This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

Examining the use of a time-varying loudness algorithm for quantifying characteristics of nonlinearly propagated noise (L)

S. Hales Swift[a] and Kent L. Gee
Department of Physics and Astronomy, N-283 Eyring Science Center, Brigham Young University, Provo, Utah 84602

(Received 25 October 2010; revised 16 February 2011; accepted 25 February 2011)

A previous letter by Gee *et al*. [J. Acoust. Soc. Am. **121**, EL1–EL7 (2007)] revealed likely shortcomings in using common, stationary (long-term) spectrum-based measures to quantify the perception of nonlinearly propagated noise. Here, the Glasberg and Moore [J. Audio Eng. Soc. **50**, 331–342 (2002)] algorithm for time-varying loudness is investigated. Their short-term loudness, when applied to a shock-containing broadband signal and a phase-randomized signal with equivalent long-term spectrum, does not show a significant difference in loudness between the signals. Further analysis and discussion focus on the possible utility of the instantaneous loudness and the need for additional investigation in this area. © *2011 Acoustical Society of America*. [DOI: 10.1121/1.3569710]

PACS number(s): 43.50.Ba, 43.25.Cb, 43.50.Nm, 43.66.Lj [PBB] Pages: 2753–2756

I. INTRODUCTION

The nonlinear propagation of broadband noise, as it relates to military jet aircraft and rockets, is a problem that comprises both physical and perceptual acoustics issues. Important features of the waveforms change as they nonlinearly steepen to the point of shock formation and then interact further. Perceptually, the acoustic shocks result in a distinct, crackle-like sound quality.[1,2] The relative importance of this feature in the overall perception of the noise has motivated the study of candidate measures that could be useful in eventually quantifying perception of the noise. This letter investigates time-varying loudness (TVL) as one such candidate measure.

Previous work by Gee *et al*.[3] investigated the response of stationary sound level, loudness, and annoyance metrics to a set of three waveforms that were included as multimedia content. (“Stationary” as used in this paper refers to use of the long-term average spectrum.) The first waveform (“nonlinear. wav”) was a broadband noise waveform which had been numerically propagated using a nonlinear numerical algorithm based on the generalized Burgers equation for spherical waves in the atmosphere. The input waveform to the numerical model was an initially Gaussian, shaped broadband noise signal with spectral features similar to that of a military jet aircraft (e.g., see Ref. 4). The second (“linear. wav”) was created using the same input and algorithm, but with the nonlinear term removed. The third (“phase-modified nonlinear. wav”) was created by randomizing the Fourier phase of the nonlinearly propagated signal in conjugate pairs to create a waveform with an equivalent long-term average spectrum, but with different time-domain behavior. Although the stationary metrics studied indicated some differences between the linearly and nonlinearly propagated signals, these tended to be small. In addition, because their average spectra were identical, none of the metrics distinguished between the nonlinearly propagated and rephased waveforms, although such differences are easily audible.[3]

Such problems are not altogether uncommon in investigating the sound quality of high-amplitude jet noise, where some important qualitative characteristics are not detectable from the power spectrum of the signal alone. (An example is the so-called “crackle” phenomenon, which has instead been described using the statistics of the time waveform[1] and more recently, the time waveform derivative.[2]) The results of the Gee *et al*.[3] study prompted further investigation to identify a metric that objectively quantifies the perceptual differences between the nonlinearly propagated signal and a signal with equivalent long-term spectrum, or, similarly, between a crackling waveform and one that has the same spectrum. After publication of the previous study,[3] it was suggested to the authors that a more suitable metric might be TVL, which incorporates both temporal and spectral features of loudness into the model. Thus, it was hypothesized that the model would respond more appropriately to the temporal features of the acoustic shocks.

In a jury-based study, Marshall and Davies[5] showed that the maximum short-term loudness of the Glasberg-Moore[6] TVL algorithm was the best of several metrics for predicting

---

[a] Author to whom correspondence should be addressed. Electronic mail: hales.swift@gmail.com

the loudness and annoyance due to a number of transient sounds including sonic booms. Although different from jet noise in many respects, sonic booms have similarities in terms of the presence of rapid, shock-like increases in pressure interspersed with more gradual decreases in pressure. Thus, this algorithm has been implemented and the short-term loudness calculated to evaluate each of the three signals from the previous paper.

The Glasberg-Moore model predicts a signal's loudness through time, accounting for temporal masking and integration, thus differing from measures based solely on long-term spectra. This is done through a series of steps. Their "short-term loudness" (L_{ST}) has been calculated by windowing the pressure waveform using windows of different lengths (longer windows for lower frequencies) and then finding the instantaneous loudness (L_{Inst}) for each 1-ms step using their now-standardized[7] stationary loudness protocol. L_{Inst} is comparable to the amount of activity in the auditory nerve during a brief time period in that, while it contains auditory information that is later processed to create a loudness impression, it is not supposed to exist in a form typically available for conscious perception as loudness. The short-term loudness is found by taking a running average of L_{Inst} using two time constants. These are meant to account for forward masking and loudness temporal integration. A shorter time-constant is used to model attacks (where an increased L_{Inst} will serve to increase the running L_{ST}) and a longer constant is used to treat decays. Interested readers are encouraged to examine Glasberg and Moore's[6] paper for a more in-depth description of the model.

II. RESULTS AND DISCUSSION

Prior to performing the calculations and based on the suggestions received, our initial expectation was that the L_{ST} of the nonlinearly propagated signal would be greater than the linearly propagated signal and also greater than the rephased signal on average. This hypothesis seemed reasonable because the presence of acoustic shocks in the nonlinearly propagated signal concentrates much of the high-frequency energy important to loudness into discrete times. Because the loudness (L_{ST}) tends to increase faster with an increase in instantaneous loudness (L_{Inst}) and decays more slowly in response to a decrease, it was expected that the concentration of the loudness in this way would lead to greater average values for the short-term loudness of the nonlinearly propagated signal than the rephased signal.

In Fig. 1, the short-term loudness level (LL_{ST}), in phon, is shown for the three signals from Ref. 3. As expected, the linearly propagated waveform has the lowest average LL_{ST} on account of the low-pass filtering effect of atmospheric absorption. The LL_{ST} values for the nonlinearly and linearly propagated waveforms follow the same trend of rise and fall because they originated from the same input waveform and differ only by the inclusion or exclusion of the nonlinear term in the propagation algorithm. Contrary to our original hypothesis, however, there are only brief occasions where the nonlinearly propagated waveform exceeds its rephased version, such that the average LL_{ST} of the rephased waveform is 0.6-phon greater than that of the nonlinearly propagated waveform. In hindsight, this likely occurs because the growth of loudness in the hearing system with input energy is compressively nonlinear, causing the L_{Inst} to grow more slowly at higher levels.[8] This effect is apparently not offset by the tendency of L_{ST} to gravitate toward the largest peaks in L_{Inst}, due to the shorter attack time-constant. The two effects are in competition; thus, the predicted loudness differences are quite small.

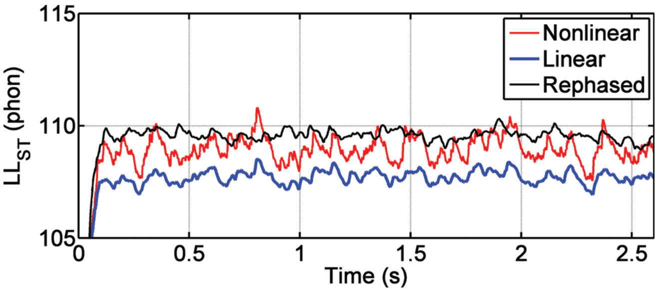

FIG. 1. (Color online) The short-term loudness represented as a loudness level (LL_{ST}) in phon of the linearly propagated, nonlinearly propagated, and rephased waveforms.[1]

The results of Marshall and Davies[5] suggest the possibility that the maximum short-term loudness could be important in predicting the loudness actually attributed to a transient sound. The nonlinearly propagated signal does have the greater maximum short-term loudness, so this might be taken to validate our original hypothesis. However, the applicability of the Marshall and Davies result, made for isolated transient signals, is unclear for these sounds. Rather than containing one or two isolated impulses (as in, for example, a sonic boom), these waveforms are composed of many closely spaced transients that may not be plainly discernable from one another. Thus, it is not entirely clear whether the average or maximum short-term loudness is more appropriate in our case. This remains an open question for this family of signals, especially because the difference in maximum LL_{ST} is only about 0.5 phon for these waveforms.

Although the maximum and average LL_{ST} differences are both less than 1 phon, one potentially significant difference between the LL_{ST} of the nonlinear signal and that of the rephased signal is that the former varies more than the latter. In order to auralize the relative magnitude of this loudness variation, an additional waveform, which we call the "modulated rephased waveform," was prepared by calculating the L_{ST} of both the rephased and nonlinearly propagated waveforms and using their ratio to amplitude-modulate the envelope of the rephased signal from Ref. 3 iteratively until its short-term loudness closely followed the nonlinearly propagated waveform. The result of the modification is shown in Fig. 2(a). Note that this modulation process caused only minor differences between the average long-term spectra of the nonlinearly propagated and modulated rephased waveforms. This new waveform (see the URL below) can be compared to Mm. 3 (the nonlinearly propagated waveform) and Mm. 4 (the rephased version) embedded in Ref. 3 by following the links provided therein.[14]

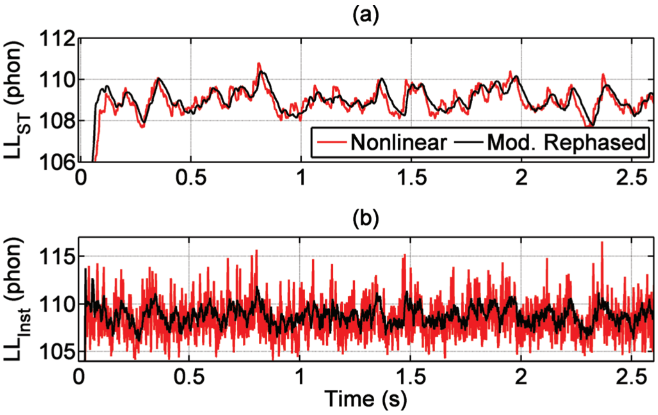

FIG. 2. (Color online) (a) LL_{ST} and (b) LL_{Inst} for the nonlinearly propagated and the modulated rephased signals.

Although the modulated rephased waveform sounds more similar to the nonlinearly propagated waveform in terms of its slow fluctuation behavior, the acoustic shocks of the nonlinearly propagated waveform continue to result in audible differences of "texture." To further investigate these differences, we examined the instantaneous loudness (represented as a level, LL_{Inst}) of both waveforms. This effectively removes the temporal integration and masking. This line of thought has been pursued despite the fact that it is presumed in the model that the listener does not directly perceive this quantity.[6] Some justification for this approach may therefore be appropriate.

Work by several authors[9–11] suggests that loudness perception may be influenced strongly by certain types of events in a waveform and that the initial attack portion of a sound may receive stronger emphasis in loudness comparisons than other portions of the waveform. Rennies *et al.*[9] indicate that this "attack effect" is stronger for signals with larger bandwidth. Pedersen *et al.*[10] suggest that mid-signal events, such as a change in spectral content, lead to a similar attack effect to that seen at signal onset. A delayed attack effect was seen by Oberfeld[11] when signals were faded in achieving steady-state levels at a time later than the signal beginning. In loudness comparisons, subjects in that experiment gave emphasis to the first signal segment achieving the full level.

The nonlinearly propagated signal contains shocks with quite rapid rise times (leading to large bandwidth concentrated at the shocks) interspersed between periods of gradual relaxation of pressure and relative quiet. The shocks thus constitute a change in bandwidth, fulfilling the condition of Pedersen *et al.*[10] for a mid-signal attack effect, a rapid increase in level, fulfilling the condition of Oberfeld[11] for a delayed attack effect, and a large bandwidth, fulfilling the condition of Rennies *et al.*[9] for an enhanced attack effect.

A second relevant process was studied by Zwicker[12] and also considered in a recent article by Rennies *et al.*,[13] Zwicker reports that for signals of short duration, less forward (temporal) masking is seen. Thus, loudness decays more rapidly after the offset of a brief signal. We therefore ask whether the acoustic shocks in the nonlinearly propagated waveform, as essentially broadband events of short duration, might contribute more to loudness increases through the attack effect mentioned than improve brevity predicted by the Glasberg-Moore[6] model and also decay more quickly thereafter (due to their brevity) leading to stronger audibility of the texture resulting from the shocks.

If these two effects do indeed play their expected roles in influencing the loudness of these signals, then we would expect the loudness to increase more quickly at the shocks and then decrease more quickly thereafter. This would, in effect, allow more of the information contained within the instantaneous loudness to become available to conscious perception. Based on the results cited, we think that the nonlinearly propagated waveform satisfies the conditions for these effects well enough to provide justification for the examination of LL_{Inst}.

Shown in Fig. 2(b), the differences in LL_{Inst} for the two waveforms—the modulated rephased and the nonlinearly propagated—are striking. For example, fluctuations in LL_{Inst} of the nonlinearly propagated signal have a standard deviation of 1.9 phon and exceed 10 phon overall. On the other hand, the standard deviation in LL_{Inst} for the modulated rephased signal is 1.0 phon. The much larger fluctuations in the instantaneous loudness of the nonlinearly propagated signal seem to match what is experienced perceptually with these signals. Further, as Fig. 3 shows, the timing of the shocks [emphasized using the time waveform derivative in Fig. 3(b)] is coincident with the peaks of the instantaneous loudness, suggesting that it is indeed the shocks that are contributing to the texture. This texture thus appears to be represented in the values of the instantaneous loudness. It may, therefore, be possible to identify and quantify the distinctive crackle-like characteristics of this waveform using the instantaneous loudness or the statistics of its distribution.

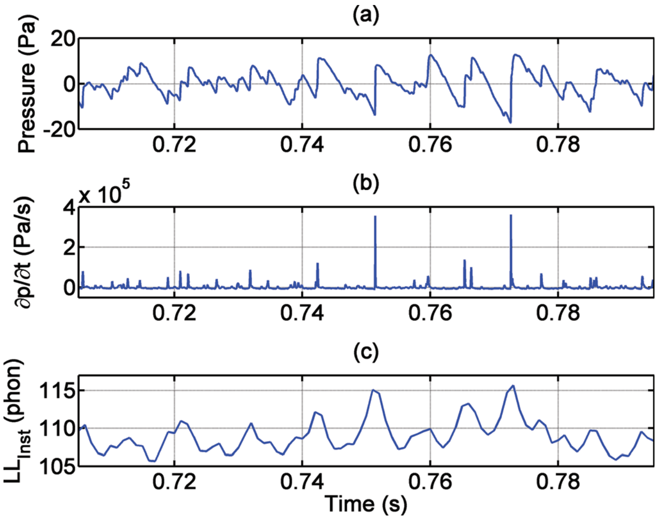

FIG. 3. (Color online) (a) The waveform of the nonlinear signal. (b) The time derivative (first difference) of the pressure waveform (notice the presence of acoustic shocks as positive spikes in the trace). (c) The instantaneous loudness level (notice the location of the peaks relative to the location of the shocks).

III. CONCLUDING REMARKS

The results of this study suggest that the Glasberg-Moore[6] short-term loudness is not an appropriate metric for quantifying the crackle-like qualitative characteristics of nonlinearly propagated broadband noise. However, the instantaneous loudness may be a more effective measure for objectively discerning and quantifying differences between nonlinearly propagated noise and sounds with similar long-term spectra but less distinctive time-domain behavior. Other time-domain or joint time-frequency domain models may be similarly or more effective in identifying these or related qualities.

In looking ahead to alternate models or psychoacoustic quantities that may be of use in this problem, a recent model of Rennies *et al.*[13] accounts for the attack effect by allowing greater spectral summation during attacks. It also allows the duration of temporal masking to vary in order to account for the abbreviated masking due to short signals described by Zwicker. It might therefore yield an advantage in dealing with some of the less common features of our signals as well as possibly other impulsive sounds. Future work should also include evaluation of a broader array of sound quality metrics, possibly including roughness and a time-varying form of sharpness. Roughness may increase for these signals because the rapid changes in instantaneous loudness due to the passage of acoustic shocks take place on the order of 50–200 Hz, well within the 15–300 Hz sensitivity range of a roughness metric. A time-varying sharpness metric would likely respond to the concentration of high-frequency energy at the shocks and to its dearth in between. Ultimately, coupled with the previous studies by Gee *et al.*,[2,3] this letter has opened the door for further investigations as to how to best describe the unique "crackle-like" quality of high-amplitude jet and rocket noise.

[1] J. E. Ffowcs Willams, J. Simson, and V. J. Virchis, "'Crackle': An annoying component of jet noise," J. Fluid Mech. **71**, 251–271 (1975).
[2] K. L. Gee, V. W. Sparrow, A. Atchley, and T. B. Gabrielson, "On the perception of crackle in high-amplitude jet noise," AIAA J. **45**, 593–598 (2007).
[3] K. L. Gee, S. H. Swift, V. W. Sparrow, K. J. Plotkin, and J. M. Downing, "On the potential limitations of conventional sound metrics in quantifying perception of nonlinearly propagated noise," J. Acoust. Soc. Am. **121**, EL1–EL7 (2007).
[4] K. L. Gee, V. W. Sparrow, M. M. James, J. M. Downing, C. M. Hobbs, T. B. Gabrielson, and A. A. Atchley, "The role of nonlinear effects in the propagation of noise from high-power jet aircraft," J. Acoust. Soc. Am. **123**, 4082–4093 (2008).
[5] A. Marshall and P. Davies "A semantic differential study of low amplitude supersonic aircraft noise and other transient sounds," in *International Congress on Acoustics*, Madrid, Spain (September 2–7, 2007), available at http://www.sea-acustica.es/WEB_ICA_07/fchrs/papers/nla-08-006.pdf (Last viewed 2/16/2011).
[6] B. R. Glasberg and B. C. J. Moore, "A model of loudness applicable to time-varying sounds," J. Audio Eng. Soc. **50**, 331–342 (2002).
[7] ANSI S3.4-2007: *Procedure for the Computation of Loudness of Steady Sounds* (American National Standards Institute, Melville, NY, 2007).
[8] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. (Elsevier Academic Press, London, 2003), pp. 133–134.
[9] J. Rennies and J. L. Verhey, "Temporal weighting in loudness of broadband and narrowband signals (L)," J. Acoust. Soc. Am. **126**, 951–954 (2009).
[10] B. Pedersen and W. Ellermeier, "Temporal weights in the level discrimination of time-varying sounds," J. Acoust. Soc. Am. **123**, 963–972 (2008).
[11] D. Oberfeld, "The temporal weighting of the loudness of time-varying sounds reflects both sensory and cognitive processes," in *Fechner Day 2009. Proceedings of the 25th Annual Meeting of the International Society for Psychophysics*, Galway, Ireland (2009), available at http://www.staff.uni-mainz.de/oberfeld/downloads/oberfeld_2009_Fechner_Day.pdf (Last viewed 2/16/2011).
[12] E. Zwicker, "Dependence of post-masking on masker duration and its relation to temporal effects in loudness," J. Acoust. Soc. Am. **75**, 219–223 (1984).
[13] J. Rennies, J. L. Verhey, J. Chalupper, and H. Fastl, "Modeling temporal effects of spectral loudness summation," Acta Acust. Acust. **95**, 1112–1122 (2009).
[14] See supplemental material at http://dx.doi.org/10.1121/1.3569710 Document No. E-JASMAN-129-008105 for modulated rephased waveform. For more information see http://www.aip.org/pubservs/epaps.html.