

## LEMMATIZATION OF MULTI-WORD TERMS IN AN ACTIVE BILINGUAL TECHNICAL DICTIONARY

Metka Brkan

### 1. Abstract

This article focuses on the problem of lemmatization of multi-word, terminological lexical units in a specialized Slovene-English dictionary. A case-study of a garment dictionary is presented. A survey was conducted among experts in the field of garment manufacturing with the aim of ascertaining where in the dictionary the informants would look up multi-word lexical units. The results of the questionnaire have shown that the majority of the Slovene informants tend to structure the notions and look up multiword specialized terms under their super-ordinate lemmata. The organization of vocabulary in alphabetical order in a prospective Slovene-English garment dictionary was recognized by the Slovene informants as an advantage as long as dictionary articles also contained elements of topical arrangement of notions. In the appendix, two model dictionary articles are drawn up for a bilingual active garment dictionary according to the findings of the questionnaire.

### 2. Introduction

In a technical encoding dictionary, one-word and multi-word terms can be lemmatized. Several theories concerning the lemmatization of multi-word units have been proposed, most of them relating to English as a source language. Some lexicographers opt for a “mechanical lemmatization principle,” which is lemmatization according to the alphabetical order of the leftmost element in a multi-word term (Nielsen 1999: 187; Svénen 1993: 207–209), while others prefer the “semantic criterium,” which is lemmatization according to the “most important word” in a word unit. Béjoint (1994: 15), Hausmann (1991: 2795), and Opitz (1990: 1509) criticize a strict alphabetic lemmatization principle because it interrupts semantically related word units. Béjoint’s study (1981) shows that users tend to look up multi-word terms under the “most important” word (for example, artificial insemination was looked up under insemination by 93% of users, false alarm under alarm by 80% of users, magnetic tape under tape by 79% of users). Hausmann (1991) suggests that a systematic approach to lemmatization should be used more frequently (for example, the word unit *Automobildach* should be sub-lemmatized under the lemma *Automobil*). Bogaards (1990) established that the principle of looking up word units is influenced by the nationality of users: Dutch students more frequently looked up the nouns in a word unit, while French students more frequently looked up the unknown word or less frequent word in a word unit, and the

verb rather than the noun. Bogaards (1996: 286) believes that a multi-word unit should be ideally lemmatized under all content words of a unit.

No survey has yet been done with respect to Slovene users' tendency in looking up multi-word units. Slovene-L2 technical dictionaries have either a systematic macrostructure (Mlakar 1989; Ogorelec 1995) or alphabetic macrostructure (Meše 1993; Humar 1996). Alphabetically arranged Slovene-L2 dictionaries follow the straight-alphabetical lemma arrangement principle (Nielsen 1994: 190), and have single- or multi-word lemmas (Meše 1993; Gams 1993), or the nest-alphabetical lemma arrangement principle, so named by Nielsen (1994: 195) (Paulin 1995; Apovnik and Karničar 1996). The principle most frequently used in recently published Slovene-L2 technical dictionaries is the straight-alphabetical lemma arrangement principle. The prevalent use of this principle can be explained by the fact that the majority of the Slovene lexicographers in special fields are field experts who, in to their opinion, do not require the assistance of a trained lexicographer.

There are, however, two exceptions to the rule—namely, the aforementioned dictionaries of Apovnik and Karničar (1996) and Paulin (1995), which are compiled according to the nest-alphabetical arrangement principle. Apovnik and Karničar (1996) is especially interesting because the dictionary articles are structured according to a consistent grammatical lemmatization principle. First, post-modified noun phrases with the lemma as a base are entered, followed by pre-modified sub-lemmas, in which the lemma functions as a base, and adjectival phrases. In Paulin (1995), however, sub-lemmas entered in the dictionary articles do not follow any special principle, but are arranged at random. One wonders why the author has not used the niche-alphabetical lemma arrangement principle (Nielsen 1994: 190–94), with sub-lemmas arranged according to the alphabet of the left-most element in a multi-word term.

The purpose of the study described in this article was to find out whether there exists a common look-up strategy among Slovene users with respect to multi-word terms in an encoding technical dictionary. This could affect the compilation of future technical dictionaries with Slovene as the source language.

## 2. Methodology

Since the purpose of this study was to find out look-up strategies of the Slovenes regarding multi-word terms, a survey was conducted among experts in the garment technology field in which twenty-three experts<sup>1</sup> were

---

<sup>1</sup> They were teachers at the secondary textile school and teachers at the Textile Department of the Faculty of Mechanical Engineering. Students were not included in the study.

asked to complete a questionnaire (Appendix I) about their prospective search of Slovene multi-word terms and about their preferences with respect to different dictionary structures.

The questionnaire consisted of four questions. Question 1 contained fifty Slovene multi-word terms from the field of garment engineering. The informants were asked to underline the word under which they would look up the term in a future Slovene-English dictionary.

In question 2, the informants were asked to specify what guided them in their search of multi-word terms. They could choose among four possible answers: (a) a search under the most important word in the multi-word term, (b) a search according to the word class, (c) a search according to the alphabetical order of the left-most member in the multi-word term, and (d) other.

Question 3 presented informants with the possibility of duplication of terms in a L1-L2 technical dictionary, because some terms can be entered under different lemmas. The informants were asked to give their opinion about possible duplication.

Question 4 of the questionnaire contained two sample dictionary articles. In the first one the multi-word terms were lemmatized according to the alphabetical order of the left-most element in the multi-word term. In the second one, the multi-word terms were lemmatized as sub-lemmas under the main lemma which was a one-word term. The informants were asked to say which type of the dictionary articles they preferred.

Data from the questionnaire were collected and processed manually. In question 1 of the questionnaire, the informants were asked to underline just one word under which they would look up the multi-word term. Some of the informants underlined two words in their answers. These answers were also considered in processing. Unanswered questions and unclear answers were considered invalid.

### **3. Results of the survey**

#### **3.1. Answers to Question 1**

Question 1 of the questionnaire contained fifty multi-word terms. The informants were asked to underline the word under which they would look the term up in a dictionary. In a questionnaire, as presented to the informants, the terms were listed at random, without being grouped according to any syntactic principle. In the results analysis below, however, the multi-word terms were grouped in nine tables, of which each contains the terms composed according to the same syntactic principle. In the tables, the Slovene terms are followed by the English counterparts for a clearer presentation. The principles of generating multi-word terms in Slovene

differ from those in English due to considerable differences in the morphology and syntax of both languages. So, to provide an adequate explanation of the look-ups, the Slovene source word is first followed by the exact English counterpart, which, in turn, is followed by the literal translation of the Slovene source word. The literal translation observes, if possible, Slovene syntactic rules. In the tables the exact English counterpart is given in inverted commas, while and the literal translation is given in double inverted commas.

In tables 1 to 9 the number of informants who would look up the multi-word terms under the first, second, or third element (abbreviated "el.") are presented with Arabic numbers. Arabic numbers are followed by a corresponding percentage in brackets.

### 3.1.1. Noun phrases pre-modified with an adjective

Tables 1 and 2 list pre-modified noun phrases in which the pre-modifier is an adjective. In Slovene, adjectives differ according to whether they determine the type of an object (so-called "type" adjectives) or the quality or characteristics of an object (so-called e. "quality" adjectives). So for example, *verižni* 'chain' in the noun phrase *verižni vbod* 'chain stitch' is a "type" adjective because it determines the type of a stitch, while *impregniran* 'impregnated' in the noun phrase *impregniran plašč* 'impregnated coat' is a "quality" adjective because it determines characteristics of the fabric the coat is made of. Sometimes, both forms of an adjective are possible, such as *kratek* and *kratki rokav* 'short sleeve', the former being a "type" adjective and the latter a "quality" adjective.

Table 1. Noun phrases pre-modified with a “type” adjective.

<i>No. and percentage of look-ups under the</i>			
	<i>first el.</i>	<i>second el.</i>	<i>several el.</i>
<i>čolnična nit</i> ‘bobbin thread’	<i>čolničen</i> ‘bobbin’ 10 (43.5 %)	<i>nit</i> ‘thread’ 13 (56.5 %)	
<i>Hoffmannova stiskalnica</i> ‘Hoffmann press’	<i>Hoffmannova</i> ‘Hoffmann’ 10 (43.5 %)	<i>stiskalnica</i> ‘press’ 13 (56.5 %)	
<i>igelno varovalo</i> ‘needle guard’	<i>igelni</i> ‘needle’ 7 (30.4%)	<i>varovalo</i> ‘guard’ 16 (69.6%)	
<i>navadni krajnik</i> ‘plain selvage’	<i>navadni</i> ‘plain’ 0	<i>krajnik</i> ‘selvage’ 23 (100%)	
<i>oblikovalno prešanje</i> ‘compression moulding’	<i>oblikovalen</i> ‘compression’ 3 (13.1 %)	<i>prešanje</i> ‘moulding’ 20 (86.9%)	
<i>verižni vbod</i> ‘chain stitch’	<i>verižni</i> ‘chain’ 4 (17.4%)	<i>vbod</i> ‘stitch’ 17 (73.9%)	<i>verižni vbod</i> ‘chain stitch’ 2 (8.7%)

Table 2. Noun phrases pre-modified with a “quality” adjective

<i>No. and percentage of look-ups under the</i>		
	<i>first el.</i>	<i>second el.</i>
<i>drobna guba</i> ‘tiny crease’	<i>droben</i> ‘tiny’ 0	<i>guba</i> ‘crease’ 23 (100%)
<i>kratek rokav</i> ‘short sleeve’	<i>kratek</i> ‘short’ 1 (4.4%)	<i>rokav</i> ‘sleeve’ 22 (95.6%)
<i>impregniran plašč</i> ‘impregnated coat’	<i>impregniran</i> ‘impregnated’ 7 (30.4%)	<i>Plašč</i> ‘coat’ 16 (69.6%)
<i>nedrsljiva podloga</i> ‘anti-slip lining’	<i>nedrsljiv</i> ‘antislip’ 1 (4.4%)	<i>podloga</i> ‘lining’ 22 (95.6%)
<i>rebrast pullover</i> ‘ribbed sweater’	<i>rebrast</i> ‘ribbed’ 2 (8.7%)	<i>pullover</i> ‘sweater’ 21 (91.3%)
<i>volnen izdelek</i> ‘woolen item’	<i>volnen</i> ‘woolen’ 14 (60.9%)	<i>izdelek</i> ‘item’ 9 (39.1%)

### 3.1.2. Noun phrases with a multi-part pre-modifier

Unlike English, multi-part pre-modifiers in Slovene can begin with or without a preposition. The informants were advised to disregard the preposition in their search. The prepositions were thus not considered in the processing of search results. The answers of the informants in which the preposition was underlined were taken as invalid. Some informants looked up a term under a certain pre-modifier element of the phrase (table 3).

### 3.1.3. Pre-modified and post-modified noun phrases

Table 4 lists pre-modified and post-modified noun phrases. In Slovene source words there is an adjective in pre-modification and a prepositional phrase in post-modification. The English counterparts, however, are only pre-modified. For a clearer presentation of the results, the literal translation of the Slovene source words is also given. The results of the look-ups are presented in table 4.

### 3.1.4. Noun phrases post-modified by a genitive

In Slovene a post-modifier may be prepositional or non-prepositional. When it is non-prepositional it has the form of a genitive. In Slovene the genitive is formed by an ending, such as *igla – igle* 'needle – of needle'. In the Slovene noun-phrases below the base of the phrase is the first noun in the phrase, i. e. the nominative, while in the English counterparts the base of the phrase is the second noun in the phrase.

### 3.1.5. Noun-phrases post-modified by a genitive and a prepositional post-modifier or two genitives

In multi-part post-modification, Slovene noun phrases can take a genitive and a prepositional item, such as *nabiranje tkanine pri transportu*, lit. "puckering of fabric in feed," or two genitives (without prepositions), such as *nastavitev dolžine vboda*, lit. "control of length of stitch." This information is necessary properly to understand the search results presented in table 6.

Table 3. Noun phrases with a multi-part pre-modifier.

Number and percentage of look-ups under the...					
	1st el.	2nd el.	3rd el.	several el.	invalid
<i>nastavljivi</i> <i>zgornji transport</i> 'variable top feed'	<i>nastavljivi</i> 'variable' 1 (4.34%)	<i>zgornji</i> 'top' 1 (4.34%)	<i>transport</i> 'feed' 15 (65.23%)	<i>zgornji</i> <i>transport</i> 'top feed' 6 (26.09%)	
<i>navzven krojeno</i> <i>krilo</i> 'flared skirt' lit. "outwards cut skirt"	<i>navzven</i> 'outwards' 0	<i>krojen</i> 'cut' 1 (4.34%)	<i>krilo</i> 'skirt' 19 (82.46%)	<i>navzven</i> <i>krojen</i> lit. "outwards cut" 1 (4.34%); <i>krojeno krilo</i> lit. "cut skirt" 1 (4.34%)	
<i>stoječ prepog-</i> <i>njen ovratnik</i> 'stiff double collar'	<i>stoječ</i> 'stiff' 0	<i>prepognjen</i> 'double' 1 (4.34%)	<i>ovratnik</i> 'collar' 20 (86.97%)	<i>prepognjen</i> <i>ovratnik</i> 'double collar' 2 (8.69%)	
<i>trajno zalikana</i> <i>guba</i> 'permanent press crease'	<i>trajno</i> 'permanent' 1 (4.34%)	<i>zalikan</i> 'press' 1 (4.34%)	<i>guba</i> 'crease' 20 (86.98%)	<i>zalikana guba</i> 'press crease' 1 (4.34%)	
<i>v gube položeno</i> <i>blago</i> 'folded fabric' lit. "into folds laid fabric"	<i>guba</i> 'fold' 12 (52.17%)	<i>položen</i> 'laid' 2 (8.69%)	<i>blago</i> 'fabric' 5 (21.75%)	<i>položeno</i> <i>blago</i> 'laid fabric' 4 (17.39%)	
<i>z zadrgo</i> <i>snemljiva</i> <i>podloga</i> 'zip-out lining'	<i>zadrga</i> 'zip' 3 (13.06%)	<i>snemljiv</i> 'removable' 6 (26.09%)	<i>podloga</i> 'lining' 12 (52.17%)	<i>snemljiva</i> <i>podloga</i> 'removable lining' 1 (4.34%)	1 (4.34%)

Table 4. Pre-modified and post-modified noun phrases.

Number and percentage of look-ups under the					
	1st el.	2nd el.	3rd el.	4th el.	several el.
<i>elastični trak v zapestju</i> 'elastic wrist band'; lit. "elastic band for wrist"	<i>elastični</i> 'elastic' 1 (4.34%)	<i>trak</i> 'band' 10 (43.48%)	<i>zapestje</i> 'wrist' 6 (26.09%)		<i>elastični trak</i> 'elastic band' 6 (26.09%)
<i>elektronsko gradiranje krojev</i> 'electronic pattern grading'; lit. "electronic grading of patterns"	<i>elektronsko</i> 'electronic' 1 (4.34%)	<i>gradiranje</i> 'grading' 13 (56.52%)	<i>kroj</i> 'pattern' 4 (17.39%)		<i>gradiranje krojev</i> 'grading of patterns' 5 (21.75%)
<i>krojilni stroj s krožnim nožem</i> 'rotary cutting machine'; lit. "cutting machine with rotary cutter"	<i>krojilni</i> 'cutting' 5 (21.75%)	<i>stroj</i> 'machine' 5 (21.75%)	<i>krožni</i> 'rotary' 0	<i>nož</i> 'cutter' 1 (4.34%)	<i>krojilni stroj</i> 'cutting machine' 11 (47.82%); <i>krožni nož</i> 'rotary cutter' 1 (4.34%)
<i>sredinska guba na hrbtnem delu</i> 'centre back pleat'; lit. "centre pleat at the back part"	<i>sredinski</i> 'centre' 0	<i>guba</i> 'pleat' 16 (69.57%)	<i>hrbten</i> 'back' 1 (4.34%)	<i>del</i> 'part' 0	<i>hrbten del</i> lit. "back part" 2 (8.60%); <i>sredinska guba</i> 'centre pleat' 4 (17.40%)
<i>stranski žep z oselcem</i> 'broad welt side pocket'; lit. "side pocket with broad welt"	<i>stranski</i> 'side' 0	<i>žep</i> 'pocket' 17 (73.2%)	<i>oslec</i> 'broad welt' 2 (8.60)		<i>stranski žep</i> 'side pocket' 3 (13.05%)
<i>vrhunska apretura proti mečkanju</i> 'creaseproof finish'; lit. "top finish against creasing"	<i>vrhunski</i> 'top' 1 (4.34%)	<i>apretura</i> 'finish' 13 (56.52%)	<i>mečkanje</i> 'creasing' 4 (17.39%)		<i>vrhunska apretura</i> 'top finish' 5 (21.75%)



Table 5. Noun-phrases post-modified by a genitive.

No. and percentage of look-ups under the		
	first el.	second el.
<i>debelina igle</i> 'needle gauge' lit. "gauge of needle"	<i>debelina</i> 'gauge' 3 (13.1 %)	<i>igla</i> 'needle' 20 (86.9%)
<i>imitacija krzna</i> 'imitation fur' lit. "imitation of fur"	<i>imitacija</i> 'imitation' 7 (30.4%)	<i>krzno</i> 'fur' 16 (69.6 %)
<i>konica ovratnika</i> 'collar point' lit. "point of collar"	<i>konica</i> 'point' 4 (17.4%)	<i>ovratnik</i> 'collar' 19 (82.6%)
<i>polaganje krojev</i> 'pattern placing' lit. "placing of patterns"	<i>polaganje</i> 'placing' 6 (26.1%)	<i>kroj</i> 'pattern' 17 (73.9%)
<i>rokav stroja</i> 'machine arm' lit. "arm of machine"	<i>rokav</i> 'arm' 9 (39.1%)	<i>stroj</i> 'machine' 14 (60.9%)
<i>zdrsniti</i> 'thread slippage' lit. "slippage of thread"	<i>zdrs</i> 'slippage' 14 (60.9%)	<i>nit</i> 'slippage' 9 (39.1%)

Table 6. Noun phrases post-modified by a genitive and a prepositional post-modifier or two genitives.

No. and percentage of look-ups under the					
	1st el.	2nd el.	3rd el.	4th el.	several el.
<i>nabiranje tkanine pri transportu</i> 'feed puckering' lit. "pucker-ing of fabric in feed"	<i>nabiranje</i> 'puckering' 10 (43.48%)	<i>tkanina</i> 'fabric' 4 (17.39%)	<i>transport</i> 'feed' 4 (17.39%)		<i>nabiranje tkanine</i> 'puckering of fabric' 5 (21.75%)
<i>nastavitev dolžine vboda</i> 'stitch length control' lit. "control of length of stitch"	<i>nastavitev</i> 'control' 2 (8.69%)	<i>dolžina</i> 'length' 3 (13.04%)	<i>vbod</i> 'stitch' 14 (60.86%)		<i>dolžina vboda</i> 'length of stitch' 4 (17.39%)
<i>obseg hlačnice na spodnjem robu</i> 'sweep' lit. "girth of leg at bottom rim"	<i>obseg</i> 'girth' 3 (13.04%)	<i>hlačnica</i> 'leg' 13 (56.52%)	<i>spodnji</i> 'bottom' 0	<i>rob</i> 'rim' 1 (4.34%)	<i>obseg hlačnice</i> 'girth of leg' 6 (26.09%)
<i>oznaka smeri niti</i> 'grain line marking' lit. "marking of grain of thread"	<i>oznaka</i> 'marking' 4 (17.39%)	<i>smer</i> 'grain' 2 (8.69%)	<i>nit</i> 'thread' 12 (52.17%)		<i>smer niti</i> lit. "grain of thread" 5 (21.75%)
<i>poraba niti na dvig igle</i> 'thread take-up' lit. "take-up of thread per path of needle"	<i>poraba</i> 'take-up' 7 (30.43%)	<i>nit</i> 'thread' 6 (26.09%)	<i>dvig</i> 'path' 0	<i>igla</i> 'needle' 3 (13.0%)	<i>dvig igle</i> 'path of needle' 2 (8.69%); <i>poraba niti</i> 'take-up of thread' 5 (21.75%)
<i>preizkušanje odpornosti proti atmosferilijam</i> 'open-air test' lit. "testing of resistance in open air"	<i>preizkušanje</i> 'testing' 6 (26.08%)	<i>odpornost</i> 'resistance' 7 (30.42%)	<i>atmosferilije</i> 'open-air' 5 (21.73%)		<i>preizkušanje odpornosti</i> 'testing of resistance' 4 (17.37%)*

\*There was 1 (4.4%) invalid answer in the search of *preizkušanje odpornosti proti atmosferilijam* 'open-air test'.

## 3.1.6. Noun phrases postmodified by a prepositional postmodifier

This section presents noun phrases post-modified by a prepositional post-modifier. The English counterparts follow a different word-formation pattern. Literal translation of the Slovene source words is given in double inverted commas. The results of the look-ups are presented in table 7.

Table 7. Noun phrases post-modified by a prepositional post-modifier.

No. and percentage of look-ups under the					
	1st el.	2nd el.	3rd el.	several el.	invalid
<i>dodatek za krojenje</i> 'trimming margin' lit. "margin for trimming"	<i>dodatek</i> 'margin' 3 (13.04%)	<i>krojenje</i> 'trimming' 19 (82.60)			1 (4.34%)
<i>izravnava po mečkanju</i> 'wrinkle recovery' lit. "recovery after wrinkling"	<i>izravnava</i> 'recovery' 7 (30.43%)	<i>mečkanje</i> 'wrinkling' 16 (69.56%)			
<i>lovilnik s centralno bobino</i> 'central bobbin shuttle' lit. "shuttle with central bobbin"	<i>lovilnik</i> 'shuttle' 20 (86.97%)	<i>centralen</i> 'central' 1 (4.34%)	<i>bobina</i> 'bobbin' 1 (4.34%)		1 (4.34%)
<i>naprava za razkodranje šiva</i> 'uncurler' lit. "attachment for the uncurling of seams"	<i>naprava</i> 'attachment' 1 (4.34%)	<i>razkodranje</i> 'uncurling' 12 (52.17%)	<i>šiv</i> 'seam' 7 (30.43%)	<i>razkodranje šiva</i> 'uncurling of seam' 3 (13.04%)	
<i>tačka za slepi vbod</i> 'blind foot stitch' lit. "foot for blind stitch"	<i>tačka</i> 'foot' 17 (73.91%)	<i>slep</i> 'blind' 0	<i>vbod</i> 'stitch' 4 (17.39%)	<i>slepi vbod</i> 'blind stitch' 2 (8.69%)	
<i>tkanina z oblikovnim spominom</i> 'shape memory fabric' lit. "fabric with shape memory"	<i>tkanina</i> 'fabric' 15 (65.21%)	<i>oblikoven</i> 'shape' 1 (4.34%)	<i>spomin</i> 'memory' 2 (8.69%)	<i>oblikovni spomin</i> 'shape memory' 3 (13.04%)	2 (8.69%)

## 3.1.7. Noun-phrases post-modified by a phrase or a clause

Table 8 lists the Slovene noun-phrases post-modified by a phrase or a clause. These Slovene terms, unfortunately, do not fulfill the requirement of the terminological theory, which states that terms should be short to ease the exchange of ideas among experts. However, the Slovene terms below have become established in the language of garment engineering in spite of their length.

Table 8. Noun phrases post-modified by a phrase or a clause.

No. and percentage of look-ups under the					
	1st el.	2nd el.	3rd el.	several el.	invalid
<i>človek, ki oblačilo nosi</i> 'wearer'; lit. "person wearing a garment"	<i>človek</i> 'person' 4 (17.39%)	<i>oblačilo</i> 'garment' 12 (52.17%)	<i>nosi</i> 'wear' 3 (13.04%)	<i>nositi oblačilo</i> 'wear a garment' 2 (8.69%)	2 (8.69%)
<i>nogavica, ki se ne para</i> 'run-resist hose' lit. "hose that does not run"	<i>nogavica</i> 'hose' 22 (95.66%)	<i>parati</i> 'run' 1 (4.34%)			
<i>sukanec, odporen proti krčenju</i> 'shrink-resistant thread' lit. "thread resistant to shrinkage"	<i>sukanec</i> 'thread' 18 (78.26%)	<i>odporen</i> 'resistant' 1 (4.34%)	<i>krčenje</i> 'shrinkage' 2 (8.69%)	<i>odporen proti krčenju</i> 'resistant to shrinkage' 2 (8.69%)	
<i>pas, ki ga stisnemo z vrvico</i> 'drawstring waist' lit. "waist drawn together with a string"	<i>pas</i> 'waist' 18 (78.26%)	<i>stisniti</i> 'draw' 0	<i>vrvica</i> 'string' 1 (4.34%)	<i>stisniti z vrvico</i> 'draw with a string' 1 (4.34%)	3 (13.04%)

3.1.8. Adjectival phrases

The Slovene adjectival phrases below (table 9) are composed of an adjective, which is the base of the phrase, and a prepositional post-modifier, consisting of a preposition and a noun. In the English counterparts, either the prepositional post-modifier or the pre-modifier is used.

Table 9. Adjectival phrases.

Number and percentage of look-ups under the...			
	1st el.	2nd el.	invalid
<i>občutljiv na pranje</i> 'sensitive to washing'	<i>občutljiv</i> 'sensitive' 4 (17.39%)	<i>pranje</i> 'washing' 17 (73.92%)	<i>na pranje</i> 'to washing' 2 (8.69%)
<i>obrobljen s čipko</i> 'lace-edged' lit. "edged with lace"	<i>obrobljen</i> 'edged' 13 (56.53%)	<i>čipka</i> 'lace' 10 (43.47%)	
<i>neodporen prot likanju</i> 'sensitive to ironing'	<i>neodporen</i> 'sensitive' 5 (21.75%)	<i>likanje</i> 'ironing' 10 (43.47%)	<i>proti likanju</i> 'to ironing' 8 (34.78%)
<i>podložen s krznom</i> 'fur-lined' lit. "lined with fur"	<i>podložen</i> 'lined' 13 (56.53%)	<i>krzno</i> 'fur' 10 (43.47%)	

3.2. Answers to Question II

In Question II of the questionnaire the informants were asked to specify what guided them in their search of multi-word terms. The answers of the informants are presented in table 10.

Table 10. Search strategy

I have searched the multi-word term			
under the most important word in the phrase	according to the word class	according to the alphabet of the left-most word in the word-phrase	Other
18 (78.3 %)	4 (17.4 %)	1 (4.3 %)	0

### 3.3. Answers to question III

In Question III of the questionnaire the informants were asked to give their opinion of the duplication of terms in the dictionary. Their answers are presented in table 11.

Table 11. Duplication of terms.

It is unnecessary to have a multi-word term lemmatized twice	It is useful to have a multi-word term lemmatized twice	Invalid
6 (26.1 %)	16 (69.6 %)	1 (4.3 %)

### 3.4. Answers to question IV

This question referred to the future Slovene-English garment dictionary. The informants were presented with two types of encoding dictionaries. In *Type A* lemmas were arranged according to the straight-alphabetic principle without any sub-lemmas. In *Type B* the multi-word sub-lemmas were entered under the single-word lemmas. The answers of the informants are presented in table 12.

Table 12. Type of the prospective dictionary

I consider the dictionary of <i>Type A</i> more user-friendly	I consider the dictionary of <i>Type B</i> more user friendly	Invalid
5 (21.7 %)	17 (74.0 %)	1 (4.3 %)

## 4. Discussion

### 4.1. Answers to Question I

Table 1 presents the noun phrases pre-modified with a “type” adjective. In these noun phrases the noun is the base of the phrase and the adjective is the modifier. The results of look-ups show that the informants tend to look up the noun more frequently than the adjective. The percentage of the look-ups under the noun is 78.2% in total. This tendency is best expressed by the noun phrase *navadni krajnik* ‘plain selvage’, which all informants (100%) said they would look up under the word *krajnik* ‘selvage’. An exception to the rule is the noun phrase *čolnična nit* ‘bobbin thread’. In this noun phrase, the adjective *čolnična* ‘bobbin’, and the noun *nit* ‘thread’, seem to be almost equally important to the informants (43.5% v. 56.5% of look-ups in favor of *nit* ‘thread’). This is also partly true for *igelno varovalo* ‘needle guard’, where the adjective *igelno* ‘needle’ shows a relatively high percentage of look-ups (30.4%). The informants seem to be

aware of the specialized meaning of the adjectives *čolnično* 'bobbin' and *igelno* 'needle' and could not have decided where exactly to search for the terms.

Table 2 lists the noun phrases formed with the so-called "quality" adjective. These were included in the questionnaire to find out whether an adjective of a more general meaning has any influence on the search results. It seems that it does influence the search, as the percentage of look-ups under the noun, or the base of the noun phrase, accounts for 81.8%, which is 9.6% higher than in the first group. Yet, what is also evident from the results is that the search is governed by a specialized rather than a general meaning of the word: the adjective *volnen* 'woolen' in the noun phrase *volnen izdelek* 'woolen item' is more specialized than the general word *izdelek* 'item'. The ratio of look-ups is 60.9% vs. 39.1% in favor of *volnen* 'woolen'. The noun phrase *drobna guba* 'tiny crease', on the contrary, shows that not a single informant would look it up under the general adjective *drobna* 'tiny', but all decided for the term *guba* 'crease'.

Table 3 presents noun phrases with a multi-part pre-modifier. Unlike in English, multi-part pre-modifiers in Slovene can begin with or without a preposition. The informants were advised to disregard the preposition in their search. The prepositions were thus not considered in the processing of search results. The results of look-ups of these noun phrases show a lower degree of correlation among the informants as compared to those of type and quality adjectives acting as pre-modifiers (tables 1 and 2). Still, the searches under nouns, which are the base of the noun phrase look-ups represented in table 3, prevail. Yet, they only account for 65.9% in total.

The percentage of look-ups under the first pre-modifier is higher when the first pre-modifier is a noun, not an adjective. Compare the noun phrase *z zadrgo snemljiva podloga*, lit. "with zip removable lining," to the first four noun phrases in table 3, in which the first element is an adjective. 13.06% of the informants would look up the phrase *z zadrgo snemljiva podloga* under *zadrga* 'zip', while the adjectives and adverbs *nastavljivi* 'variable', *navzven* 'outwards', *stoječi* 'stiff', *trajno* 'permanently' in the noun phrases *nastavljivi zgornji transport* 'variable top feed', *navzven krojeno krilo*, lit. "outwards cut skirt," *stoječ prepognjen ovratnik*, lit. "stiff double collar," and *trajno zalikana guba* 'permanent press crease' did not attract the informants' attention. An exception is the noun phrase *v gube položeno blago*, lit. "in folds laid fabric," which would be looked up under the pre-modifier *gube* 'folds' by 52.1% of informants and under the base *blago* 'fabric' by 21.7% of the informants. Again, the reason may be the specialized word *guba* 'fold' compared to the more general *blago* 'fabric'. Relative to the first two groups, the third group shows a more frequent tendency of look-ups under two elements of the noun phrase. The reason

may be that the informants consider the phrases *prepognjen ovratnik* 'double collar', *zalikana guba* 'pressed crease', *snemljiva podloga* 'removable lining', and especially *zgornji transport* 'top feed' (contrasted to *spodnji transport* 'bottom feed') as one notion and expect the term to be lemmatized under a two-word lemma. But the reason may also be that the informants were simply not sure about their search and thus underlined two words (especially *navzven krojen*, lit. "outwards cut," and *krojeno krilo*, lit. "cut skirt").

The results of the searches for pre- and post-modified noun phrases show a lower total percentage of look-ups under the base of the phrase (table 4). The search under the base only accounts for 53.6% in total. This result is greatly influenced by the noun phrase *krojilni stroj s krožnim nožem*, lit. "cutting machine with rotary cutter," which shows little search correlation among the informants. Five informants (21.75%) reported they would look it up under *krojilni* 'cutting' and *stroj* 'machine', respectively, and eleven (47.82%) under *krojilni stroj* 'cutting machine'. Interestingly, the informants understand this noun phrase differently from other noun phrases in this group. While it is evident from the look-ups of other five noun phrases in this group that the informants understand them as a type of the super-ordinate term (such as a type of band, pleat, pocket, finish, and a way of grading), this reasoning is no longer true for *krojilni stroj s krožnim nožem* 'rotary cutting machine'. The informants expect this noun phrase to be lemmatized under *krojilni stroj* 'cutting machine' (47.82% of look-ups), and not under the base *stroj* 'machine'. *Krojilni stroj* 'cutting machine' is such an important term that most of the informants expect to find it lemmatized as an individual lemma, and not as a sub-lemma under *stroj* 'machine', which seems to be too general in meaning. The percentage of look-ups under the general words is low in this group, too. One informant (4.3%) would look up the words *elastični* 'elastic', *elektronski* 'electronic', or *vrhunski* 'top', respectively, while no informant would look up the words *sredinski* 'centre' and *stranski* 'side'. A considerable high percentage of look-ups under two words is difficult to interpret. Again, some of the informants may have understood the two-word unit as one notion, such as *elastični trak* 'elastic band', *gradiranje krojev* 'grading of patterns', and *stranski žep* 'side pocket', which resulted in the search under the two-word lemma.

Table 5 lists noun phrases post-modified by a genitive. In these noun phrases (for example *debelina igle* 'needle gauge', lit. "gauge of needle") it is the noun in the nominative (for example *debelina* 'gauge'), which is the base of the phrase. It is evident from the results of the search that the informants would choose the post-modifiers more frequently than the bases in looking up these phrases (62.8% in total). This percentage is greatly reduced by the look-up results of *zdrs niti*, lit. "slippage of thread," that fourteen informants (60.9%) would look up under the base *zdrs*



‘slippage’, not under the post-modifier *nit* ‘thread’. Again, it is the specialized word *zdrs* ‘slippage’, not the more general *nit* ‘thread’, that directs the informants in their search. In other noun phrases in this group, the opposite searching strategy can be seen, i.e. they would be more frequently looked up under the post-modifier. This is a valuable finding, showing that users expect subordinate terms (part of something, such as *konica ovratnika*, lit. “point of collar,” or property of something, such as *deblina igle* lit. “gauge of needle”) to be entered under the super-ordinate term, in this case *ovratnik* ‘collar’ and *igla* ‘needle’. However, a relatively high percentage of look-ups under the base should not be neglected. The words *debelina* ‘gauge’, *imitacija* ‘imitation’, *konica* ‘point’, *polaganje* ‘placing’, and *rokav* ‘arm’ are all terms and should as such be also lemmatized in their corresponding alphabetical order.

Table 6 groups noun phrases post-modified by a genitive and a prepositional post-modifier or two genitives, such as for example *nabiranje tkanine pri transportu*, lit. “puckering of fabric in feed,” or *nastavitev dolžine vboda*, lit. “control of length of stitch.” The results of prospective look-ups of noun phrases in this group show the least correlation among the informants. Different words or phrases were underlined by the informants. Partially, the post-modifier of the noun phrase, i.e. the genitive, prevails in the search, such as *hlačnice* ‘of leg’ (56.52%) in *obseg hlačnice na spodnjem robu*, lit. ‘girth of leg at bottom rim’, or *odpornosti* ‘of resistance’ (30.42%) in *preizkušanje odpornosti proti atmosferilijam*, lit. “testing of resistance in open-air.” In case of two genitives, the informants would look up the second genitive, such as *vboda* ‘of stitch’ (60.86%) in *nastavitev dolžine vboda*, lit. “control of length of stitch,” or *niti*, lit. “of thread” (52.17%) in *oznaka smeri niti*, lit. “marking of grain of thread.” Partially, the base of the noun phrase prevails in the search, such as *nabiranje* ‘puckering’ (43.47%) in *nabiranje tkanine pri transportu*, lit. “puckering of fabric in feed,” and *poraba* ‘take-up’ (30.43%) in *poraba niti na dvig igle*, lit. “take-up of thread per pass of needle.” One possible reason that the informants selected the words *hlačnica* ‘leg’, *odpornost* ‘resistance’, *vbod* ‘stitch’, or *nit* ‘thread’ in their search may have been the fact that they considered these words to be the most important in the noun phrases, being the basic terms. The same may be true of *nabiranje* ‘puckering’. In the search of the noun phrase *poraba niti na dvig igle*, lit. “take-up of thread per pass of needle,” there was the least correlation among the informants: the percentage of look-ups under *poraba* ‘take-up’ and *nit* ‘thread’ is almost equal (30.43% v. 26.09%, or 7 vs. six informants, respectively, in favor of *poraba* ‘take-up’), which shows that there was no prevalent decision on the most important word among the informants.

Table 7 groups noun phrases post-modified by a prepositional post-modifier. The search under the base of the noun phrase prevails in three examples, i.e. under *lovilnik* ‘shuttle’ (86.97%) in the noun phrase

*lovilnik s centralno bobino*, lit. “shuttle with central bobbin,” *tačka* ‘foot’ (73.91%) in the noun phrase *tačka za slepi vbod*, lit. “foot for blind stitch,” and *tkanina* ‘fabric’ (65.21%) in the noun phrase *tkanina z oblikovnim spominom*, lit. “fabric with shape memory.” The informants understand the above noun phrases as subordinate terms to *lovilnik* ‘shuttle’, *tačka* ‘foot’, and *tkanina* ‘fabric’. In *tkanina z oblikovnim spominom*, lit. “fabric with shape memory,” the percentage of the look-ups under the base is the lowest: six informants (26.07%) all together would prefer to look up the phrase under the presumably unknown words *oblikovni* ‘shape’ (one informant or 4.34%), *spomin* ‘memory’ (two informants or 8.69%), and the phrase *oblikovni spomin* ‘shape memory’ (three informants or 13.04%). The term *naprava za razkodranje šiva*, lit. “attachment for uncurling of seams,” was not understood by the informants as a subordinate term (i.e. type of attachment): the majority of the informants (52.1%) would look the term up under the word *razkodranje* ‘uncurling’ which is a specialized word. The noun phrases *dodatek za krojenje*, lit. “margin for trimming,” and *izravnava po mečkanju*, lit. “recovery after wrinkling,” would be looked up by the majority of informants under the post-modifier, that is under *krojenje* ‘trimming’ (82.6%) and under *mečkanje* ‘wrinkling’ (69.5%). The reason may again be the more specialized word.

Table 8 groups noun phrases post-modified by a phrase or a clause. All noun phrases in this group, with the exception of the first one, show the prevalent searching strategy under the base of the noun phrase, such as under *nogavica* ‘hose’ (95.66%), and under *sukanec* ‘thread’ and *pas* ‘waist’ (78.2% of the informants, respectively). A very important English term in garment technology, a “wearer,” has a Slovene counterpart *človek, ki oblačilo nosi*, lit. “person wearing a garment.” We will not go into a discussion of the appropriateness of this term here, as this is not the topic of this article. Let us mention, however, that it is not always possible in Slovene to use short derivatives instead of complex phrases as terms. The Slovene derivative *nosilec*, from the verb *nositi* ‘to wear’, ‘to carry’, ‘to bear’, would sound strange if used in garment technology. *Nosilec* is a general word denoting someone who “carries something” or “promotes something,” and a specialized civil engineering term, a ‘beam’. Its use in the meaning of “someone who wears something” may have been justified only if used with a post-modifier, for example *nosilec oblačila*, lit. “wearer of a garment,” although the *Dictionary of the Slovene Language* (Bajec 1994) does not allow this use of the word. Therefore the long term *človek, ki oblačilo nosi*, lit. “person wearing a garment,” was adopted in garment manufacture. In our questionnaire only four informants would search this phrase under the general word *človek* ‘person’, most of them would look it up under *oblačilo* ‘garment’ (twelve informants or 52.1%, respectively). The percentage of look-ups under *nositi* ‘to wear’ and *nositi oblačilo* ‘to wear a garment’ accounts for 21.73% in total and is not negligible. In an

encoding dictionary, the term could be also entered under the verb *nositi* 'to wear'.

Table 9 lists adjectival phrases. In these phrases the base is the adjective, for example *občutljiv* 'sensitive', *obrobljen* 'edged', *neodporen* 'sensitive', and *podložen* 'lined'. As evident from the results of prospective look-ups, the informants would either choose the base or the modifier in their search for the word. The adjectives *obrobljen* 'edged' and *podložen* 'lined' seem to be more important to the informants than the nouns *čipka* 'lace' and *krzno* 'fur' which are in the post-modifying position in Slovene. The difference in percentage of look-ups, however, is not very significant, namely 56.53% v. 43.47% in favor of the adjective. The nouns *čipka* 'lace' and *krzno* 'fur' seem to be almost as important to the informants as the adjectives *obrobljen* 'edged' and *podložen* 'lined'. This is understandable as all the above words are terms and the informants feel that they should be entered in a specialized dictionary.

In the other two adjectival phrases presented to the informants, i.e. *občutljiv na pranje* 'sensitive to washing' and *neodporen proti likanju* 'sensitive to ironing', the search under the noun (i.e. the modifier) prevails, namely under *pranje* 'washing' (73.92%) or under *likanje* 'ironing' (43.47%). The search of these two phrases under two elements, i.e. *na pranje* 'to washing' and *proti likanju* 'to ironing' is illogical as the first elements are prepositions and the informants were advised not to observe them. It may be that the informants wanted to underline the noun in these two cases too, but were simply careless in doing so, adding the preposition to the noun. If the percentage of look-ups under two elements is added to the percentage of look-ups under the noun, we get 82.61% for *pranje* 'washing' and 78.25% for *likanje* 'ironing'. The average percentage of look-ups under the noun (i.e. the modifier in this group) is thus 61.95%. With respect to the adjectives, the informants give preference to more specialized adjectives, such as *obrobljen* 'edged' and *podložen* 'lined' as compared with *občutljiv* 'sensitive' or *neodporen* 'sensitive', which only four (17.39%) or five (21.75%) informants, respectively, would look up.

From the results of Question I some common findings can be drawn, valid for the search of Slovene multi-word units by the Slovene informants. The Slovene informants would look up:

- the noun phrases pre-modified by an adjective or adjectival phrase under the noun or the base, respectively, for example *navadni krajnik* 'plain selvage' under *krajnik* 'selvage', *trajno zalikana guba* 'durably pressed crease' under *guba* 'crease', *stoječ prepognjen ovratnik* 'stiff double collar' under *ovratnik* 'collar';
- pre-modified and post-modified noun phrases under the noun or the base respectively, for example *stranski žep z oslecem* 'broad welt side pocket', lit. "side pocket with broad welt," under *žep* 'pocket',

*vrhunska apretura proti mečkanju* 'creaseproof finish', lit. "top finish against creasing," under *apretura* 'finish', *elektronsko gradiranje krojev* "electronic pattern grading," lit. "electronic grading of patterns," under *gradiranje* 'grading';

- noun phrases post-modified by a prepositional post-modifier under the noun or the base, respectively, such as *lovilnik s centralno bobino* 'central bobbin shuttle', lit. "shuttle with central bobbin," under *lovilnik* 'shuttle', *tačka za slepi vbod* 'blind stitch foot', lit. "foot for blind stitch," under *tačka* 'foot', and noun phrases post-modified by a phrase or a clause, such as *nogavica, ki se ne para* 'run-resist hose', lit. "hose that does not run," under *nogavica* 'hose', *pas, ki ga stisnemo z vrvico* 'drawstring waist', lit. "waist drawn together with a string," under *pas* 'waist';
- noun phrases post-modified by a genitive (i.e. non-prepositional post-modifier) under the noun or post-modifier, respectively; for example, *rokav stroja* 'machine arm', lit. "arm of machine," under *stroj* 'machine', *konica ovratnika* 'collar point', lit. "point of collar," under *ovratnik* 'collar', *polaganje krojev* 'pattern placing', lit. "placing of patterns," under *kroj* 'pattern'.

Analysis of the questionnaire showed that the majority of the informants structured their knowledge and arranged the notions into a system. The informants would search subordinate terms under their super-ordinate term, and under the term that means the whole, not the part. They would also search for specialized words rather than for general words. We can say that the informants used a "logical approach" in their search of multi-word terms, concentrating on "the most important word" in the phrase.

#### 4.2. Answers to Question II

Within the questionnaire, the informants were also asked to specify what led them in the search of the multi-word units. Eighteen (78.3%) informants answered that they searched the multi-word unit under the most important word in the unit, four (17.4%) under the word class and one (4.3%) alphabetically, under the left-most element in the unit. Mihaljević (1991) argues that the "most important" word differs for each user and that the semantic criterion is a subjective criterion. Yet, the results of the questionnaire showed that in some cases at least there was strong agreement (even a 100%) between the informants as to the most important word in the multi-word unit (for example *navadni krajnik* 'plain selvage'). It seems that for the informants the most important word in a multi-word unit is the word that denotes the super-ordinate term, not the subordinate, for example *stiskalnica* 'press' in the unit *Hoffmannova stiskalnica* 'Hoffmann press', and the word that denotes the whole, not the part of a unit, for example *stroj* 'machine' in the word-unit *glava stroja* 'machine head'. Considering this

definition of “the most important” word in a word unit, the agreement between the informants’ answer (i.e. “I looked up the multi-word units under the most important word in the unit”) and their actual search was 96.1% for pre-modified noun phrases at least (i.e. the multi-word terms in tables 1, 2, and 3).

The questionnaire results also showed that the search of multi-word units according to the word class is not frequent. Namely, only four informants (17.4%) answered that they looked-up the multi-word units according to the word class, i.e. under the noun, or under the first noun if there were two nouns, or under the adjective if there was no noun in the unit (cf. Mihaljević 1991: 139). Interestingly, their answers do not agree with their actual searching principle. Namely, the informants looked up the multi-word units *imitacija krzna*, lit. “imitation of fur,” *dodatek za krojenje*, lit. “margin for trimming,” *naprava za razkodranje šiva*, lit. “attachment for uncurling of seams,” *krojilni stroj s krožnim nožem*, lit. “cutting machine with rotary cutter,” and *impregniran plašč* ‘impregnated coat’ under the words *krzno* ‘fur’, *krojenje* ‘trimming’, *razkodranje* ‘uncurling’, *krojilni* ‘cutting’, and *impregniran* ‘impregnated’, which are not the first nouns in the multi-word units, or nouns at all (some words are adjectives, such as *krojilni* ‘cutting’ and *impregniran* ‘impregnated’). It is evident from the answers that the informants are not familiar with word classes. One informant (4.3%) answered that he/she looked up the multi-word units under the leftmost element in the multi-word unit. This answer, too, does not agree with the informant’s actual search: he/she underlined the words *nit* ‘thread’, *prešanje* ‘moulding’, *stroj* ‘machine’ and *kroj* ‘pattern’ in the multi-word units *čolnična nit* ‘bobbin thread’, *oblikovno prešanje* ‘compression moulding’, *rokav stroja*, lit. “arm of machine,” *polaganje krojev*, lit. “placing of patterns,” which are not the leftmost elements in the unit.

### 4.3. Answers to Question III

The third question in the questionnaire referred to the frequency of lemmatization of multi-word terms. Sixteen informants (69.6%) thought it useful to have a multi-word term lemmatized twice in a dictionary, and only six informants (26.1%) considered this principle an unnecessary reduplication. In practice, however, it is not possible to lemmatize all multi-word terms twice or even three times in a dictionary. As to the cross-references, they are generally used to direct the user to the dictionary article that contains additional information about the term. All multi-word terms cannot be cross-referenced; this would overburden the dictionary. A sensible criterion would thus be to adopt the practice of lemmatizing “all important words” (i.e. semantically meaningful words) of a multi-word term.

#### 4.4. Answers to Question IV

The last question in the questionnaire referred to the prospective L1-L2 garment dictionary. The informants were presented with two types of prospective dictionaries in which terms were lemmatized according to different lemmatization principles. In Type A, single-word and multi-word terms were arranged according to the straight-alphabetic principle. Type B contained single-word lemmas and multi-word sub-lemmas entered in the article in sub-sections, so that post-modified terms containing the lemma were followed by pre-modified terms with the lemma at the second place. The informants were asked to choose the dictionary they would prefer in active work. The informants were only given the Slovene terms. The English equivalents were added in this paper for clarity.

The analysis of the informants' answers shows that five informants (21.7%) would prefer Type A of a dictionary, while seventeen informants (74.0%) would prefer Type B of a dictionary. One answer was invalid. Yet, a curious disagreement was observed between the answers of those informants who opted for Type A of the dictionary and their way of looking-up multi-word terms. Saying that they preferred Type A of the dictionary, they would have had to underline all initial words in the presented word units. This was however not the case. Out of fifty multi-word terms, the informant number 12 underlined only twelve initial terms, the informant number 14 only seven initial terms, the informant number 19 only nineteen initial terms, the informant no. 20 only 10 initial terms, and the informant number 22 only sixteen initial terms. The reason for such a disagreement between their preference for a dictionary and their prospective way of looking-up multi-word terms is not known. But it can be guessed that when presented with two types of dictionary arrangements, the informants consciously decided for Type A, but they were led by some subconscious reason, possibly the tendency to systematization, in their search.

#### 5. Conclusion

Twenty-three informants, specialists in the field of garment manufacture, were asked to complete the questionnaire on prospective search of multi-word terms in a Slovene-English specialized dictionary. The results of the survey have shown that the majority of the informants would look up the Slovene multi-word terms under the noun, or under the base of the noun-phrase, respectively. The only exception was the group of Slovene noun phrases post-modified by genitives. These noun phrases would be looked up more frequently under the post-modifier, not the base of the noun phrase, for example *konica ovratnika*, lit. "point of collar," under *ovratnik* 'collar', which is the post-modifier in this case.

The analysis of the questionnaire also showed that the majority of the informants structured their knowledge and arranged the notions into a system. The informants would search subordinate terms under their super-ordinate terms, and under the term that means the whole, not the part. We can say that the informants used a 'logical approach' (i.e. semantic criterion) in their search of multi-word terms, concentrating on 'the most important word' in the phrase. This finding could have a direct influence on the principles of compiling Slovene-foreign language specialized dictionaries.

Another finding of the study was that the informants looked-up specialized words more frequently than general words. This should also be considered in compiling a Slovene-foreign language specialized dictionary. For example, the noun-phrase *zdrs niti* 'thread slippage', lit. "slippage of thread," should be lemmatized under *nit* 'thread', according to the principle of subordination of terms, and under *zdrs* 'slippage', according to the principle of specialization of terms. The principle of entering in a dictionary all Slovene specialized words from a multi-word term is valid especially for highly complex multi-word terms. An example of such is the multi-word term *z zadrgo snemljiva podloga* 'zip-out lining', which should be lemmatized under *zadrga* 'zip' and *podloga* 'lining'.

The semantic principle of lemmatization (i.e. the lemmatization of subordinate terms under their super-ordinate terms, and terms expressing parts under the term expressing the whole) is applicable, yet it does not yield totally error-free searching results. It is impossible to completely foresee the searching strategy of the user. Therefore, the user should be advised in the user's guide to the dictionary to look up a multi-word term under the second element if he/she is unable to find it under the first one. Double lemmatization is not contrary to the results of the questionnaire because 69% of the informants opted for it. The semantic lemmatization principle, however, has advantages over a strictly "mechanical" lemmatization principle (i.e. the lemmatization of a multi-word term according to the alphabetical order of the first element in a term). This is especially true in case of great variability of terms or the lack of standardized terms in a field. Slovene textile terms often exhibit several variations of a term for one notion, with only slight lexical differences in the term's left-most element, yet with the same semantic meaning. Such is the case, for example, with the weaving term *izmenjava/zamenjava/menjava čolničkov* 'shuttle changing', lit. "changing of shuttles." The user, not knowing the exact lexical form of this term's first element, would be greatly helped if the term was lemmatized under the super-ordinate term *čolniček* 'shuttle'.

## Appendix A: Questionnaire

### Question 1

Below you will find fifty Slovene multi-word terms from the field of garment engineering. Please, underline the word under which you would look up the term in a future Slovene-English dictionary. Example: You want to find the English counterpart of the multi-word term *igelni žleb*. If you choose to look it up under the lemma *igelni*, underline igelni; if not, underline žleb.

*čolnična nit* ‘bobbin thread’ \* \*  
*vrhunska apretura proti mečkanju* ‘creaseproof finish’  
*dodatek za krojenje* ‘trimming margin’  
*navadni krajnik* ‘plain selvage’  
*zdrs niti* ‘thread slippage’  
*nogavica, ki se ne para* ‘run-resist hose’  
*stranski žep z oslecem* ‘broad welt side pocket’  
*navzven krojeno krilo* ‘flared skirt’  
*pas, ki ga stisnemo z vrvico* ‘drawstring waist’  
*nedrsljiva podloga* ‘anti-slip lining’  
*obrobljen s čipko* ‘lace-edged’  
*oblikovalno prešanje* ‘compression moulding’  
*impregniran plašč* ‘impregnated coat’  
*Hoffmannova stiskalnica* ‘Hoffmann press’  
*nastavljivi zgornji transport* ‘variable top feed’  
*sredinska guba na hrbtnem delu* ‘centre back pleat’  
*imitacija krzna* ‘imitation fur’  
*konica ovratnika* ‘collar point’  
*neodporen proti likanju* ‘sensitive to ironing’  
*trajno zalikana guba* ‘durable crease’  
*človek, ki oblačilo nosi* ‘wearer’  
*tkanina z oblikovnim spominom* ‘shape memory fabric’  
*volnen izdelek* ‘woolen item’  
*elektronsko gradiranje krojev* ‘electronic pattern grading’  
*izravnava po mečkanju* ‘wrinkle recovery’  
*debelina igle* ‘needle gauge’  
*kratek rokav* ‘short sleeve’  
*rebrast pullover* ‘ribbed sweater’  
*občutljiv na pranje* ‘sensitive to water’  
*drobna guba* ‘tiny crease’  
*oznaka smeri niti* ‘grain line marking’

---

\* Only Slovene multi-word terms were presented to the informants, the English counterparts were added here for clarity. The same is true for Question 4.



*oznaka smeri niti* 'grain line marking'  
*nabiranje tkanine pri transportu* 'feed puckering'  
*rokav stroja* 'machine arm'  
*naprava za razkodranje šiva* 'uncurler'  
*polaganje krojev* 'pattern placing'  
*poraba niti na dvig igle* 'thread take-up'  
*tačka za slepi vbod* 'blind stitch foot'  
*v gube položeno blago* 'folded fabric'  
*nastavitev dolžine vboda* 'stitch length control'  
*elastični trak v zapestju* 'elastic wrist band'  
*lovilnik s centralno bobino* 'central bobbin shuttle'  
*sukanec, odporen proti krčenju* 'shrink-resistant thread'  
*igelno varovalo* 'needle guard'  
*stoječi prepognjeni ovratnik* 'stiff double collar'  
*podložen s krznom* 'fur-lined'  
*obseg hlačnice na spodnjem robu* 'sweep'  
*verižni vbod* 'chain stitch'  
*krojilni stroj s krožnim nožem* 'rotary cutting machine'  
*z zadrgo snemljiva podloga* 'zip-out lining'  
*preizkušanje odpornosti proti atmosferilijam* 'open-air test'

---

## Question 2

In question 1 you underlined the words under which you would look up multi-word terms. Please indicate which principle led you in your search.

I looked up multi-word terms:

- under the most important word in the multi-word term;
- according to the word class, i. e. (1) under the noun, (2) under the first noun if there were two or more nouns, (3) under the adjective if there was no noun, (4) under the verb if there was no noun or adjective;
- according to the alphabet of the left-most member in the multi-word term;
- other. Please specify .....

## Question 3

If multi-word terms were entered twice or three times in an encoding dictionary (i.e. under all important words of a word-phrase), this would be, according to your opinion:

- unnecessary duplication, because I always search the term under the other element if I do not find it under the first one searched;
- welcome, as I could find the term more quickly.

**Question 4**

Below you will find two types of the prospective Slovene-English technical dictionary. Type A lemmatizes one-word and multi-word lemmas which are arranged according to the alphabet of the leftmost element in the multi-word term. In Type B multi-word terms are entered as sub-lemmas under the main lemma which is a one-word term. Which type of the dictionary do you prefer?

***Type A***

*gumb button*

*gumb iz biserne matice mother-of-pearl button*

*gumb ovratnika collar button*

*gumb s pecljem shank button*

*gumb s stblom shank button*

*gumb z dvema luknjicama two-hole button*

*gumb z luknjicami pierced button*

*gumb z obodcem shank button*

*gumb za izbiro cikcak šiva zigzag control*

*gumb za izbiro vboda stitch selector*

*gumb za nastavitev vbodov stitch adjusting screw*

*gumb za posteljno perilo linen button*

*gumbna letvica button facing*

*gumbna prevleka button covering*

*gumbna zanka button loop*

*gumbnica buttonhole*

*gumbnična letvica buttonhole facing*

***Type B***

*gumb button*

*gumb iz biserne matice mother-of-pearl button*

*gumb ovratnika collar button*

*gumb s pecljem shank button*

*gumb s stblom shank button*

*gumb z dvema luknjicama two-hole button*

*gumb z luknjicami pierced button*

*gumb z obodcem shank button*

*gumb za izbiro cikcak šiva zigzag control*

*gumb za izbiro vboda stitch selector*

*gumb za nastavitev vbodov stitch adjusting screw*

*gumb za posteljno perilo linen button*

*manšetni gumb cufflink*

*neobdelani gumb blank button*

*oblečeni gumb fabric button*

*okrasni gumb fancy button*  
*okrogli gumb ball button*  
*pritisni gumb press stud*  
*slepi gumb decorative button*  
*srajčni gumb shirt button*  
*surovi gumb blank button*  
*zatični gumb toggle fastening*

***gumbni***

*gumbna letvica button facing*  
*gumbna prevleka button covering*  
*gumbna zanka button loop*

***gumbnica buttonhole***

*obrojena gumbnica bound buttonhole*  
*ravna gumbnica straight buttonhole*  
*s trakom podložena gumbnica gimped buttonhole*  
*slepa gumbnica dummy buttonhole*

***gumbnični***

*gumbnična letvica buttonhole facing*

## Appendix B

In Appendix B two entries are suggested for the prospective Slovene-English garment dictionary. The dictionary articles contain (1) a section of post-modified noun phrases with the lemma as the base, (2) a section of pre-modified noun phrases with the lemma as the base, and (3) a section of post-modified noun phrases with the lemma as a post-modifier. The collocation field is added. The polysemy of the lemma is indicated by Arabic numbers, the equivalents contain grammatical information (number in nouns if it differs from Slovene, attributive and predicative use in adjectives, verb patterns), examples of use (when the use of the English equivalent differs from that of the Slovene one), encyclopedic information (glosses, field qualifiers, explanations, definitions of terms), and cross-references, as suggested in lexicographical literature (Scholfield 1980: 98-109, Christiansen, Duva and Laursen 1994: 269-284, Bergenholtz 1994b: 421-440, Bergenholtz and Pedersen 1994: 352-384, Haensch 1991: 2937-2948).

- guba** *sam* 1. (pregib blaga) fold (of cloth) 2. (namerno preganjena) pleat, plait: *krilo z gubami* pleated skirt / *Ko izdelujemo gube, preganemo blago po črti pregiba*. When forming pleats the fabric is folded along the foldline. 3. (nabrana) gather 4. KROJ (pri polaganju plasti) fold 5. (okrasna, prečna, na rokavih) tuck 6. (rob pri hlačah) crease: *zalikati gubo pri hlačah* to iron a crease into one's trousers. 7. (zaradi mečkanja) crease, wrinkle 8. (zaradi napačnega krojenja) gather, pucker, wrinkle, inadvertent pleat: *Od mednožja čez stegno se delajo gube*. Wrinkles extend from the crotch across the thigh.
- **guba pri ravnem pliseju** accordion pleat
  - guba pri sončnem pliseju** sunburst pleat
  - **enostranska guba** knife pleat
  - koračna guba** kick pleat short pleat at bottom edge of a skirt [kratka guba na spodnjem robu krila]
  - ležeča guba** *glej enostranska guba*
  - navidezna guba** *glej slepa guba*
  - notranja guba** inverted pleat double pleat formed when two knife pleats meet on the right side [dvostranska guba, pri kateri sta leva in desna guba položeni ena proti drugi na zunanji strani] ► *prim zunanja guba*
  - notranja podložena guba** inverted pleat with separate underlay
  - slepa guba** mock pleat a pleat that does not open [guba, ki se ne odpira]
  - sredinska guba na hrbtnem delu** centre back pleat
  - trajna guba** permanent crease
  - trajno zalikana guba** 1. (na krilu) heat-set pleat, permanent press pleat  
2. (rob na hlačah) durable crease, permanent crease
  - varčna guba** spaced pleat
  - votla guba** *glej zunanja guba*

**vzdolžna guba** lengthwise crease

**zalikana guba** (pri hlačah) crease

**zunanja guba** box pleat double pleat formed by two knife pleats pressed away from each other at the right side [dvostranska zunanja guba, pri kateri sta leva in desna guba položeni ena proti drugi na notranji strani] ► *prim notranja guba*

■ **polaganje gub** pleating of fabric, laying of pleats; (enostransko) knife pleating; (okrasnih, prečnih, v celoti ali delno pošitih) tucking [up] [of fabric]

**tvorjenje gub** (nenamerno, mečkanje) wrinkling, formation of wrinkles

**v gube položeno blago** pleated fabric

■ **guba se oblikuje/tvori** the pleat forms (*Gube se tvorijo lepše in so bolj obstojne, če so položene vzporedno s krajnikom*. Pleats form more readily and have more staying power when they are laid in parallel alignment to the selvages.); **guba z ostrim robom** sharp pleat; **drobna guba** tiny tuck; (zaradi mečkanja) tiny crease; **mehka guba** soft pleat; **pošita guba** topstitched pleat; **delati gube** (nenamerno) to wrinkle (*Obleka dela gube čez hrbet*. The garment wrinkles across the back); **izdelati gubo** to form a pleat; **izgladiti gube** to take out creases; **položiti gube** (*na tkanini*) to pleat (the fabric), to lay the pleats; (drobne, okrasne) to tuck [up] [the fabric]; **zatrditi/pošiti gubo** to stitch down a pleat, to topstitch a pleat; (delno) to part-stitch a pleat, stitch a pleat part way; (po vsej dolžini) to stitch a pleat down its whole length

**oblačilo** *sam* garment, piece of clothing, item of clothing, clothes item; **oblačila** garments, clothes *mn*, clothing, AM apparel, AM wearing apparel; (ločeni, neformalni kosi oblačil, npr. hlače, puloverji, krila, itd.) separates *mn*; (iz istega blaga ali barvno usklajena oblačila) coordinates *mn*: *Oblačila se proizvajajo v serijah*. Garments are produced in lots. / *Oblačila so še vedno eden najvidnejših statusnih simbolov*. Clothing is still one of the most visible status symbols. / *Oblačila so eden najosebnejših masovnih izdelkov*. Clothing is one of the most personal of mass-produced products. / *Pletena oblačila so enostavna za vzdrževanje in se ne mečkajo*. AM Knit apparel is easily cared for and does not wrinkle. (OPOMBA: Samostalnika **clothing** in **apparel** nimata množine. Vežeta se z glagolom v ednini. Termina **garment(s)** in AM **apparel** sta v konfekciji pogostejša kot termina **clothes** in **clothing**.) *Oblačila za dež so vodo-odporna ali vodo-odbojna*. Rainwear is water-repellent or waterproof. (OPOMBA: Zloženke z besedo **wear**, kot npr. **babywear**, **beachwear**, **vacation wear**, nimajo množine. Vežejo se z glagolom v ednini).

■ **oblačila s svetlobnimi odbojniki** signal clothing

**oblačila z obešalnika** (konfekcijska) *neform* reach-me-downs *mn*

**oblačila za dež** rainwear, AM rain apparel

**oblačila za dojenčke** baby garments, babywear

**oblačila za dom** home wear

**oblačila za mladostnike** juvenile clothing

- oblačila za plažo** beachwear, beach wear  
**oblačila za počitnice** vacation wear  
**oblačila za prosti čas** leisure wear  
**oblačila za ulico** outdoor clothing, outdoor wear  
 ■ **azbestna oblačila** asbestos clothing  
**dekliška oblačila** girls's garments, girl's wear  
**deška oblačila** boy's garments, boy's wear  
**civilna oblačila** plain clothes *mn*  
**dizajnerska oblačila** designer clothes *mn*  
**fantovska oblačila** *glej* **deška oblačila**  
**formalna oblačila** formal dress, formal wear, formal clothes; (za slovesnosti) full dress  
**konfekcijska oblačila** ready-made clothing, ready-made garments *mn*, ready-to-wear clothes *mn*, AM ready-to-wear apparel  
**moška oblačila** men's clothing, men's wear  
**neformalna oblačila** casual wear, casuals *mn*  
**nosečniška oblačila** maternity wear  
**oprijeta raztegljiva oblačila** (za telovadbo in prosti čas) bodywear  
**otročka oblačila** (do treh let) infant's garments, infant's wear; (od treh do dvanajst let) children's clothing, children's wear, juvenile clothing  
**otročka spalna oblačila** children's sleepwear  
**poletna oblačila** summer clothes, summerwear  
**spalna oblačila** sleepwear, slumber wear  
**večerna oblačila** evening dress, evening wear, *form* evening attire  
**vrhnja oblačila** outer garments *mn*, outerwear, outer-wear  
**vojaška oblačila** (za vaje, manevre) fatigue clothes *mn*, fatigues *mn*  
**zaščitna oblačila** protective clothing  
**zaščitna športna oblačila** protective athletic clothing  
**zaščitno oblačilo** protective garment  
**zimska oblačila** winterwear  
**žalna oblačila** mourning wear  
 ■ **kopija oblačila** copy of a garment design, *neform* knock-off, *mn* knock-offs: *Kopija oblačila je veliko cenejša verzija originala.* A knock-off is a much less expensive version of an original.  
**pristajanje oblačila** fit of a garment ► *glej tudi* **pristajanje**  
**videz oblačila** appearance of a garment  
 ■ **kakovostno oblačilo** garment of quality; **najkakovostnejše oblačilo** the highest quality garment; **nekakovostno oblačilo** garment of low/poor quality; **ohlapno oblačilo** loose-fitting garment; **tesno oblačilo** close-fitting garment; **oblačilo slabše kakovosti** garment of lower/poorer/inferior quality, second quality garment; **oblačilo z napako** faulty garment; **oblačilo dobro stoji (na telesu)** the garment sets properly (on the body); **oblačilo gre za med neform** the garment is selling like hot cakes *ali* the garment is walking out of the stores; **oblačilo leze navzgor** the garment shifts upwards; **oblačilo lepo pada** the garment drapes

nicely; *oblačilo pada* (*od telesa, enako spredaj in zadaj*) the garment falls/hangs/extends (from the body the same distance front to back); *izdelati oblačilo* to make a garment; *(iz)meriti oblačilo* to take measurements of a garment; *kopirati oblačilo* to copy a design of a garment, *neform* to knock off a garment; *kreirati oblačilo* to design a garment; *pregledati kakovost oblačila* to inspect a garment for quality; *proizvajati oblačila* to manufacture/produce garments; *sparati oblačilo* to take a garment apart at seams, to disassemble a garment; *vzeti komu mero za oblačilo* to take measure for someone's clothes; *zapreti oblačilo* to close a garment

## Works Cited

## A. Dictionaries

- Apovnik, Pavel and Ludvik Karničar. 1996. *Slovar pravnega in ekonomskega jezika. 2. del. Slovensko-nemški*. Dunaj: Manz.
- Bajec, Anton, Janko Jurančič, Mile Klopčič, Lino Legiša, Stane Suhadolnik, and France Tomšič, ed. 1994. *Slovar slovenskega knjižnega jezika*. Ljubljana: DZS.
- Brkan, Metka. 2001. *Angleško-slovenski slovar konfekcijske tehnologije*. Maribor: Univerza v Mariboru, Fakulteta za strojništvo.
- Gams, Mataž, ed. 1993. *Računalniški slovarček: angleško-slovenski, slovensko-angleški*. 3rd ed. Ljubljana: Cankarjeva založba.
- Hohenadel, Paul and Jonathon Relton. 1984. *A Modern Textile Dictionary*. Vol. 1, *English-German*. Wiesbaden: Oscar Brandstetter Verlag.
- Humar, Marjeta, ed. 1996. *Papirniški terminološki slovar*. Ljubljana: Znanstvenoraziskovalni center SAZU.
- Meše, Pavel. 1993. *Angleško-slovenski, slovensko-angleški slovar telekomunikacij*. Ljubljana: Smart Com.
- Mlakar, Frank. 1989. *Mednarodni slovar osnovnih in splošnih izrazov s področja metrologije*. Ljubljana: Slovensko društvo za merilno in procesno tehniko, Elektrotehniška zveza Slovenije.
- Ogorelec, Anton, ed. 1995. *Slovenski elektrotehniški slovar*. Ljubljana: Sloko CIGRÉ.
- Paulin, Andrej. 1995. *Tehniški metalurški slovar: slovensko-angleško-nemški*. Ljubljana: Mladinska knjiga.

## B. Other literature

- Béjoint, Henri. 1981. The foreign student's use of monolingual English dictionaries: A study of language needs and reference skills. *Applied Linguistics* 2: 207–22.
- . 1994. *Tradition and Innovation in Modern English Dictionaries*. (Oxford Studies in Lexicography and Lexicology). Oxford: Clarendon Press.
- Bergenholtz, Henning. 1994a. Fachsprache und Gemeinsprache: Lemmaselektion im Fachwörterbuch. In Schaeder and Bergenholtz 1994, 285–304.
- Bergenholtz, Henning. 1994b. Beispiele in Fachwörterbüchern. In Schaeder and Bergenholtz 1994, 421–32.
- Bergenholtz, Henning and Jette Pedersen. 1994. Grammar in bilingual LSP dictionaries, with a special view to technical English. In Schaeder and Bergenholtz 1994, 351–83.



- Bergenholtz, Henning and Sven Tarp, eds. 1995. *Manual of Specialized Lexicography. The Preparation of Specialized Dictionaries*. Amsterdam: John Benjamins.
- Bogaards, Paul. 1990. Où cherche-t-on dans le dictionnaire? *International Journal of Lexicography* 3: 79–102.
- . 1996. Dictionaries for learners of English. *International Journal of Lexicography* 9: 277–320.
- Bujas, Željko. 1974. A manual conversion procedure in bilingual (Croatian-English) lexicography. *Studia Romanica et Anglica Zagradiensia* 37: 187–203.
- Christiansen, Lisbeth Maidahl, Grete Duvå, and Anna-Lise Laursen. 1994. Das Translationswörterbuch für Fachsprachen: Ein intergriertes Konzept. In Schaeder and Bergenholtz 1994, 269–82.
- Grinev, Sergei Victorovic. 1989. Some problems of terminological dictionaries and their classification. *UNESCO. ALSED-LSP Newsletter* 11(2): 3–9.
- Haensch, Günther. 1991. Die zweisprachige Fachlexikographie und ihre Probleme. In Hausmann et al. 1991, vol. 3: 2937–48.
- Hausmann, Franz Josef. 1991. Die Paradigmatik im zweisprachigen Wörterbuch. In Hausmann et al. 1991, vol. 3, 2794–96.
- Hausmann, Franz Josef, Oscar Reichmann, Herbert Ernst Wiegand, and Ladislav Zgusta, eds. *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Vol. 2, 3. Berlin/New York: de Gruyter.
- Hausmann, Franz Josef, and Reinhold Werner, R. O. 1991. Spezifische Bauteile und Strukturen zweisprachiger Wörterbücher: eine Übersicht. In Hausmann et al., 1991, vol. 3, 2729–69.
- Iannucci, James E. 1985. Sense discrimination and translation complements in bilingual dictionaries. *Dictionaries* 7: 57–65.
- Kalverkämper, Hartwig. 1978. Die Problematik von Fachsprache und Gemeinsprache. *Sprachwissenschaft* 3: 406–44.
- Mihaljević, Milica. 1991. Višerječne natuknice i podnatuknice u jednojezičnom općem rječniku hrvatskoga jezika. *Rasprave Zavoda za hrvatski jezik* 17: 133–44.
- Nielsen, Sandro. 1994. *The Bilingual LSP Dictionary: Principles and Practice for Legal Language*. Tübingen: Narr.
- Opitz, Kurt. 1990. The Technical Dictionary for the Expert. In Hausmann et al. 1991, vol. 2, 1505–12.
- Schaeder, Burkhard and Henning Bergenholtz, eds. 1994. *Fachlexikographie: Fachwissen und seine Repräsentation in Wörterbüchern*. Tübingen: Narr.
- Scholfield, Philip. 1980. Evaluating selection policy and grammatical and semantic information in an EST dictionary. *Fachsprache* 2: 98–109.
- Svensén, Bo. 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.

### UVRŠČANJE VEČBESEDNIH ZVEZ V STROKOVNE SLOVENSKO-ANGLEŠKE SLOVARJE

*Članek obravnava problematiko uvrščanja večbesednih zvez v strokovne slovensko-angleške slovarje na primeru konfekcijskega slovarja. Analiza na osnovi vprašalnika, ki so ga izpolnjevali strokovnjaki za oblačilno tehnologijo, je pokazala, da večina slovenskih strokovnjakov pojme strukturira v sistem nadrejenega pojma ter temu pomensko podrejenih pojmov. To ima za posledico iskanje večbesednega termina pri nadrejenem enobesednem terminu. Slovenski strokovnjaki menijo, da je abecedna slovarska mikrostruktura ustrezna, če vsebujejo slovarski zapisi podpomenke ter tista podgesla, ki se leksikalno navezujejo na glavno geslo. V prilogi sta predstavljena modela slovarskih zapisov za bodoči slovensko-angleški konfekcijski slovar, sestavljena na osnovi rezultatov vprašalnika.*