inter.noise 2000
The 29th International Congress and Exhibition on Noise Control Engineering
27-30 August 2000, Nice, FRANCE

I-INCE Classification: 6.3

STRUCTURE OF AUDITORY SIGNS: SEMIOTIC THEORY APPLIED TO SOUNDS

B. Dürrer*, U. Jekosch**

* Institut für Kommunikationsakustik, Ruhr-Universität Bochum, Geb. IC 1/132, 44780, Bochum, Germany

** FB 3 Literatur- und Sprachwissenschaften, Universität GH Essen, Universitätsstraße 12, 45517, Essen, Germany

Tel.: +49-234-32-23481 / Fax: +49-234-32-14165 / Email: duerrer@ika.ruhr-uni-bochum.de

Keywords:
MAN-MACHINE INTERFACE, SEMIOTICS, SOUND QUALITY

ABSTRACT
Most technical equipment or machinery produces sound as an inevitable side effect of its function; in addition, sounds (auditory signs) are emitted as a conveyor of information. Both types of sounds are subject to research in the field of sound quality, for instance to enhance the overall product quality (e. g. product sound design) or to improve the efficiency of an auditory man-machine interface (e. g. warning sound design). For man-machine interfaces, apart from conveying the desired information, the designer has to face the problem of distinctiveness of auditory signs. In order to investigate the process of interpretation of non-verbal auditory signs, listening experiments have been performed where properties of the presented sounds have been varied. The performance of the test subjects is discussed with respect to a model based on semiotic theory.

1 - INTRODUCTION

One of the oldest definitions of signs is *aliquid stat pro aliquo* [1]: This definition only considers the sign vehicle (i. e. the sign's representation in the physical world) and the object or concept to which it refers. De Saussure's semiotic concept is based on a similar dyadic relation which distinguishes between the signifier and the signified: However, in de Saussure's concept the signifier is not the sign vehicle in the physical world (e. g. the sound of a spoken word), but its mental representation in the mind of an interpreter, as result of the perception of the sign vehicle [2].

In order to draw conclusions for an improved design of non-verbal auditory signs, a more comprehensive model is required: As the physical sign vehicle is the outcome of the design process, it has to be included. According to Peirce, a sign is constituted by the physical sign vehicle (*representamen*) which creates an equivalent or more developed sign (*interpretant*) in the mind of a person. This interpretant stands for something, its *object* [3]. The combination of these three constituents in the process of *semiosis* results in a logical interpretant, which is the essential effect of the sign on the interpreter. This logical interpretant may be involved in a further process of semiosis *ad infinitum.* In Peirce's semiotic theory, a sign is rather a mental process than a fixed relation between certain entities. It also has to be pointed out that Peirce's theory is not restricted to linguistic signs, but comprises signs of any kind in any modality.

2 - SEMIOTIC ANALYSIS OF SOUNDS

In the case of auditory signs, a sound takes the role of the sign vehicle or representamen: As a prerequisite for sound design it is necessary to analyze the relation between the sound, the listener (in whose mind forms the interpretant) and the object to which the sign refers. Peirce defined three trichotomies of signs: As the sound designer's aim is to convey information, most interesting for him is the trichotomy which classifies the sign (or, rather the process of semiosis) according to its relation to the object it refers to.

- An *index* is a sign which is really affected by its object, e. g. the sound of a starting engine.

- An *icon* has some characteristics by which it refers to its object, e. g. a recording of the sound of a starting engine is played back in a computer interface when a process starts.
- A *symbol* is arbitrarily related to its object, by virtue of a law, e. g. Morse symbols.

As can already be seen from the first example, these are not fixed categories: A sign may be a complex which involves several indexes, icons and/or symbols. A spoken sentence consists of words which are mostly symbols, but some are also icons (e. g. onomatopoetic words) or indexes (e. g. pronouns); furthermore, this spoken sentence also contains prosodic information which may be an index for the emotional state of the speaker. In a similar way, the attributes of a non-verbal sound can be analyzed and proactively used by a sound designer.

In a man-machine interface, it is an important aspect of auditory warnings that they allow the operator to prioritize his reactions according to the hazards that these warnings represent. The characteristic of a warning sound which will determine this prioritization has also been called "perceived urgency". Edworthy, Loxley and Dennis have investigated the impact of several warning sound parameters on perceived urgency [4], e. g. the amplitude envelope: A slow onset envelope was found to be more urgent than a slow offset envelope. The authors' interpretation of this result is that "a slow onset envelope is typical of an approaching object, whereas a slow offset envelope is typical of an object moving away from the listener, which may be perceived as less urgent."

In this example, the amplitude envelope is a sign which stands for the priority which should be given to the auditory warning. In a natural environment, the perceived sound level varies with the distance between source and receiver and therefore is a sign of the type "index": The listener can deduce whether the object emitting the sound is approaching and poses a threat. One might argue that − if this amplitude envelope is applied to an arbitrary sound − the type of the sign is rather "icon" than "index", as the object does not really affect the sign (the sound source is not moving): The sound mimics a characteristic of a potentially dangerous object. The important point however is that a sound as sign vehicle may contain several signs as components, encoded in its parameters (e. g. amplitude envelope, timbre, speed, rhythm). These component signs may be of different types (index, icon, symbol) which also means that they will cause different cognitive workload for the interpreter.

3 - LISTENING EXPERIMENT

In order to investigate how different sound parameters affect the recognition of auditory signs, a listening experiment with several stages was performed. Stimuli were based on 12 sequences consisting of 4 tones each. Half of these sequences were based on arbitrarily selected musical motifs; it was assumed that these might be well known to the test subjects (e. g. motif from a TV show theme); the other half began with the musical tone C, followed by a combination of the musical tones E, G and A (resulting in 6 sequences). All tones used were in the range of one octave (261.6Hz to 523.2Hz); the lowest tone of each sequence was C.

Twelve test subjects (aged 25 to 35) without any known hearing losses participated; all test subjects had experience with psychoacoustic experiments, but were naive to the purpose of the experiment. In addition, the test subjects were asked whether they had a musical talent (according to their own judgement).

Each test subject participated in three listening experiments from a series of five experiments; all test subjects participated in the fifth and final experiment. Apart from the stimuli, the procedure in the first four experiments was identical; the progress of the experiment was controlled by a PC software which also recorded the responses of the subjects.

The task of the test subjects was to identify a stimulus (out of twelve different stimuli in total). At first, the test subject had the opportunity to become familiar with the whole set of stimuli; no limit was enforced upon the duration of this training phase. By mouse-clicking on a text on the computer screen, the test subject could trigger the presentation of a stimulus. Afterwards, the default text ("Stimulus 1" to "Stimulus 12") could be replaced by entering a new description (up to 20 characters). With respect to the semiotic model this means that the test subject had to define an object for each sound (i. e. the sign vehicle).

Stimuli could be played as often as desired by the test subject; each time the text could be changed (if desired). After each stimulus had been played at least once, the identification task could be started: The test subject was presented with a stimulus and had to identify it by clicking on the corresponding description. Stimuli were presented in randomized order; each stimulus was presented ten times (resulting in 120 stimuli presentations during the course of the experiment).

In the first experiment, a stimulus consisted of a sequence of four pure tones of equal duration and amplitude; half of the test subjects participated in this experiment. The other half participated in the second experiment: Again, a stimulus consisted of a sequence of four pure tones of equal duration, however tones were accentuated by varying the amplitude.
Results show that test subjects with a musical talent were significantly more successful in identifying stimuli (94% vs. 53%) − though their training efforts were considerably lower. The difference in the identification scores becomes more understandable by analyzing the descriptions which were associated with the stimuli: Subjects with musical talent either wrote down musical notes or used technical terminology of music (e. g. "jazz bass" or "chromatic") while the other subjects either used onomatopoetic notation (e. g. "dudidado") or tried to describe the pitch contour of the sequence (e. g. "up and down"). Obviously, this type of listener was not able to associate an object or a concept with the given sign vehicle (due to missing musical expertise) and consequently failed to find an appropriate verbalization.
For the third listening experiment, each of the 12 sequences was played using one of three timbres of musical instruments (flute, piano, guitar); each of these timbres was applied to 4 sequences. The results of the former two listening tests were incorporated in this assignment of timbres to sequences: Sequences that had been confused very often were assigned different timbres. Again, the 4 tones of a sequence were of equal duration.
Test subjects immensely benefited from the additional cue provided by timbre: The identification score for subjects with musical talent improved to 100%; the result for the other subjects was not uniform. Two of these subjects did not use the timbre information for their descriptions of the stimulus: Consequently, their identification scores were almost identical with the results from the first two tests. The other subjects without a musical talent reached an identification score of 98%, which is a great improvement compared with the first two tests.
For the fourth listening test, pure tones have been used but their temporal structure has been varied resulting in a rhythmic version of the 12 sequences: Again, the results of the first two tests were used to assign different rhythmic structures to stimuli which had been confused often in the previous tests. The additional cue helped to improve the performance of the subjects without musical talent: They identified 78% of the stimuli correctly, the subjects with musical talent 96%. As the improvement of the identification score is lower than in the third experiment, this may indicate that the contribution of timbre to distinctiveness is greater than that of rhythmic structure − however, with the same right it may be said that it is just more difficult for a designer to find rhythmic structures that have a comparable effect.
In the fifth and final experiment, the number of distinct stimuli was reduced from 12 to 9: Again each stimulus consisted of 4 tones. One of the three timbres used in the third listening test was assigned to each of the stimuli; each of the three timbres was assigned to three of the stimuli. The stimuli had the same individual rhythmic structure that had already been used in the fourth listening experiment.
However, in this test the subjects were not free to enter their own descriptions for the stimuli: Instead, they had to choose a description from a list of 18 predefined terms, all describing events that may occur during daily work with a personal computer (e. g. "virus detected" or "printer job finished"). The identification scores were 89% for subjects with and 80% for subjects without a musical talent.
All subjects applied a similar strategy in choosing the description for the stimuli: Timbre was used to group events connected with a common equipment or concept (e. g. "printer" or "computer virus"). Subjects reported that they tried to map the severity of the event to the "mood" that was felt when hearing the stimulus (which is comparable to the concept of perceived urgency mentioned earlier [4]). This also means that the test subjects tried to use different characteristics of the sound (i. e. the sign vehicle) to convey different parts of the information: The sound was not used as a single sign, but as a compound of several signs which are encoded in its perceptible parameters.
If the use of timbre as a sign is classified according to Peirce's categories index, icon and symbol, two different phases have to be distinguished: The design phase (when a certain timbre was assigned, for example, to printer) and the employment phase. In the design phase the assignment was arbitrary and therefore the relation between printer and timbre would be symbolic. In the employment phase, the user might as well interpret this relation (and the variation of the sound according to the associated event) as an index (in the same way as he may judge from the sound of a real printer whether it is operating normally or has a paper jam.)

4 - CONCLUSION

As can be seen from this example, a design strategy for non-verbal auditory signs should aim to distribute the information to be conveyed to several components of a sound. However, further knowledge is required to identify how the perceptible parameters of a sound contribute to the process of interpretation: For "perceived urgency", some parameters have been identified and their influence has been empirically quantified [4]. A model based on semiotic theory may assist the designer in the identification, analysis and further optimization of sign constituents.

REFERENCES

1. **Nöth**, *Handbuch der Semiotik*, 2000
2. **De Saussure**, *Grundfragen der allgemeinen Sprachwissenschaft*, 1931
3. **Peirce**, *Collected Papers of Charles Sanders Peirce*, 1960
4. **Edworthy, Loxley and Dennis**, Improving Auditory Warning Design: Relationship between Warning Sound Parameters and Perceived Urgency, *Human Factors*, Vol. 33(2), pp. 205-231, 1991