

## Types of Language Nomination: Universals, Typology and Lexicographical Relevance

František Čermák

Institute of the Czech National Corpus, Faculty of Arts, Charles University  
nám. J. Palacha 2, Praha 1, 110 00, Czech Republic  
Frantisek.Cermak@ff.cuni.cz

### Abstract

Against the background of the Prague tradition of research into language nomination, known through V. Mathesius, some basic aspects of the field are examined. S. Ullmann's remarks [1966] have been taken up and examined statistically on ten common nouns in English, Finnish and Czech (*head, nose, eye, tongue, heart, dog, cat, tree, stone, water* and their counterparts); data for them have been taken from very large dictionaries in each case. For its framework, three of Skalička's five universal language types, isolating, agglutinative and inflectional, have been chosen, while introflexional and polysynthetic types had to have been left aside, for a number of reasons. From the point of view of the language nomination, the fourth and major type, that of collocation (combination), had to be left out, too. It was found that the extent of polysemy is smallest for inflectional languages, being largest for isolating languages. On the other hand, derivation is represented in an almost converse way, having a very small representation for an isolating language, while being very large in an inflectional language. Finally, it is the agglutinative language which has the highest use of compounding as against its limited use in English and Czech, which are basically similar. However, to round up the picture, the fourth type of nomination has to be studied, too. On the basis of figures obtained, the traditional, rather vague, statements can thus be given some basis and the most salient oppositions in each of the languages clarified: polysemy versus derivation (English and Czech) and compounding versus derivation (Finnish). Two tentative universals, referring back to Ullmann, could be reformulated.

### 1 Introduction: Dictionaries and Their Lemmata

Since most dictionaries are either bilingual or monolingual, descriptively catering for typological features of the languages covered, the whole scope of language nomination types is hardly ever apparent, let alone their interrelationship. Generally, it is taken for granted that large dictionaries should offer a decent coverage, or a principled selection, of all types of stable nominations in each language, and not just words.

In principle, language **nomination** draws either on morphemes (as in Vietnamese) and their combinations or words and their combinations (in most languages). Depending on the language in question, these might appear as single-word or multi-word lexemes, but in some cases as stable sentences (e.g. proverbs), too, a question which I propose not to go into here. Language nomination is taken here, following the Prague Circle linguistic tradition, specifically that of Vilém Mathesius, to broadly mean a language name (form) for an item of extralinguistic reality or mental content. The scope of possible nominations is limited here to stable linguistic nominations only, however (see, e.g. Mathesius 1975 and Čermák 2001). *Internal* nominations is the label used here for those types which arise inside the language, in contradistinction to *external* nominations, basically loan-words, which come into the language from elsewhere. It is obvious that different types of nominations have a different scope of meaning expressed, too, this being one of the main questions this paper is concerned with. Though evident, no attempt will be made here to go into types of explicit

and implicit expression of the meaning involved, as that would require a way how to measure it.

Impressionistically, dictionaries approach, in their coverage, the variety of nomination types in a rather uneven way. While multi-word nominations may often be underrepresented and undertreated [Čermák 2000], single-morpheme and single-word nominations seem to get a rather fair representation in dictionaries. Yet this is no longer quite true should we switch, as far as morphemes are concerned, from Vietnamese, Chinese etc. to, for instance, inflectional languages, such as Czech. Here, it is obvious that there is no standard or consensual approach applied, and that mapping, for example, of all productive prefixes as separate entries etc. is not regularly and systematically done.

In the following, some of these problems will be taken up and examined against standard large dictionaries.

## 2 Types of Nomination

There has never been much attention paid to the interrelationship of derivation, compounding, collocation and polysemy across languages, these being the standard linguistic names used for processes and their results of various types of language nomination. While the first three, being of a formal nature, have been mentioned above and have to be examined separately, their obvious semantic counterpart, polysemy, is common to all.

In the sixties, S. Ullmann [Ullmann 1966, 224, 232] has summed up, in his search for potential semantic universals, what has been repeatedly taken for granted before, also in the Prague school. In doing this, he established a link between three of the four major formal internal types of nomination formation, or, rather, in his terms of the time, of word formation. In his view, some languages use (A) *derivation* and (B) *compounding* for the formation of new words, while other, using these two sporadically, exhibit a tendency to fill gaps in the vocabulary by resorting to (C) *polysemy*, i.e. by addition of new meanings to the existing ones. Although this is no straightforward rule, being limited and modified by a number of factors (including borrowings), our practical every-day experience with different languages and language types may basically suggest this, too, without any prior research. However, should one want to be more specific, he or she is faced with problems, the most difficult one being polysemy.

It is common knowledge, that dictionaries wildly differ in their treatment of *polysemy* and there is no consensus here whatsoever. How does one, then, measure and view meaning of an item, let alone count number of its meanings in a reliable way? There seems to be no good answer, the problem being seemingly aggravated, in fact, by the presence of large corpora, showing a semantic continuum, based on scales of collocability of items, where traditional lexicographers and their customers were used to expect comfortably clear-cut segmentation of meaning. This may, eventually, turn out to be a major challenge for future lexicographers working with corpora, namely search for reliable and transferable criteria for the meaning segmentation and its documentation as it appears in various types of collocations and combinations of diminishing frequency. Insights brought in by research into

regular polysemy and desiderata coming from computational linguistics [Kilgarriff 1992] are promising, being, however, far from exhaustive of the field.

Nevertheless, let us use what is available, however problematic and pragmatic it may appear, namely the current dictionary practice. Whatever reservations about their vague and imprecise criteria applied one might have, dictionaries do represent both a *fait accompli* and a kind of meaning representation lexicographers are able to come up with. As an illustration and suggestion for a further research, it is as good as any. In fact, it is the only one available.

For the above reasons, an analysis of ten typical nouns in (very) large dictionaries of three languages will be undertaken. Linguistic reasons for their choice are to be found in their having rich polysemy, derivation and compounding and in the fact that they are rather frequent in use; semiotically they have a similar extension and denotate, being limited to concrete nouns in each case. Lexicostatistically, they are relatively constant and stable, being very old. None of them is derived and, being based on a single root in each case, they may easily serve as bases for subsequent derivation or compounding.

In order to get some insight into the questions raised above and to delimit a suitable framework of reference, representatives from three language types have been chosen, one from the predominantly *isolating* or analytical type (English), *agglutinative* type (Finnish) and *inflectional* type (Czech). The background typology applied here is the Praguian typology, usually identified with Vladimír Skalička [e.g. Skalička 1979]. To be within comparable limits, dictionaries of similar size have been chosen for exemplification and subsequent analysis, ranging in size between 100 000 and 200 000 entries (lexemes), namely *New Shorter Oxford Concise Dictionary on Historical Principles I-II (NSCD)* for English, *Nykysuomen sanakirja I-III (NS)* for Finnish and *Slovník spisovného jazyka českého I-IV (SSJČ)* for the Czech language.

The fourth language type, the *introflexional* one (found in Arabic, usually), is viewed as a subtype of the inflectional type sometimes. This, a different morpheme status in this language type and a lack of a suitable dictionary for analysis were instrumental for its omission. Finally, neither the last (fifth) type of Skalička's typology, namely the *polysynthetic* type, is represented here.

The reason for this may be sought in different views on what exactly *polysynthesis* might be, Skalička's view being different from the current approaches [Drossard 1997]. Although polysynthesis is based on, but also confused with, incorporation (found in Amerindian, Papuan or Eskimo languages etc.), it is, basically, different from the three types included into the analysis here in that it belongs to *la parole* (mostly or only?). Polysynthetic constructions, made up of very long chains ("words") of formal elements which are definable semantically, whether they are sentential, where the whole long textual "word" is identical with a sentence, or non-sentential (with the pronominal subject standing apart), are textual formations, clearly. Because of lack of large and reliable dictionaries for this language type, one may only wonder about a number of questions one might have, so far. Thus, it is not known what is the character and extent, if any, of stable collocations here, what are rules for agglutination of various lexical and grammatical morphemes, including roots, to their bases

(usually verbal ones) etc. Among other things, it seems that the vexing problem of collocations does not, basically, exist here. But all of this must become subject of a further research.

Having thus abandoned the polysynthetic type, so far, and not having sufficient data, which only start emerging with large language corpora now, the fourth major type of language nomination, namely (D) that of *collocation* has to be left to desiderata only. By collocation, all types of multi-word stable combinations are broadly meant here, primarily. Although no reliable and representative investigation is available in any language, our intuition suggests that this seems the primary and richest type of nomination, in fact (including the vast realms of multi-word terms).

There is also an obvious and long tradition of perceiving a correlation between the *physical length* of the word, whether it is viewed only mechanically or as made up of a number of morphemes used, and the *extent of its meaning*, a fact often mentioned by V. Mathesius in Prague. This fact will be taken up in conclusions.

Having all this in mind, three distinct groups of nouns have been selected for further investigation, namely those of (a) body parts, (b) animals, and (c) nature objects, in the three languages:

<b>a</b> <i>hlava, nos, oko, jazyk, srdce;</i>	<b>b</b> <i>pes, kočka;</i>	<b>c</b> <i>strom, kámen, voda</i>
<b>a</b> <i>head, nose, eye, tongue, heart;</i>	<b>b</b> <i>dog, cat;</i>	<b>c</b> <i>tree, stone, water</i>
<b>a</b> <i>pää, nenä, silmä, kieli, sydän;</i>	<b>b</b> <i>koira, kissa;</i>	<b>c</b> <i>puu, kivi, vesi</i>

Since a reasonable correlation between these words, representing one-to-one equivalents in the three languages, was sought, some other limitations had to be imposed. These, then, excluded such words as the English *leg* and *foot*, being two nominations used here for the semantic span covered by single terms in both Czech and Finnish, i.e. *noha* and *jalka*, respectively. Other discrepancies, although observed, were not taken into account, however, such as the Finnish *puu* corresponding both to the English *tree* and *wood*. A minor overlapping may also be observed in the case of the English *tree*, which is used for *wood* sometimes, too, and corresponds to a single word used here in Czech, namely *strom*, in this meaning. But the line had to be drawn somewhere and these might be viewed as minor differences only.

Finally, it must be stressed that the system approach chosen and advocated here, as mirrored by the dictionaries used, tells us nothing about the frequency of phenomena which have been inspected. This may only become possible if a corpus or rather, corpora, are used, contributing, thus, to a more reliable picture of the nature of the languages in question, of their preferences and typical features in this field.

### **3 Polysemy and Its Dictionary Representation in Three Language Types**

Typologically, the greatest extent of polysemy is to be found in the analytical languages. Those analytical languages, where the tendency to increase polysemy too much conflicts with (phonologically) too short and limited form available and a further increase would not,

then, be economical, have developed a way out of this, namely in tones. This is to be observed in, for example, Chinese, which did not have any tones originally. Thanks to their existence, there is now scope for mere 1542 acoustic syllables to accommodate quite a lot of meanings, although this scope does not seem to offer much more "free space", so to say. Needless to say that this has nothing to do with the number of characters, which is much higher and which one is traditionally used to take for the starting point. Specifically, if, say, a single Chinese syllable *li* is linked with some 30 meanings, an auxiliary means, such as tones, has to be used, since the syllable seems to be functionally overburdened. But any serious research along these lines is still missing. Obviously, such complex and complicated problems might have far-reaching consequences in more than one direction, including syntax, and ways for resolution of rich polysemy have to be sought also here.

The analysis of polysemy in the words examined offers more than one conclusion. Generally, it is possible to take for granted that the smallest extent of polysemy is to be found in inflectional languages, while its extent is greatest in isolating (analytical) languages, the ratio being almost 1:7 (5,1 : 34,9). There seems to be an observable correlation between great polysemy and animate nouns or nouns related to human beings. The anthropomorphic principle is very strong in all three languages, giving rise to typical polysemy, including symbolic meanings; the three languages seem to have developed in a similar way. Thus, as far as polysemy is concerned, words of the first group (a) are manifested in a similar way, the list (in the descending order) being headed, in each case, by *head* and its equivalents.

	English (isolating)	Finnish (agglutinative)	Czech (inflectional)	
<b>a</b>	<i>head</i>	50 + 22V	<i>paä</i> 33	<i>hlava</i> 14
	<i>hear</i>	32 + 4V	<i>sydän</i> 29	<i>srdce</i> 8
	<i>tongue</i>	27 + 12V	<i>kieli</i> 6	<i>jazyk</i> 3
	<i>eye</i>	13 + 1V	<i>silmä</i> 14	<i>oko</i> 5
	<i>nose</i>	11 + 14V	<i>nenä</i> 5	<i>nos</i> 2
<b>b</b>	<i>dog</i>	29 + 9V	<i>koira</i> 7	<i>pes</i> 5
	<i>cat</i>	13	<i>kissa</i> 7	<i>kočka</i> 7
<b>c</b>	<i>tree</i>	18 + 5V	<i>puu</i> 10	<i>strom</i> 1
	<i>stone</i>	33 + 8V	<i>kivi</i> 9	<i>kámen</i> 4
	<i>water</i>	22 + 16V	<i>vesi</i> 21	<i>voda</i> 2
Average:	34,9	13,6	5,1	

Table No. 1

Some comment is in place here. Even such a limited sample (corresponding to a sizeable part of the printed dictionary, however) reveals some inconsistencies, which, in theory, should not be there, and which may be ascribed to lexicographers, perhaps. Thus, the absence of polysemy of the Czech *strom* is conspicuous, if the word is compared with other words examined (see also the note above, however). Similarly, the English *cat* seems to be lacking the conversive verbal type of meaning, frequently found with other nouns. Other dictionaries do record this verbal type of meaning, however. Due to other reasons, there has been no attempt made here to record the most frequent type of conversion between the noun and

adjective, which is impossible in Czech or Finnish. Should this be taken into account, the extent of the English polysemy would grow considerably. However, Noun-Verb conversion in English is recorded (see V above).

#### **4 Derivation and Its Dictionary Representation in Three Language Types**

In order to be able to cope with derivation in the same sample of the three languages, a pragmatic decision was made to include only those derivatives which are formed on the right side of the stem/root of the basic word, i.e. suffixes only. One of the reasons for the exclusion of prefixes might be seen in their uneven distribution across these languages, not really offering a good basis for comparison. In fact, prefixes of nouns are not to be found very often here.

Also here, a number of conclusions may be drawn. A major contrast is to be seen between the isolating and inflectional languages (i.e. English and Czech), though in a reverse order to that found for polysemy, being almost 1:3,5 (7,9 : 27). The agglutinative Finnish lies in between the two. It is to be observed, however, that the Czech derivation (suffixation), in its high occurrence of allomorphs and indistinct morpheme boundary, is typologically an agglutinative feature; hence, though prominent, it is not quite typical for the inflectional type of language.

It may be of some interest to further observe, that for Czech the degree of derivation seems to be much higher with concrete inanimate nouns than with the animate ones. For English, this ratio seems to hold almost the other way round, however, while in Finnish there does not seem to be any prominent tendency to be observed. The conspicuously high derivation of the Finnish *pää* "head" is due to two reasons, mostly. Finnish does not form adjectives from this type of nouns very often, hence *pää* is used in the "adjectival", or rather, attributive function in compounds with the meaning "upper, main". A second reason is to be seen in that this attributive function is often of a prominently pragmatic, evaluative nature, a function which is very much in need and use. Both reasons explain, hopefully, the rather unusual nature of the Finnish *pää*.

	English (isolating)		Finnish (agglutinative)		Czech (inflectional)	
<b>a</b>	<i>head</i>	7	<i>pää</i>	52	<i>hlava</i>	37
	<i>heart</i>	10	<i>sydän</i>	8	<i>srdce</i>	24
	<i>tongue</i>	7	<i>kieli</i>	11	<i>jazyk</i>	5
	<i>eye</i>	8	<i>silmä</i>	19	<i>oko</i>	13
	<i>nose</i>	10	<i>nenä</i>	3	<i>nos</i>	20
<b>b</b>	<i>dog</i>	17	<i>koira</i>	5	<i>pes</i>	6
	<i>cat</i>	2	<i>kissa</i>	1	<i>kočka</i>	20
<b>c</b>	<i>tree</i>	6	<i>puu</i>	7	<i>strom</i>	23
	<i>stone</i>	6	<i>kivi</i>	14	<i>kámen</i>	43
	<i>water</i>	6	<i>ves</i>	18	<i>voda</i>	54
Average:		<b>7,9</b>		<b>13,8</b>		<b>27</b>

Table No. 2

Quite a few of the results obtained make one curious, prompting questions, but in lack of an extensive research it would be too risky to even formulate a conjecture. Thus, one may only ask why there is almost no derivative for the Finnish *kissa* "cat", which stands out if compared with *koira* "dog" (*koiramainen* etc.). Or, alternatively, why the corresponding Czech *kočka* "cat" seems to be relatively rich in this feature? More sense, however, can be made for these and other questions, as soon as one takes into account the complementary character of derivation and compounding.

One may, however, question the marginal nature of some of the derivatives found, some of which suggesting potentiality of formation rather than real existence. Marginal character of some of these, such as the English *tonguey* or Finnish *sydämettämyys*, may readily be perceived by native speakers or corroborated by a large corpus. More of these less frequent or infrequent, if not experimental, cases may be seen in a selection in the lists for English, Finnish and Czech below.

<b>head:</b>	<i>header, headless, heady</i>
<b>dog:</b>	<i>doggish, doggy, doggo</i>
<b>eye:</b>	<i>eyeful, eyable</i>
<b>heart:</b>	<i>heartful, hearten, heartless</i>
<b>nose:</b>	<i>nosy, nosing</i>
<b>stone:</b>	<i>stonelike, stoner</i>
<b>tongue:</b>	<i>tonguelet, tonguey</i>
<b>tree:</b>	<i>treeful, treen, treey</i>
<b>water:</b>	<i>watery, waterage, waterer</i>
<b>puu:</b>	<i>puutua</i> (no record of * <i>puullinen</i> , however)
<b>nenä:</b>	<i>nenäkas</i>
<b>kivi:</b>	<i>kivellinen, kivettyä, kives, kivikko</i>
<b>kieli:</b>	<i>kieleke, kiellinen, kielevä, kielistö</i>
<b>silmä:</b>	<i>silmikko, silmitellä, silmitön</i>
<b>sydän:</b>	<i>sydämistyä, sydämetön, sydämettämyys</i>

<b>hlava:</b>	<i>hlavečka, hlávek, hlavnice</i>
<b>srdce:</b>	<i>srdit se, srdnatec</i>
<b>oko:</b>	<i>očisko</i>
<b>nos:</b>	<i>nosan</i>
<b>pes:</b>	<i>pesan, peský, psotý, psovina, psout</i>
<b>strom:</b>	<i>stromstvo</i>
<b>kámen:</b>	<i>kaméneček, kameněný, kamenisko</i>

There is a number of problems related to any straightforward division of derivation and compounding, regardless of the many existing theoretical approaches and theories, often language-dependent. As it is virtually impossible and, in fact, useless to find a way in a mass of often conflicting approaches, a simple way had to be found. However, a simple binary approach applied must of necessity be somewhat arbitrary. Thus, principles used for the analysis and classification of data obtained, if more than one suffix was found to be attached to the base words, include the following scale:

if the root/base word is followed by

- (a) two suffixes (derivation), these are considered to be derivatives
- (b) a suffix and a root (derivation and compounding), these are considered to be compounds
- (c) one or more roots, these are considered to be compounds
- (d) a root and a suffix, these are considered to be compounds

Obviously, no account has been made of compounds here in the sense outlined.

## 5 Compounds and Their Representation in Three Language Types

However, compounds have been treated separately. As Finnish abounds in hundreds of compounds here and because approaches to compounds and their identification are widely different in various languages, the analysis has been limited, for pragmatic reasons, to 3 typical nouns only.

There is a further factor which had to be taken into account, namely the fact that, in English, compounds are often passed for a type of collocations and there seems to be no consensus. As collocations are not exhaustively covered in dictionaries, this appeared to be an effective limitation also in this research.

However limited the basis chosen for analysis might be, the general nature of the results obtained may hardly be questioned. It seems to confirm both theoretical premisses and practical experience, or intuition.

	ENGL. (isolating)	FIN. (agglutinative)	CZE. (inflectional)
<b>a</b>	<i>head</i> 17	<i>pää</i> 914	<i>hlava</i> 6
<b>b</b>	<i>dog</i> 12	<i>koira</i> 97	<i>pes</i> 6
<b>c</b>	<i>wate</i> 49	<i>vesi</i> 660	<i>voda</i> 61
Average:	26	557	24,3

Table No. 3



Evidently, it is the agglutinative Finnish that is prominent and conspicuous here, in which compounding appears to be one of the most typical features. This stands out clearly in comparison with English and Czech, which, surprisingly, do not differ to any high degree. On the average, the Finnish compounding seems to be 22 times higher (26:557, or, 24,3:557). To draw any other conclusions is not possible here.

One must, yet again, remind oneself that, should English and Czech appear to be almost identical in this respect, these two languages do widely differ in the fourth type of nomination (see above), which is not dealt with here, namely in collocations. Typologically, it is the analytical (isolating) languages which suffer from lack of criteria which might distinguish, in any reliable way, compounds and collocations (this happens to a traditional problem of Chinese linguistics, for example). It seems that the reasons for this are to be sought in our insufficient knowledge of the stableness or fixity problem, on which any judgement of lexicalization is subsequently based. This might also be due to that type of tradition obsessed with syntax of words, not really being able to observe something lying between the two, namely collocations and combinations as a third entity. Reluctance of grammars to consistently admit existence of multi-word units in all of the word-classes might be a proof of this. A ridiculous and, perhaps, the most prominent, example is to be seen in multi-word prepositions in, say, either English or Czech. Despite their being far more numerous than single-word prepositions, they are, invariably, missing or under-represented in grammars.

## **6 Conclusions: General and Lexicographic**

The above results should be viewed as mere illustrations, not a research proper (which should follow), and it is not really difficult to come up with somewhat different results based on different data, should one wish to. However, it seems that the typological framework (of Skalička's) suggested here is useful in general and that some preliminary conclusions may be made.

The nominative potential of the three languages is put to use in a number of different ways. Thus, in terms of languages covered, *English* seems to be strongest in polysemy and weakest in derivation, while difference between derivation and compounding is small. *Finnish*, on the other hand, seems to be strongest in compounding (for some words, about a thousand of compounds may be found) and weakest in derivation, where the difference between derivation and polysemy may not be really great. Let us observe that the exact status of compounds, being far from consensual, must be viewed against the langue-parole distinction, too. So far, it is far from clear where to draw the line between stable and textual compounds.

*Czech* is clearly strongest in its derivation, while polysemy is its weakest feature. This corroborates our earlier results, based on a much larger data [Čermák 1990], where the average number of derivations per root was found to be at least 30,7 in Czech.

All four types of nominations are complementary, obviously, although in different proportions in different languages. To be able to achieve a complete functional picture of the whole field of nominations, one must find out their mutual proportions in each case as well

as their interplay. Despite our not being able to include any tangible figures for collocations and the fourth type (polysynthetic), even the data for the first three types do, hopefully, offer some suggestions. Thus, on a more specific level, it seems that there is a basic complementarity to be found between derivation and compounding on the one hand and polysemy on the other. Needless to say, all of the results should be viewed as mere suggestions and ideas for a further, real research, which might and should bring improved and modified insight into the field.

On a more abstract level, one can hopefully conclude that the present results have corroborated the original and somewhat loose idea of Ullmann. Thus, one may tentatively suggest two *implicational universals* here, which are both mutually complementary and inversely proportional, while complementary is meant here to include overlapping, too. It seems to hold that

(1) *The shorter an average lexeme is, the more polysemy the language has (and vice versa).*

(2) *The longer an average lexeme is, the more derivation (in inflectional languages) or compounding (in agglutinative languages) the language has (and vice versa).*

Apart from limitations mentioned above, one might further specify the two general suggestions in more detail, should such distinctions as frequency and register be brought into consideration.

Evidently, on a lexicographical level, these conclusions might help dictionary planning and distribution of work as well as deciding different levels of elaboration given to the phenomena studied here. Specifically, this should become the more prominent factor in the dictionary design, the more typologically distant the languages covered in a bilingual dictionary are.

## References

- [Apresjan 1974] Apresjan J. D., 1974. Regular Polysemy, in: *Linguistics* 142, 5-32.
- [Čermák 1978] Čermák, F., 1978. Typology of the Germanic Languages with Special Reference to Dutch, in: *Folia Linguistica* XII: 1/2, 1978, 65-106.
- [Čermák 1990] Čermák F., 1990. *Syntagmatika a paradigmatika českého slova 2: Morfologie a tvoření slov*. Univerzita Karlova Praha (Syntagmatics and Paradigmatics of the Czech Word: Morphology and Word Formation).
- [Čermák 2000a] Čermák, F., 2000a. Combination, Collocation and Multi-Word Units, in: *Proceedings of The Ninth Euralex International Congress EURALEX 2000*, eds. U. Heid, S. Evert, E. Lehmann, C. Rohrer, Institut für Maschinelle Sprachverarbeitung Universität Stuttgart, Stuttgart, pp. 489-495.
- [Čermák 2000b] Čermák, F., 2000b. Jazyková nominace: povaha a souvislost některých univerzálií v sémantice a morfologii (polysémie, derivace a kompozice), in: *Slovo a Slovesnost* 61, 249-256 (Language Nomination: Character and Relationship of Some Universals in Semantics and Morphology).
- [Čermák 2001] Čermák, F., 2001a. Substance of Idioms: Perennial Problems, Lack of Data or Theory?, in: *International Journal of Lexicography* 14, 2001, 1-20.

- [Čermák 2001b] Čermák F., 2001b. *Jazyk a jazykověda. Přehled a slovníky*. 3. vyd. Karolinum Praha (Language and Linguistics. A Survey and Dictionaries, 3rd ed.).
- [Drossard 1997] Drossard W., 1997. Polysynthesis and polysynthetic languages in comparative perspective, in: *Typology: prototypes, item orderings and universals*. Proceedings of LP'96, ed. B. Palek, Karolinum Prague, 251-264
- [Filipec-Čermák 1985] Filipec J., F. Čermák, 1985. *Česká lexikologie*. Academia Praha (Czech Lexicology).
- [Kilgariff 1992] Kilgariff A., 1992. *Polysemy*. University of of Sussex (unpublished doctoral dissertation, <http://www.itri.bton.ac.uk/>).
- [Mathesius 1975] Mathesius V., 1975. *A Functional Analysis of Present Day English on a General Linguistic Basis*. Academia Prague.
- [Nykysuomen sanakirja 1967] *Nykysuomen sanakirja* I-III, 1967. Werner Söderström, Porvoo Helsinki 1967 (cca 201 000 entries, 2381p., =NS, Dictionary of New Finnish).
- [Quirk-Greenbaum-Leech-Svartvik 1985] Quirk R., S. Greenbaum, G. Leech, J. Svartvik, 1985. *A Comprehensive Grammar of the English Language*. Longman London.
- [Skalička 1951] Skalička V., 1951. *Typ češtiny*. Slovanské nakladatelství Praha (Type of Czech).
- [Skalička 1979] Skalička V., 1979. *Typologische Studien*. Viewegh Braunschweig.
- [Slovník spisovného jazyka českého 1957-1961] *Slovník spisovného jazyka českého* I-IV, 1957-1961. Naklad. Čs. akademie věd Praha (cca 182 000 entries, 4593p., =SSJČ, Dictionary of the Literary Czech Language).
- [New Shorter Oxford Concise Dictionary 1993] *The New Shorter Oxford Concise Dictionary on Historical Principles* I-II, 1993. Clarendon Press Oxford (cca 125 000 entries, 3767p., =NSOCD).
- [Ullmann 1962] Ullmann S., Semantic Universals. in: J. Greenberg, ed., 1966 (2. vyd.), *Universals of Language*. M.I.T. Press Cambridge, 217-262.
- [Vesikansaa 1977] Vesikansa J., 1977. *Johdokset*. W. Söderström Porvoo-Jelsinki-Juva.