EARCONSAMPLER: A TOOL FOR DESIGNING EMOTIONAL AUDITORY DRIVER-VEHICLE INTERFACES

Pontus Larsson

Volvo Technology
Humans, Systems and Structures
6310, M1.6, SE-40508 Göteborg, Sweden
pontus.larsson@volvo.com

ABSTRACT

EarconSampler is a simple tool for designing and modifying auditory driver-vehicle interfaces. It allows for creating melodic patterns of wav-snippets and easy adjustment of parameters such as tempo and pitch. It also contains an analysis section where sound quality parameters, urgency and emotional response to the sound is calculated / predicted, so that the user directly can see how a certain parameter affects perception and emotional response.

1. BACKGROUND

Designing sounds for use in automotive human-machine interfaces (HMIs) is not an easy task. A modern car or truck is a highly complex product involving many different electronic warning and information systems which must be interfaced to the driver in the most efficient way to ensure safety and usability. When correctly designed however, HMI sound has a great potential in improving traffic safety by e.g. reducing reaction times, increasing understandability and awareness and improving attention direction. Nonetheless, there are a number of different aspects to deal with when designing sounds for an automotive HMI. The aspects of automotive HMI sound design we consider as being the most important ones in this context are summarized below.

Urgency. A central characteristic of any sonic alarm is that it should convey the appropriate level of urgency, i.e. an "… indication from the warning itself as to how rapidly one should react to it." ([1], p207). A sound which e.g. signals that the vehicle is low on windshield wiper fluid should not make the driver depress the brake as soon as possible while an imminent collision warning system sound in most cases should. Too urgent sounds may, apart from lead to the wrong behavior, also cause severe annoyance which in turn may lead to that the sound is switched off (if possible). Too low urgency may lead to unnecessarily slow reaction or that the warning is neglected. When it comes to sonic alarm design there are a number of auditory parameters which have been found to influence the perceived urgency. Hellier et. al. [2] found that speed, number of repeating units, fundamental frequency, and inharmonicity all affect the perceived urgency of a warning sound, with speed being the most effective in altering the urgency and inharmonicity being the least effective. Moreover, fast onset of tones, unpredictability and irregularities have also been found to increase urgency. Other musical attributes of the signals such as pitch contour have been shown to have little effect on urgency but is suggested to sounds easier to discriminate and to recall. Of several possible parameters available, one has found that loudness is one of the stronger cues for urgency [3-4]. However, the range within which loudness can be varied before the sound becomes un-ergonomic is in practice rather small; the sound should of course be loud enough to be heard over the background noise in the operator's environment and quiet enough not to cause annoyance or hearing impairment.

Emotional response. We believe that the driver's emotional response is important to consider in the design of automotive HMI sounds. Emotional response is also tightly related to the perceived degree of urgency. For this reason, the Emotion Reaction Model (ERM) framework [5] was developed. The ERM framework builds on neuropsychological research showing that the human brain automatically reacts to certain sound properties, either in a very fundamental way (approach/avoidance reactions linked to survival) or by activating associative networks (primes/memories from previous exposures to situations where the sound was experienced). Most importantly, the framework suggests that emotional or affective reactions are the driving force of behavior or action. Thus for a warning or information sound to elicit the "correct" action, it needs to induce an emotional reaction.

Quality. Naturally, a certain degree of sound quality is often desired – especially for often reoccurring sounds – to give the driver an overall high quality impression of the product. Which parameters control quality or how a high quality impression can be achieved is difficult to say in general, but a good starting point may be to quantitatively evaluate sound quality measures such as sharpness and roughness. One also often has to consider the playback system's quality. The amplifier(s) / loudspeaker(s) through which the sounds are played may be of better or worse quality in terms of frequency response, distortion, max SPL etc. Moreover, storage capacity and sound format parameters such as sampling frequency and bit depth may also be limited. Unless the audio system is full range, sounds have to be designed with the audio system in mind.

Consistency with visual and other displays. Sounds in driver-vehicle interfaces are often used in conjunction with visual and sometimes haptic displays. Usually the same or similar information is then presented over the displays, to enhance the message, improve reaction time and to accommodate to hearing impaired drivers. The look and feel of the interface across modalities must then be consistent.

2. THE EARCONSAMPLER SOFTWARE

EarconSampler is a simple sample sequenser/player application implemented in Matlab™ which can be used to test and design/modify earcon sound snippets (brief musical sounds that may be used for HMI purposes). The purpose of this software is to provide a tool with which prototype sounds can be easily and rapidly developed and/or modified for further use in user studies or for the final product. Although EarconSampler was designed primarily for automotive applications, it is not by any means limited to these kinds of displays but can used as a general sound design/analysis tool.

EarconSampler's built-in algorithms will calculate objective sound quality values (so far loudness and sharpness has been implemented) of the designed sound and predict the emotional reaction (activation/valence) and urgency to the sound. It is the intention that with this tool, the designer will be more guided by research results and can easily get an understanding how the designed sound will work in reality. Moreover, people who are not familiar with sound design may be involved in the total HMI design and may therefore have opinions on how the sound should sound. With the tool, they can easily try different variations to a certain sound, and give suggestions on modifications to the sound designer.

Short wav file segments ('patches') are required as input to the software as well as a 'melody' scheme which resembles a midi-file – e.g. it contains information on how the 'patches' should be played. The patches can be created in any sound synthesis software, recorded or obtained by other means. Different modifications to the melody (i.e. the earcon itself) can be easily done via EarconSampler's GUI: tempo, pitch, the patch's decay time, level and the length of the patch. EarconSampler also has an analysis functionality which visualizes the earcon in temporal and frequency domains.

2.1. 2. The Graphical User Interface

The Graphical User Interface (GUI) of EarconSampler, shown in figure 1 below, was implemented using Matlab's GUIDE interface design tool. It contains basically two sections; the upper one containing five sliders, some radio buttons and a drop down menu is the earcon design/modifcation section, and the middle/lower one containing graphs and push buttons is the render/analysis section. There is also a menu section which allows loading new patches or melodies and setting some plot options. Via the File menu, one can also save the current earcon to a wavefile.

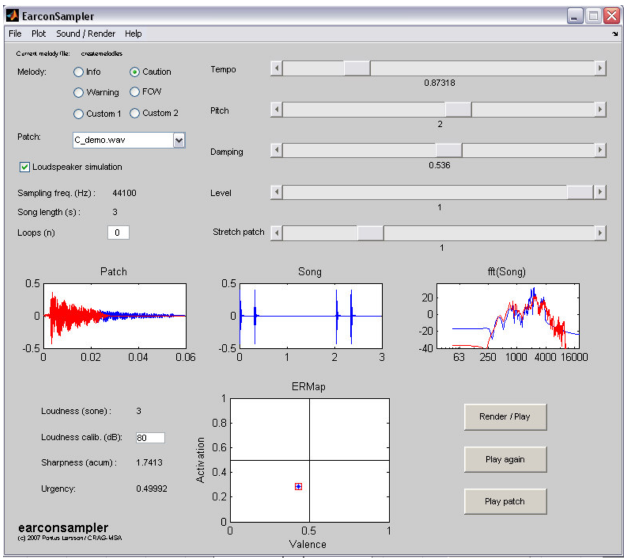

Figure 1. *Screenshot of the main GUI*

2.2. The design/modification section

In the upper left corner of the GUI one finds six different radio buttons with which one can set the melody scheme for the earcon (the 'song'). It is possible to change the melodies themselves by editing textfiles containing melody information which can be loaded via the File menu. Below the radiobuttons there is a drop down menu where one selects the desired wavefile which will be used for the 'patch'. A default patch list is always loaded but it is possible to add/change patches by using the File/Add patch menu item. The sampling frequency of the selected patch and the resulting song length (which naturally varies with tempo, see next paragraph) is also shown.

The four sliders found to the right in the design/modification section control Tempo, Pitch, Damping, Level and Patch stretch (the length of the patch) respectively. The current value of each parameter is also shown below each slider.

Tempo adjusts the time intervals between each 'note on' in the melody scheme. A tempo of 1, which is the default value, corresponds to unaltered tempo (e.g. as given in the melody scheme), higher or lower values scales the note onset intervals with the inverse of those values.

The pitch slider adjusts the fundamental pitch of the melody in semitone steps. Consequently, a pitch value of –12 or 12 lowers or raises the pitch by one octave. A pitch value of 0 gives an unaltered pitch. The algorithm which handles these pitch shifts also retains the length of the patch/melody.

Damping is a parameter which acts upon the patch itself. This parameter decreases the decay time of the patch by multiplying the patch with an exponentially decaying window. This can be useful if the patch is too long, making consecutive notes sound "on top" of each other. In figure 1 can be seen the result of such damping; the red curve in the "patch" graph shows the patch after windowing. The damping slider ranges from 0 to 1 where 0 corresponds to no window (no damping) and 1 corresponds to a decay of 60 dB / patch length.

The level slider naturally controls the level of the melody and ranges from 0-1 where 0 means no sound at all and 1 means full amplitude.

Finally, the stretch patch slider alters the length of the patch without changing the pitch, which could be useful for e.g. speech patches.

2.3. The render/analysis section

In the lower half of the GUI one first of all finds three different plots which display the temporal view of the selected patch and the result of the damping operation (plotted in red on top of the original patch), the whole song (middle plot), and the FFT of the whole song (right plot). The plots are updated after each press on the "render/play"-button.

Below these plots and to the left are shown some sound quality metrics (Loudness, Sharpness) and Urgency and as with the plots, these values are updated each time the "render/play" button is pressed. The loudness calculation (and hence the sharpness calculation, which is based on loudness) requires a calibration value to produce correct values, which is entered into the edit box below the loudness. If no calibration value is available and if one simply wants to compare loudness/sharpness of different earcons, any reasonable value could be given.

In the middle, lower half of the GUI the predicted emotional response from the sound in terms of Valence/Activation for the current earcon is plotted. The Valence axis goes from 0 (positive) to 1 (negative) and the Activation goes from 0 (low activation) to 1 (high activation). The calculation of the V/A values is based on previous research and extrapolations of this research [5-6]. In some cases there is an exact correspondence between the current settings (patch/melody/tempo etc) and settings used in previous studies; in these cases, the blue '*' marker is emphasized by a red square marker.

Urgency, shown below Loudness/sharpness is calculated as a linear combination of Valence/Activation values (see next paragraph) – i.e. activating and negative sounds will result in higher urgency values than non-activating, positive sounds.

Finally, to the right one finds four buttons with which one can 1) create and play the song with the current settings ('render/play'), 2) play the most recently rendered song ('play again'), and 3) play only the patch ('play patch') –without damping.

3. CONCLUSIONS

The EarconSampler application has so far been successfully used both in focus groups meetings for the purpose of evaluating and improving sound designs, and for actual editing and design of sounds (in conjunction with sound synthesis software). In a next version we also intend to include a more refined editing module in which parameters such as sample envelope (Attack/Decay/Sustain/Release), sample starts/stops, and frequency filtering can be easily adjusted. Another useful addition would be a visual vehicle interior module of the where the sound can be played together with the visual information presented by e.g. LEDs, information displays and HUDs and perhaps also some traffic scenarios can be simulated. This would provide a first indication whether the sound fits the context and other information and allow for matching the sound to visual impressions.

4. ACKNOWLEDGEMENTS

This research was supported by the Swedish Foundation for Strategic Research (SSF).

5. REFERENCES

[1] J. Edworthy and E. Hellier, "Complex nonverbal auditory signals and speech warnings", in M. S. Wogalter (ed.) *Handbook of Warnings*. Lawrence Erlbaum Associates, Mahwah, NJ, 2006.

[2] E. J. 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, no. 4, pp. 693-706, 1993

[3] K. L. Momtahan, *Mapping of psychoacoustic parameters to the perceived urgency of auditory warning signals.* Unpublished master's thesis. Carleton University, Ottawa, Ontario, Canada, 1990.

[4] E. C. Haas and J. G. Casali "Perceived urgency and response time to multi-tone and frequency modulated warning signals in broadband noise," *Ergonomics*, vol. 38, no. 11, pp. 2313-2326, 1995.

[5] D. Västfjäll, A. Sköld, A. Genell, P. Larsson, A. Tajadura, A. Väljamäe, and M. Kleiner "Emotional acoustics: A new framework for designing and measuring the effects of information and warning sounds," Manuscript submitted for publication.

[6] A. Sköld "Integrative Analysis of Perception and Reaction to Information and Warning Sounds in Vehicles", Chalmers University of Technology, Ph D Thesis, 2008.