INCREASING EFFECTIVENESS OF TRAIN HORNS WITHOUT INCREASING INTENSITY

Frank A. Russo
Department of Psychology
University of Toronto
Mississauga, Ontario
L5L 1C6
frusso@utm.utoronto.ca

Michael E. Lantz
Department of Psychology
Queen's University
Kingston, Ontario
K7L 3N6
lantzm@psyc.queensu.ca

Gordon W. English
Transys Research
682 Milford Drive
Kingston, Ontario
K7M 6B4
transys.rtg@cogeco.ca

Lola L. Cuddy
Department of Psychology
Queen's University
Kingston, Ontario
K7L 3N6
cuddyl@psyc.queensu.ca

ABSTRACT

This study assessed methods of increasing the effectiveness of train horns without increasing intensity. Special attention was paid to the problem of masking by car noise. In Experiment 1, the audibility of pure tones presented in car noise was assessed across a broad range of frequencies. In Experiment 2, the urgency of pure tones presented in car noise was assessed across a broad range of frequencies. In Experiment 3, urgency of novel train horns was assessed in quiet. Novel train horns varied with regard to spectral centroid and musical dissonance but not intensity. Findings suggested that train horns could be made more effective by ensuring substantial mid-frequency energy, shifting the spectral centroid higher, and increasing musical dissonance.

1. INTRODUCTION

Train horns are an important instrument of public safety. They serve as both an early warning and emergency alarm; correspondingly, there are two basic requirements of train horns. First, that they be audible from enough distance so that pedestrians and motorists receive ample warning. Second, that they convey enough urgency so that said warning is compelling enough to motivate corrective action[1] (i.e., stop, curb or reverse trajectory).

Both these requirements have traditionally been met, in large part, through intensity. Beyond the basic requirement that train horns sound similar to the steam whistles from which they evolved, intensity levels are often the only feature of train horns that is federally mandated (e.g., the federal standard in the U.S. is a minimum of 96 dB A at 100 feet[2]).

Despite the potential public safety advantages that would be conferred by increasing the minimum intensity level of train horns, there are at least two good reasons for not doing so. First, an increase in intensity would result in a rise in the already high incidence of hearing impairment amongst train crews[3]. Second, an increase in intensity would agitate growing concerns about community noise[4]. An alternative means of increasing effectiveness of train horns involves manipulating the spectral characteristics of the train horn (e.g., spectral centroid and musical dissonance). The current study sought to investigate what spectral changes might confer the best advantage on audibility and urgency of train horns.

Special attention was paid to the problems associated with car noise. A U.S. National Transportation Safety Board study[5] revealed that 27 of 85 train/car accidents in 1985 were caused in part by car noise compromising train horn effectiveness.

2. EXPERIMENT I: PURE-TONE AUDIBILITY IN CAR NOISE

2.1 Introduction

Audibility of pure tones was assessed in the presence of car noise. Signals were pure tones at eight frequencies and six intensity levels relative to the car noise. The eight frequencies spanned the frequency range in which substantial energy (>75 dB SPL) may be generated by currently manufactured train horns (261-4000 Hz).

2.2. Method

2.2.1. Participants

Nine participants were recruited from the Queen's University community. These participants included four females and five males ranging in age from 18 to 40 years with a mean age of 25.8 years. All participants reported normal hearing and were able to detect all test signals when presented in quiet. Participants received course credit or cash payment for their participation.

2.2.2. Apparatus

Testing took place in a sound attenuated chamber. Participants were seated on an adjustable stool positioned in the middle of the chamber and in front of a response terminal. The car noise was carried by two loudspeakers positioned on stands 130 cm from the participant at 45 degrees to the left and right of the neutral plane (the participant's line of fixation), and 45 cm above the floor. The test signal was carried by a single loudspeaker positioned on a stand 150 cm from the participant at 90 degrees to the left of the neutral plane, and 75 cm above the floor.

The car noise (mask) was modeled after the average spectrum of car noise as determined from recordings made inside seven cars. These recordings were heavily biased toward lower frequencies, with an approximate roll-off of 6 dB per octave. The car noise was created with an IBM Pentium III running CoolEdit[6] software by iteratively filtering white noise until the generated spectrum at the participant's head approximated the average spectrum to within 0.5 dB SPL in any 1/3-octave band in the frequency range spanning 31.5-10000 Hz. For playback, the car noise was generated on a Macintosh Power PC running SoundEdit16[7] software and projected through a NAD 3020A amplifier connected to B&W loudspeakers. The overall intensity level of the car noise at the participant's head was approximately 85 dB SPL.

2.2.3. Signals

Fourty-eight 1-sec pure-tone signals spanning eight frequencies (261, 400, 500, 562, 630, 1000, 2000, 4000 Hz) and six intensity levels relative to the car noise (0, -3, -6, -9, -12, -15 dB) were tested. Signals were presented using Judge[8] software running on an IBM Pentium with output via 16-bit sound card to a Fostex 6301 speaker monitor.

2.2.4. Procedure

The height of the stool was adjusted such that the distance from the ground to the midpoint of the participant's head was approximately 120 cm. Participants were told that they would be detecting signals in car noise. To familiarize participants with the task, the experimenter provided feedback regarding a sample of signal and no-signal trials that were presented with the car noise. Participants were told that they would be asked whether a signal was present after each trial, and to expect an equal number of trials with and without a signal. Responses were entered on the response terminal. Trials were in random order for each participant with each of the 48 signals presented 10 times (480 signal trials), plus 480 no-signal trials ("catch trials"), for 960 trials.

2.4. Results and Discussion

The false alarm rate (i.e., ratio of "yes" responses to no-signal trials) was very low (M = .025) with no participant exceeding .01. Because of the low false-alarm rate, subsequent analyses only considered hit rates (i.e., ratio of "yes" responses to signal trials).

In Figure 1, the hit rate collapsed across participants is plotted as a function of frequency and signal-to-noise ratio. Hit rates were subjected to a within-subject analysis of variance. The main effect of intensity was significant, $F(5, 40) = 19.30$, $p < .0001$. As expected, hit rates were best for higher signal-to-noise ratios. The main effect of frequency was significant, $F(7, 56) = 5.89$, $p < .0001$. As seen in Figure 1, hit rates were best for mid-frequency signals, peaking at 562 Hz.

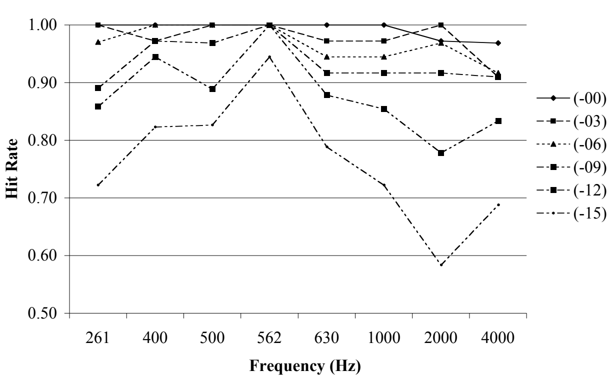

Figure 1: *Hit rate for pure tones presented in car noise plotted as a function of frequency for each signal-to-noise ratio*

There was also a significant interaction between intensity and frequency, $F(35, 280) = 2.51$, $p < .0001$. This interaction confirms what may clearly be observed in Figure 1; the advantage of mid-frequency signals is amplified for low signal-to-noise ratios.

Mid-frequency signals may have had an advantage in low signal-to-noise ratios because of a compromise between two psychoacoustic factors. First, threshold levels are lower for high-frequency signals, particularly those that fall between 2000-4000 Hz.[9] Second, high-frequency signals are particularly vulnerable to masking in the presence of broadband noise.[10,11] Such masking would have been further amplified by the spectrum of car noise, which was heavily weighted toward the lower frequencies. Thus, mid-frequency signals may have led to higher hit rates because they were more audible than low-frequency signals but less vulnerable to masking than high-frequency signals. An implication of this finding is that train horns with a strong mid-frequency component may be more audible in car noise than train horns of comparable intensity without such a component.

3. EXPERIMENT II: PERCEIVED URGENCY OF PURE-TONES IN CAR NOISE

3.1. Introduction

Previous tests have shown that the perceived urgency of pure-tone signals is positively correlated with a signal's frequency and intensity[12]. The purpose of the current experiment was to assess such variability in the presence of car noise. Signals were pure tones at four frequencies (500, 1000, 2000, 4000) and four intensity levels at or above threshold in car noise. Train horns were only presented at super-threshold levels in order to examine effects on urgency independent of masking.

3.2. Method

3.2.1. Participants

Eleven participants were recruited from the Queen's University and Kingston (Ontario) area. These participants included 8 females and 3 males ranging in age from 17 to 39 years with a mean age of 25.5 years. All participants reported normal hearing and were able to detect all test signals when presented in quiet. Participants received course credit or cash payment for their participation.

3.2.2. Apparatus

The testing environment and car noise were identical to those described in Experiment 1. The 12 test signals consisted of all combinations of four frequencies (500, 1000, 2000, 4000 Hz) and three intensity levels relative to threshold (0, 3, 6 dB). The criterion for threshold at a given frequency was a hit rate of .9 or better as determined in Experiment 1. For example, a review of Figure 1 reveals that threshold for a 1000 Hz signal was –9 dB.

3.2.3. Procedure

As in Experiment 1, the height of the stool was adjusted such that the distance from the ground to the midpoint of the participant's head was approximately 120 cm. The task was free-modulus magnitude estimation for urgency[13]. Responses were made using an adjustable slider presented on a video display terminal. The slider continuum was labeled with low urgency on the far left and high urgency on the far right. Participants were asked to make their slider adjustments proportional to urgency as it was perceived. There were two trials for each signal. The order of trials was independently randomized for each participant.

3.3 Results and Discussion

All ratings were given an integer value between 1 and 100 relative to the final position of the slider on a given trial. The urgency ratings were subjected to a within-subject analysis of variance. The main effect of frequency was significant, $F(3, 30) = 7.55$, $p < .001$. Higher ratings of urgency were associated with higher frequencies. The main effect of intensity relative to threshold was also significant, $F(2, 20)$. Higher ratings of urgency were associated with higher intensities relative to threshold. The interaction between the two factors was not significant. An implication of these findings is that if train horns are presented at intensity levels above threshold, urgency should be positively correlated with spectral centroid.

4. EXPERIMENT III: PERCEIVED URGENCY OF NOVEL TRAIN HORNS

4.1 Introduction

The objective of this experiment was to determine how perceived urgency varies in novel train horns. Train horns varied with regard to spectral centroid, dissonance, and number of flutes, but were controlled with regard to intensity and bandwidth. Train horns were presented in quiet in order to examine effects on urgency independent of masking.

4.2. Method

4.2.1. Participants

Fourteen participants were recruited from the Queen's University and Kingston (Ontario) area. These participants included 9 females and 5 males ranging in age from 17 to 45 years with a mean age of 24.5 years. All participants reported normal hearing. Participants received course credit or cash payment for their participation.

4.2.2. Apparatus

The testing environment was identical to that described in Experiment 1. Seven stockyard flute recordings provided the raw stimulus ingredients. All flutes had comparable spectra but varied with respect to fundamental frequency: 311, 370, 415, 440, 470, 494, 622 Hz. The intensity of each flute was adjusted such that intensity from the vantage point of the participant was equalized (approximately 66 dB C). The intensity-adjusted flutes were mixed together to form 23 novel horns. The 23 novel horns represented all 3, 4 and 5 flute combinations that were possible if the frequency bandwidth of horns was fixed. The frequency was fixed by imposing a simple constraint: the lowest flute in any novel horn was always 311 Hz and the highest flute in any novel horn was always 622 Hz. Although novel horns were equalized with respect to loudness and bandwidth, they varied with respect to other important characteristics – in particular, spectral centroid, musical dissonance, and the number of flutes.

The spectral centroid for each horn was determined by calculating the amplitude-weighted mean of its average spectrum. Musical dissonance of each novel horn was determined in three steps. First, the component intervals generated by the novel horn were determined. Second, a dissonance value for each component interval was calculated by collapsing across the corresponding values from four scales of musical dissonance[14]. Third, the average value of musical dissonance, collapsed across values for component intervals, was interpreted as the musical dissonance value for the novel horn. To corroborate the validity of the musical dissonance value, four musically trained listeners were asked to rate the musical dissonance in each novel horn. The average value of their ratings was well correlated with the computed musical dissonance values, $r(21) = .82$, $p < .0001$.

4.2.3. Procedure

The procedure was identical to that described in Experiment 2.

4.3. Results and Discussion

All ratings were given an integer value between 1 and 100 relative to the final position of the slider on a given trial. Mean urgency ratings ranged from a minimum of 37.40 (SE = 6.12) to 72.02 (SE = 4.68). The relative influence of spectral centroid and musical dissonance on perceived urgency was evaluated by subjecting mean urgency ratings to a multiple regression analysis. The resulting model was significant accounting for 58% of the variance, $F(2, 20) = 13.54$, $p < .001$. This model suggests that perceived urgency in train horns may be increased by increasing the spectral centroid and increasing musical dissonance. The relative influence of spectral centroid (BETA = .67) and musical dissonance (BETA = .47) on perceived urgency is depicted in Figure 2. Adding number of flutes to the 2-factor model only led to a marginal increase in explained variance (4%).

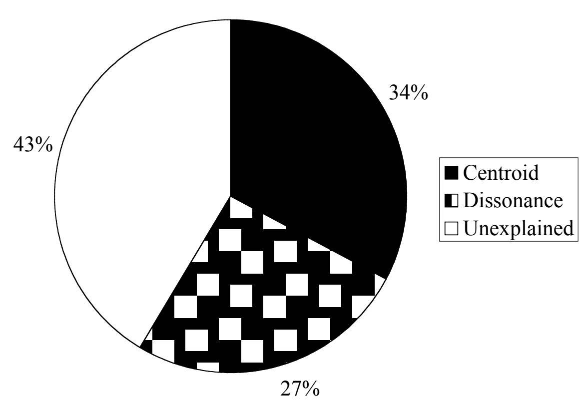

Figure 2: *Relative influence of spectral centroid and musical dissonance on perceived urgency in novel train horns*

To ensure that the results of this experiment were not due to properties of the test chamber or localization of sound, the experiment was replicated using binaural presentation over headphones. All flutes were equalized for intensity over headphones (approximately 66 dB C) and remixed to create 23 novel horns. The 2-factor model described above (spectral centroid and dissonance) was highly significant, accounting for 73% of the variance.

5. CONCLUSIONS

In Experiment 1, we found that audibility for pure tone signals presented in car noise was best for mid-frequency signals. The advantage of mid-frequency signals was amplified for low signal-to-noise ratios. In Experiment 2, we found that urgency for pure tone signals presented in car noise was positively correlated with both frequency and intensity. In Experiment 3, we found that urgency for novel train horns was positively affected by spectral centroid and musical dissonance.

In our view, these data suggest two basic proposals that may increase the effectiveness of train horns without increasing intensity. We recommend that the first proposal be implemented in any train horn and that the second be reserved for an auxiliary train horn to be used in emergencies.

Ensuring that one of the train horn's flutes has a fundamental frequency of approximately 562 Hz (to be adjusted for speed of train and corresponding Doppler effect) may improve its effectiveness. In addition to being particularly audible in car noise, energy in this frequency region is easily localized[15] and low enough to resist substantial absorption by the atmosphere[16]. Adoption of this proposal may assist in meeting the early warning requirement of train horns.

Raising the train horn's spectral centroid and increasing musical dissonance may improve its effectiveness. Adoption of this proposal may assist in meeting the urgency requirement of train horns. We would however, not recommend adopting this proposal in the standard horn -- doing so, would involve substantial changes to timbre, which may invite unintended consequences to do with aesthetics and familiarity. Rather, we would recommend that this second proposal be realized in the form of an auxiliary horn, to be sounded only in case of driver-identified emergencies.

This study was funded by the Natural Science and Engineering Research Council of Canada and by Transport Canada under its Direction 2006 grade-crossing safety initiative.

6. REFERENCES

[1] J. Edworthy, S. Loxley, and I. Dennis, Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors*, vol. 33, pp. 205-231, 1991.

[2] U.S. Department of Transportation, Federal Railway Administration, *Draft Environmental Impact Statement, Proposed rule for use of locomotive horns at grade crossings.* Washington, D.C., 1999.

[3] Atchison Topeka & Santa Fe Railway Company v. Stonewall Insurance Company, *Case No. 94-CV-1464*, Kansas Dist. Ct. July 24, 2000.

[4] Word Health Organization Report: *Community Noise*, edited by B. Berglund, T. Lindvall, and D. H. Schwela, 1999.

[5] U.S. National Transportation Safety Board. *Passenger/Commuter train and motor vehicle collision at grade crossing*. Safety Study, NTSB/SS-86/04. Washington, D.C., 1986.

[6] Syntrillium Software, *CoolEdit*, P.O. Box 62255, Phoenix, AZ 85082-2255, U.S.A., 2000.

[7] Macromedia Inc., *SoundEdit 16*, 600 Townsend Street, San Francisco, CA 94103, 2000.

[8] Judge SOUNDSWELL Signal Workstation, Nyvalla DSP AB, Åldermansvägen 19-21, SE-17148 Solna, Sweden, 1998.

[9] H. Fletcher and W. A. Munson, Loudness, its definition, measurement and calculation. *Journal of the Acoustical Society of America*, vol. 5, pp. 82-108, 1933.

[10] J. P. Egan and Hake, H. W., On the masking pattern of a simple auditory stimulus. *Journal of the Acoustical Society of America*, vol. 22, pp. 622-630, 1950.

[11] E. Zwicker, H. Fastl, *Psychoacoustics – Facts and Models*. Springer-Verlag: Berlin, Germany, 1990.

[12] E. Hellier, J. Edworthy, and I. Dennis, Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors*, vol. 35, pp. 693-706, 1993.

[13] E. Hellier, J. Edworthy, and I. Dennis, A comparison of different techniques for scaling perceived urgency. *Ergonomics*, vol. 38, pp. 659-670, 1995.

[14] C. L. Krumhansl, *Cognitive Foundations of Musical Pitch*. Oxford University Press: Oxford, England, 1990.

[15] S. S. Stevens and E. B. Newman, The localization of actual sources of sound. *American Journal of Psychology*, vol. 48, pp. 297-306, 1936.

[16] Acoustical Society of America, American National Standard: *Method of the calculation of the absorption of sound by the atmosphere*, ANSI S1.26, 1978.