

## THE STRUCTURE OF ESTONIAN CONCEPTS OF EMOTION: A SELF-ORGANIZATIONAL APPROACH

Ene Vainik<sup>1</sup> and Toomas Kirt<sup>2</sup>

<sup>1</sup>*Institute of the Estonian Language, Tallinn*

<sup>2</sup>*Tallinn University of Technology*

**Abstract.** The aim of the study was to test the hypothesis that the way of approaching Estonian concepts of emotion may influence their emergent structure. One hundred participants assessed the semantics of 24 words in two different tests that provided access to the concepts from two different levels of representation of knowledge of emotions. The first task addressed the semantic interrelations of emotion terms (synonymy and antonymy) and the second addressed the qualities of emotional experience as measured on seven scales. The results were visualized by self-organizing maps, revealing two visually different topological layouts, where coherence of the conceptual structure was recognizable only in general terms. Two main clusters, positivity and negativity, were found in both maps. In addition, a cluster of alertness emerged while the concepts were accessed through the interrelations of emotion words. There was much less coherence for local neighbourhood relations. That is, relations which coincided on both maps were identified only in approximately half of the cases.

**DOI:** 10.3176/tr.2008.4.02

**Keywords:** emotion concepts, self-organizing maps, conceptual representation of emotions, Estonian language

### 1. Introduction

Studies of the emotion lexicon with the aim of discovering the structure of emotions have enjoyed great popularity in recent decades. Most of the studies rely on a default assumption of lexical isomorphism (e.g., Church, Katigbak, Reyes, and Jensen 1998, Fontaine, Poortinga, Setiadi, and Markam 2002, Frijda, Markam, Sato, and Wiers 1995, Johnsn-Laird and Oatley 1989, Russell, Lewicka, and Niit 1989, Shaver, Schwartz, Kirson, and O'Connor 1987, Zammuner 1989). There are, however, other authors who claim that the study of words is not sufficient for

discovering the structure of affect (e.g., Cacioppo and Berntson 1999, Feldmann Barrett 2004, Fillenbaum and Rapoport 1971, Ortony, Clore, and Collins 1988), or that its sufficiency cannot be taken for granted (e.g., Haslam 1995, Storm and Storm 1987). In the recent literature it is not obvious what is really under observation in the studies of the emotion lexicon, that is, whether emotions or words. Another unclear issue is whether the structure of emotional experience and emotion vocabulary should coincide or not.

In the present paper we assume that the general representations of emotions, accessible by means of different statistical methods of data analysis, are conceptual by nature. Concepts of emotion may not be independent entities but rather abstractions from experience or words. As such, they can be understood and studied both from the level of experience and from the level of verbal expression. These two alternative approaches may influence the outcome of data analysis. When emotional experience is taken as a starting point (like in emotional self-ratings), the results tend to reveal the representation of an emotional-motivational system, where positive and negative affects exist in a mutually non-exclusive way as two kinds of activation (Watson, Wiese, Vaidya, and Tellegen 1999). However, when emotion words are taken as a starting point (as in word similarity judgements), the results are best represented by an emotional circumplex of two bipolar crossing dimensions of activation and valence (Russell 1980, Russell et al. 1989).

In this paper, we focus directly on a set of Estonian concepts of emotion and study them from both these levels. Our purpose was to find out whether the level of access might affect the emergent structure of concepts of emotion as far as can be inferred from self-organising maps (SOM). Two lexical tasks were carried out which addressed the set of concepts firstly on the basis of their relation to the quality of experience, and secondly the semantic interrelations of words. The results revealed that the way the data was gathered partly affected its layout in SOM. In both cases we found two main clusters, that of positivity and negativity among the studied concepts. In addition, a cluster of alertness emerged while the concepts were accessed through the level of interrelations of emotion words. In terms of local neighbourhood relations, a partial heteromorphism of the two structures was detected.

## **2. A case study of Estonian concepts of emotion**

Previous research into the Estonian emotion lexicon has pointed to the possibility of differences in the experiential and semantic knowledge of emotions (Vainik 2006). Some structural divergence has also been found in tasks which address emotional self-report (Allik 1997, Allik and Realo 1997, Veski 1996) and word similarity judgments (Kästik 2000, Russell et al. 1989). The results of previous research are consistent as far as the general division of the lexicon into two subcategories of positive and negative emotions is concerned. However, there

are inconsistencies in the details, and no satisfactory resolution has been found for the question of whether the dimensions of positivity and negativity should be treated as mutually exclusive bipolar opposites, or as possibly independent and co-occurring measures of two unipolar dimensions.

We suspected that the level of access (lexicon-related vs. experience-related) and different statistical methods (FA, MDS) applied to the results might have influenced the outcome. In factor analysis (FA) the results are provided in terms of factors which explain as large a proportion of the variance in the data as possible, with as small a number of factors as possible. Interpretation of such general factors is tempting because they might gather data units with a wide range of specifications. In multidimensional scaling (MDS) the results are provided in terms of dimensions which reflect the best distributions of data in a multidimensional space, where the most dissimilar items are kept as far apart from each other as possible. As such, it reveals best the global dissimilarities in data and in the case of emotional phenomena being analysed, tends to end up in a circular representation determined by two crossing dimensions.

In this study we decided to take both lexicon-related and experience-related approaches simultaneously, and to use SOM as the statistical method, as it enables us to explicate the relations of similarity between concepts in terms of local neighbourhoods. This method of analysis uses an unsupervised learning algorithm which is claimed to partly simulate the self-organizing processes which take place in the human brain (Kohonen 2000:104). SOM is also homologous with the theory of Conceptual Spaces (Gärdenfors 2000), a geometrical model of conceptual representations where the similarity of concepts is represented by their spatial closeness. Our aim was to discover and compare the possible alternative self-organizational structures.

We hypothesized that if the conceptual structure depends on the level of access, the SOM outcomes of different lexical tasks addressing those levels should not be identical. We expected the emergent structures to be comparable only in general terms. In order to test this hypothesis, a set of concepts was selected and two lexical tasks were designed which provided information about concepts of emotion either through the semantic interrelations of emotion terms (i.e. synonymy and antonymy), or through their relation to instances of emotional experience.

In the first task we relied on the speakers' intuitive knowledge about the similarities and dissimilarities of the concepts. The smallest and biggest difference between the underlying concepts is expressed on the lexical level of a language in the relations of synonymy and antonymy (Cruse 2000). For example, *irritation* can be considered as a near synonym of *anger* because the conceptual difference between them is minimal, whereas *irritation* is considered an antonym of *calmness*, because the conceptual difference is large.

In the second task we relied on the participants' intuitive evaluations of certain semantic features which may be attributed to instances of emotional experience. The semantic features under investigation were selected from the characteristics of emotions which are most discussed in the literature, and were formulated as clearly

Plate 1

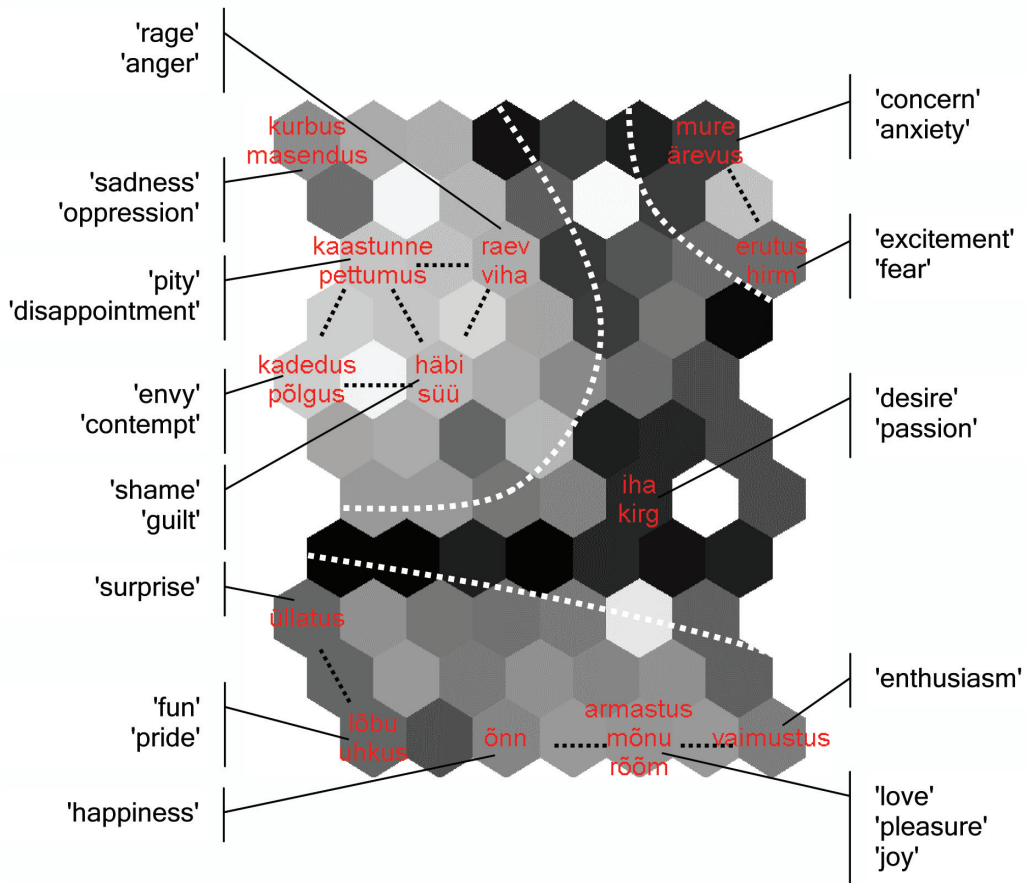


Plate 1. SOM of the results of Task A. A self-organizing map of concepts of emotion based on the relations of similarity and oppositeness.

Plate 2

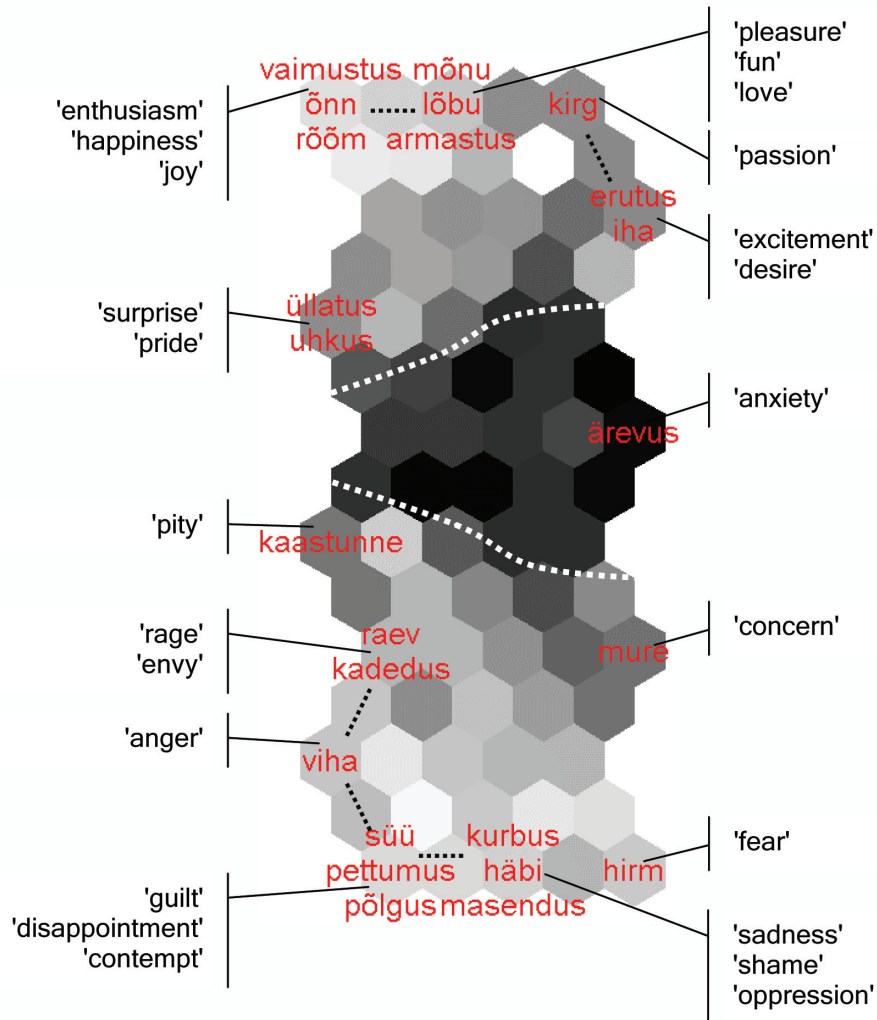


Plate 2. SOM of the results of Task B. A self-organizing map of concepts of emotion based on the evaluations on seven scales.

as possible. The semantic features were evaluated by means of various different scales. We included a scale of subjective evaluations in terms of *pleasantness vs. unpleasantness* which is considered the most pervasive characteristic of emotion terms (e.g., Fillenbaum and Rapoport 1971, Larsen and Diener 1992, Russell 1980, Storm and Storm 1987). In order to measure the extent to which the parameters of intensity of emotion and duration of emotion may influence concept formation, we asked the informants to assess each emotional state on a scale of *strong vs. weak* and *long duration vs. short duration*, respectively. Intensity of emotion has been claimed to be one of the main prototypicality markers of emotion terms (e.g., Niedenthal, Auxiette, Nugier, Dalle, Bonin, and Fayol 2004, Shaver et al. 1987, Zammuner 1989). The scale of action readiness i.e., *increasing vs. decreasing*, was meant to measure another universal and important component of emotion structure (Frijda 1987, Frijda et al. 1995). The distinction *felt in the mind vs. felt in the body* was used to determine the degree to which bodily feelings and conscious cognition influence perceptions of emotional experience, since it has been argued that there is a distinction between “higher cognitive emotions” and the evolutionally rooted, purely biological ones (e.g., Griffiths 1997:100–136). Another scale, *depends mostly on others vs. depends mostly on oneself*, was included so as to measure the relevance of intrapersonal versus interpersonal attribution as a possible distinction between social emotions (e.g., Kemper 2000) and basic ones (e.g., Ekman 1992). The final scale was designed to measure the extent to which the temporal and causal sequences of events belong to the cognitive structure of emotion terms (cf. Wierzbicka 1999): participants were asked whether a given emotion was typically perceived as *preceding an event* or *following an event*, not actually specifying what the event was, to allow the informants to specify whether their conceptualized focus of attention fell either on antecedent events or on an emotional episode following the event by which it was elicited.

The purpose of providing participants with scales of semantic features was to target their experiential and possibly subconceptual knowledge of emotions, while the really relevant hidden dimensions were expected to emerge in the process of data analysis.

### 2.1. Selection of emotion terms

Similar to any other language, Estonian has many words which refer to and differentiate between the qualitative and quantitative aspects of emotional experience. As the aim was to use the intuitions of unprepared participants on a voluntary basis, the list of stimulus terms was kept fairly short. After a pilot test the number of terms was fixed at 24 (see Appendix B). All the words chosen were nouns, as they are more abstract than adjectives.

Term selection was based on the results of tests of free listing (Vainik 2002), word frequencies in corpora (see the frequency data in Appendix B), and a comparison with word lists used by some earlier studies of Estonian emotion terms (Kästik 2000, Veski 1996). As a result, the list contained the Estonian basic emotion terms (*viha* ‘anger, hate’, *armastus* ‘love’, *rõõm* ‘joy’, *kurbus* ‘sadness’),

as well as terms which were experientially similar, but of a different degree (*raev* 'rage', *kirg* 'passion', *vaimustus* 'enthusiasm', *masendus* 'oppression'). Based on the author's intuition, some higher 'cognitive' (*süü* 'guilt', *pettumus* 'disappointment') and 'social' emotion terms (*kaastunne* 'sympathy', *kadedus* 'envy') were also included. As the purpose of the study was not to analyse the whole emotion category, but rather to explore whether there was any coherent structure to the emotion lexicon, the numbers of 'positive' and 'negative' terms were balanced. Some supposedly neutral or ambivalent terms were also included (*üllatus* 'surprise', *kaastunne* 'pity, compassion'). We believe that the selected lexical items form a small but representative set of the core of the emotion category of Estonian lexicon, and that they are sufficient for comparing the structure of concepts of emotion which emerge from the two different lexical tasks.

## 2.2. Subjects and method

The experiments were carried out during the summer of 2003 at different locations in Estonia (Tallinn, Tartu, Võru county, Western Estonia). A total of 115 questionnaires were distributed. Fifteen of them turned out to be incomplete, or were not returned. The final number of participants was 100 (50 men and 50 women), age range 14–76 ( $M = 40.2$ ,  $SD = 18.61$ ), all native speakers of Estonian.

There were separate sheets in the questionnaire for each of the 24 concepts and two tasks had to be completed. Task A was a free listing task. Participants were provided with a blank space to write down as many synonyms and antonyms as came to mind for the presented item. In Task B, participants had to evaluate the meanings of the same 24 emotion terms against a set of seven scales consisting of seemingly opposite values, described above.

The scales task was inspired by Osgood's method of semantic differentials (Osgood, Suci, and Tannenbaum 1975), consisting of three degrees on each side and a neutral zero value in the middle. The degrees on both sides were labelled with positive values (1, 2, 3), and thus were treated as two independent semantic features. The participants were instructed to mark their first opinion about a given concept on one side of the scale, indicating the degree of relevance of a specific feature. In the case of semantic irrelevance, a zero value was suggested, and in the case of ambivalent relevance an additional mark for secondary opinion was available on the other side of the scale.

Participants were provided with questionnaires to complete at their leisure in their own time. Since no time limit was given, the duration of the two tasks was reported to vary from twenty minutes up to two hours. Most of the participants reported some difficulty in the lexical task, probably because it turned out to be quite hard to produce synonyms and antonyms. Partially or totally incomplete answers were excluded from the subsequent analysis.

## 3. Analysis by self-organizing maps (SOM)

SOM is a widely used application of the general principles of self-organization in the field of data analysis. Self-organization is a process characterized by an

increase of order in the system without any external control: it is only the components of the system itself and their interactions which are responsible for organizing it (Heylighen 2001). As with any other reductional analysis, the purpose of SOM is to reduce the dimensionality of input data and to reveal its hidden structure. It performs two tasks: (i) clustering, i.e. reducing the amount of data by grouping similar data items together, and (ii) projection, i.e. reducing the dimensionality of the data by projecting the input data into a lower-dimensional space in such a way that the data items located close to each other in multidimensional space will appear as nearby units in a two-dimensional map. The results are presented on a topological map called the Unified Distance Matrix (U-matrix) where not only dimensions but also clusters and neighbourhood relations are relevant.

The map is generated in a gradual process of unsupervised learning, i.e. it is self-organizing in the sense that there is no prior knowledge about the expected result or mapping. By way of illustration, assume that there is a set of input variables defined as a real vector  $\mathbf{x}$ , which can also be called the input space. The output of the SOM method is a grid of vectors  $\mathbf{m}_i$  which is of the same size as the input vector. To start with, all the vectors of the output grid are initialised randomly. The SOM algorithm has two main basic steps which are reiterated a number of times. First, one random input vector  $\mathbf{x}(t)$  is chosen, and compared with all output vectors  $\mathbf{m}_i$  to find the closest vector on the output grid. Second, this best match or winning vector and its neighbourhood are changed to become closer to the input vector, using the following equation

$$\text{Weight(New)} = \text{Weight(Old)} + \text{Learning rate} \times \text{Neighbourhood function} \\ \times (\text{Input value} - \text{Weight(Old)}).$$

During the learning process the learning rate and the neighbourhood function shrink. As a result of this process the output becomes ordered and all the output vectors are valued so that the total distance from the input vectors is minimized.

The distance between each pair of map units is represented in a U-matrix by their location and also with colour coding. A light colour corresponds to a small distance between two map units and a dark colour represents a bigger difference between the map units. The points on the output map that lie in the light area belong to the same group or cluster, while the dark area shows the borders between the clusters.

Self-organizing maps are related to multidimensional scaling (Duda, Hart, and Stork 1997, Kaski 1997). Both perform the projection of a data set into a lower-dimensional space, but there are some differences, too. For example, multidimensional scaling tries to preserve the metrics of the original spaces, whereas the SOM tries to preserve the topology, i.e. the local neighbourhood relations (Kaski 1997).



### 3. Results of task A

#### 3.1. Data handling

The elicitation of similar words resulted in 4068 lexical items and the elicitation of opposite concepts resulted in 3694 lexical items. In order to let the lexical information self-organize it was first quantified. Every single event of listing similar or opposite concepts was treated as an instance of free listing and an index of relative cognitive salience was calculated for every relation mentioned by at least three persons. The salience index (S) proposed by Sutrop (2001) takes into account the frequency (F), mean position (mP) and number of participants (N), and is calculated as follows

$$S = F / (N \times mP),$$

where  $mP = (\sum R_j) / F$  (the mean position (mP) of an item is a quotient of the sum of all individual ranks ( $\sum R_j$ ) and the frequency (F) of an item in a given list task).

The salience indices varied from .88 to .01. Appendix A presents the data of the 30 most salient relations among the emotion terms. For generating a SOM, relations with indices greater than or equal to the average ( $S_{ave} = .07$ ), were taken into account (219 relations out of 488). For the SOM input the indices of similarity ( $S_s$ ) were transformed into those of theoretical closeness between the concepts ( $1 - S_s$ ), and the indices of oppositeness ( $S_o$ ) were transformed into theoretical distances with polar values ( $0 - S_o$ ). Some examples of the input values are presented in the rightmost column of Appendix A.

#### 3.2. The self-organizing map: task A

The data from Task A was processed by a program (Alhoniemi, Himberg, Parhankangas, and Vesanto 2005) which triggers a self-organizing process. The hidden conceptual structure emerged as visible on the U-matrix (Plate 1). White dotted lines have been added for clarity to show the different clusters. There is a general vertical alignment of positive versus negative concepts, and a noticeably darker row of nodes is aligned horizontally, separating those two categories of unequal size. There are in total three clusters though, as in the upper right hand side of the graph there is a darker area which distinguishes the cluster of *erutus* 'excitement', *ärevus* 'anxiety', *mure* 'concern' and *hirm* 'fear'. What these terms have in common is their reference to states of alertness or vigilance (arousal accompanied by a state of anticipation). From Appendix B it can be seen that these concepts were characterized as 'strong' and as 'preceding (rather than following) an event'.

Inside the clusters the concepts are positioned in such a way as that the most similar items are collocated in the same node. Concepts somewhat less similar to each other appear as neighbours having only one node of the same or lighter colour between them. Black dotted lines have been added as connectors between neighbours to explicate the fragments of conceptual networks. In cases where there is a relatively dark node between different concepts, the similarity of the concepts is smaller, and no connector was added.

### 3.3. Discussion

The results of Task A indicate that there are three predominant aspects in the meanings of the terms under observation: negativity, positivity and alertness. These are the aspects which tend to attract concepts to clusters in self-organization. This result confirms the findings of Feldman Barrett (1995) about the possible alternative foci for arousal versus valence in the structure of affective experience. In her study, those foci appeared as participants' individual preferences based on selective attention. From our results it seems that selective attention is not only a property of every single experiencer or conceptualizer, but it also influences the very content of the concepts themselves. The impact of selective attention on linguistic units, various semantic structures included, is not a new idea in the field of cognitive linguistics (cf. Langacker 1987).

The three semantic foci could also be called dimensions, and could be argued to provide support to the suggestion that valence and activation are the main organizing dimensions of emotion terms (Russell 1980, Russell et al. 1989). However, in this self-organization we do not see the circumplex, and there is no reason to treat positivity and negativity as two mutually exclusive characteristics. For example, we found the concept of *erutus* 'excitement' in the upper part of the graph which was otherwise reserved for negative concepts. It appears that in the case of *erutus* 'excitement' it is the aspect of arousal, or alertness, which is in focus, rather than valence. The terms *iha* 'desire' and *kirg* 'passion', however, were located outside the clusters, perhaps because there was a conflict in their semantic specifications, or because no single semantic aspect was clearly in focus.

Unlike the results which favour a circumplex model, in our study no cluster occurs for states of low activation or alertness. We believe this is because our list of terms consisted only of prototypical emotion words which included a default feature of some level of activation; therefore no large variation was shown. It is likely that a cluster of low activation would have occurred, if the list had contained terms such as *sluggishness*, *calmness*, *fatigue*, etc.

The SOM also explicated fragments of conceptual networks based on neighbourhood relations. The concepts which are collocated in nodes are either versions of qualitatively the same emotion differing in degree (e.g., *kurbus* 'sadness' vs. *masendus* 'oppression', and *viha* 'anger' vs. *raev* 'rage'), or emotions occurring naturally in the same situational context (e.g., *häbi* 'shame' vs. *süü* 'guilt'; *kadedus* 'envy' vs. *põlgus* 'contempt'; *iha* 'desire' vs. *kirg* 'passion', etc). Some scholars have claimed that episodic co-occurrences of emotions are one of the possible predictors of similarity judgements of the emotion terms (Schimmack and Reisenzein 1997).

The results of this study, however, are not quite comparable with the findings of other studies of emotional similarity judgements (e.g., Russell et al. 1989, Schimmack and Reisenzein 1997, Shaver et al. 1987). In our experiment the informants had to list lexical items (an open task) instead of sorting given terms into piles, or judging the similarity of given pairs of terms. In tasks of free listing it

is the active vocabulary and active part of knowledge that tends to prevail over the passive ones (Sutrop 2001, Zammuner 1998).

## 4. Results of task B

### 4.1. Data handling

As a result of Task B, a data cube of 33600 items was gathered (i.e. the evaluations of 100 informants about 24 emotion terms on 14 possibly independent but semantically opposite binary features). For all 24 concepts, one end of each scale was marked more highly than the other. We regarded those features as dominant and joined both ends of each of the seemingly bipolar scales to create a new, joint scale measuring the dominant feature. For example, the feature 'strong' was exploited more heavily than 'weak', so in the joint scale *strong (vs. weak) emotion*, strength should be regarded as the dominant feature, while weakness (given in parentheses) is its less used counterpart.

As the next step, the joint scales were transformed from having +/- values into positive scales of 7–1, starting from 7 as the maximum value of the dominant or default feature (through 4 indicating the irrelevance of the scale, and down to 1 as the minimum value, corresponding to the maximum of the opposite feature). Subsequently the evaluations of a word which accumulated on opposite sides of the same scale were summarized so that a 'second opinion' was given half the value of its actual score. Finally, that half-score was subtracted from the score of the dominant counterpart. For details about the average loadings and standard deviations of the concepts on joint scales see Appendix B.

### 4.2. The self-organizing map: task B

A SOM program processed the data of Task B and the hidden conceptual structure emerged as visible on another U-matrix (Plate 2). As in Task A, there is a general vertical alignment of negative and positive concepts observable and a darker area in the middle separates these two clusters. The shape of the darker area is wider on its right side and it restricts the clusters of truly positive and truly negative concepts to the left uppermost and bottom corners, respectively. For the sake of clarity white dotted lines have been added to mark the two main clusters. One concept, *ärevus* 'anxiety', was located outside of these two clusters. This concept apparently cannot be identified as either positive or negative, or else it may have conflicting specifications with respect to valence. Inside the clusters there are also identifiable chain-like structures of relatively more closely located (and thus more similar) items. The black dotted lines on Plate 2 refer to neighbourhood relations which all appear chain-like. Concepts which are very highly similar are collocated in the same nodes.

It can also be seen in Plate 2 that concepts which had somewhat varying valence but which were evaluated as preceding rather than following an event were aligned at the right edge of the graph (see Plate 2). These terms include

*erutus* 'excitement', *iha* 'desire', *kirg* 'passion', *ärevus* 'anxiety', *mure* 'concern', and *hirm* 'fear', that is, terms which refer to situations of anticipation, accompanied by perceptible bodily sensations (see Appendix B). However, these concepts do not form a separate cluster in the results of Task B. We conclude therefore that the anticipation of an event (alertness) appears as a less significant horizontal dimension of the SOM.

#### 4.3. Discussion

The results of Task B reveal a bilaterally symmetrical structure with two clearly separated clusters of positive and negative concepts of emotion. This finding is consistent with some previous studies of the Estonian lexicon of emotion (Allik and Realo 1997), as well as with the division found in folk psychology (Vainik 2002). Our result can be interpreted as a contribution to Watson and Tellegen's model of two unipolar dimensions of Positive and Negative Affect, which is claimed to account for between 50–75% of semantic variation in emotion terms in various languages (Watson and Tellegen 1985). As the overall shape of the map is extended in one direction it can be concluded that valence probably plays the central role in the conceptual structure of emotions. This might be related to the preference for focus to be on valence (Feldman Barrett 1995) when our participants estimated the qualities of their individual concepts of emotion.

Since all the concepts are situated on the edges of the graph, one might also recognize an oval and claim that these findings are consistent with the circumplex of valence and activation as two bipolar dimensions organizing emotion terms (e.g., Russell 1980, Russell et al. 1989). However, positivity and negativity do not exclude each other in our results, nor is there actually any measurable variance of activation/arousal along the horizontal axis.

Our results from Task B are not quite comparable with the results of emotional self-ratings either in Estonian or any other language (e.g., Allik and Realo 1997, Watson and Tellegen 1985). In our task, the participants had to evaluate the implicit qualities of the individual concepts which were presented to them, not the extent of having experienced a certain emotion. Many of the participants took the opportunity to provide both primary and secondary opinions about some of the qualities. This means that the specifications of the concepts may vary not only across the concepts or the individuals, but also across the different conceptualisations which might have been formed by essentially similar kinds of emotional experience. For example, many people estimated that *viha* 'anger' decreased their action readiness, but there were also a number of people, mostly men, who found a quality of increasing action readiness in the same concept. The experiential record of actual changes in an 'angry' person's action potential is a different issue, and may vary from situation to situation. The rates of standard deviations in our Appendix B reflect the variance and reflexivity of the measured semantic specifications.

## 5. Comparison of the results of the two tasks

The SOMs of our two tasks do not look identical either visually nor in absolute terms (see Plates 1 and 2). On both maps we find clusters of positive and negative concepts, but their spatial alignment is reversed. We assume that what is really important in the conceptual structures is the presence of some specifications such as positivity and negativity. These appeared as possible semantic foci in the meanings of the emotion terms. It is interesting to note that the cluster of negative emotion concepts is bigger in the results of both tasks, although the number of positive and negative stimulus terms was balanced (see Selection of Emotion Terms). Apparently, there was relatively bigger variation in both their estimated qualities and interrelations with other concepts. The reason might be that negative concepts of emotion occur as semantically more differentiated from each other and across individuals than positive ones.

A slightly bigger difference was found between the two maps in alertness (related to event-anticipating situations): alertness was less important in the case of the evaluations of individual concepts (Plate 2) compared to similarity and oppositeness (Plate 1). In both tasks, perceptions of causal and temporal sequences of events turned out to be the second most important semantic determiner of emotion concepts besides valence. Its influence can be seen as a cluster in Task A, and as a dimension in Task B. This could be related to the tendency for emotional experience to be structured by dimensions, and for emotion words to be structured by clusters in a self-organizational layout.

In addition to the general visual comparison of SOMs, an analysis of the neighbourhood relations of the concepts inside the local clusters was carried out in order to find equivalent structures in local conceptual networks. The following criteria were used to define *neighbours*: (i) concepts were collocated in the same node, (ii) concepts were located no further than one node apart, and (iii) concepts were not separated by a darker shade of colour. It appeared that the two SOMs shared 46 of neighbourhood relations. This makes 51% out of the total number of 89, which was found in the results of Task A, and 48% out of the total number of 96, which was found in the results of Task B. An analysis in terms of more detailed conceptual networks thus revealed that the level of access to concepts (either lexicon-related vs. experience-related) influenced the outcome to some extent. It seems that the dimensions used in Task B for describing the concepts of emotion did not fully capture the semantic gestalts that people operate with while providing synonyms and antonyms in lexical tasks.

Thus, we can conclude that the coherence of the conceptual structures in our two tasks is recognizable only in general terms. This supports our hypothesis that the level of access to concepts of emotion can influence its structure, at least to some extent.

## 6. General discussion

Like in emotionology in general, the main question about Estonian terms of emotion has been whether their structure is determined by two unipolar dimensions of positive and negative affect (Allik 1997, Allik and Realo 1997, Veski 1996), or by a circumplex of valence and activation/deactivation (Kästik 2000, Russell et al. 1989). The results of the present study confirm neither of these theories. We found that different emotion terms can focus on two semantic aspects, positivity and negativity, which are not mutually exclusive, and also on an aspect of situation anticipating alertness. These aspects serve as semantic potential which can be realized selectively.

The main finding that the structure obtained through data reduction tends to be dependent on the way the data is gathered and handled is unlikely to be a peculiarity of Estonian and it has wider implications. Many authors have pointed to the possibility of methodological bias in the studies of the emotion lexicon (e.g. Alvarado 1998, Feldman Barrett 2004, Feldman Barrett and Fossum 2001, Fillenbaum and Rapoport 1971, Larsen and Diener 1992, Reisenzein 1995), and this was the reason why we turned to self-organizing maps as a method of analysis. Still, our results show that tasks which start from the level of the interrelations between words, and those which start from the level of measuring the qualities of emotional experience, have resulted in two different structures, which agree only in general terms. This might be related to partial heteromorphism of experiential and semantic knowledge of emotions, found in earlier studies of Estonian lexicon of emotion (Vainik 2006). This interpretation would be consistent with a recent study by Feldman Barrett (2004), in which systematic differences were found in self-reported feelings versus the meanings of emotion words. Her studies of emotion knowledge include various methods and are carried out with English-speaking informants.

If it is the case that there is a partial heteromorphism of experiential and semantic knowledge of emotions, a more general question arises. Namely, are there one or two structures of emotion representation in the mind, and is either of them isomorphic with the structure of our emotions? The distinction between episodic and semantic knowledge (and memory systems, see Tulving 1972) seems to support the hypothesis that there are two parallel systems. That is, the experience-related representations of emotions may be related to episodic knowledge, and purely linguistic representations to semantic knowledge. However, some researchers claim that episodic and semantic knowledge of affect (and memory) are not as distinct as believed, and argue that there is an active interaction between them, especially during the early years of development (Feldman Barrett and Fossum 2001). Evidence for the distinction of the semantic and episodic knowledge was provided by some of our Estonian informants, who intuitively seemed to distinguish between their individual (experience-related, episodic) and collective (linguistically shared, semantic) contents of emotion words. They asked us how to fill in the questionnaire, whether according to their own experience (i.e. 'the way the meanings of the emotion words are *for me*'), or to lexical meaning (i.e. 'the way they are *in*

*general*'). They were instructed to respond with what came to mind first, while in the case of contradictory evaluations a secondary opinion was allowed (see Subjects and Method).

Some authors even argue for three levels of emotion representation, including an intermediate level of concepts. A model consisting of lexical, conceptual and somatic levels has been proposed (Niedentahl, Setterlund, and Jones 1994), as well as a more general model of representations consisting of symbolic, conceptual and sub-conceptual levels (Gärdenfors 2000). The latter is applicable to a broader range of cognitive domains than emotions alone. In attempting to interpret our results in terms of this three-level model, we assume that the parts which were found to be shared between our two structures (the two non-mutually exclusive aspects of positivity and negativity) most likely belong to the conceptual level of representation. The details of the local neighbourhood relations, which were not shared between our two structures, as well as the cluster-like versus dimension-like occurrence of situation related alertness, could be interpreted as belonging only to the symbolic and sub-conceptual levels, respectively.

The conceptual realm of emotions has gained considerable support in recent theories which claim that emotions can be understood as states that are both affective and conceptual at the same time (Feldman Barrett 2006). Any study of concepts of emotion is complicated, though, by our inability to access them directly. They have to be approached through different lexical tasks, possibly leading to partly divergent structures which can be compared only in general terms. With respect to emotion–lexicon isomorphism, the question is rather in the expected degree of isomorphism of alternative reductional representations of those phenomena, achieved through various methods and data analysis.

### Acknowledgements

This study was supported by the Estonian Science Foundation Grant No 7149. Our heartfelt thanks belong Merilin Miljan and Catherine Dickie for their commitment to correct the text of this paper.

#### Addresses:

Ene Vainik  
Institute of the Estonian Language  
Roosikrantsi 6  
10119 Tallinn  
Estonia

E-mail: Ene.Vainik@eki.ee

Toomas Kirt  
Laboratory of Phonetics and Speech Technology  
Institute of Cybernetics at Tallinn University of Technology  
Akadeemia tee 21  
12618 Tallinn  
Estonia

E-mail: Toomas.Kirt@mail.ee

## References

- Alhoniemi, E., Himberg, J., Parhankangas, J., and Vesanto J. (2005) *SOM Toolbox* (Version 2.0). [Computer software and manual]. Retrieved November 11, 2005, from <http://www.cis.hut.fi/projects/somtoolbox/>
- Allik, J. (1997) "Eesti keele emotsioone väljendava sõnavara tähendus". [On the meaning of Estonian lexicon expressing emotions.] In *Psühholoogia lihtsusest*, 131–150. [The Simplicity of Psychology.] Tartu, Estonia: Tartu Ülikooli kirjastus.
- Allik, J. and Realo, A. (1997) "Emotional experience and its relation to the five-factor model". *Journal of Personality* 65, 3, 625–645.
- Alvarado, N. (1998) "A reconsideration of the structure of the emotion lexicon". *Motivation and Emotion* 22, 4, 329–344.
- Cacioppo, J. T. and Berntson, G. G. (1999) "The affect system: architecture and operating characteristics". *Current Directions in Psychological Science* 8, 133–137.
- Church, A. T., Katigbak, M. S., Reyes, J. A. S., and Jensen, S. M. (1998) "Language and organization of Filipino emotion concepts: comparing emotion concepts and dimensions across cultures". *Cognition and Emotion* 12, 1, 63–92.
- Cruse, A. (2000) *Meaning in language. An introduction to semantics and pragmatics*. Oxford: Oxford University Press.
- Duda, R. O., Hart, P. E., and Stork, D. G. (2001) *Pattern classification*. 2nd ed. New York: John Wiley and Sons.
- Ekman, P. (1992) "An argument for basic emotions". *Cognition and Emotion* 6, 3-4, 169–200.
- Feldman Barrett, L. (2006) "Solving the emotion paradox: categorization and the experience of emotion". *Personality and Social Psychology Review* 10, 1, 20–46.
- Feldman Barrett, L. (2004) "Feelings or words? Understanding the content in self-report ratings of experienced emotion". *Journal of Personality and Social Psychology* 87, 2, 266–281.
- Feldman Barrett, L. (1995) "Valence focus and arousal focus: Individual differences in the structure of affective experience". *Journal of Personality and Social Psychology* 69, 153–166.
- Feldman Barrett, L. and Fossum, T. (2001) "Mental representations of affect knowledge". *Cognition and Emotion* 15, 3, 333–363.
- Fillenbaum, S. and Rapoport, A. (1971) *Structures in the subjective lexicon*. New York: Academic Press.
- Fontaine, J. R. J., Poortinga, Y. H., Setiadi, B., and Markam, S. S. (2002) "Cognitive structure of emotion terms in Indonesia and the Netherlands". *Cognition and Emotion* 16, 1, 61–86.
- Frijda, N. H. (1987) "Emotion, cognitive structure, and action tendency". *Cognition and Emotion* 1, 115–143.
- Frijda, N. H., Markam, S., Sato, K., and Wiers, R. (1995) "Emotions and emotion words". In *Everyday conceptions of emotion: An introduction to the psychology, anthropology and linguistics of emotion*, 121–144. J. A. Russell, J.-M. Fernandez-Dols, A. S. R. Manstead, and J. C. Wellenkamp, eds. Dordrecht: Kluwer Academy Press.
- Gärdenfors, P. (2000) *Conceptual spaces. The geometry of thought*. London: The MIT Press.
- Griffiths, P. E. (1997) *What emotions really are: the problem of psychological categories*. Chicago: University of Chicago Press.
- Haslam, N. (1995) "The discreteness of emotion concepts: categorical structure in the affective circumplex". *Personality and Social Psychology Bulletin* 21,10, 1012–1019.
- Heylighen, F. (2001) "The science of self-organization and adaptivity". In *The encyclopedia of life support systems (EOLSS): knowledge management, organizational intelligence and learning, and complexity*. L. D. Kiel, ed. Oxford: Eolss Publishers. Available from <http://www.eolss.net>
- Johnson-Laird, P. N. and Oatley, K. (1989) "The language of emotions: An analysis of a semantic field". *Cognition and Emotion* 3, 2, 81–123.
- Kaski, S. (1997) *Data exploration using self-organizing maps*. (Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series, 82.) Helsinki: Helsinki University of Technology.



- Kästik, L. (2000) *Emotsioone väljendavate terminite subjektiivne ruum*. [The subjective space of emotion terms.] Unpublished diploma paper. Tallinn: Pedagogical University of Tallinn.
- Kemper, T. (2000) "Social models in the explanation of emotions". In *Handbook of emotions*. 2nd ed., 45–58. M. Lewis and J. Haviland-Jones, eds. New York: The Guilford Press.
- Kohonen, T. (2000) *Self-organizing maps*. 3rd ed. Berlin: Springer.
- Langacker, R. (1987) *Foundations of cognitive grammar I. Theoretical prerequisites*. Stanford: Stanford University Press.
- Larsen, R. J. and Diener, E. (1992) "Problems and promises with the circumplex model of emotion". *Review of Personality and Social Psychology* 13, 25–59.
- Niedenthal, P. M., Auxiette, K., Nugier, A., Dalle, N., Bonin, P., and Fayol, M. (2004) "A prototype analysis of the French category 'émotion'". *Cognition and Emotion* 18, 289–312.
- Niedenthal, P. M., Setterlund, M. B., and Jones, D. E. (1994) "Emotional organization of perceptual memory". In *The hearts eye. Emotional influences in perception and attention*. P. M. Niedenthal and S. Kitayama, eds. 87–113. New York: Academic Press.
- Ortony, A., Clore, G. L., and Collins, A. (1988) *The cognitive structure of emotions*. Cambridge: Cambridge University Press.
- Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1975) *The measurement of meaning*. Urbana and Chicago: University of Illinois Press.
- Reisenzein, R. (1995) "On Oatley and Johnson-Laird's theory of emotion and hierarchical structures in the affective lexicon". *Cognition and Emotion* 9, 4, 383–416.
- Russell, J. A. (1980) "A circumplex model of affect". *Journal of Personality and Social Psychology* 39, 1161–1178.
- Russell, J. A., Lewicka, M., and Niit, T. (1989) "A cross-cultural study of a circumflex model of affect". *Journal of Personality and Social Psychology* 57, 5, 848–856.
- Schimmack, U. and Reisenzein, R. (1997) "Cognitive processes involved in similarity judgments of emotions". *Journal of Personality and Social Psychology* 73, 4, 645–661.
- Shaver, P., Schwartz, J., Kirson, D., and O'Connor, C. (1987) "Emotion knowledge: further exploration of a prototype approach". *Journal of Personality and Social Psychology* 52, 6, 1061–1086.
- Storm, C. and Storm, T. (1987) "A taxonomic study of the vocabulary of emotions". *Journal of Personality and Social Psychology* 53, 4, 805–816.
- Sutrop, U. (2001) "List task and a cognitive salience index". *Field Methods* 13, 289–302.
- Tulving, E. (1972) "Episodic and semantic memory". In *Organization of memory*. 381–403. E. Tulving and W. Donaldson, eds. New York: Academic Press.
- Vainik, E. (2002) "Emotions, emotion terms and emotion concepts in an Estonian folk model". *Trames* 6, 4, 322–341.
- Vainik, E. (2006) "Intracultural variation of semantic and episodic emotion knowledge in Estonian". *Trames* 10, 2, 169–189.
- Veski, E. (1996) *Emotsionaalseid seisundeid kirjeldav sõnavara eesti keeles ja selle seos isiksuse omadustega*. [The vocabulary of emotional states in Estonian and its relation to personality traits.] Unpublished diploma paper. Tartu: University of Tartu.
- Watson, D., Clark, L. A., and Tellegen, A. (1988) "Development and validation of brief measures of Positive and Negative Affect: the PANAS Scales". *Journal of Personality and Social Psychology* 47, 127–144.
- Watson, D., and Tellegen, A. (1985) "Toward a consensual structure of mood". *Psychological Bulletin* 98, 219–235.
- Watson, D., Wiese, D., Vaidya, J., and Tellegen, A. (1999) "The two general activation systems of affect: structural findings, evolutionary considerations, and psychological evidence". *Journal of Personality and Social Psychology* 76, 5, 820–838.
- Wierzbicka, A. (1999) *Emotions across languages and cultures: diversity and universals*. Cambridge: Cambridge University Press.
- Zammuner, V. L. (1998) "Concepts of emotion: 'Emotionness', and dimensional ratings of Italian emotion words". *Cognition and Emotion* 12, 2, 243–272.

## Appendix A

## 30 Most salient relations between Emotion terms

Stimulus word		Target word		R	F	mP	S	I
<i>raev</i>	'rage'	<i>viha</i>	'anger'	S	97	1.17	.88	.12
<i>kurbus</i>	'sadness'	<i>rõõm</i>	'joy'	O	74	1.15	.71	-.71
<i>rõõm</i>	'joy'	<i>kurbus</i>	'sadness'	O	66	1.15	.63	-.63
<i>viha</i>	'anger'	<i>raev</i>	'rage'	S	65	1.17	.61	.39
<i>ärevus</i>	'anxiety'	<i>rahu</i>	'peace'	O	59	1.15	.56	-.56
<i>mure</i>	'concern'	<i>rõõm</i>	'joy'	O	52	1.15	.53	-.53
<i>masendus</i>	'oppression'	<i>rõõm</i>	'joy'	O	55	1.15	.53	-.53
<i>kurbus</i>	'sadness'	<i>nukrus</i>	'wistfulness'	S	49	1.06	.52	.48
<i>hirm</i>	'fear'	<i>kartus</i>	'alarm'	S	62	1.29	.51	.49
<i>kirg</i>	'passion'	<i>iha</i>	'desire'	S	53	1.20	.49	.51
<i>hirm</i>	'fear'	<i> julgus</i>	'courage'	O	49	1.15	.48	-.48
<i>õnn</i>	'happiness'	<i>rõõm</i>	'joy'	S	50	1.14	.47	.53
<i>rõõm</i>	'joy'	<i>õnn</i>	'happiness'	S	52	1.35	.47	.53
<i>õnn</i>	'happiness'	<i>õnnetus</i>	'unhappiness'	O	48	1.15	.44	-.44
<i>kadedus</i>	'envy'	<i>lahkus</i>	'kindness'	O	43	1.15	.42	-.42
<i>erutus</i>	'excitement'	<i>rahu</i>	'peace'	O	40	1.15	.41	-.41
<i>lõbu</i>	'fun'	<i>kurbus</i>	'sadness'	O	40	1.15	.40	-.40
<i>lõbu</i>	'fun'	<i>rõõm</i>	'joy'	S	47	1.29	.40	.60
<i>kirg</i>	'passion'	<i>ükskõiksus</i>	'indifference'	O	38	1.15	.38	-.38
<i>iha</i>	'desire'	<i>ükskõiksus</i>	'indifference'	O	34	1.15	.37	-.37
<i>erutus</i>	'excitement'	<i>ärevus</i>	'anxiety'	S	37	1.16	.37	.63
<i>kaastunne</i>	'pity'	<i>ükskõiksus</i>	'indifference'	O	35	1.15	.36	-.36
<i>raev</i>	'rage'	<i>rahu</i>	'peace'	O	35	1.17	.33	-.33
<i>häbi</i>	'shame'	<i>piinlikkus</i>	'embarrassment'	S	29	1.06	.33	.67
<i>iha</i>	'desire'	<i>kirg</i>	'passion'	S	31	1.16	.32	.68
<i>armastus</i>	'love'	<i>viha</i>	'anger'	O	32	1.15	.31	-.31
<i>põlgus</i>	'contempt'	<i>viha</i>	'anger'	S	35	1.28	.30	.70
<i>üllatus</i>	'surprise'	<i>ootamatus</i>	'unexpectedness'	S	31	1.29	.28	.72
<i>masendus</i>	'oppression'	<i>kurbus</i>	'sadness'	S	32	1.30	.27	.73
<i>mõnu</i>	'pleasure'	<i>nauding</i>	'enjoyment'	S	34	1.50	.26	.74

Note. R – relation, s – similarity, o – opposition, F – frequency, mP – mean position, S – index of cognitive salience, I – input values.

## Appendix B

Estonian emotion terms, their frequency data and average semantic profiles against the set of seven joint scales in task B

Term	Gloss	F1	F2	Joint scales						
				Strong (vs. weak) emotion	Follows (vs. precedes) an event	Felt in the mind (vs. body)	Long (vs. short) duration	Depends mostly on oneself (vs. others)	Increases (vs. decreases) action readiness	Unpleasant (vs. pleasant)
<i>armastus</i>	'love'	194	520	6.58(1.01)	4.89(1.96)	5.11(1.74)	6.49(0.95)	4.37(2.03)	6.51(1.12)	1.17(0.40)
<i>erutus</i>	'excitement, arousal'	6	49	6.45(0.89)	2.97(1.85)	3.63(1.90)	3.75(1.92)	4.46(2.06)	5.26(1.69)	3.01(1.71)
<i>hirm</i>	'fear'	44	217	5.84(1.78)	2.86(1.81)	4.36(2.08)	4.27(1.97)	4.65(2.16)	2.25(1.55)	6.51(1.34)
<i>häbi</i>	'shame'	3	92	5.94(1.44)	6.10(1.40)	5.21(1.86)	4.87(1.76)	5.12(1.89)	2.10(1.41)	6.56(0.84)
<i>iha</i>	'lust, desire'	-	59	5.93(1.60)	2.25(1.59)	3.51(1.91)	4.46(1.93)	4.37(2.04)	5.97(1.46)	2.14(1.30)
<i>kaastunne</i>	'pity, sympathy, compassion'	29	33	5.39(1.61)	5.88(1.36)	5.88(1.41)	4.71(1.87)	3.98(2.31)	4.82(1.69)	4.50(1.76)
<i>kadedus</i>	'envy'	31	39	4.47(1.96)	5.00(1.72)	5.85(1.56)	4.56(2.16)	4.32(2.41)	3.90(2.02)	6.24(1.45)
<i>king</i>	'passion, heat'	18	51	6.55(0.85)	3.00(1.86)	3.76(1.93)	4.72(1.82)	4.42(1.94)	6.19(1.20)	1.73(0.97)
<i>kurbus</i>	'sadness'	185	127	5.51(1.79)	5.69(1.47)	5.33(1.78)	5.01(1.78)	4.53(2.00)	1.92(1.03)	6.20(1.22)
<i>lõbu</i>	'pleasure, fun'	30	87	5.66(1.36)	4.94(1.59)	4.63(1.72)	4.52(1.79)	4.26(1.86)	6.03(1.07)	1.48(0.70)
<i>masendus</i>	'oppression'	10	10	5.84(1.60)	5.85(1.35)	4.68(2.03)	5.27(1.88)	4.88(2.08)	1.39(0.63)	6.74(0.80)
<i>mure</i>	'concern, worry'	26	86	5.88(1.62)	3.56(2.02)	5.20(1.89)	5.44(1.69)	4.25(1.98)	2.89(1.82)	6.40(1.19)
<i>mõnu</i>	'pleasure'	8	74	5.99(1.36)	5.54(1.68)	3.20(1.91)	4.55(1.80)	4.64(1.73)	5.62(1.62)	1.31(0.56)
<i>pettumus</i>	'disappointment'	13	64	5.49(1.78)	6.27(1.31)	5.46(1.75)	4.39(1.92)	3.38(1.97)	1.97(1.05)	6.62(0.87)
<i>põlgus</i>	'contempt'	6	23	5.00(1.87)	5.45(1.52)	5.81(1.34)	4.56(1.97)	3.29(2.09)	3.22(1.61)	6.34(1.06)
<i>raev</i>	'rage, fury'	52	33	6.40(1.40)	6.22(1.23)	4.74(2.03)	3.21(1.95)	3.83(2.11)	3.80(2.34)	6.54(0.99)
<i>rõõm</i>	'joy, gladness'	301	274	6.38(1.06)	5.74(1.37)	5.18(1.71)	5.34(1.57)	4.69(1.76)	6.50(0.95)	1.20(0.44)
<i>süü</i>	'guilt'	3	172	5.73(1.75)	6.24(1.38)	5.43(1.82)	5.35(1.86)	5.65(1.64)	2.34(1.68)	6.56(1.19)
<i>uhkus</i>	'pride'	3	135	5.53(1.56)	4.98(1.83)	5.78(1.33)	5.48(1.48)	5.36(1.69)	5.64(1.45)	2.80(1.82)
<i>vaimustus</i>	'enthusiasm'	6	98	6.28(1.18)	5.37(1.76)	5.53(1.77)	4.58(1.94)	4.67(2.11)	6.36(0.95)	1.67(1.20)
<i>vihla</i>	'anger, hate'	190	159	6.18(1.59)	6.15(1.31)	4.98(1.99)	4.01(1.99)	3.79(2.12)	3.86(2.17)	6.47(0.96)
<i>õnn</i>	'happiness, luck'	29	249	6.56(0.87)	5.41(1.77)	5.11(1.89)	5.58(1.69)	4.72(1.79)	6.64(0.81)	1.17(0.72)

Term	Gloss	F1	F2	Joint scales						
				Strong (vs. <i>weak</i> ) emotion	Follows (vs. <i>precedes</i> ) an event	Felt in the mind (vs. <i>body</i> )	Long (vs. <i>short</i> ) duration	Depends mostly on oneself (vs. <i>others</i> )	Increases (vs. <i>decreases</i> ) action readiness	Unpleasant (vs. <i>pleasant</i> )
<i>ärevus</i>	'anxiety'	14	76	5.88(1.53)	2.46(1.65)	3.94(1.94)	4.88(1.79)	5.33(1.78)	3.70(2.01)	5.12(1.62)
<i>üllatus</i>	'surprise'	24	226	6.06(1.29)	5.69(1.85)	5.30(1.78)	2.99(1.95)	3.47(2.10)	5.61(1.35)	2.21(1.30)

Note. F1 – summary frequency in a series of seven specific list tasks (Vainik 2002), F2 – frequency in a text corpus of approximately 1 million words. The average ratings are given on a descending scale (7–1), value 4 pointing to the irrelevance of a scale and value 1 pointing to the maximum value of the opposite feature. Standard deviations are given in parentheses.