



Published in final edited form as:

Proc Meet Acoust. 2011 October ; 12: 050005–050010.

Amplitude modulation detection by human listeners in sound fields

Pavel Zahorik*, Duck O. Kim, Shigeyuki Kuwada, Paul W. Anderson, Eugene Brandewie, and Nirmal Srinivasan

Abstract

The temporal modulation transfer function (TMTF) approach allows techniques from linear systems analysis to be used to predict how the auditory system will respond to arbitrary patterns of amplitude modulation (AM). Although this approach forms the basis for a standard method of predicting speech intelligibility based on estimates of the acoustical modulation transfer function (MTF) between source and receiver, human sensitivity to AM as characterized by the TMTF has not been extensively studied under realistic listening conditions, such as in reverberant sound fields. Here, TMTFs (octave bands from 2 – 512 Hz) were obtained in 3 listening conditions simulated using virtual auditory space techniques: diotic, anechoic sound field, reverberant room sound field. TMTFs were then related to acoustical MTFs estimated using two different methods in each of the listening conditions. Both diotic and anechoic data were found to be in good agreement with classic results, but AM thresholds in the reverberant room were lower than predictions based on acoustical MTFs. This result suggests that simple linear systems techniques may not be appropriate for predicting TMTFs from acoustical MTFs in reverberant sound fields, and may be suggestive of mechanisms that functionally enhance modulation during reverberant listening.

1. Introduction

One now classic approach to quantifying the temporal resolution of the human auditory system involves measurement of detection thresholds for envelope amplitude modulation (AM) for various modulation frequencies. The function that relates AM threshold to modulation frequency is known as a behavioral temporal modulation transfer function (TMTF), and has been shown to have a low-pass characteristic (Viemeister, 1979). This suggests that the human auditory system has limited temporal resolution and as a result, cannot follow variations in sound amplitude if they occur too rapidly. One of the principal advantages of this approach is that, once the behavioral TMTF is known, standard techniques from linear systems analysis can then be used to predict how the auditory system will respond to arbitrary patterns of temporal variation in envelope amplitude. Such methods have been successfully used to predict human sensitivity to non-sinusoidal amplitude modulation patterns (Viemeister, 1979), and form the basis for a now standard method of predicting the intelligibility of speech signals: the speech transmission index, or STI (Houtgast and Steeneken, 1985). In addition to source signal characteristics, it is well-known that various environmental factors, such as room reverberation or background noise, can profoundly influence the AM reaching a listener's ears (Houtgast and Steeneken, 1985). To our knowledge, the influence of room acoustic effects, however, has not been extensively

studied under realistic listening conditions. The classic work of Houtgast and Steeneken (1985), for example, did not involve any measurements of real room acoustics, only models based on a number of fairly unrealistic assumptions, such as: perfectly diffuse reverberant energy, input to only one ear, and indeterminate source distance. Here we characterize the acoustical MTFs for humans under more complex and realistic simulated sound field listening conditions in which the reverberation and sound source location are varied. We then relate acoustical MTFs to human behavioral TMTFs.

2. Methods

2.1 Subjects

Five listeners (4 male, 1 female, age range 23 – 41 years) participated in the experiment. All had normal hearing, and four were authors of this study.

2.2 Sound Field Simulation

Two types of sound field listening conditions were simulated using virtual auditory space techniques. One condition simulated listening in an anechoic sound field using non-individualized head-related impulse responses (HRIRs). A second condition simulated the sound field in a reverberant room with dimensions of $5.7 \times 4.3 \times 2.6$ m (L \times W \times H) and broadband reverberation time, T_{60} , of approximately 1.8 s. This simulation used binaural room impulse responses (BRIRs) modeled using techniques described by Zahorik (2009) and utilized the same HRIR sets as in the anechoic condition. Two source locations were tested in each sound field at a fixed distance of 1.4 m: directly in front of the listener (0°), and opposite the listener's right ear (90°). A third diotic listening condition that contained no HRIR or BRIR processing was also tested as a control.

2.3 Acoustical MTF Determination

Two methods of determining the acoustical MTFs for each sound field listening condition were used. The first method determined the MTF directly by numerically computing the modulation depths in a sinusoidally modulated broadband Gaussian noise signals pre and post convolution with HRIRs or BRIRs at octave-spaced modulation frequencies ranging from 2 to 512 Hz. The second method used a technique described by Schroeder (1981) for computing the MTF from the squared impulse response normalized by its total energy.

2.4 Stimuli and Procedure

To facilitate rapid behavioral TMTF data collection, a method of adjustment procedure was used in which subjects were instructed to adjust the modulation depth of sinusoidal AM so that it was just audible. These adjustments were conducted for 9 modulation frequencies from 2 – 512 Hz in octave steps and implemented in real time via MAX/MSP software (Cycling '74) using the Freeverb3 VST plug-in for HRIR/BRIR convolution. The source signal was broadband noise. A graphical depiction of the stimulus generation and method of adjustment procedure is shown in Figure 1

3. Results

Figure 2 displays acoustical MTFs for both anechoic and reverberant sound fields at both 0° and 90° . Consistent with past research, the reverberant room exhibits a low-pass characteristic, although its shape is both considerably different and more complex than that suggested by classic prediction (Houtgast and Steeneken, 1985). Good agreement between the two methods of MTF determinations is observed.

Figure 3 displays mean behavioral TMTFs for 5 listeners, along with a stylistic representation of classic monaural behavioral TMTF data (Viemeister, 1979) using a broadband noise carrier for comparison. Modulation thresholds are displayed in units of $20\log_{10}(m)$, where m is the modulation depth. Both diotic and anechoic data are in good agreement with Viemeister's data. The room data, however, exhibit considerably higher AM thresholds (increase of 4 to 5 dB) at modulation frequencies above 16 Hz. Such a result is generally expected given the low-pass characteristic of the room acoustical MTF shown in Figure 2. However, the increases in psychophysical threshold are less than would be predicted from the acoustical MTFs (particularly at modulation frequencies above 32 Hz). This result is both surprising and important because it suggests that simple linear systems techniques may not be appropriate for predicting behavioral TMTFs from acoustical MTFs in all situations, such as evaluating the effects of the acoustic environment. Although further work is needed to determine the precise cause of this effect, including additional analyses based on different audio frequency regions of the stimulus, it may be indicative of neural mechanisms that serve to functionally enhance modulation processing in reverberant environments.

4. Conclusions

Human sensitivity to AM in reverberant sound fields is greater than predicted based on the broadband acoustical MTF alone. Of course humans may not equally attend to the entire range of audio frequencies afforded by the broadband noise carrier signal used in this study. As such, additional acoustical analyses and behavioral testing with narrowband carrier signals centered at various frequency regions may provide further insight into the apparent gap between the acoustical and behavioral MTFs in reverberant sound fields.

Acknowledgments

Work supported in part by the NIH-NIDCD (R01 DC008168, Zahorik; R01 DC002178, Kuwada).

References

- Houtgast T, Steeneken HJM. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J. Acoust. Soc. Am.* 1985; 77:1069–1077.
- Schroeder MR. Modulation transfer functions: Definition and measurement. *Acustica.* 1981; 49:179–182.
- Viemeister NF. Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am.* 1979; 66:1364–1380. [PubMed: 500975]
- Zahorik P. Perceptually relevant parameters for virtual listening simulation of small room acoustics. *J. Acoust. Soc. Am.* 2009; 126:776–791. [PubMed: 19640043]

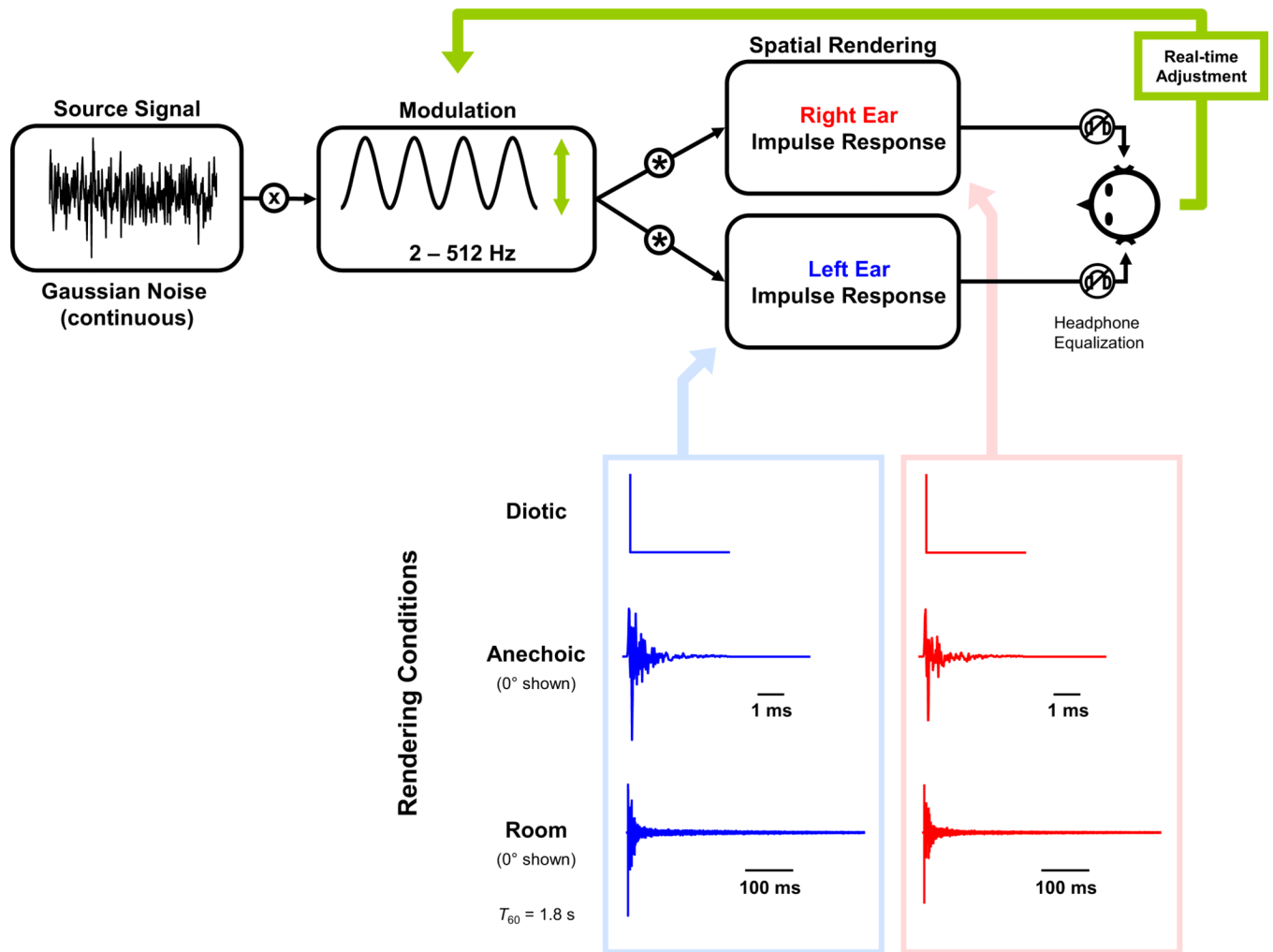


Fig. 1. Schematic of stimulus presentation and method-of-adjustment response procedures used in this study. Three listening conditions (diotic, anechoic, and reverberant room) were tested. For the sound field conditions, two source azimuth locations (0° and 90°) were tested at a distance of 1.4 m.

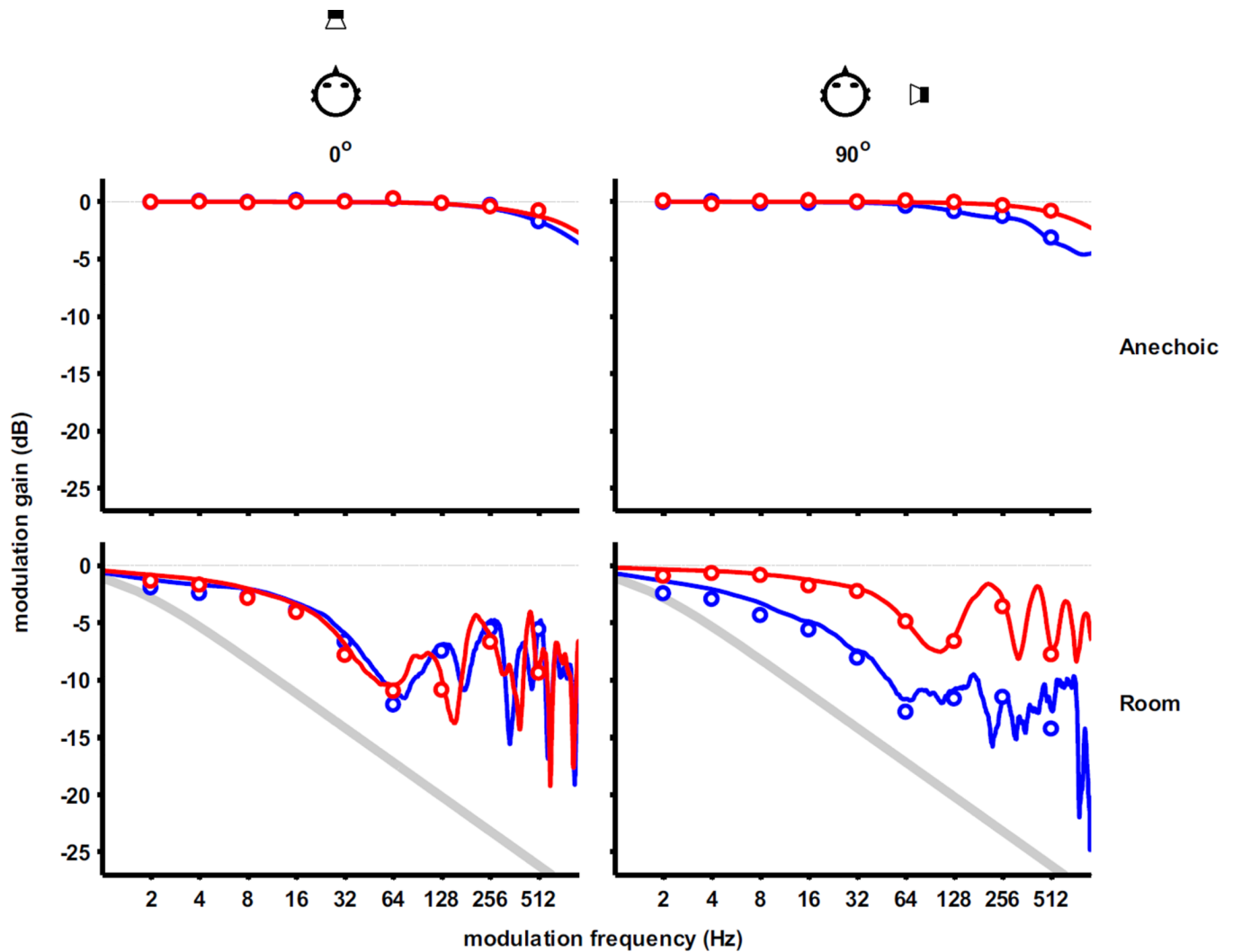


Fig. 2.

Acoustical MTFs for the anechoic and reverberant room sound fields used for testing in this study. MTFs are displayed for both source orientations (0° and 90°) and for both left (**blue**) and right (**red**) ears; the distance was 1.4 m. The curves represent MTFs computed from HRIRs / BRIRs and the symbols represent results of direct methods of MTF determination (see text for details). Predicted acoustical MTFs for the reverberant room are also displayed based on an algorithm described by Houtgast and Steeneken (1985) assuming exponentially decaying diffuse reverberation, with $T_{60} = 1.8$ s.

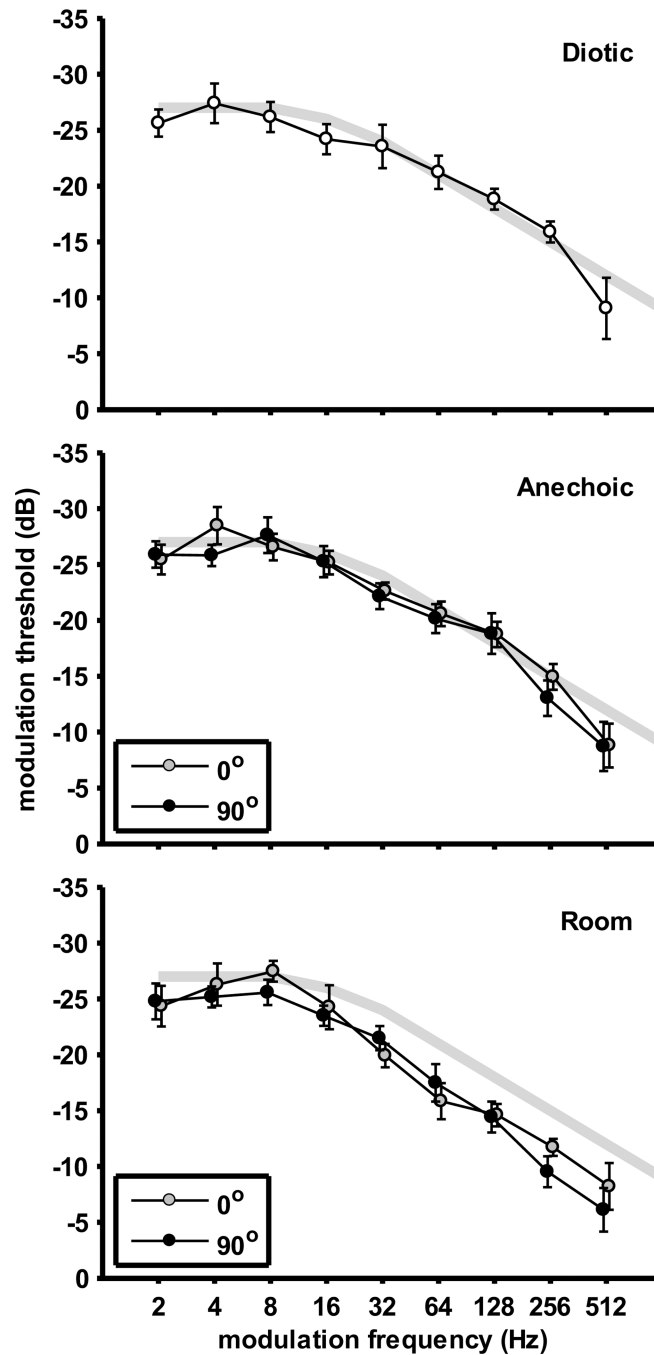


Fig. 3. Psychophysical TMTFs corresponding to threshold modulation depth for detecting AM based on data from 5 human listeners (mean \pm 1 SEM) in 3 listening conditions (diotic, anechoic, and room) and 2 source positions (0° and 90° at a distance of 1.4 m). The solid gray line is a stylistic representation of classic monaural TMTF data using a broadband noise carrier (Viemeister, 1979).