

# akup

ARBEITEN DES KÖLNER UNIVERSALIEN - PROJEKTS

Nr.16

Russell Ultan

DESCRIPTIVITY GRADING OF  
FINNISH BODY-PART TERMS

Juni 1975

## ABSTRACT

Three quantificational approaches to the measurement of lexical descriptivity are proposed, based on: the semantic sum of the parts of a lexeme is equal to the whole, paraphrase-term and term-paraphrase congruence, explicitness of semantic elements of a construction. Combination of all possible values into tripartite sets and then into equipollent groups results in a system composed of 12 grades. This system was tested with a semantic domain of the Finnish lexicon: body-part terms. The descriptivity indices for each lexical item were correlated with natural divisions of the body, construction-motivation types (form, function, location), grammatical construction types (endo- and exocentric compounds, derived forms, metaphors), and loanwords. These comparisons result in a number of grade profiles whereby specific descriptivity grades are characteristically associated with one or more types of body section, construction motivation, and grammatical construction. Diachronic and synchronic evidence points overwhelmingly to a process of semantic narrowing in the development of descriptive words and labels from phrases or sentences.

## DESCRIPTIVITY GRADING OF FINNISH BODY-PART TERMS

Russell Ultan

### 1. Introduction

One of the basic problems immediately encountered in the determination of relative descriptivity of lexical items in a given language is the lack of a specific and consistent means of accounting for different degrees of descriptivity. Are, for example, terms such as: bookshelf, bookworm, butcher, and butterfingers to be lumped together as more or less descriptive as opposed to labels<sup>1</sup> such as book or house or can we be more specific as to the degrees of lexical descriptivity that are to be found in natural language? In the present paper, I will discuss three different quantificational approaches which when combined yield a reasonably precise method for measuring the degree of descriptivity of a given lexical item. The first is based on a familiar sort of simple "semantic arithmetic", namely that the sum of the (meanings of the) parts of a construction equal to the whole (meaning), the second on paraphrase and paraphrase reversibility possibilities, and the third on the presence of explicit or implicit semantic elements in the construction.

Another important consideration in evaluating lexicon for relative descriptivity is the productivity of the construction type (in the case of descriptive, rather than labeling, terms). Thus, while a term may be formally and semantically analyzable, an unproductive or even mildly productive construction type may interfere with the otherwise isomorphic relation between form and content, e.g. lifeboat, an unproductive type of compound whose elements are unambiguously analyzable but are in themselves insufficient to fully account for the meaning of the term.

In order to test this descriptivity-indexing system, a pilot study of Finnish terms for body parts was made. While relatively comprehensive in terms of the semantic domain as a whole, the list of 235 items is hardly exhaustive. In particular, some terms for internal organs (different kinds of blood cells, specific muscles, nerves, tissues, etc.) were omitted, the general guiding principle of selection being inclusion of most of the items which would normally occur in the speech of lay speakers of Finnish, that is, essentially nontechnical vocabulary.

The second step involved grammatical and semantic analysis of each item, prerequisite to application of the grading criteria of the three metrics. Once these values are determined, the term can be assigned a descriptivity grade. The entire corpus can then be subjected to various comparisons based on relative descriptivity which may be expected to yield information on the internal lexical patterning -- both grammatical and semantic -- of the lexical domain. In the present instance, correlations between descriptivity and the following distributions were examined: terms by section of the body (arm, leg, head, etc.), descriptive motivation (form, function, location, etc.), metaphors, and loanwords.

An additional goal of this study was to trace directions of change in terms of relative descriptivity, to note any correlations between the development of labels from descriptive terms (or vice versa) and specific semantic or grammatical areas of the lexicon.

## 2. The Metrics

### 2.1 The Sum of the Parts Equals the Whole

Semantic interpretation of linguistic forms based on the deceptively simple notion that the sum of the parts of a construction equals or does not equal the whole presents a number of problems. Expressions like cheekbone or kneecap

(sum of parts equals whole) on the one hand and butterfingers or sparrowgrass (sum of parts does not equal whole) on the other are readily analyzable in terms of such a metric. But what of cases like rider 1. 'one who rides' 2. 'proviso appended to legal document' or bookworm 1. 'worm (that eats, lives in, etc.) books' 2. 'avid reader', where one form has two quite different readings, or only partially analyzable forms such as cranberry or butcher? Then too, the sum of the parts may be equal to more than the whole as in backside or forefinger. Furthermore, in the former example both parts denote more than the whole (cf. background, backlog and underside, topside) whereas in the latter only fore- (cf. forefront, forefather) is a more inclusive term.

The function and relative productivity of the construction type (attributive, possessive, agentive, and so on) is also an important factor in determining how to add the parts. Thus blackbird, madman and darkroom represent a productive attributive type of compound, tabletop and eyelid a productive possessive or partitive type, while lifeboat and sunflower are unique or unproductive types. The function of the construction type in blackbird or eyelid, which contain no overt marker of the relationship between the two parts of the compound, must be viewed as implicit or unmarked since the type is productive in both cases. However, with examples like bookworm the additional information 'eats, lives in', is neither explicit nor unmarked in terms of the construction type. In other words, it must be supplied, in which case the isomorphism between form and meaning is incomplete and the sum of the parts cannot exactly equal the whole.

In some instances, one or more of the parts of the term may have more than one basic, but related, meaning, e.g. head (of body) vs. head (principal end of an object, as in head of a cane, wellhead, etc.). If the basic meaning is applied in the interpretation of wellhead, for example, the sum of the parts will not equal the whole; if, however, the

secondary meaning is applied the compound is isomorphic. Similarly but in another dimension, if the current meaning of the form is applied in cases like meat (in sweetmeat), the sum of the parts does not equal the whole; but if the former meaning (food) is applied, it does.

And how are we to interpret constructions composed of term plus inflectional morpheme? In the present corpus, there are several examples of terms with the plural suffix. Some of these optionally occur without the suffix: nivus ~ nivukset (pl.) 'groin', ien ~ ikenet (pl.) 'gums', while others appear only with the suffix: aivot 'brain' (but aivo- in compounds), vyötäiset 'waist'. The question here is: does the inflectional morpheme function more like a derivational unit thus contributing to the lexical meaning of the term or is it irrelevant to the semantic interpretation? In some of the Finnish examples, the addition of the plural morpheme may originally have involved a corresponding semantic addition (other than its grammatical value) to the construction but the dynamics of such a process can only be speculative at this point. Because of this, the plural morpheme in such cases has -- at least for the present -- not been taken into account in the interpretation of the form.

Summarizing these observations, five major types are discernable constituting as a whole a scale of relative descriptivity. Constructions covered by the first four are analyzable in varying degree; those subsumed under type 5 are not, i.e. are pure labels. The arbitrary values are from one to five (from most to least descriptive):

1.  $a + b = \sum(a + b)$ . The sum of the parts is equal to the whole. (unambiguous):  
korvakalvo 'eardrum' (korva 'ear', kalvo 'membrane';  
N + N, attributive compound, productive)  
lantio 'pelvis' (lant- 'hip, haunch', -io 'vessel';  
N-s, denominative noun, productive)

2.  $a + b = / \neq \xi (a + b)$ . The sum of the parts is or is not equal to the whole. (ambiguous):  
nielu 1. 'throat, pharynx' (niel- 'swallow', -u 'means'; N-s, deverbative noun, productive) 2. 'entrance'  
nivel 1. 'joint (body)' (niv- 'plait, join', -el locative; N-s, deverbative noun, mildly productive)  
2. 'articulation, node'
3.  $a + b = / \neq \overset{?}{\xi} (a + b)$ . The sum of the parts may or may not be equal to the whole. (one or more of the formally analyzable elements cannot be semantically identified):  
solisluu 'collarbone' (solis- ? , luu 'bone')  
emätin 'vagina' (emä 'womb', -t ? , -in instrument/agent)
4.  $a + b \neq \xi (a + b)$ . The sum of the parts is not equal to the whole:  
liikavarvas 'corn (toe)' (liika 'excess', varvas 'toe'; N + N, attributive compound, productive)  
munuainen 'kidney' (muna 'egg', -uainen noun suffix; N-s, denominative noun, unproductive)
5. The term is not analyzable (label):  
nilkka 'ankle'  
maha 'stomach'.

As noted above, Value 1 could be further refined in terms of various criteria. In descending order of presumed descriptivity these would be (for the Finnish data at least):

- a. Complete isomorphism between corresponding parts of form and content:  
hiusjuuri 'hair root' (hius 'hair', juuri 'root'; N + N, attributive compound, productive). See also korvakalvo above.

- b. Simple narrowing: the sum of the parts is equal to more than the whole, representing a referent class that includes that of the whole. This situation reflects semantic narrowing of an originally perfectly isomorphic term. Note, however, that the term remains unambiguous:

jalkaholvi 'arch (foot)' (jalca 'foot', holvi 'vault, vault-shaped object'; N + N, attributive compound, productive)

kovettuna 'collus' (kova- 'hard', -ttu passive participle, -na locative/diminutive; N(V-s)-s, denominative noun, productive)

- c. As (b) but with the application of a secondary rather than primary meaning to one or more of the components of the term:

sormenpää 'fingertip' (sorme- 'finger', -n genitive, pää 1. 'head' 2. 'end, tip'; N-g + N, possessive compound, productive)

silmäripsi 'eyelash' (silmä 'eye', ripsi 1. 'fringe' 2. 'cilia, eyelashes' (with plural only); N + N, attributive compound, productive)

- d. Complex narrowing: as (b) but with more than one term component representing a more inclusive referent class:

takapuoli 'behind' (taka 'back (general)', puoli 'side, part (general)'; N + N attributive compound, productive)

- e. Simile:

hauislihas 'biceps' (hauki 'pike', -s similative, lihas 'muscle'; N-s + N, similative compound, ? productive)

kives 'testicle' (kive- 'stone', -s similative).

One instance of semantic widening was noted:

häpykarva 'pubic hair' (häpy 'vulva', karva 'hair').



There were also several tautologies such as silmäripsi above and olkapää 'shoulder' (olka 'shoulder', pää 'end'). These were assigned Value 1.

As one might expect, the great majority of Value 1 examples were cases of narrowing, instances of perfect isomorphism and simile being of relatively rare occurrence, of widening unique. Value 1 terms as a whole accounted for over half of the analyzable forms.

Examples of Value 2 were rare, including the two mentioned above and possibly a third: syntymämerkki, literally 'birthmark' but also used to specify the inclusive concept 'mole'.

Of the 32 terms assigned Value 3, my impression is that the average native speaker of Finnish would probably feel that at least four of them, including the only two compounds in the group, are somehow more than labels.

Most -- and possibly all -- of the Value 4 terms are metaphors.

## 2.2. Paraphrase-Term and Term-Paraphrase Relations

The second metric is based on logical paraphrase relations between term and its most appropriate paraphrase where the latter must include the semantic components of the former. Furthermore, reversibility of term and paraphrase is a feature of this metric. Both paraphrase and term are subjected to an all-some-no test. Thus, for example, the term tabletop may be tested for paraphrase-term congruence:

all tops of tables are tabletops (true)

or for term-paraphrase congruence, the reversibility condition:

all tabletops are tops of tables (true).

At first glance, it may seem that paraphrase-term and term-paraphrase relations for a given set are merely mirror images of one another and that if, say, the logical quantifier all applies to one then it must also apply to

the other. This, however, is not necessarily so as may be shown by examples like:

- darkroom: some rooms which are dark are darkrooms  
but all darkrooms are rooms which are dark
- bearhug : all hugs of bears are bearhugs  
but some bearhugs are hugs of bears.

This approach yields nine possible combinations of paraphrase-term and term-paraphrase congruence relations:

<u>Paraphrase</u>	<u>Term</u>		
	<u>ALL</u>	<u>SOME</u>	<u>NO</u>
ALL	1	2	
SOME	3	4	
NO			5

Note that only five of the nine possible relations actually occur: The figures in the grid form the metric, from one to five in descending order of descriptivity:

	<u>Paraphrase</u>	<u>Term</u>	<u>Term</u>	<u>Paraphrase</u>
1.	all a + b	ab;	all ab	a + b
2.	all a + b	ab;	some ab	a + b
3.	some a + b	ab;	all ab	a + b
4.	some a + b	ab;	some ab	a + b
5.	no a + b	ab;	no ab	a + b

Examples:

- poskiluu 'cheekbone' (poski 'cheek', luu 'bone')  
jalkapohja 'sole (foot)' (jalka 'foot', pohja 'bottom').
- häpykarva 'pubic hair' (see p. 6)
- välikalvo 'diaphragm' (väli 'between, interval', kalvo 'membrane').  
siitin 'penis' (siit- 'cause to give birth, produce',  
-in instrument).
- nielu 'throat' (see p.4.)  
sierain 'nostril' (siera- 'chapped', -in result).

5. silmämuna 'eyeball' (silmä 'eye', muna 'egg')  
nielurisa 'tonsils' (nielu 'throat', risa 'rag').

As for the first metric, Value 1 accounts for the major share of the analyzable terms. With one exception, these are all compounds,<sup>2</sup> endocentric in nature, and primarily exhibit simple narrowing. There is only one example of Value 2, häpykarva, the unique instance of widening referred to above. Value 3 is characterized by complex narrowing, also containing the few examples of simile. Of the five forms assigned Value 4, all but one are cases of complex narrowing. Value 5 terms consist of exocentric compounds and comparable derived forms.

### 2.3. Explicit vs. Implicit

The third metric is designed to evaluate the role of the function of the construction type in terms of explicit and implicit features. The construction is assigned Value 1 (most descriptive) if all components necessary for an unambiguous reading are explicit. Although not always "explicit" in the literal sense, the internal relational or functional values of productive construction types (e.g. attribute-head, possessor-possessed, etc.) are also subsumed under Value 1. This includes such constructions as tabletop (top of a table) and toaster (that which toasts (something)).

Value 2 is assigned if some of the necessary components are implicit as inherent features of one or more of the explicit components or of the components in terms of the construction as a whole. This value would be used to characterize words like breadknife (a knife that cuts bread) and housecoat (a coat worn in the house). Note that while "cuts" and "worn" have to be added to appropriately paraphrase the terms, cutting is an inherent feature of knives and coats are normally worn.

Value 3 covers all other cases, that is, those that cannot be regarded as either explicit or implicit in the sense employed here. Thus whitecap is not a cap but a wave, a concept which is inherent in neither white nor cap nor the construction as a whole. Similarly for sparrowgrass which is neither grass nor does it have anything to do with sparrows.

The analyzable forms of the present corpus are about equally divided between values 1 and 3 with Value 2 accounting for only two or possibly three terms. With one exception Value 1 forms are characterized by simple narrowing or complete isomorphism while Value 3 is principally composed of instances of complex narrowing, similes and metaphors. Value 1 consists almost entirely of compounds, all endocentric. Value 3 contains all the exocentric compounds in the corpus.

Some examples of the three values follow:

1. virtsarakko 'bladder' (virtsa 'urine', rakko 'bladder, vesicle')  
kielenkärki 'tonguetip' (kiele- 'tongue', -n genitive, kärki 'tip').
2. rintalasta 'breastbone' (rinta 'breast', lasta 'spatula (-shaped object), splint')  
lonkkamaljakko 'hip socket' (lonkka 'hip', maljakko 'vase (-shaped object)' - lit. 'little bowl').
3. kurkunpää 'larynx' (kurku- 'throat, pharynx', -n genitive, pää 'head, end')  
nimetön 'ring finger' (nime- 'name', -tön privative, 'without, -less').

#### 2.4 Descriptivity Index

By assigning equal weight to the values of each metric, then combining the three values for each item, we can obtain a relative grade or index of descriptivity for the items in any given corpus, regardless of the semantic domain. The total number of possible sets of values is 75 (5 x 5 x 3). Of these, many can be automatically eliminated.

Thus the 15 sets containing a value of five for the first metric include only labels and may therefore be combined into a single set. Similarly, the 15 sets containing a value of three for the first metric (one or more of the formally analyzable elements cannot be semantically identified) preclude further analysis and consequently also assignment of values for the remaining two metrics. Such items are either not gradable or may be given a provisional grade of \*0 (see below) to indicate that they are provisionally classified as labels but are potentially descriptive terms. This leaves us with the 47 possible sets listed below. By totaling the values for each set, these can be further reduced to ten grades, each grade containing one or more equipollent set of values, plus 0 (label) and \*0. Abbreviations are Anal for the first metric, Para for the second, and Expl for the third:

<u>Grade</u>	<u>Anal</u>	<u>Para</u>	<u>Expl</u>	<u>Grade</u>	<u>Anal</u>	<u>Para</u>	<u>Expl</u>
0	+ 5	--	--	6	4	2	1
*0	+ 3	--	--		4	1	2
1	+ 4	5	3		+ 2	4	1
2	4	5	2		2	3	2
	? 4	4	3		2	2	3
3	4	5	1		1	5	1
	4	4	2		1	4	2
	? 4	3	3		+ 1	3	3
	2	5	3	7	4	1	1
4	4	4	1		2	3	1
	4	3	2		2	2	2
	4	2	3		2	1	3
	2	5	2		1	4	1
	+ 2	4	3		1	3	2
	1	5	3		? 1	2	3
5	4	3	1	8	2	2	1
	4	2	2		2	1	2
	4	1	3		1	3	1
	2	5	1		1	2	2
	2	4	2		1	1	3
	+ 2	3	3	9	2	1	1
	1	5	2		1	2	1
	+ 1	4	3		+ 1	1	2
				10	+ 1	1	1

The scale is in increasing order of descriptivity: from 0 (label) to 10 (maximally descriptive). The plus sign (+) to the left of a set indicates occurrence of forms in the present corpus with those values; the question mark (?), of doubtful occurrence.

### 3. Distribution of Body-Part Terms by Descriptive Grades

Now that we have quantified each item in our list of body-part terms by relative descriptivity grade within the framework of the three metrics described above, the question is: what does this tell us about those terms? In this section, I will attempt to answer that question in terms of various distributions: semantic, grammatical, and socio-linguistic, as they relate to the different grades of descriptivity.

But first, let us take a brief look at the overall grade distribution of the Finnish list:

<u>Grade</u>	<u>Number of Terms</u>	<u>Percentage of Total</u>
0	97	41.3
*0	32	13.6
1	20	8.6
2	? 1	.4
3	? 1	.4
4	1	.4
5	2	.8
6	27	11.5
7	? 1	.4
9	3	1.3
10	50	21.3

Percentages are rounded off to the nearest tenth. Grade 0 includes: 85 labels; three terms synchronically labels but historically analyzable with the respective grades \*2, \*3, and \*6; and nine labels which optionally occur as heads of tautological constructions. Grade \*0 includes: 31 Anal-3 terms and one at present unclassifiable term, haven 'beard'

(cf. \*hap- 'hair', -n genitive), which appears to be the result of earlier ellipsis. Grade 1 includes: 19 terms plus one which, if historically analyzed, would be assigned Grade 10. And finally, Grade 10 includes: 39 fully descriptive terms and 11 tautologies.

### 3.1 Terms by Body Sections

One way of examining the descriptive content of a semantic domain is to divide the terms according to some natural internal order and then note any correlations in descriptivity between these subdivisions and the terms used to designate the members of each of them. For body parts, the human body was divided into external (based chiefly on form) and internal (based chiefly on function) sections, the external section including: head, neck, upper torso, lower torso, reproductive organs, arms, legs and integument (skin and hair), and the internal section including: digestive, circulatory, excretory, secretory, skeletal, nervous, respiratory and muscular systems. In addition, a small general category was set up. Of course, some items may be considered as belonging to more than one of these categories as, for example, kulmakarva 'eyebrow' which can be classed with both head (eye) and integument (hair). Then too, some components of items in different categories are identical as, for example, kalvo 'membrane' in: korvakalvo 'eardrum' (head:ear), verkkokalvo 'retina' (head:eye), välikalvo 'diaphragm' (respiratory system). Both of these factors were taken into account in the evaluation procedure.

While somewhat irregular, a certain amount of patterning is evident. In some instances, more or less systematic descriptivity applies to a subcategory or major body part, in others to an entire category. The various categories and subcategories thus fall into roughly four groups in terms of relative descriptivity and internal taxo-

nonic systematicity. In descending order of descriptivity, these are the terms for parts of the:

1. eye (head), finger (arm), reproductive system (both male and female), skeletal system;
2. arm (other than finger, hand), ends of different body parts (kantapää 'heel', kielenpää 'tonguetip', kyynärpää 'elbow', et al), foot (leg), hair (integument), head (other than face, ear), membranes, neck;
3. butt (lower back), ear (head), excretory system, digestive system, joints, leg (other than foot), mouth (head), muscular system, respiratory system, secretory system, skin (integument), trunk (general);
4. torso (other than butt), circulatory system, face (head), general, hand (arms), nose (head).

All four high-descriptivity areas are particularly essential and physiologically specialized. They are also visually well-defined forming clear-cut complexes. However, the same might just as well be said of some low-descriptivity areas like ear or mouth. Whether this distribution is peculiar to Finnish, reflecting chance development, or contains some features common to other or all languages is, of course, a question that can only be answered in light of future comparative studies of this sector of the lexicon in other languages.

### 3.2. Grammatical and Motivational Types

In this section, I will go into the various types of grammatical constructions encountered and their descriptive semantics. Note that differences in distribution and frequency are not especially attributable to language-specific characteristics since Finnish makes extensive use of both nominal compounding and derivation.

Descriptive terms (Grades 1-10) in Finnish are formed either by compounding or derivation. At least for body-part terms, nominal compounds consist of noun + noun<sup>3</sup> in the



order determiner (or modifier, attribute) + determined. The determiner is usually in the nominative or unmarked case form but there are a few examples of genitive determiners. As regards body-part terms, it is often a moot question as to whether the determiner-determined relationship is one of attribute to head or possessor to possessed. This pattern is so predominant that compounds that have evolved from other syntactic constructions also tend to be remodeled as, for example, was the case with silmäluomi 'eyelid' (silmä 'eye'; luo- 'make, create; cast, turn toward', -ma action, -i instrument), presumably derivable from a nominalized object + verb construction (cf. luoda silmänsä maahan 'cast one's eyes down'). But alongside silmäluomi, we also find luomi with the same meaning. As a result, the former is now felt to be an attributive compound. Derived forms, denominative and deverbative, are found in a number of different functional or semantic categories: agent, result, instrument, location and diminutive, to name but the commonest. Compounds are in general more descriptive than derived forms and compounding is preferred over derivation. Derivation is largely deverbative.

Almost all of the descriptive terms on the list can be analyzed in terms of three motivational categories: form, function or location. The first of these lays emphasis on the outer form of the object as in: umpisuoli 'appendix', lit. 'closed gut or intestine', mykiö 'lens', lit. 'convex (-shaped) vessel or instrument'; the second on its function: nivelside 'ligament', lit. 'joint-binding instrument', päkiä 'ball of foot', lit. 'pressing agent'; and the third on its location: takapuoli 'behind', lit. 'back side', ohimo 'temple', lit. 'place (that is gone) by or past'. The great majority of the body-part terms belong to the formal category; the smallest number to the locational.

Combining the results of the above grammatical and motivational analyses, we find that formal terms tend to be represented more by compounds and are hence in general more

descriptive than functional terms which favor derivation over compounding. The evidence for locational terms is too limited to allow for such generalizations at this point. There also appears to be a parallel relationship -- although somewhat less marked -- between the major motivational and derivational types. That is, functional terms are more often represented by deverbatives while formal terms favor denominatives.

Of the 20-odd descriptive metaphors<sup>4</sup> found in the corpus, 15 were assigned Grade 1 and most were form-motivated. There were only three similes, all Grade 6. In addition, there was the one above-mentioned case of apparent ellipsis (haven 'beard') and possibly two or three other ellipses. While korva 'ear', synchronically unanalyzable but diachronically perhaps \*kar- 'protrude' + -va object (participle), and nimetön 'ring finger' (see p. 10), lit. 'nameless', are normally adjectives, they may occur as nominals (as predicate for example) and, as such, could be interpreted as metaphors: \*korva 'the protruding one', nimetön 'the nameless one'. The third case, kurkunpää 'larynx' (see p.10), could be analyzed as a locational ellipse: '(organ, etc. at) end of throat', or as a metaphor: 'end of throat'.

All 22 identifiable loanwords are labels, although some also appear as elements of compounds alongside native roots or constructions or, in one instance, as a derivational base. Logically, this is precisely what one would expect since all loanwords must be unanalyzable, at least to unsophisticated speakers of the target language. It will be interesting to determine, through multilingual comparison, whether this is truly a universal phenomenon or whether it may be conditioned by general or language-specific rules.

Although there were relatively few tabu words in the corpus, there is perhaps evidence of some sort of patterning as regards the relative descriptivity of tabu (i.e. vulgar) vs. corresponding polite form. While the five vulgar terms are all labels, three of their polite counterparts are descriptive and at least one of these has three different

descriptive polite alternatives.

The foregoing analysis statistically associates certain descriptivity grades with specific semantic, motivational and grammatical word types. These more or less consistent associations may be said to constitute descriptivity profiles. Thus, regardless of motivational type, Grade 10 usually points to a compound and Grade 6 to a derived form, including similes; regardless of construction type (compound, derived form), Grade 1 usually points to a form-oriented metaphor; and Grade 0, which must by definition be a label, also includes all loanwords. The question remains: are these profiles valid for or characteristic of other semantic domains in the Finnish lexicon and of body-part and other domains in languages other than Finnish?

#### 4. Dynamics of Descriptivity

If we ask ourselves how labels evolve, several possible lines of development come to mind. Some labels have undoubtedly always been labels as, for example, gas. But in many cases labels are the remains of original descriptive constructions. Forms change and may in the process give rise to corresponding changes or loss of meaning as was apparently the case with koura 'fist, closed hollow of the hand' < kop(a)ra 'hoof' < Slav. \*kapa + -ra diminutive and luomi 'eyelid' (see p. 15). Conversely, meanings may change as in koura, selkä 'back' < \*sel- '(be) long, extended' + -kä an unproductive noun suffix, or korva 'ear'. A productive construction may become unproductive incurring a loss of function or meaning as with selkä and probably chimo (see p. 15). A loanword, descriptive in the source language, is borrowed as a label in the target language as, for example, pupilli 'pupil (eye)' < Lat. pupilla (pūpa 'girl, doll', -illi diminutive). In all of these cases, narrowing of some sort is involved. We may thus posit the following chain of development from descriptive to labeling term:

narrowing                      (further narrowing)

descriptive phrase > descriptive word > label  
(or sentence)

A preterminal stage, often symptomatic of the impending final or labeling stage, is the presence of tautologies. This would be true of luomi and silmäluomi 'eyelid' and probably of näppi and sormennäppi 'fingertip' (cf. näp- 'with the fingers', -i unproductive noun suffix; sorme- 'finger') and perhaps two or three others of the nine tautological expressions found in the present corpus.

If labels can evolve from descriptive terms, is the reverse development possible? Yes, there is at least one way this happens. Folk etymologies, which are based on loanwords borrowed from another language or dialect are cases in point. Dialectal sparrowgrass, noted earlier, from standard English asparagus illustrates this kind of development. Another interesting type is represented by Rotten Row from French Route du Roi where both starting and end points are descriptive terms. Here we must posit an intermediate labeling stage on the premise that all loans are (borrowed as) labels. Thus the historical sequence here in terms of descriptivity grades would be:

10 > \*0 > 1

The unidirectional change (outside of folk etymologies): descriptive > labeling term, is diachronically supported by the few such examples in the Finnish corpus for which historical information or reconstruction is available (see selkä and korva above). There are, in addition, a number of terms which are probably or possibly applicable depending to a certain extent on whether or not they are still felt by some speakers of the language to be analyzable. For instance, is ohimo still associated with ohi or nivus 'groin' with niv- 'join' by some speakers? Then, of course, there is the abundant synchronic evidence in the form of narrowing discussed above (2.) which points toward this direction of change.

## 5. Summary and Conclusions

In this paper, I have proposed a method for measuring different degrees of descriptivity in lexical items. This consists of three logically related metrics. The first, composed of five possible degrees, evaluates words on the basis of the extent to which the meaning of the sum of the parts equals that of the construction as a whole. For the Finnish material explored in this study, Value 1 (most descriptive) terms include endocentric compounds (isomorphic, simple narrowing) and derived forms (similes, complex narrowing); Value 3 (not fully analyzable) terms are all derived forms; and Value 4 (least descriptive; Value 5 = label) terms are exocentric compounds and derived forms (primarily metaphors).

The second metric, also composed of five degrees, evaluates words on the basis of paraphrase-term and term-paraphrase truth values. Value 1 (most descriptive) terms include primarily endocentric compounds (isomorphic, simple narrowing); Value 3 terms consist of endocentric compounds (similes, complex narrowing); Value 4 terms are endocentric compounds (complex narrowing); and Value 5 (least descriptive) terms include exocentric compounds and derived forms (complex narrowing).

The third metric, composed of three degrees, evaluates words on the basis of explicit vs. implicit isomorphism between form and meaning. Value 1 (most descriptive) terms are endocentric compounds (isomorphic, simple narrowing). Value 3 (least descriptive) terms include all others.

Note that for all three metrics Value 2 terms are extremely rare.

Of the 12-grade descriptivity scale arrived at by combining values for the three metrics into equipollent sets, approximately 55% proved to be labels, 21% maximally descriptive and 9% minimally descriptive, with the remainder largely concentrated in the mid-range.

A distribution by body sections revealed that the terms for the subsystems: eye, finger, reproductive system and skeleton, rate highest on the descriptivity scale while those pertaining to the torso in general, circulatory system, face (exclusive of the eyes) and hand (exclusive of the fingers) rate lowest. The first three high-descriptivity areas are perhaps especially suggestive of some extra-linguistic motivation since they are more obviously and importantly functional for man and his survival. Also -- and this could be said of the fourth high-descriptivity area as well -- they constitute complexes of unequivocally and visibly related parts.

An examination of the corpus in terms of construction motivation indicated that three major types exist, the most favored displaying an emphasis on external form, the second on function and the third on location. Of these, the first appears chiefly in the form of compounds and is generally more descriptive than the others. The second includes many derived forms and is less descriptive than the first.

Metaphors fall mainly within the range of Grade 1 (minimally descriptive) and similes within that of Grade 6. Loanwords are all labels.

These various correlations between grades and semantic and grammatical features tend to produce characteristic descriptive profiles. Thus a Grade 10 item will more than likely be a compound, Grade 6 a derived form including similes, Grade 1 a form-oriented metaphor, and Grade 0 will include all loans.

Diachronic and synchronic evidence supports the thesis that descriptive phrases or sentences go through a process of semantic narrowing as they become descriptive words and eventually labels.

## NOTES

- \* Diese Arbeit entstand im Rahmen des Kölner Universalien-Projekts.
1. The terms "descriptive" and "label" refer to Seiler's "deskriptiv" and "Etikett" respectively (see Seiler 1974a and van den Boom 1974).
  2. Two of these contain the same prefix as the initial element but, since this and the few other prefixes found in Finnish may also occur as nominal roots with limited distribution and are not subject to vowel harmony (as most suffixes are), they may be regarded as functional roots here.
  3. There is one exception to this statement. The first element of pikkusormi 'little finger' is an adjective which is otherwise unique in that it is indeclinable.
  4. There were also, of course, several instances of metaphorical labels such as: suoni 1. 'sinew'  
2. 'vein'.

## BIBLIOGRAPHY

- Aikio, Annukka (ed.). 1969. Uusi sivistyssanakirja.  
Helsinki:Otava
- Collinder, B. 1932. Die urgermanischen Lehnwörter im  
Finnischen. Uppsala
- \_\_\_\_\_ 1955. Fenno-Ugric vocabulary: an etymo-  
logical dictionary of the Uralic languages.  
Uppsala:Almqvist & Wiksell
- Donner, O. 1874. Vergleichendes Wörterbuch der finnisch-  
ugrischen Sprachen. Helsinki:Frenckel ja pojan
- Donner, K. 1937. Verzeichnis der etymologisch behandelten  
finnischen Wörter. AASF, Ser. B, Vol. 36. Helsinki
- Erämetsä, E. 1953. Verzeichnis der etymologisch behandelten  
finnischen Wörter, Bd. II. AASF, Ser. B, Vol. 77.  
Helsinki
- Hakulinen, L. 1968. Suomen kielen rakenne ja kehitys.  
Helsinki:Otava
- Károly, S. 1968. Über die Entwicklung der Komposita in den  
finnisch-ugrischen Sprachen.  
in: Ravilla (ed.) 1965:244-250
- Oehmann-Helsinki, E. 1954. Die ältesten germanischen  
Lehnwörter im Finnischen. NAWG, I. Phil.-Hist.  
Klasse, Nr 2
- Penttilä, A. 1963. Suomen kielioppi. Porvoo/Helsinki:  
Söderström
- Ravila, P. (ed.). 1965. Congressus Secundus Internationalis  
Fenno-Ugristarum, Pars I. Helsinki
- Rufener, J. 1971. Studies in the motivation of English  
and German compounds. Zürich:Juris Druck u. Verlag



- Seiler, H. 1974a. Die Prinzipien der deskriptiven und der etikettierenden Benennung (zum Verhältnis von Prädikat und Argument). Unpub. Ms.
- \_\_\_\_\_ 1974b. Die Prinzipien der deskriptiven und der etikettierenden Benennung (Neufassung von Abschnitt 2.1. - 2.4.). Unpub. Ms.
- Tauli, V. 1966. Structural tendencies in the Uralic languages. Indiana University Publications, Uralic and Altaic Series 17
- Toivonen, Y.H., E. Itkonen, and A. Joki. 1955, 1958. Suomen kielen etymologinen sanakirja. Helsinki
- van den Boom, H. 1974. Zum Verhältnis von Logik und Grammatik am Beispiel des neuinterpretierten Carnapschen -Operators. Unpub. Ms.
- Ullmann, S. 1951. Words and their use. London

### Liste der bisherigen AKUP

1. Seiler, H. 1973, "Das Universalienkonzept", LW I  
6 - 19
2. Lehmann, C. 1973, "Wortstellung in Fragesätzen", LW I  
20 - 53
3. Ibañez, R. 1973, "Programmatische Skizze: Intonation  
und Frage", LW I, 54 - 61
4. Brettschneider, G. 1973, "'Sexus' im Baskischen: Die  
sprachliche Umsetzung einer kognitiven Kategorie" LW I,  
62 - 72
5. Stephany, U. 1973, "Zur Rolle der Wiederholung in der  
sprachlichen Kommunikation zwischen Kind und Erwach-  
senen", LW I, 73 - 98
7. Seiler, H. 1974, "The Principle of Concomitance: Instru-  
mental, Comitative and Collective (With Special Reference  
to German)", LW II, 2 - 55
6. Seiler, H. 1974, "The Principle of Concomitance in Uto-  
Aztecan", LW II, 56 - 68
8. Lehmann, C. 1974, "Prinzipien für 'Universal 14'"  
LW II, 69 - 97
9. Lehmann, C. 1974, "Isomorphismus im sprachlichen  
Zeichen", LW II, 98 - 123
10. Seiler, H. 1975, "Die Prinzipien der deskriptiven und  
der etikettierenden Benennung", LW III, 2 - 57
11. Boom, H. van den 1975, "Zum Verhältnis von Logik und  
Grammatik am Beispiel des neuinterpretierten  
 $\lambda$  - Operators", LW III, 58 - 92
12. Untermann, J. 1975, "Etymologie und Wortgeschichte"  
LW III, 93 - 116
13. Lehmann, C. 1975, "Strategien für Relativsätze"  
LW III, 117 - 156
14. Ultan, R. 1975, "Infixes and their origins" LW III,  
157 - 205

15. Stephany, U. 1975, "Linguistic and extralinguistic factors in the interpretation of children's early utterances", LW III, 206 - 233

LW I : Seiler, H. (ed.), Linguistic Workshop I, Vorarbeiten zu einem Universalienprojekt.  
München: Fink 1973

LW II: Seiler, H. (ed.), Linguistic Workshop II, Arbeiten des Kölner Universalienprojekts 1973/74

LW III: Seiler, H. (ed.), Linguistic Workshop III, Arbeiten des Kölner Universalienprojekts 1974.  
München: Fink 1975

Herausgeber: Prof. Dr. Hansjakob Seiler

Adresse: Universalienprojekt  
Institut für Sprachwissenschaft  
D - 5 Köln 41  
Universität