Spectral integration versus multiple looks in the sample discrimination of intensity

Robert A. Lutfi^{a)} and Lynn Gilbertson

*Auditory Behavioral Research Laboratory, Department of Communicative Disorders,
University of Wisconsin, Madison, Wisconsin 53706
ralutfi@wisc.edu, lrjohnson1@wisc.edu*

Abstract: A fundamental property of hearing is that signals become more detectable as their bandwidth is increased. Two models have been proposed to account for this result. The integration model assumes that detection is mediated by the output of a single frequency channel matched in bandwidth to the signal. The multiple-looks model assumes that detection is based on the combination of outputs from multiple channels matched to the individual frequencies of the signal. Results are reported supporting the integration model.

© 2012 Acoustical Society of America

PACS numbers: 43.66.Ba, 43.66.Fe [QJF]

Date Received: September 6, 2011 **Date Accepted:** November 2, 2011

1. Introduction

Most speech and everyday sounds are distinguished by changes that occur over a wide frequency range. Thus, a preoccupation of psychoacoustics has been to understand how listeners accumulate information across the broad spectrum to improve their ability to detect and discriminate such sounds. Past studies investigating this problem have focused on the listener's ability to detect a level increment in a single band of noise or in combinations of tones as the bandwidth of these signals increased (Green, 1960; Berg, 1987; Berg and Robinson, 1987; Lutfi, 1989). The results of these studies generally agree in showing a restricted range over which increases in bandwidth improves detectability; however, accounts of the results differ for the two types of signals. In early studies, where single bands of noise are used, the data are described using a classic energy-integration model (Green, 1960). This model assumes that detection is mediated by an increment in energy integrated over the output of a single frequency channel matched to the bandwidth of the noise. In more recent studies, where multi-tone complexes have been used, the intensities of tones are sampled at random (sample discrimination) and the results are more typically described using a multiple-looks model. Unlike the integration model, the multiple-looks model assumes that detection is based on a weighted sum of the outputs of independent frequency channels tuned to the individual frequencies of the tones (Berg, 1990; Berg and Robinson, 1987; Lutfi, 1989).

A similar dichotomy exists in the literature on the temporal processing of signals. Here, early experiments gave rise to the classical view that detection results from energy integration over a single, relatively long, temporal window (Zwislocki, 1960), analogous to the single wideband spectral window of the spectral integration model. This was the prevailing view until counter evidence was presented in the seminal study of Viemeister and Wakefield (1991). Viemeister and Wakefield (1991) asked listeners to detect an increment in a pair of brief tone pulses separated in time and presented within temporal gaps of a band of noise. Consistent with the classical view, thresholds for detection were found to be lower for the two tones over either one. However, thresholds were also found not to be affected by the intensity of the noise within the

^{a)} Author to whom correspondence should be addressed.

interval between the tones, a result inconsistent with the operation of a single, long temporal-integration window. Viemeister and Wakefield (1991) interpreted their results as support for a multiple-looks model in which the information from the two tones is combined from brief samples separated in time (however, see also Dai and Wright 1995, 1996).

In the present study we take a different approach to testing spectral integration and multiple-looks models that exploits a well-documented effect for tone-in-noise detection described early on by Greenberg and Larkin (1968). These authors showed that when the frequency of the tone presented on any trial deviates from what the listener expects, detection performance is reduced, the size of reduction varying systematically with the size of the deviation from the expected frequency. Greenberg and Larkin (1968) attribute the frequency-uncertainty effect to the response properties of an auditory filter tuned to the expected frequency, the idea being that as the tone deviates from the expected frequency it is effectively reduced in level by the filter's attenuation characteristic, resulting in poorer detection performance. Following this logic, consider the predictions of the two models when the frequencies of a multitone complex deviate at random from trial to trial about their expected values. The deviation should have no effect on the level of the tones at the output of a broadband filter associated with an integration model. The integration model, therefore, predicts no effect of the deviation on performance. In contrast, the effect of the deviation for the multiple-looks model will be to induce variability across trials in the level of the tones at the outputs of the auditory filters. This variability is expected to decrease performance by making it more difficult for the listener to detect the level increment on tones.

2. General method

2.1 Stimuli

The stimuli were N -tone complexes, where N ranged from 1 to 6. The frequencies of the tones were spaced at equilog intervals from 250 to 7049 Hz. The ratio of adjacent frequencies f_{i+1}/f_i was 1.95. This spacing ensured negligible mutual masking among tones (Patterson, 1976). For each of the three conditions, the number of tones was increased beginning with the lowest frequency tone and ending with the highest. For the fixed frequency condition (F) the tones maintained their nominal frequencies throughout each trial block. For the random-frequency conditions the tones deviated at random about their nominal values from trial to trial. The deviation was uniform over a range of 15% above and 15% below the nominal frequency. It was the same for all tones for the perturbation same (PS) condition and it was selected independently for each tone for the perturbation different (PD) condition.

All tone complexes were played at a 44 100-Hz sampling rate with 16-bit resolution using a MOTU 896 audio interface. From the interface the sounds were buffered through a Rolls RA62c headphone amplifier and then delivered to the right ear of listeners over Beyerdynamic DT 990 headphones. A loudness balancing procedure was used (see Lutfi *et al.*, 2008) to calibrate the level of individual tones to be approximately 67 dB SPL on average at the eardrum. Listeners were seated individually in a double-walled, IAC sound-attenuated chamber. The tone complexes were gated on and off with 5-ms, cosine-squared ramps for a total duration from 0 voltage points of 100 ms.

2.2 Procedure

A sample discrimination procedure was used to measure the listener's ability to discriminate changes in the level of the multitone complexes (cf. Lutfi, 1989). On each trial the individual levels of the tones were selected independently and at random from one of two, equivariate, normal distributions differing in mean. The means of the two distributions were 65 and 69 dB SPL, the standard deviation was $\sigma = 2$ dB. The listener heard a single N -tone complex on each trial and was required to indicate by button press whether the levels of the tones were selected from the low- or high-level

distribution. The value of N was fixed for each block of 50 trials and feedback was given after each trial. Hit and false alarm rates were obtained from a total of 250 trials per subject, per condition and were converted to values of d' (Swets, 1964). No significant response bias was observed across listeners and so bias measures are not reported. The recording of trial-by-trial data, the generation of stimuli, and all other experimental events were computer controlled.

Five female students from the University of Wisconsin-Madison participated as paid listeners in the study. The ages of the students ranged from 21 to 35 years. All listeners had normal hearing as determined by standard audiometric screening (ANSI, 1996), and all had extensive experience in psychoacoustic experiments. Listeners ran a single 2-h session per day, five days out of each week, with sessions including frequent breaks. All listeners received two or more weeks of training in the task before data collection began.

2.3 Simulation

To ensure that the conditions selected would produce effects large enough to distinguish between models, we undertook a simulation of the results, replicating the precise conditions of the study with the same number of trials. We chose a general form of the multiple-looks model that has proven successful in accounting for the results of past sample-discrimination studies (Berg and Robinson, 1987; Lutfi, 1989). Other forms of the model would make similar predictions for these data provided the observations are independent. The decision rule for the model is

$$\text{Respond SN iff } \sum_{i=1}^N (\log G_i - A_i + e) > C, \quad (1)$$

where G_i is the output power of a Gaussian auditory filter at the i th frequency for the fixed frequency condition (cf. Patterson, 1976), A_i is the dB attenuation at the output of the filter resulting from the perturbation in frequency, e is a normal random deviate representing internal noise, and C is a fixed response criterion. For the $\pm 15\%$ jitter in frequencies in PD and PS conditions, the values of A_i ranged from 0 to 10 dB. The standard deviation of the internal noise e was chosen to have a value of 1 dB so as to be consistent with the data for the discrimination of pure-tone intensity (cf. Jesteadt *et al.*, 1977). For integration model the decision rule was chosen to have a form consistent with the early integration model described by Green (1960),

$$\text{Respond SN iff } \log \sum_{i=1}^N R_i + e > C, \quad (2)$$

where R_i is the output power at the i th frequency of a single rectangular filter matched in bandwidth to that of each N -tone complex.

The results of the simulation are given in Fig. 1. The data for PS are given by squares in the left panels, for PD by diamonds in the right panels and for F by circles in both panels. The continuous curves shown in the figure are linear least-squares fits to the data for each condition. The dashed line represents ideal performance, growing as the square root of N . The two models [Figs. 1 (top) and 1 (bottom)] yield essentially identical results for F, with performance growing at a rate slightly less than the square root of N due to the internal noise. For the multiple-looks model performance is significantly poorer for PS and PD, with performance growing at the fastest rate with N for PD. This is as expected since for PD there is less shared variance across filter outputs and so greater benefit from additional observations. For the integration model performance differs little across the three conditions, also as expected. The results of the simulation indicate that the conditions of the study should be adequate to decide between the two models.

Fig. 1. Simulation results for exact conditions of the study for (top) multiple-looks and (bottom) integration models. Conditions are fixed-frequency, F (circles), perturbation same, PS (squares) and perturbation different, PD (diamonds). Continuous curves are linear least-squares fits to the data for each condition. The dashed line represents ideal performance, growing as the square root of N .

3. Results and discussion

Figure 2 gives the individual results for the five listeners (panel rows) plotted in the manner of Fig. 1. Results for the fixed-frequency condition (circles) replicate earlier findings of Lutfi (1989) and Neff *et al.* (1994). These authors also measured sample

Fig. 2. Results from 5 listeners (rows) are plotted in the same manner as Fig. 1.

intensity discrimination for fixed-frequency tone complexes and report performance to improve with the number of tones at less than an optimal rate. In the [Lutfi \(1989\)](#) study, d' is described as growing approximately as the cube root of N . This would require that the slopes of the linear curves fitted through the data have a slope of 0.33. The slopes deviate about this value somewhat from one listener to the next, but the average value is, indeed, 0.33. This is also the average value of the slopes generated by repeated simulations of the multiple-looks [Eq. (1)] and integration [Eq. (2)] models for the fixed-frequency condition. The more important outcome, however, is that the results for the PS and PD conditions differ very little from those of the fixed-frequency condition. This outcome clearly supports the integration model.

The present results are difficult to reconcile in terms of the simple multiple-looks model given by Eq. (1). This model was originally proposed to account for the data from fixed-frequency conditions in sample-discrimination experiments, which it does quite well ([Lutfi, 1989](#); [Berg and Robinson, 1987](#)), but it has not until now been tested in conditions that would yield predictions fundamentally different from those of a traditional integration model for these experiments. Generally, multiple-looks models can be difficult to reject because they entail a large number of free parameters that can be chosen to accommodate different aspects of the data. Had we chosen, for example, a different filter shape, different weights on filter outputs or a different model of internal noise in Eq. (1), we could have generated different predictions for the form of the function relating d' to N in the fixed-frequency condition. However, whatever the form of the function, if the multiple looks are to be largely independent, as required by such models, then a trial-by-trial perturbation in frequency in PD and PS conditions should have detrimentally affected performance by increasing the variance of these looks. We see little sign of such an effect in the conditions of this study.

Acknowledgment

This research was supported by a NIDCD grant R01 DC001262-20.

References and links

- ANSI (1996). ANSI S3.6-1996, American National Standards Specification for Audiometers (American National Standards Institute, New York).
- Berg, B. G. (1987). "Internal noise in auditory decision tasks," Doctoral dissertation, Indiana University, Bloomington, IN.
- Berg, B. G. (1990). "Observer efficiency and weights in a multiple observation task," *J. Acoust. Soc. Am.* **88**, 149–158.
- Berg, B. G., and Robinson, D. E. (1987). "Multiple observations and internal noise," *J. Acoust. Soc. Am.* **81**, S33.
- Dai, H., and Wright, B. A. (1995). "Detecting signals of unexpected or uncertain durations," *J. Acoust. Soc. Am.* **98**, 798–806.
- Dai, H., and Wright, B. A. (1996). "The lack of frequency dependence of threshold for short tones in continuous broadband noise," *J. Acoust. Soc. Am.* **100**, 467–472.
- Green, D. M. (1960). "Auditory detection of a noise signal," *J. Acoust. Soc. Am.* **32**, 121–131.
- Greenberg, G. Z., and Larkin, W. D. (1968). "Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: The probe-signal method," *J. Acoust. Soc. Am.* **44**, 1513–1523.
- Jesteadt, W., Wier, C. C., and Green, D. M. (1977). "Intensity discrimination as a function of frequency and sensation level," *J. Acoust. Soc. Am.* **61**, 169–177.
- Lutfi, R. A. (1989). "Informational processing of complex sound. I. Intensity discrimination," *J. Acoust. Soc. Am.* **86**, 934–944.
- Lutfi, R. A., and Liu, C. J., Stoelinga, C. N. J. (2008). "Level dominance in sound source identification," *J. Acoust. Soc. Am.* **124**, 3784–3792.
- Neff, D. L., Kessler, C. J., and Sullivan, R. E. (1994). "Sample discrimination of intensity differences for mixtures of target and context tones," *J. Acoust. Soc. Am.* **95**, 2942.

- Patterson, R. D. (1976). "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. **59**, 640–654.
- Swets, J. A. (1964). *Signal Detection and Recognition by Human Observers* (Wiley, New York), pp. 147–198.
- Viemeister, N. F., and Wakefield, G. H. (1991). "Temporal integration and multiple looks," J. Acoust. Soc. Am. **90**, 858–865 (1991).
- Zwislocki, J. J. (1960). "Theory of temporal auditory summation," J. Acoust. Soc. Am. **32**, 1046–1060.