

MEANINGS CONVEYED BY SIMPLE AUDITORY RHYTHMS

Henni Palomäki

University of Jyväskylä
Department of Computer Science and Information Systems
P.O. Box 35 (Agora), FIN-40014 University of Jyväskylä, Finland
hepapalo@cc.jyu.fi

ABSTRACT

In this article we concentrate on perception of non-musical rhythm. The purpose of this study has been to find possible meanings related to simple auditory rhythms. Meanings were examined using semantic scales. 26 subjects rated nine different rhythm samples according to adjective pair scales. We also identify some preliminary design suggestions as to how rhythm can be used in sonification and discuss duration limitation when composing earcons.

1. INTRODUCTION

This article addresses the question of meanings conveyed by rhythm. The examination is done in the context of earcon design. Earcons are non-verbal sound messages used in the user-computer interface to provide information and feedback to the user about computer object, operation or interaction [1]. They cover all non-speech audio signs that do not directly imitate sounds of real-world events [2]. The benefits of the earcons are based on the fact that they can convey all the information that the user needs to understand a message. Although earcons are symbolic and their meanings are not assumed to be known without learning, it is only a relief to the users if the mapping between meaning and signal is as natural and intuitive as possible. The conveyed message should also be able to be interpreted consistently among users.

Rhythm has been seen as an important element in earcon design and some guidelines are also presented [1, 3]. However, thorough conceptualization of rhythm is difficult. Definitions are presented by many authors but none of them are widely accepted. This difficulty is probably derived from the fact that rhythm is a compilation of many unclear constituents. An attempt can be made to understand rhythm both in terms of the physical properties of the stimulus, from the perspective of rhythm perception or as a property of the notation system [4]. We may conceptualize rhythm as the perception of the grouping and ordering of elements. In a rhythmic auditory pattern some elements are temporally redundant. According to Fraisse [5] there is a rhythm when we can predict on the basis of what is perceived what will follow.

Psychoacoustic dimensions that are often identified to be characteristics of sound are pitch, duration, timbre and loudness. Rhythm is the most strongly related to the dimension of the duration and its opposite silence because they segment elements in time. Two close sounds are perceived to belong to the same group if the interval is under approximately 1.8-2 sec. When intervals are increased above that, there are only isolated notes and the rhythm disappears. [6]

Also other dimensions interact with rhythm perception. The rhythmic grouping can also be a result of loudness, pitch and timbre [4]. Explanations of the formation of the rhythm can be found from grouping principles proposed by the Gestalt psychologists. Both Moore [7] and Deutsch [8] have given an account of applying these to the arrangement of individual auditory elements into “chunks”. Also guidelines for rhythmic organization derived from rhythm studies [4] are helpful when designing earcons.

2. RHYTHM AS A DESIGN ELEMENT OF EARCONS

Existing taxonomies that could describe meanings conveyed by the non-musical rhythm are piecemeal. Suggestive results can be derived from academic experiments where rhythm has been only a secondary subject of the study. There are a lot of studies on meaning in the area of music but these are only partly applicable to the earcons. Meaning in music concerns mainly the music piece as whole and larger-scale properties of the form.

In spite of the slight body of research, rhythm has been seen as one important design element of earcons. For example, Blattner, Sumikawa and Greenberg [1] see rhythm as one of the most prominent building elements of the earcons because listeners can respond to it more readily than to any other musical parameter. Rhythm is also seen as a means to differentiate earcons. Confusion can be avoided by having the designer put a different number of sounding elements in every created earcon. [9]

One important source of the studies related to meaning of symbolic sounds can be found in the area of auditory alarm studies. Patterson [10] noticed that the warning should be composed of a distinctive temporal pattern to minimise the probability of confusion among different warnings. The rhythm and tempo or speed of the signals has been noticed to have also a clear impact on the perceived urgency of the alarm [11, 12, 13]. Acoustically more urgent warnings produce faster responses to important, hazardous situations and events than those that are less acoustically urgent [11]. Signals with no inter-pulse intervals (0 ms.) were rated most urgent of all [13]. Other factors affecting perceived urgency suggested were speed and predictability of the structure of the sound [11].

3. EXPERIMENT

The experiment was designed to provide information about the meanings that different rhythmic sounds can convey. The experiment utilized the Semantic Differential developed by

Osgood [14]. It is originally a method for measuring the meaning of concepts [15]. In this study the method was used to clarify the semantic properties of the simple rhythmic auditory stimulus.

The method uses adjective pairs to represent polarities in a semantic space which subjects are asked to rate their view on sound that is being studied (Figure 1). A scale like this measures both directionality of a reaction (e.g., familiar versus strange) and its intensity (slight through extreme).

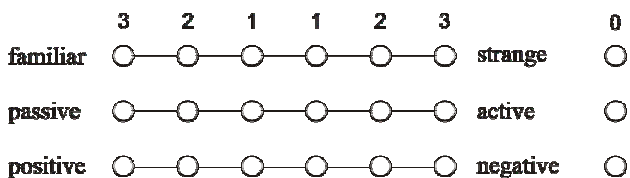


Figure 1. Example of scale used in Semantic Differential.

The design phase consisted of two parts: selecting polar adjective pairs for the scales and composing sample sounds. 13 adjective pairs were selected as potential characteristics of the sounds. These polar adjective pairs were: familiar – strange, passive – active, positive – negative, uneven – even, pleasant – unpleasant, intense – mild, easy – difficult, excited – calm, upward – downward, dangerous – safe, stable – dynamic, backward – forward and fast – slow. Adjectives were selected to represent meanings that can be used in user interface design. What adjectives or meanings are suited, for example, to help users in navigation? User interface is considered here freely. We did not want to be restricted here only to desktop computer applications. Also adjectives used in past research on similar types of sounds were also taken into the account [16, 17, 18, 13]. The scale position marked 3 meant "extremely", the position 2 "quite," and position 1 were labelled "slightly". Position 0 was labelled as "irrelevant". The "irrelevant" category was provided for each scale so that subjects could mark an adjective as irrelevant if they felt that it did not apply to a stimulus at all. The order of the scales on the scale sheet was randomized so that there were four formats of the scale sheets.

3.1. Stimuli

Nine sounds were created for the experiment using Ableton Live software. We decided not to use complex musical rhythms in the experiment because it would hinder the drawing of a conclusion between acoustic properties and perceived meaning. Thus test sounds were very simple and uninteresting from a musical point of view. Signals represented different rhythms and they were roughly based on the grouping principles suggested by Handel [4] and studies of rhythm perception conducted by Fraisse [5, 6]. However, these gave only general advice. When choosing rhythms, trial and error was used because there were no precise rules on which a choice could be based. In general, we aimed to create sounds that

- are clearly diverse
- are as short as possible
- have different tempos and
- use different rhythm perception rules

Only rhythmic characteristics of the sounds were varied: pitch, timbre and loudness were kept constant. In all sounds the timbre was a General MIDI instrument number 12, a vibraphone. The vibraphone was chosen because it can be assumed that it is not familiar to subjects and thus it does not convey any particular meaning to the subjects. It does not have clear pitch but the duration of the elements can be perceived (how long a note is pressed down).

Loudness was set at a comfortable listening level and it was also kept constant during the whole sound. Subjects' grouping was supported by the length of elements and rests, not with intensity accentuation. Pitch was also kept at the level of note number 60 (C5). Finally Midi notes were converted to wav files. Test signals are depicted in more detail in Table 1.

Duration of the signals varied between 1.65 and 5.2 seconds. According to guidelines suggested by Brewster, Wright and Edwards [3] earcons should be kept as short as possible. Short sounds can keep up interaction. We also endeavoured to follow this rule when creating test sounds.

At first sight it seems that there are practically no limitations on possible musical rhythms. However, if duration of the sound is restricted to a few seconds, alternatives will diminish substantially. A human being needs repetition of elements to perceive rhythmic structure and that need lengthens the duration of the sound. Especially in slow tempo rhythms the amount of repetitive elements reduces and perception of the structure becomes more difficult. But as has already been pointed out, limited duration of the sounds is practical because in many user interfaces, sounds have to be short. Users have not time to listen to long sounds because the auditory channel and attention are needed soon for other tasks.

Sound label	Tempo	Duration
"Accelerating"	~ 4/ sec.	2 sec.
"Decelerating"	~ 4/ sec.	2 sec.
"Steady and slow"	~0.77/ sec.	5,2 sec.
"Steady"	~ 1.8/ sec.	2,25 sec.
"Steady fast"	~ 5.4/ sec.	1,65 sec.
"No rests"	~ 6/ sec.	2 sec.
"Irregular"	~ 5/ sec.	2 sec.
"Front stressed"	~ 3,5/ sec.	2,25 sec.
"Back stressed"	~ 3,3/ sec.	2,25 sec.

Table 1. List of the nine rhythms used in the study. Tempo describes how many singing elements the sound has per second. Graphical descriptions of the sounds are not to a common scale.

3.2. Subjects

Twenty-six subjects participated in the experiment: 16 females and 10 males with an average age of 31.5 years, ranging from 13 to 58 years. Participants had different backgrounds but all of the subjects were non-musicians. A musician was defined as a person who has performed music more than ten years or whose profession or studies aiming at a profession are related to music. All subjects were asked whether they were aware that they had any hearing difficulties and none reported such problems. No further audiometric tests were performed; they were believed to be unnecessary, in that the stimuli were presented at a constant intensity level.

3.3. Procedure

Subjects were tested one at a time. At the beginning of the experiment, the participant filled out basic demographic information including age and sex. Then he was read the instructions and it was verified that the task had been understood. The subject was instructed to rate the test sounds on adjective scales according to how he perceived it or felt toward it. The ratings were filled in on a paper questionnaire. He was asked to mark the point on each scale where he thought the sound belongs. The scales consisted of adjective pairs that were presented in an earlier section. The instructions emphasized that the participant should rate each sound independently of previously rated sounds. He was also told that there was no right or wrong way to rate sounds.

In the experiment the participant sat in front of a computer monitor that displayed a web page with buttons for every sound. The participant could hear the test sounds by clicking the buttons with the mouse. The sounds were listened to through American Acoustic 770M headphones at a comfortable loudness level. The participant was allowed to listen to all sounds as many times as he wished while he evaluated the stimuli. When the task was completed, the participant was asked some general questions about how he had found the task.

Sounds were presented to the subjects in four different orders. Also the adjectives were arranged on the questionnaire in four different ways. Each subject thus judged each of the 9 sounds along each of the 13 scales. Thus all participants made 117 evaluations in total. The experiment took from 13 to 25 minutes to carry out.

4. RESULTS

The answers on individual scales were converted to numerical quantities and treated statistically. Consistency between answers among participants was studied with Standard Deviation. Also correlation techniques were tried in order to determine interrelationships between ratings but only weak results were found. It was then decided to concentrate the analysis only on the adjectives most clearly associated with the sounds.

4.1. Consistency of answers between subjects

When trying to detect how consistently people can interpret a particular sound, one way is to look at the dispersion of the ratings. When sounds are used to convey information, it is very important that they are interpreted at least mainly equally.

Standard deviation was counted for every sound and scale. The smaller the SD, the more consistent was the interpretation

of the sound. In this data, if all subjects had rated the same sound on the scale 'easy – difficult' to be very, quite or slightly easy, the SD value would be 0.82. That could be considered a very consistent result in this kind of study because all the answers are then situated at the same end of the scale and considered to convey the same meaning in general.

Most consistently subjects rated the sound "Steady and slow" on the scales 'stable – dynamic' (0.697), 'fast – slow' (0.703), 'uneven – even' (0.838), 'excited – calm' (0.87). Other consistent ratings were given to the sound "Steady and fast" on the scale 'fast – slow' (0.578), the sound "Irregular" on the scale 'intense – mild' (0.732) and the sound "No rests" on the scale 'fast – slow' (0.752).

There also seemed to be some problems in the use of some scales. The biggest Standard Deviations occurred with the sound "Steady and fast" on the scales 'uneven – even' (2.321) and 'stable – dynamic' (2.2), the sound "Back stressed" on the scales 'stable – dynamic' (1.985), 'passive – active' (1.919) and 'uneven – even' (2.2), the sound "Accelerating" on the scales 'upward – downward' (1.976) and the sound "No rests" on the scales 'easy – difficult' (1.919) and 'pleasant – unpleasant' (2.015).

Respondents were allowed to evaluate sound as irrelevant in a scale. The irrelevant scales tell that the adjectives are not suitable descriptors for the sound. Some of these scales can even be called inadequate as a descriptor for simple auditory rhythms. The scale 'upward – downward' was marked as irrelevant for most of the sounds. Many participants commented that the scale was inappropriate because there were no changes in pitch of any of the sounds. It seems that the scale would be better when evaluating stimuli that have changes in pitch. The other two scales that were not useful in evaluating many sounds were 'familiar – unfamiliar' and 'backward – forward'. All in all some adjectives used in this rhythm study should be reconsidered.

4.2. Strongest meanings linked with sounds

Median values (Md) of the scales were calculated for every sound. We noticed that the median is better descriptor of the typical value than the mean because the median is less sensitive to extreme scores and distribution of the data was skewed to the right or left on every scale. Only the scales where the SD was smaller than 1.3 were accepted for following analysis. Also the scales that were marked irrelevant by more than 15% (N=26) of the participants were rejected.

The median value reveals the most typical association with the particular sound. (In the following description medians are: 3=very, 2=quite, 1=slightly). The rhythm "No rests" was evaluated to be fast (Md=3), excited (2.5), active (2), intense (2) and forward (2). Quite similar ratings were given to the rhythm "Irregular": it was described as active (2), uneven (2), intense (2), excited (2), forward (2) and fast (2). "Accelerating" was rated to be dynamic (2.5), uneven (2) and fast (2). "Decelerating" got only one clearly consistent judgment: It was rated to be uneven (2) as was its mirror sound "Accelerating". "Back stressed" was associated with upward (1), forward (2) and fast (2). Its mirror sound "Front stressed" did not receive any consistent ratings.

The rhythms "Steady and slow", "Steady" and "Steady fast" are very simple rhythms where one constant element is repeated after a constant interval. These kinds of rhythms can also be depicted as *isochronous pulse trains* [4]. It is interesting to compare ratings between these rhythms because they are performed at different tempos. Rests between sounding

elements were 0.5, 0.25 and 0.05 seconds. Figure 2. presents a visual profile analysis of these sounds. The profile shows that there is to some extent a consistent trend in association when the tempo increases or decreases. The faster the tempo, the stranger, more active, intense, excited, dynamic, dangerous and faster sound was associated.

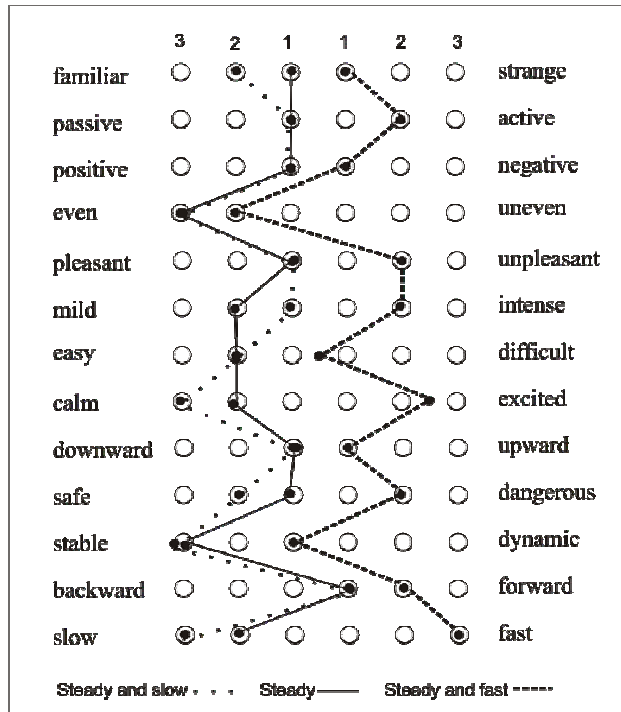


Figure 2. Profile of medians of three test sounds: “Steady and slow”, “Steady” and “Steady and fast”.

5. DISCUSSION

The experiment revealed a fair number of consistent ratings between subjects. However, it is safe to assume that there are types of rhythms which are more appropriate to indicate specific meanings than others. The rhythms might be perceived more consistently if they were longer. Now the duration of the sound was mainly restricted to a couple of seconds and interpretation of the rhythmic structure may be imperceptible to some listeners. The duration limitation of the created test sounds might have an impact in that most of the associations were related to meanings that are typically related to sounds with a fast tempo. But this study shows that many positive associations are related to slow tempo (less than 0.8 sounding elements per second) rhythms like familiarity, calmness and safety. This means that slow tempo rhythms are very important in sonification. As Blattner et al. [1] said “The length of an earcon should be sufficient to convey the message effectively, but no longer.”

This experiment concentrated only on simple elementary rhythms. A further experiment would be needed to test more complex sounds where rhythms are combined with some other auditory dimension like pitch or loudness. Although it would be all too naïve to assume that meaning conveyed by combinations of two auditory dimensions (for example rhythm and pitch) will be a sum of their semantic space studied individually, some

combinations may have a strengthening effect and make interpretation of the sound easier and more consistent.

The present study can give some preliminary ideas on how rhythm can be utilized when creating earcons. A summary of the associations related to each studied rhythm is presented in Table 2. When characteristics of the sounds are known designers can find information suggestive of how to convey a particular meaning.

This study utilized Semantic Differential. One problem in using semantic scales is the exhaustingness of the testing for the subjects. Rating is quite a monotonous task and might easily make subjects put marks without proper consideration. There is also another weakness in Osgood's method – there is a tendency to assume that the adjectives chosen mean the same to everyone. This study also showed that the most consistent results were given by adjectives that are simple to interpret like slow – fast and even – uneven.

6. CONCLUSIONS

This experiment examined the relationship between 26 adjectives and nine simple rhythmic sounds. The sounds varied in their tempo, amount of sounding elements and rate of predictability. The study showed that some adjectives like excited, active, intense and fast, were more often associated with fast tempo rhythms and some adjectives like easy, calm, safe and slow, with slow tempo rhythms. However some of the adjectives seemed to be poor at describing rhythmic sounds, like upward, downward, backward, forward, familiar and unfamiliar. This was also rather apparent from the subjects' verbal description.

	No rests	Irregular	Steady and slow	Steady	Steady and fast	Accelerating	Decelerating	Front stressed	Back Stressed
familiar	24	8	44	26	16	15	9	20	20
strange	16	39	7	11	18	27	27	18	19
passive	2	2	31	31	4	8	11	15	12
active	50	51	11	14	45	44	29	28	43
positive	15	7	34	30	7	17	14	18	24
negative	25	36	10	11	35	24	17	21	14
uneven	13	51	1	4	17	54	52	8	26
even	36	3	71	60	46	3	4	43	17
pleasant	27	5	37	30	7	15	13	19	20
unpleasant	19	51	9	9	48	28	27	20	19
intense	47	48	11	9	54	29	35	33	38
mild	3	0	33	30	2	10	8	11	9
easy	35	4	56	48	21	25	10	38	30
difficult	14	46	2	2	22	18	33	8	14
excited	56	52	1	4	60	40	39	30	44
calm	1	2	62	43	2	7	7	10	4
upward	19	21	8	7	20	35	10	9	19
downward	4	3	16	12	5	13	42	12	6
dangerous	26	39	4	8	42	24	21	18	24
safe	15	9	49	31	8	16	12	22	12
stable	36	6	68	61	34	3	5	47	30
dynamic	13	49	0	2	18	59	49	9	18
backward	2	2	7	10	8	4	36	11	2
forward	41	36	22	19	36	56	14	23	41
fast	62	55	0	10	67	37	31	34	44
slow	0	2	63	37	0	4	9	10	2

Table 2. Sum of the ratings of 26 subjects. The highest value in a row has been bolded (max. 78). "Front stressed" and "Back stressed" sounds do not seem to convey any particular meaning.

7. REFERENCES

- [1] M. Blattner, D. Sumikawa and R. Greenberg, "Earcons and icons: Their structure and common design principles," *Human Computer Interaction*, vol. 4 no. 1, pp. 11-44, 1989.
- [2] A. Pirhonen, E. Murphy, G. McAllister and W. Yu, "Non-speech sounds as elements of a use scenario: A semiotic perspective," in *Proceedings of ICAD 2006*, in press.
- [3] S.A. Brewster, P.C. Wright and A.D.N. Edwards, "An evaluation of earcons for use in auditory human-computer interfaces," in S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel & T. White, Ed., *Proceedings of InterCHI'93*, Amsterdam, ACM Press, Addison-Wesley, pp. 222-227, 1993.
- [4] S. Handel, *Listening: An Introduction to the Perception of Auditory Events*. MIT Press, Cambridge, Massachusetts, 1989.
- [5] P. Fraise, "Rhythm and tempo," in D. Deutsch, Ed., *The Psychology of Music*, Academic Press, New York, pp. 149-180, 1982.
- [6] P. Fraise, "Time and rhythm perception," in E.C. Carterette and M.P. Friedmans, Ed., *Handbook of Perception*, Academic Press, New York, vol. 8, pp. 203-254, 1978.
- [7] B.C.J. Moore, *An Introduction to the Psychology of Hearing*. Academic Press, San Diego, 2001.
- [8] D. Deutsch, "Grouping mechanisms in music," in D. Deutsch, Ed., *The Psychology of Music*. Revised edition. Academic Press, San Diego, pp. 299-348, 1999.
- [9] S.A. Brewster, "Non-Speech Auditory Output," in J.A. Jacko and A. Sears, Eds., *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, Lawrence Erlbaum Associates, Mahwah, N.J, 2003.
- [10] R.D. Patterson, *Guidelines for auditory warning systems on civil aircraft*, C.A.A. Paper 82017, Civil Aviation Authority, London, 1982.
- [11] J. Edworthy, S. Loxley, and I. Dennis, "Improving auditory warning design: Relationship between warning sound parameters and perceived urgency," *Human Factors*, vol. 33, no. 2, pp. 205-232, 1991.
- [12] 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.
- [13] J. Edworthy, E.J. Hellier and R. Hards, "The semantic associations of acoustic parameters commonly used in the

- design of auditory information and warning signals,” *Ergonomics*, vol. 38 no. 11, pp. 2341–2361, 1995.
- [14] C.E. Osgood, “The Nature and Measurement of Meaning,” *Psychological Bulletin*. Vol. 49, no. 3, pp. 197–237, 1952.
- [15] S. Isaac and W.B. Michael, *Handbook in Research and Evaluation: a Collection of Principles, Methods, and Strategies Useful in the Planning, Design, and Evaluation of Studies in Education and the Behavioral Sciences*. Edits, San Diego, 1987.
- [16] L.N. Solomon, “Semantic Approach to the Perception to Complex Sounds,” *The Journal of the Acoustical Society of America*, vol. 30, no. 5, pp. 421–425, 1958.
- [17] L.N. Solomon, “Search for Physical Correlates to Psychological Dimensions of Sounds,” *The Journal of the Acoustical Society of America*, vol. 31, no. 4, pp. 492–497, 1959.
- [18] L.N. Solomon, “Semantic Reactions to Systematically Varied Sounds.” *The Journal of the Acoustical Society of America*, vol. 31, no. 7, pp. 986–990, 1959.