the OHCs and IHCs. To model elevation of absolute threshold, it is assumed that the hearing loss at a specific frequency, HL TOTAL (dB), can be partitioned into two parts, caused by loss of function of OHCs and loss of function of IHCs/neurons:

$$HL_{OHC} + HL_{IHC} = HL_{TOTAL}$$

The value of HL TOTAL is determined from empirical measurements at the standard audiometric frequencies (plus 10, 12 and 14 kHz for the most recent version of the model). Linear interpolation is used to calculate values at intermediate frequencies. The value of HL OHC at a given CF may be estimated for a specific individual using measures of frequency selectivity and/or forward masking (Moore et al., 1999b). Alternatively, default (average) values may be used, based on the value of HL TOTAL. In the model, the value of HL OHC is not allowed to be greater than 57.6 dB.

(2) A reduction in or loss of the compressive nonlinearity in the input-output function of the BM, which is assumed to be determined by HL OHC. This is implemented in the model by using the value of HL OHC to modify the assumed low-level gain of the cochlear amplifier.

(3) Loss of frequency selectivity, which is again associated with HL OHC. For CFs above 500 Hz, it is assumed that the ERB of the auditory filter is broadened by a certain factor, B, relative to normal, where

$$B = 10^{0.01 HL_{OHC}}$$

The constant 0.01 is chosen so that B has a value of 3.8 for HL OHC = 57.6 dB. For values of HL OHC greater than 57.6 dB, the value of B is set to the value that would be obtained for HL OHC = 57.6 dB, i.e., 3.8. The maximum value of B corresponds roughly to the tuning of the BM in “dead” cochleas (Sellick et al., 1982).

(4) Reduced IHC and neural function. In an early version of the model (Moore and Glasberg, 1997), the effect of loss of IHC and/or neural function was modeled by a CF-dependent attenuation of the excitation level. For example, if HL IHC = 20 dB, the excitation level was attenuated by 20 dB. However, damaged IHCs may respond weakly to low-level inputs, but respond in a more normal way when the input signal is well above the threshold value. In a later version of the model (Moore and Glasberg, 2004), this was simulated by progressively reducing the attenuation applied to the excitation level as that level was increased.

(5) Complete loss of IHCs or functional neurons at certain places within the cochlea; effectively there may be a “dead” region that does not respond at all (Schuknecht, 1993; Moore, 2001). This is modeled by setting the
excitation to a very low value (effectively zero) over the range of CFs corresponding to the dead region.

A model based on these ideas (Moore and Glasberg, 2004) has been shown to give reasonably accurate predictions of a variety of data on the perception of loudness by hearing-impaired people, including the effects of loudness recruitment (Moore, 2004) and reduced loudness summation (Bonding, 1979).

In the models described above, auditory filtering and compression are treated as separate processes. However, this is unrealistic (Robles and Ruggero, 2001). Chen et al. (2011a) described a model for calculating excitation patterns and loudness for normal hearing based on a nonlinear filterbank, in which each filter is the sum of a broad passive filter and a sharp active filter. The gain of the active filter at a given CF is controlled by the output of the passive filter at that CF (Glasberg and Moore, 2000). Loudness is calculated from the area under the excitation pattern when plotted in intensity-like units on an ERBn-number scale; no transformation from excitation to specific loudness is required. Chen et al. (2011b) described an extension of this model to cases of hearing loss. Again, it is assumed that the overall hearing loss can be partitioned into two parts, HL\textsubscript{OHC} and HL\textsubscript{HLIC}. OHC loss is modeled by decreasing the maximum gain of the active filter, which results in increased absolute threshold, reduced compressive nonlinearity and reduced frequency selectivity. IHC loss is modeled by a level-dependent attenuation of the excitation level. This model also makes accurate predictions of the effects of loudness recruitment and reduced loudness summation for impaired ears.

APPLICATION OF LOUDNESS MODELS TO THE FITTING OF HEARING AIDS

Models of loudness have been used to develop methods for the initial fitting of hearing aids with multi-channel compression. In such aids, the input signal is filtered into several frequency channels, and level-dependent gains are applied in each channel; generally the gain for a given channel decreases as the input level increases, so a large range of input levels is compressed into a smaller range of output levels. The methods of hearing aid fitting described below are based on the concept of applying frequency- and level-dependent gains so as to achieve one or more goals in terms of the shapes of the specific loudness patterns and the overall loudness. The results are used to derive rules relating the gains to the degree of hearing loss at each frequency.

Two fitting methods were developed based on the loudness model of Moore and Glasberg (1997). One is called CAMEQ (Cambridge loudness equalization, Moore et al., 1999a). The goal of CAMEQ is to give a “flat” specific loudness pattern over the range 500–4000 Hz for a sound with the long-term average spectrum of speech, for a range of levels from 65 to 85 dB SPL. The overall loudness is also intended to be the same as normal for speech over this range of levels. The average speech spectrum for low-to-medium levels has more power at low frequencies than at high frequencies (ANSI, 1997). Hence, for a normal ear, the specific loudness pattern evoked by speech at 65 dB SPL has its highest values around 500 Hz. With the gains recommended by CAMEQ, the specific loudness pattern for an impaired ear is flatter than normal for low to medium input levels, so the timbre is expected to be somewhat “sharper” or “brighter” than normal. This is illustrated in Fig. 1.

![FIGURE 1](image)

FIGURE 1. The solid line shows the specific loudness pattern for a normal ear for a signal with the long-term average spectrum of speech with a level of 65 dB SPL. The other lines show specific loudness patterns for different amounts of sloping hearing loss when the signal has been amplified using the frequency-specific gains prescribed by the CAM2 fitting method (which is similar to CAMEQ for CFs up to 6 kHz).
The other procedure is called CAMREST (Cambridge loudness restoration, Moore, 2000). This aims to restore the specific loudness pattern to normal for speech-like stimuli, over a wide range of sound levels. This should make both the overall loudness and the timbre of speech similar to normal. For a speech-shaped noise signal with a level of 65 dB SPL, CAMREST prescribes slightly more low-frequency gain and slightly less mid-frequency gain than CAMEQ.

Evaluations of the CAMEQ and CAMREST methods, using digital multi-channel hearing aids with fast-acting compression (Moore et al., 2001; Alcántara et al., 2004; Marriage et al., 2004), showed that the two methods were roughly equally effective; the mean gain adjustments required to achieve satisfactory fits were close to zero for both methods, as illustrated in Fig. 2. In contrast, for the DSL[i/o] fitting method (Cornelisse et al., 1995) that was included for comparison, the mean gain adjustments were consistently negative, more so at high frequencies.

Moore et al. (2010) described a modified version of the CAMEQ method, intended to be applicable to the fitting of wide-bandwidth hearing aids that have been introduced commercially. The new method, now called CAM2, differs from CAMEQ in the following ways: (1) CAM2 gives recommended gains for center frequencies up to 10 kHz, whereas the upper limit for CAMEQ is 6 kHz; (2) CAM2 is based on the assumption that the user may wish to hear sounds from many directions, and uses a diffuse-field-to-eardrum transfer function, whereas CAMEQ uses a free-field-to-eardrum transfer function for frontal incidence; (3) CAM2 is based on the loudness model for impaired hearing described by Moore and Glasberg (2004), whereas CAMEQ was based on an earlier version of the model (Moore and Glasberg, 1997); (4) CAM2 is based on recent wideband measurements of the average spectrum of speech (Moore et al., 2008). In addition, the current version of CAM2 incorporates slightly reduced gains for inexperienced users, based on experimental evidence that such gains are preferred (Keidser et al., 2008).

A field trial of wide-bandwidth hearing aids with multi-channel slow-acting compression, fitted using CAM2, showed that the CAM2 method led to satisfactory loudness in most everyday listening situations, apart from some situations involving transient sounds (Moore and Füllgrabe, 2010). A laboratory trial comparing preferences for a simulated five-channel compression hearing aid fitted using either CAM2 or NAL-NL2 (Keidser et al., 2011), a currently popular fitting method, showed small but significant overall preferences for CAM2 for most participants, for both speech and music signals (Moore and Sek, 2012), as illustrated in Fig. 3. This was the case for both slow-acting and fast-acting compression (For a discussion of the pros and cons of slow versus fast compression see Moore, 2008).
LIMITATIONS OF CURRENT LOUDNESS MODELS AS APPLIED TO HEARING-AID FITTING

There are several limitations of loudness models as applied to hearing-aid fitting. The models do not take into account effects of: (1) Adaptation to hearing loss or acclimatization to hearing aids (Formby et al., 2003); (2) Activation of the efferent system (Guinan, 2006); (3) “Higher-level” processes. For example, the loudness of speech can be influenced by the apparent vocal effort of the talker (Lehiste and Peterson, 1959) and by the visually determined distance of the sound source (Mershon et al., 1981).

A fourth limitation is that the methods for hearing aid fitting described above do not take into account whether or not the excitation at a given CF is caused mainly by components in the input signal with frequencies close to that CF, called here “local” components. The intelligibility of speech can be predicted using the articulation index (AI, French and Steinberg, 1947) or the speech intelligibility index (SII, ANSI, 1997), which represents the proportion of the speech spectrum that is audible, with weighting for the relative importance of different frequency regions. The audibility of speech in different frequency regions can be estimated from the excitation pattern of a sound with the same long-term average spectrum as the speech (Moore and Glasberg, 1998). However, the assumptions underlying the SII calculation are only valid if the excitation at a given CF is dominated by local frequency components. If the excitation at a given CF is largely produced by “remote” components, then little or no information about the local components is transmitted to the brain. Essentially, the local components are masked by the remote components, an effect called interband masking.

French and Steinberg (1947) included a crude way of allowing for interband masking in their method of calculating the AI. Interband masking is also included in the SII procedure, but only very approximately, and only for normal hearing. To reduce interband masking, many hearing aid fitting formulae incorporate reduced gains at low frequencies. However, the amount of gain reduction is often chosen without any firm empirical or theoretical basis.

One way of assessing whether excitation at a given CF is dominated by local or remote components in the input signal is to compare the calculated excitation at that CF for a signal with the long-term average spectrum of speech and that same signal but with a band of frequencies around the CF removed. If the calculated excitation drops when the frequency band is removed, this indicates that the excitation at that CF was at least partly caused by local components. Further research is needed to establish quantitatively what proportion of the excitation needs to be determined locally for that excitation to be useful for speech perception.
ACKNOWLEDGMENTS

This work was supported by the Medical Research Council, UK (grant G0701870). I thank Brian Glasberg, Michael Stone, Christian Füllgrabe, Akeksander Sek and Zhangli Chen for their collaboration in this work and thank Hedwig Gockel for helpful comments and corrections.

REFERENCES


