

Lexical Knowledge of Personality Traits

Heili Orav

Dept. of General Linguistics, University of Tartu
Liivi 2-308, Tartu, Estonia
heili.orav@ut.ee

Abstract

This article studies the vocabulary and concepts of personality traits in Estonian. The choice of character words and concepts that are in active use can give us an idea which traits are considered important for Estonians and how the concepts of personality traits are organized on 'the map of character landscape'. Whereas most common personality traits are expressed by adjectives, I will focus here mostly on lexical semantics of personality adjectives and examine them in accordance with the principles of WordNet.

This research was supported by Estonian Science Foundation Grant No 5534.

Introduction

The compilation of the Estonian WordNet (EstWN) started in 1997 and the work is still in progress. In 1998 Estonia joined the EuroWordNet project. The aims of the project were to compile a language-specific database of nouns and verbs – adjectives were not included in this project. Thus, the need for the present research which would also add adjectives to the thesaurus became obvious at this compositional stage of the Estonian WordNet (Orav 2002).

One of the lines of research in lexical semantics which is oriented from theory to practice deals with semantically coherent fields, known as semantic fields in formal terminology. The structure of semantic fields reflects the understanding of the particular culture of the structure of this world fragment and the existing regularities in there that the semantic field covers (Õim 1997). The present article focuses on the concepts of character terms in Estonian, which are mostly adjectives and which form a specific domain in a language.

The study of the adjective lexicon in English by Gross, Fischer and Miller (1989) seems to confirm that about half of the English adjectives studied by them can be applied to persons, their behaviour, experience and appearance. More generally, personality seems to be one of the few domains that are lexicalised as adjectives (if the language has any). Indeed, Dixon (1982) includes 'human propensity' terms as one of the semantic types that languages universally express with adjectives.

In other words, an understanding of the dimensions of personality description in Estonian must begin with an analysis of Estonian character terms and create a language-

specific taxonomy that would aim to add new concepts to the Estonian thesaurus. The present article summarises the research done towards reaching that goal.

1 Estonians are hard-workers

The category of personality is attributed to people intuitively. That can be seen in the fact that there are many words representing character traits¹ in a language and that people can use these words the same way as other words and decide whether a particular usage of a word is linguistically (essentially) correct or not. Hence, character terms are terms that can be used to describe the permanent characteristics of a person's behaviour.

The aim of the case study performed by the author of the article in the years 2003–2004 was to determine the active vocabulary of character terms in Estonian. All in all 100 persons between the ages of 14–90 were interviewed using the field method (Sutrop 2001) in the study. There were equal numbers of men and women participating in the questionnaire. The questionnaire consisted of four parts where Estonians were asked to name personality traits that in their opinion characterise the following: 1) the person's character in general, 2) the character of a pleasant/unpleasant acquaintance, 3) the character of an Estonian, 4) the character of himself/herself. There were 5461 answers altogether. The results of the field-method include several words and expressions whose meaning may give rise to problems of understanding. The words whose connection to human character traits remained obscure were left out in the analysis. Hence, only a small fraction of the concepts that the informants in the study actually used are analysed here – 1271 different words/expressions. The main representatives of character terms belong to the word class of adjectives. The present research thus focuses on adjectives, although during the test nouns (223 times) and verb phrases (26 times) were also mentioned.

The most frequently mentioned words also mark important character traits of a particular culture. The most frequently named character trait in all the tests was *diligence* – the adjective *töökas* 'hardworking' was mentioned 150 times. Table 1 below gives an overview of the frequently

¹In addition to the character terms, the model of personality is also linguistically expressed in e.g. phraseological expressions, phrases and other grammatical structures.

mentioned words and illustrates other aspects of human behaviour and qualities which were noticed by the informants.

Table 1: Frequency count of experiment of listing (with frequency 30+n).

Adjective	Frequency
<i>töökas</i> 'hard-working'	150
<i>sõbralik</i> 'friendly'	109
<i>tark</i> 'wise'	86
<i>abivalmis</i> 'helpful'	85
<i>kade</i> 'envious'	82
<i>laisk</i> 'lazy'	78
<i>kinnine</i> 'private, introverted'	68
<i>aus</i> 'honest'	60
<i>lahke</i> 'kind'	60
<i>kuri</i> 'cruel, wicked'	51
<i>edasipüüdlük</i> 'ambitious'	47
<i>tagasihoidlik</i> 'modest'	47
<i>rahulik</i> 'peaceful'	46
<i>hea</i> 'good'	45
<i>õel</i> 'evil'	35
<i>rumal</i> 'stupid'	34
<i>ahne</i> 'greedy'	33
<i>rõõmsameelne</i> 'cheerful, jovial'	33
<i>hooliv</i> 'regardful'	32
<i>lõbus</i> 'merry'	32
<i>julge</i> 'courageous'	30

Some researchers (e.g. Goldberg 1993; John 1990) have claimed that during some period of time in social communication all the important human character traits are documented in language – the more prominent and important the quality, the more likely it is that there exists an independent word for encoding this quality and the more conceptions are associated with it. People consider these character traits as important ones, which is shown by the fact that they are recalled faster and that they belong to everyday language usage.

2 Taxonomy of character terms

It is believed that the concepts of a specific domain, or actually all concepts, and the words that represent them are not chaotically located in a person's head. As Cruse (2000:179) puts it: "The vocabulary of a language is not just a collection of words scattered at random throughout the mental landscape." He (ibid.) believes that vocabulary is at least partly structured. Some authors (e.g. Viberg 1994:170-171) go even further in arguing that words and concepts are located in quite a systematic way in the human brain, which facilitates categorization and classification. If this statement is true, it should be easy to present words by categories – even inside one specific domain.

2.1 Synonymy

In the experiments, only indirect data was collected for words which are considered equivalent and synonymous by Estonians.

As to the character concepts, there was a tendency for subjects to use several semantically related units that varied a little lexically or morphologically. For instance, the concept VAIKNE 'SILENT/STILL' was referred to as *vaik-ija* 'a person who is quiet' (noun), *vaikne* 'quiet, silent' (adjective), *vaikiv* 'being quiet' (verb); *enesesetõmbunud-endassetõmbunud* stylistic variants of 'introvert' (both adjectives); *valetaja* 'liar' (noun), *valelik* 'untruthful' (adjective), *valetav* 'lying' (verb)²; *alaväärne* 'inferior' (adjective) *alaväärsuskompleksidega* 'with an inferiority complex' (noun), etc. Items that are closely related semantically but lexically different were also named quite often. For example *otsekohene-sirgjooneline* 'straight(forward)', *viisakas-distsiplineeritud* 'polite/disciplined', *julge-kartmatu-vapper* 'bold/brave', etc.

Thus, a character concept might not be linked to one rigid character term but to a "family of terms".

3 Hierarchy of character terms

Besides the Princeton WordNet, there are other projects, such as GermaNet [] and SIMPLE [], that have considered adjectives thoroughly. These projects have worked out a detailed theoretical basis for presenting different word classes, including adjectives. In these projects, adjectives are hierarchically divided according to semantic fields, including human properties.

With respect to character traits, there have been attempts to apply the claim that the mental lexicon is hierarchically organised (Aitchison 2003). Many researchers (Hampson, John and Goldberg, 1986) have tried to describe the vocabulary of personality via hierarchical relations. For example, the hyponym of the word *talented* is *musical* and the hyponym of *trustworthy* is *punctual*, etc. The latter example is however exceptional and not entirely correct. It is possible to be punctual in an extreme way; in this case a person is not suitable, for instance, as an undercover agent and thereby sacrifices his/her credibility (trustworthiness).

In analysing Estonian, the first problem that arises is that the number of compounds is indefinite. It is easy for a speaker of Estonian to create new compounds (also adjectives) that are not listed in any dictionary but are, nevertheless, easily understood. In compounds, the first element functions as an attribute and the second as the head. For example³³: *arenemisvõimeline* 'develop-capable', *otsustusvõimeline* 'decision-capable', *õpivõimeline* 'learn-capable', *armastusvõimeline* 'love-capable', *vastutusvõimeline* 'responsibility-capable', etc. Does it mean that the second element of these words, i.e., – *võimeline* 'capable', acts as a hyperonym in Estonian? The answer is no, because in these cases it is only via the first element of the words that the conceptually correct lexical relation is expressed. Thus, for example *armastusvõimeline* 'love-capable' is synonymous with *hooliv* 'regardful', and *vastutusvõimeline* 'responsibility-capable' is antonymous to *vastutustundetetu* 'irresponsible', etc.

²The word stem in Estonian is *vale* – 'lie'.

³³These compounds are given a literal word-by-word translation here.

At the same time the examples below show that some concepts can act as hyponyms for others where the situation is precisely opposite. Consider the following examples:

kade 'envious' -> hyponym is *armukade* 'jealous',
ahne 'greedy' > hyponym is *võimuahne* 'power-greedy',
 hyponym is *rahaahne* 'money-greedy/money-grubber', hy-
 ponym is *kasuahne* 'benefit-greedy';
lahke 'kind' -> hyponym is *külalislahke* 'guest-kind/hos-
 pitable', etc.

The examples above point to the fact that only some part of character terms in Estonian is hierarchical and show that even the lexical relations themselves can be language- or culture-specific.

One solution could be that nouns which, among other things, relate to adjectives, act as top concepts as in EuroWordNet. Lyons (1977: 299) calls such relation between a noun and an adjective **quasi-paradigmatic**. He states (ibid.) that "if we include quasi-hyponymy with hyponymy as a relation in terms of which vocabularies are structured hierarchically, the hypothesis that the vocabulary in all languages is structured hierarchically under a relatively small set of lexemes of very general sense is rather more plausible." The problem is that the constraints on quasi-hyponymy have to be determined. In WordNet this relation is the **attribute** relation; in EuroWordNet the relations between nouns and adjectives are called **be_in_state** and **state_of**. Thus the previous examples are then the following: LAHKUS (KINDNESS) be_in_state *lahke* (kind) and *külalislahke* (hospitable) or TUNDLIKKUS (SENSIBILITY) be_in_state *ilitundlik* (hy-persensitive) or *solvuv* (touchy).

3.1 Opposition

As the hyponymy relation does not appear to be very relevant to adjectives, the Princeton WordNet uses antonymy as "the basic semantic relation among adjectives" (Gross & Miller 1990: 267). WordNet's notion of antonymy presents a number of problems. First, in the Princeton WordNet indirect antonyms are used. Some researchers (e.g. Murphy & Andrew 1993) have found that indirect antonyms are more difficult to verify than direct antonyms. Secondly, it is surprising to treat antonymy as a basic semantic relation on a par of synonymy (Murphy & Andrew 1993).

The vocabulary of personality traits raises the following questions:

1. Do the adjectives representing character traits occur as antonym pairs?
2. Does every adjective representing character traits have an antonym?
3. Does natural language maintain consistency in this part of the lexicon?
4. What kind of scales form in the vocabulary of character traits?

The results of the list task indicate that there are few truly antonymic relations in the Estonian character vocabulary. These are more evident on the conceptual than on the lexical

level. The results indicate that word meanings are mentally represented as concepts. The recent accounts of antonymy refer to the conceptual structure of adjective meaning as well (Cruse 1991).

According to linguistic criteria, the importance of polarity as an essential factor was shown by the occurrence of a great number of expressions where a character trait was defined by a non-existent word marking the opposite quality. Derivation and compounding were often applied in order to create a required concept. Some "artificial" constructions were also made up, that is, a person could not find an appropriate word(form) and used purely formal means such as negation as in words like **ebaustav* 'untrustworthy', **ebasihikindel* 'undetermined', **ebaustav* 'unloyal', **ebatolerantne*⁴ 'intolerant', etc., or he or she marked the opposite quality using adverbs (like *vähe* – 'little') or the negation of verbs (like *pole*, *ei ole* – 'does not'). These formations show that language users tend to think in terms of opposites but this is not systematic and opposites are not fully lexicalized.

The domain of character terms seems to be organised differently: at the experiential level positive and negative character terms can be self-reported and mentally operated while not mutually excluding one another, whereas at the conceptual level (which is influenced by the forms of social cognition like a model of culture), the positive and negative character term concepts are treated as opposites and related to each other through antonymity relations on the lexical level.

3.2 Gradation

Another remarkable aspect of character terms is their gradient nature. Research on semantic theory has dealt with gradation a lot, as gradation is one of the most typical features of adjectives. Gradation is considered a fundamental semantic feature and it presumes the presence of scales in the structure of adjectives (Sapir 1944).

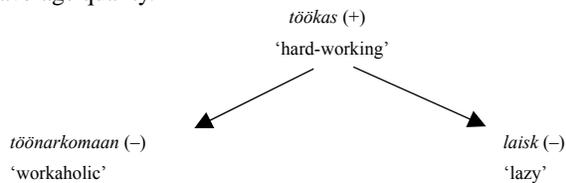
One of the important qualitative aspects of the vocabulary of personality is the scalarity of personality concepts (the quality can appear less or more) and the nature of the scales. The scales usually break down into negative and positive values. For example, concepts of spatial relations always possess the following kinds of scales: up-down, high-low, etc. Scalarity is not so obvious as applied to personal qualities. One cannot name many of the opposites of the human qualities, although the implicit scale can be present. It is probably hard to define the antonym of words such as *envious*, *muddle-headed*, and *gossipy*. Or, if the lexicalization process has not occurred, then the language possesses its own means of expression to measure the rate of qualities in the form of adverbs/quantifiers (*very*, *quite*, *too*, *a little*, *awfully*, *few*, etc).

Secondly – can the scales of qualities of personality be characterized as positive and/or negative? For example, self-confidence is overall a positive quality, the lack of self-confidence is negative, but a person can also be too self-

⁴The prefix *eba* – in Estonian is a typical marker of negative meaning.

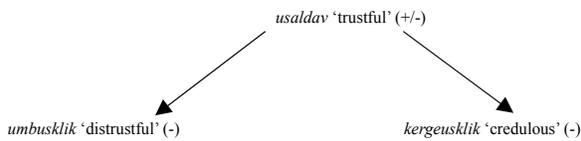
confident (if the rate of self-confidence exceeds the average or what is considered normal in a particular situation) and this is considered as a rather negative quality. Another example is that there is a problem with plus-minus headed scales: where is one to place word combinations like *too hard-working*. The problem derives from the attribute *too* which expresses the notion extreme, incomprehensible, unrelated to the usual ideas, even inhuman, and which is considered deplorable. Hence, the language community has begun to use new words to express the notion 'too hard-working' – *over-working* and *workaholic*, which show extremity while working and act like negative qualities.

From this, two scales form, based on the questionnaire's data of work-related words: *workaholic* – *hard-working* and *hard-working* – *lazy*, in which the positive – *hard-working* – is in the middle and the ends – *workaholic* and *lazy* – are negative. Comparing the extremes at the ends of the scale, the quality *hard-working* turns out to be the normative average quality.



It is remarkable that the property of working hard evoked ambivalence: from one side it is seen as positive. From the other side, the quality of being too hard-working is viewed as a pathology which deviates from a social norm and the quality of being too hard-working is denounced with linguistic labels. Probably, the double standards of opinions for yourself as opposed to those applied to others in general are culturally specific.

Scales similar to the scale of diligence can also be formed for other character trait concepts, for example



Here it is interesting that *usaldav* 'trustful' may be observed as a trait with a plus or minus – it depends on one's personal evaluation.

These are the most interesting examples of the inner structure of the qualities of personality. These kinds of concepts, where the nature of the scales is different, exist nonetheless. Different types of scales, which became obvious during the analysis of the vocabulary of character terms, are:

1. traditional cases of true antonyms, for example *naiselik* 'feminine' – *mehelik* 'masculine'; *ebaõiglane* 'unrighteous' – *õiglane* 'righteous'; *ebaprofessionaalne* 'unprofessional' – *professionaalne* 'professional', etc.
2. gradable qualities, for example *liiga arg* 'too shy' – *arg* 'shy' –> *julge* 'courageous' –> *külmavereline*

line 'cold-blooded' –> *liiga külmavereline* 'too cold-blooded' or, as another example: *frigüidne* 'frigid' –> *seksikas* 'sexy' –> *himur* 'lecherous'

3. traits without scales, for example *kade* 'envious'
4. scales of evaluation (negative – neutral – positive), for example: *igav* 'dull' – *tavaline* 'normal/ordinary' – *huvitav* 'interesting'.

Since this conceptually important relation of gradation does not play a central role in the organization of adjectives, it has not been coded in WordNet (Miller 1998). But as I have shown it cannot be completely excluded either.

4 Other semantic relations applying to personality traits

A language user's intuitive model of personality is expressed particularly in the definite organisation of aspects of the personal qualities.

As these aspects give important information about the understanding of some concepts or the usage of a particular word, it is necessary for them to be encoded in some way in a thesaurus or in some other representation. These aspects should be regarded as types of semantic relations. For example, considering that some qualities of personality appear in human behaviour in general and some in communication, then this kind of distinction refers already to a hierarchical distribution – the activity of communication is one of the sorts and possibilities of human activity.

A large number personal qualities were studied and the following aspects were considered as particularly important:

- 1) The concepts describing a person, so to speak, alone, which refer to his/her overall activity, but do not apply to other people:
 - 1a) qualities that appear in externally visible behaviour (*brave, shy, determined*, etc) – i.e., those involving action
 - 1b) qualities that do not appear in behaviour, and are more like inner attitudes (*inferior, faint, self-confident*, etc) – those involving attitude
- 2) Qualities of personality that appear in human relations. Also:
 - 2a) qualities appearing in behaviour; without direct communication (*aggressive, mercenary, sociable*, etc) – those involving interaction
 - 2b) qualities appearing in communication (*imperious, talkative, saucy*, etc) – those involving communication
 - 2c) qualities appearing in attitudes which are addressed to others and which do not expect actual (nor communicational) activity (*envious, distrustful*, etc) – those involving attitude.

It is well known that positive and negative personality traits may occur in one person at the same time and quick

inspection shows that the above-mentioned aspects of personality traits can also be involved in one concept. For example, *ahne* ‘greedy’ can be connected with one’s external and internal action. Indeed, this is evidence that aspects and dimensions of person-descriptive terms overlap and have fuzzy boundaries.

In general, it has been claimed about the mental lexicon that concepts are located in certain areas in the human brain. It is possible to make the same claim with respect to the present empirical data – the human’s model of personality consists of areas, dimensions and aspects, from some of which are more important than others. The importance is determined by the number of words or expressions which can vary grammatically (for example different word classes) but not conceptually.

Conclusion

It is clear that WordNet opened up a new aspect of language research which tries to unite its conceptual, lexical-semantic and lexicographic dimensions. But the theoretical as well as practical work in the development of this line is only midway still, as it seems. It is not clear how to describe the lexical data on the organisational level (e.g. semantic relations) and how deep we should go on the conceptual level (e.g. how important the word-class differences are). For instance, the well-known FrameNet project (<http://framenet.icsi.berkeley.edu>), which is clearly inspired by WordNet, goes much deeper into the conceptual organisation of languages. Could the approaches of WordNet and FrameNet be combined? In my PhD thesis I am trying to tackle this problem too, but in this article I wanted to concentrate on the problems “inside” WordNet. The primary aim of this article was not to propose new relations, rather, the aim was to display facts and, refer to tendencies in order to finally set up a hypothesis and offer some solutions.

References

- Aitchison, J. (2003) *Words in the Mind: An Introduction to the Mental Lexicon*. 3rd ed. Oxford: Blackwell Publishing.
- Dixon R.M.W. (1982) *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin: Mouton.
- Cruse A. (1991) *Antonymy revisited: Some thoughts on the relationship between words and concepts*. – Frames, fields, and contrasts: New essays in semantic and lexical organization. Ed. by A. Lehrer & E.F. Kittay; Hillsdale, NJ: Erlbaum, 289–306.
- Cruse, A. (2000) *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.
- Goldberg L. R. (1993) *The Structure of Phenotypic Personality Traits*. “American Psychologist”, 48, 26–34.
- Gross D., Fischer U., Miller G.A. (1989) *The organization of adjectival meanings*. Journal of Memory and Language, 28, 92–106.
- Gross D., Miller K.J. (1990) *Adjectives in WordNet*. International Journal of Lexicography, 3, 265–277.
- Hampson S.E., John O.P., and Goldberg L.R. (1986) *Category breadth and hierarchical structure in personality: Studies of asymmetric in judgements of trait implications*. Journal of Personality and Social Psychology, 51: 37–54.
- John O.P. (1990) *The “Big Five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires*. – Handbook of personality: Theory and Research. Ed. by L.A. Pervin. New York: The Guilford Press, 66–100.
- Lyons J. (1977) *Semantics I*. New York: Cambridge University Press.
- Miller K.J. (1998) *Modifiers in WordNet*. – An WordNet. Electronic Lexical Database. Ed. by Ch. Fellbaum. The MIT Press, 47–69.
- Murphy G.L., Andrew J.M. (1993) *The Conceptual Basis of Antonymy and Synonymy in Adjectives*. Journal of Memory and Language, 32, 301–319.
- Orav, H. (2002) *Adjectives in wordnet-type thesaurus: Estonian experience*. In Proceedings of the 1st International Global WordNet Conference, Central Institute of Indian Languages, Mysore, India, 22–25.
- Õim, H. (1997) *Eesti keele mentaalse maailmapildi allikaid ja piirjooni*. – Pühendusteos Huno Rätsepale 28. 12. 1997. (= Tartu Ülikooli eesti keele õppetooli toimetised 7). Tartu: Tartu Ülikooli Kirjastus, 255–268.
- Sapir E. (1944) *Grading: A Study in Semantics*. Philosophy of Science. 11, 83–116.
- Sutrop U. (2001) *List task and a cognitive salience index*. – Field Methods 13, 289–302.
- Viberg A. (1994) *Vocabularies. Bilingualism in deaf education*. International Studies on Sign Language and Communication of the Deaf, 27, 169–199.
- FrameNet (<http://framenet.icsi.berkeley.edu>), GermaNet (<http://www.sfs.nphil.uni-tuebingen.de>) SIMPLE (<http://www.ub.es/gilcub/SIMPLE>)