

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/322153404>

# Frazeologie a idiomatika česká a obecná. Czech and General Phraseology

**Book** · December 2017

CITATION

1

READS

62

## 1 author:



**František Čermák**

Charles University in Prague

**106** PUBLICATIONS **339** CITATIONS

[SEE PROFILE](#)

**Some of the authors of this publication are also working on these related projects:**



Function of Sentence Idioms [View project](#)



Language Periphery. Monocollocable words in English, Italian, German and Czech [View project](#)

Publikace vychází v rámci výzkumného záměru MSM002162084  
*Základy moderního světa v zrcadle literatury a filozofie,*  
řešeného na Filozofické fakultě UK

# Obsah

<b>Předmluva</b> .....	7
<b>Preface</b> .....	9
<b>Úvod:</b> Idiomatika a frazeologie ve vztahu k jiným disciplínám • <i>Introduction: Relationship of Idiomatics and Phraseology and Other Disciplines</i> .....	13

## A TEORIE a PODSTATA • THEORY and SUBSTANCE

Základy a zásady frazeologie a idiomatiky .....	23
Foundations and Principles of Phraseology and Idiomatics .....	73
Linguistic Units and Text Entities: Theory and Practice • <i>Jazykové jednotky a textové entity: teorie a praxe</i> ...	129
Idiomatics • <i>Idiomatika a frazeologie</i> .....	141
Substance of Idioms: Perennial Problems, Lack of Data or Theory? • <i>Podstata frazémů a idiomů: věčné problémy, nedostatek dat nebo teorie?</i> .....	149
Aktuální problémy frazeologie a idiomatiky a jejich odraz v slovníkové praxi (Nad bulharskou sbírkou přirovnání) • <i>Current Problems of Phraseology and Idiomatics and Their Reflection in Dictionaries</i> .....	167
On the Identification of Idioms • <i>Identifikace frazémů a idiomů</i> .....	177
Statistical Methods for Searching Idioms in Text Corpora .....	192
Statistické metody hledání frazémů a idiomů v korpusech .....	200
Funkce frazémů a idiomů • <i>Function of Idioms</i> .....	212
Termín a frazém. Příklad překrývání a periférie dvou nominativních oblastí • <i>Terms and Idioms. A Case of Overlapping and Periphery of Two Areas of Nomination</i> .....	220
Somatic Idioms Revisited • <i>Znovu o somatických frazémech</i> .....	228
On the Nature of Universality in Idioms .....	237
Povaha univerzálního ve frazeologii a idiomatice .....	245
Universals, Typology and Idioms • <i>Univerzálie, typologie a frazémy</i> .....	252

## B TYPY a OBLASTI • TYPES and AREAS

České lexikální frazémy a idiomy • <i>Czech Lexical Idioms</i> .....	263
Idioms and Morphology • <i>Frazémy, idiomy a morfologie</i> .....	277
Grammatical Idioms • <i>Gramatické frazémy a idiomy</i> .....	289
České frazémy a idiomy gramatické • <i>Czech Grammatical Idioms</i> .....	308
České frazémy a idiomy nominální • <i>Czech Nominal Idioms</i> .....	326
České frazémy a idiomy modifikační • <i>Czech Modification Idioms</i> .....	334
České frazémy a idiomy verbální • <i>Czech Verbal Idioms</i> .....	343
Česká přirovnání • <i>Czech Similes</i> .....	384
Binomiály • <i>Binomials</i> .....	414
Propositional Idioms .....	430

Propoziční frazémy a idiomy v češtině .....	438
České frazémy a idiomy propoziční • <i>Czech Propositional Idioms</i> .....	446
Frazeologie a idiomatika češtiny: stručný přehled historie • <i>Phraseology and Idiomatics of Czech: A Short Survey of Its History</i> .....	482

## C STUDIE • STUDIES

Abstract Nouns Collocations: Their Nature in a Parallel English-Czech Corpus • <i>Kolokace abstraktních substantiv a jejich kolokace v anglicko-českém paralelním korpusu</i> .....	491
Frazeologie novinového úvodníku: úzus, funkce a aktualizace • <i>Idioms in Newspaper Editorials</i> .....	503
Frazeologie a biblická etika: Kolokabilita lexémů <i>dobrý</i> a <i>zlý</i> a povaha přísloví • <i>Phraseology and the Biblical Ethics: Collocability of good and bad and the Nature of Proverbs</i> .....	514
Verbální frazémy s komponentem <i>být</i> v češtině • <i>Verbal Idioms Containing the Verb být in Czech</i> .....	527
What One Can Do with Proverbs in Text • <i>Co se dá v textu dělat s příslovím</i> .....	536
Text Introducers of Proverbs and Other Idioms • <i>Textové introduktory přísloví a dalších frazémů a idiomů</i> .....	549
Usage of Proverbs in Today's Czech language: What the Czech National Corpus Seems to Suggest • <i>Užívání přísloví v dnešní češtině: Co naznačuje Český národní korpus</i> .....	569
Jazyková variabilita: případ přísloví • <i>Language Variability: The Case of Proverbs</i> .....	584
Paremiological Minimum of Czech: The Corpus Evidence • <i>Paremiologické minimum češtiny: svědectví korpusu</i> .....	597

## D FRAZEOGRAFIE • PHRASEOGRAPHY

Frazeografie • <i>Phraseological Lexicography (Idiomography, Phraseography)</i> .....	617
Idioms and Lexicography: In Search of Criteria • <i>Frazémy a lexikografie: hledání kritérií</i> .....	638
Český frazeologický slovník .....	646
A Dictionary of Czech Phraseology and Idiomatics .....	655
Onomaziologické systémy u propozic a proverbíí • <i>Onomasiological Systems for Propositions and Proverbs</i> .....	666
Zásady zpracování a užívání <i>Slovníku české frazeologie a idiomatiky (výrazy slovesné)</i> • <i>Principles of Compilation of Slovník české frazeologie a idiomatiky (výrazy slovesné)</i> .....	674

<b>Bibliografie • Bibliography</b> .....	707
--	-----

# Předmluva

Tato kniha představuje autorovy vybrané příspěvky z oblasti frazeologie a idiomatiky (IF) české a obecné, zveřejňované v různých zdrojích v průběhu posledních 35 let většinou česky nebo anglicky; některé z nich jsou ale nové a zde vycházejí poprvé. Kniha podává napřed přehled teorie IF, a to jednak v prezentaci jejích základů (**A Teorie a podstata**) a jednak v podobě monografických analytických studií všech jednotlivých oblastí české frazeologie, které takto představují ucelený a v podstatě vyčerpávající popis celé oblasti (**B Typy a oblasti**). Tyto části knihy doplňují vybrané analýzy systému frazeologie i studie textové (**C Studie**). Poslední část je věnována frazeografii (**D Frazeografie**), zvláště její teoretické a koncepční podstatě i vlastní frazeografické praxi, uplatňované léta v SČFI (*Slovník české frazeologie a idiomatiky*). Ty texty, které sem byly přejaty z dřívějších vydání, doznaly menších úprav; všechny texty nové jsou psány podle stejné metodologie ve snaze ukázat na jednotu přístupu a celistvost popisu. Příspěvky zde zařazené vycházejí většinou v původních jazycích, zvl. česky a anglicky, a pro uživatele jen jednoho z nich obsahuje každý z nich na konci obsáhlé anglické **Summary** nebo české **Shrnutí**. Některé příspěvky vyšly s menšími modifikacemi v obou jazycích; informace o jejich vztahu jsou tu uváděné pak s příslušným odkazem taky.

Vhled do problémů tohoto tehdy nového oboru vznikál postupně a pod tlakem potřeb praxe daných rozhodnutím popsat českou frazeologii a sestavit její relativně vyčerpávající slovník. Počátečních diskuzí nad problémy a při pozdější frazeografické práci promítající se do dokončeného čtyřsvazkového SČFI (*Slovník české frazeologie a idiomatiky* 1983–2007) se účastnila řada kolegů a spolupracovníků. Teorie české disciplíny vznikala paralelně s praxí a je výsledkem úvah a zkušeností autora; ten měl ovšem cennou zpětnou vazbu ve svých kolezích, v potřebách řešit problémy vznikajícího slovníku a šíře i v možnosti se kriticky vymezit proti některým zahraničním přístupům. Dnes je frazeologie už obecně uznávanou disciplínou a po rozmanitých (a ne vždy přesvědčivých) zahraničních počátcích se etablovala jako samostatná výzkumná a vědecká disciplína, mj. i díky *Evropské asociaci pro frazeologii* či *Europhras*, založené až v 90. letech (1999), která pravidelně vydává i své kongresové sborníky.

Obecněji se ovšem tato kniha snaží ukázat více, mj. i to, že v jazyce není všechno syntaktické (gramatické) zlato, co se třpytí, že systematickému rozboru textů stávající gramatiky a syntaxe nestačí, že v jazycích jsou primární jiná pravidla, sémantická, která vždy těm

formálním předcházejí, i to, že naše promluvy se ve znepokojivé míře skládají z hotových kombinací, stereotypů, které se tradičními přístupy smysluplně analyzovat nedají. Žádná z gramatik toto dosud nepřiznává a tváří se, že nabízejí recept na popis vyčerpávající, který ovšem nedovedou. Matoucí N. Chomsky dokonce jiná pravidla než ta, která jsou ad hoc vymyšlená a neustále proměňovaná, omezená navíc jen na syntax, nezná a reálnému univerzu jazykových promluv se vymyšlením hypotetických příkladů stále víc vzdaluje. Moderní korpusová lingvistika se naproti tomu odmítá zabývat pouze bezbřehým a bezprincipiálním generováním, protože texty a reálné jazykové promluvy se tak analyzovat nedají; zabývá se naopak tím, jak v textech kombinací forem vzniká význam a jak se tyto kombinace utvářejí, podle pravidel i bez nich. Soustřeďuje se na uзуální souvřyskyty tvarů slov, různé dlouhé, a pod názvem kolokace se je snaží popsat neselektivně všechny; otevřeně přitom přiznává, co vše se tu dosud nedaří. Už z dílčích poznatků studia velkých korpusů je jasné, jak málo dosud například známe, co je v jazyce pravidelné a jak je leckde např. obtížné vést jasnou hranici mezi langue a parole; přesto všude tam tradiční gramatiky i generativisté předstírají, že to dovedou a znají.

Do obtížné oblasti kolokací, stále teprve poznávané, v níž je jádro jazykových stereotypů, nepochybně patří i frazémy a idiomy. Jejich specifikum je však v tom, že právě zde jako v jediné oblasti jazyka běžná pravidla sémantická a pak gramatická (v tomto sledu) nikdy neplatí úplně, popř. neplatí vůbec, že u nich vždy jde o nějakou formu anomálie, která se pod dnešní deskriptivní či generativistickou formulku či algoritmus nevejde; odtud i jejich náročnost a potřeba jiného přístupu, ale i rezignace na to, co tradiční lingvistické přístupy dosud hlásaly. Odtud také zdůrazňovaná opozice frazeologie jako oblasti takto anomálního, nepravidelného jazyka a naproti tomu jazyka pravidelného, tedy toho, kde platí tradiční pravidla sémantická a gramatická. Jazykové kombinace, resp. kombinatoriku, dnes nově si vynucující řešení aspektů syntagmatických, lze tudíž seriózně a přiznaně vnímat jen jako kontinuální univerzum, kde jejich jednu část tvoří se svými pravidly jazyk pravidelný, na který komplementárně navazují tam, kde pravidelný jazyk končí, kombinace anomální, frazeologické; směšovat obojí nelze.

Dnes je stále jasnější, že velká část našich textů a promluv není generovaná ve smyslu plně originálních produktů svých autorů a mluvčích, lidé tak kreativní zase nejsou, a že texty jsou do značné míry prefabrikované, využívající vhodných a už hotových kombinací, frazémů aj. Často se tu proto mluví volně o stereotypech. Ať už je poměr stereotypů a frazeologie jakýkoliv, ta jejich část sledovaná zde jako frazémy a idiomy je významná a její poznání snad přispěje i k lepšímu chápání toho, co je v jazyce tvořené a závislé na pravidlech a toho, co není a co je třeba v textech vnímat jako hotové celky. V konečném smyslu půjde i o zpřesnění našeho vnímání rozhraní parole a langue. Významným ohledem při budoucím bádání v tomto základním směru jsou nové možnosti systematického poznání syntagmatiky jazyka, jak je dnes v podstatě poprvé umožňují moderní jazykové korpusy.

# Preface

This book presents the author's selected papers from the field of Czech and general idiomatics and phraseology (IP) which have been published mostly in Czech or English in a variety of journals and publications over the course of the last 35 years. Some of these, however, are new and appear here for the first time. At the beginning, the book offers a survey of theories of IP by presenting, on the one hand, the foundations of IP (***Theory and Substance***), and on the other, by presenting comprehensive analytical studies covering the whole range of Czech phraseology, it provides a well-rounded and basically exhaustive description of the whole field (***Types and Areas***). This part of the book is supplemented by selected analyses of both the IP system and some textual phenomena (***Studies***). The last part is devoted to the description of phraseology in dictionaries (***Phraseography***), specifically to its theoretical and conceptual basis and actual phraseographic practice as they have been applied in the compilation of *Slovník české frazeologie a idiomatiky* (*Dictionary of Czech Phraseology and Idiomatics*, or DCPI) for many years. The texts which were taken over and included here from previous publications have been slightly modified; all the other texts, i.e. the new ones, have been written using the same methodology, as applied elsewhere, in an attempt to achieve a unified and comprehensive approach. The present papers appear in their original language versions, especially Czech and English; readers unfamiliar with one or the other language are offered, at the end of each text, either a comprehensive English ***Summary*** or a Czech ***Shrnutí***. Some of the articles have already come out, with minor modifications, in both languages. Information about this is also provided, together with the respective references.

Insight into the problems of this new field has been gained slowly, mostly under the pressure of the practical needs arising from the decision to describe Czech phraseology and compile a relatively exhaustive dictionary. A number of colleagues and collaborators took part in the initial discussions over the problems both before and during the compilation of the recently finished four-volume DCPI dictionary (1983–2007). This Czech theory of IP, as is now adopted by others, has developed in tandem with its practical application and is the result of the author's reflections and experience. Naturally, he benefited from valuable feedback from his colleagues, from the eternal struggle to resolve problems surrounding the dictionary-making process and, on a more general level, from the opportunity to define his approach critically against some of the approaches in other countries. Today, phraseology and idiomatics is a generally recognized discipline which, after its tortuous (and not always

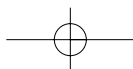
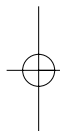
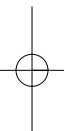
convincing) beginnings in various countries, has become established as an independent scientific discipline and area of research. In part this is also due to the *European Association for Phraseology* or *Europhras*, founded as late as the nineties (1999, to be exact), which nowadays organizes its own congresses and publishes their proceedings.

In a broader sense, this book strives for more than just the presentation of IP: it aims to show that not all that glitters syntactically (grammatically) is gold, that the existing grammars and theories of syntax are not quite up to the task of a systematic analysis of text, that there are other rules in language, rules which are primary and which always precede those of grammar, namely semantic rules. It also aims to show that, to a disconcerting degree, our utterances consist of ready-made combinations, stereotyped expressions which resist meaningful analysis relying on traditional approaches. None of the grammars in use at present has admitted this fact so far. They pretend instead they have a formula for an exhaustive description, which, however, they do not. Chomsky's contribution has been more confusing than helpful, as he has no use for rules other than the ad hoc rules of grammar, moreover rules limited only to syntax, which, to make things worse, he constantly changes. Thus he has become more and more removed from the actual universe of language utterances, especially by inventing hypothetical examples. Contrary to this, modern corpus linguistics refuses to limit itself to endless and unbridled generation, since real texts, actual language utterances cannot be analyzed in this way. By contrast, it examines the ways in which meaning arises in text by combinations of linguistic forms and how these combinations are formed, whether by rules or not. It concentrates on the habitual and usual co-occurrences of word forms of various lengths, and, using the coverall term collocation, attempts to offer a non-selective description of all their types. At the same time it states clearly the instances where it has not yet been successful. Still, even these tentative and partial generalizations based on the study of large corpora clearly indicate how precious little we know to date, for example, of what is regular in language and how difficult it is to draw a clear line between 'la langue' and 'la parole'. All of this at a time when the authors of traditional grammars and generativists alike are still busy pretending that they are privy to these distinctions and know how to make them.

This tricky area of collocations, with which we are only gradually coming to grips and which incorporates the very core of language stereotypes, clearly includes phrasemes and idioms too. Their hallmark consists in that they form the only area of language wherein the standard language rules, semantic and grammatical (in this order), never apply in full or at all, that there is always some kind of anomaly to be found which does not fit in with the current descriptive or generativist formula or algorithm. That is where their complexity stems from and the need for a different approach. This need can only be met by abandoning the tenets of traditional linguistic approaches. Likewise, this is the reason for stressing the importance of phraseology as an area of language anomalies and irregularities which stands in opposition to regular language, i.e., that part of language which follows the traditional semantic and grammatical rules. Linguistic form combinations, or combinatorics, urgently demanding novel solutions of syntagmatic aspects of language today, can be seriously and consciously viewed as belonging to a continuous universe one part of which is rule-based, formed by regular language, whilst the other, complementary, part, which begins where regular language ends, is comprised of anomalous and phraseological combinations. These two parts cannot be mixed up.



It is becoming increasingly evident today that a substantial part of natural language texts and utterances are not generated in the sense of being the fully original product of their authors and speakers, after all people are not that creative. It appears instead that texts are prefabricated to a considerable degree, drawing on suitable and ready-made combinations, idioms, etc. It is therefore often described in terms of stereotypes, to use a rather loose cognitive allusion. Regardless of what the proportion of stereotypes and phraseology may be, that part of prototypes explored here under the label of idioms and phrasemes is of enormous importance, and through gaining more knowledge of this area we may better contribute to our understanding of what is and what is not rule-based and rule-dependent in language, thus allowing us to view the second group as ready-made complexes. Ultimately, the objective is to arrive at a sharper understanding and discrimination of the interface between 'la parole' and 'la langue'. In so doing, an important aspect of the future basic research along these lines will be to explore the new possibilities of a systematic study of language syntagmatics, which is something made possible, for the first time ever, by modern language corpora.



# Úvod: Idiomatika a frazeologie ve vztahu k jiným disciplínám

## 1. Idiomatika a frazeologie: jazykové disciplíny

Jev idiomatiky a frazeologie (IF) je omezený na kombinace takových jednotek jazyka, které jsou v zásadě nadány významem. Samy frazémy a idiomy, které ho vytvářejí, mají svůj význam a/n funkci, jsou však vždy nějak anomální a vymykají se tedy běžným jazykovým pravidlům sémantickým a/n gramatickým (o vztahu frazémů a pravidelného jazyka viz dál). Protože jsou frazémy a idiomy vždy součástí jazykového systému, patří všechny coby svérázná pojmenování do široce pojímaného lexikonu; všechny ty další roviny jazyka, jejichž jednotky mají význam (ne tedy fonologie a fonetika), do oblasti frazeologie však dílčím způsobem porůznu zasahují taky. Ten způsob je však specifický a vždy limitovaný hranicemi anomálnosti.

Mylně a v příliš zúženém pohledu, který oblast jevu IF nevyčerpává, některá pojetí spatřují ve frazeologii pouze ty kombinace, kde lze mluvit o jakési obraznosti či metafoře; její pojetí je ovšem velmi rozplzlé a pro různé lidi znamená bohužel dosti různé věci. Takto by ovšem všechny kombinace „nemetaforické“, pokud bychom byli schopni se shodnout na tom, které to jsou, protože literární přístupy tu jsou notoricky nepoužitelné, při takovém pohledu zůstaly mimo. Srov. typy

(1a) *otevřít dokořán, dát najevo, být hin,*

(1b) *co do, na úkor někoho, v souvislosti s něčím, kór když, co kdyby,*

(1c) *Že by?, To zrovna!, Všechno má svůj čas., Blázni a děti mluví pravdu., Sedni si, ať ti kapsa nevisí! aj.*

Zastánci takového nepřijatelného přístupu, pokud si ovšem toto omezení vůbec připouštějí, už přitom neříkají, jak takové kombinace řešit, kam je řadit a jak je popisovat a tváří se obvykle, že ani neexistují. Shazováním nepohodlných případů se stolu se ovšem problémy nevyřeší.

Do centrálního **lexikonu**, který se tu chápe především jako depozitář jazykových frazémových pojmenování více druhů, daný kombinacemi slovních tvarů s různou funkcí, se tedy ukládají jako hotová pojmenování nejružnější ustálené kombinace většinou tvarů slov, ale i prvků a jednotek jiných. Na jejich tvorbě v minulosti se účastnila jak **syntax** tak i morfolgie a jejich anomálie vznikala postupně, díky specifické funkci a ustalujícímu se úzu frazémů i změně v jazykovém systému samotném. Že jde v případě dominantní anomální morfolgie (tvoření slov) o jedinou oblast úzce propojenou s oblastí kombinatoriky slov, což starší pohledy často opět nevnímaly, ukazují případy jako

(2a) *svěhlavý* vs. *mít svou hlavu*, *pecivál* vs. starší *válet se za pecí* nebo novější *měkkosrdcatý* vs. *mít měkké srdce*,

mezi kterými je transformační propojení, nebo případy jako

(2b) *beztoho, načase*, *nadranc* versus alternující, popř. starší podoba *bez toho*, *(být) na čase*, *(být) na dranc*,

které nedomyšlené ortografické reformy prezentují jako dvě různé věci, jakkoliv jde jen o varianty téhož. V obou případech vždy první tři příklady, starší teorií neuznávané, ukazují na frazémy „uvnitř“ hranic slova (což bývalo dříve neslýchané pomýšlení a představa) proti druhým třem příkladům „řádné“ frazeologie; „řádné“ proto, že tu jde už o „vyšší“ a intuitivně a ovšem i naivně přijatelnější kombinace slov, jakkoliv mezi oběma trojicemi zásadní rozdíl není.

Je zjevné, že arbitrární a nesmyslná pravopisná reforma zasahující jazyk zpravidla necitlivě nemůže a ani nesmí určovat, kde jsou hranice frazeologie. V případě (2b) se tedy zároveň ukazuje, že frazémy existují už na úrovni kombinací morfémů, v **morfologii**, i když může jít o řadu dalších a odlišných případů, srov.

(3) *najít* (kde *na* ≡ „nahore“ + *jít* ≡ „pohybovat se“ ≠ „objevit“, a proti tomu stojící angl. *find*, fr. *trouver*, švéd. *finna*, fin. *löytää* aj., což jsou všechna bezprefixální verba, jejich sufix, pokud tu je /není v angl./, má povahu jen gramatickou, a taková anomální představa jako v českém *najít* tu tedy vzniknout nemůže), srov. dál ještě i *zákeřný*, *vykutálený*, *černokněžník*, *veršotepec*, *mateřídouška*, *levhart* aj.

Případ *mateřídouška* výše ilustruje zároveň

(4) frazeologické (**kvazi**)**termíny**, srov. mj. ještě víceslovné *zlatý déšť* (laburnum), *čapí nos* (pelargonie, obojí botanika), *vlčí zuby*, *mrtvý chod* (technika), *černá díra* (astronomie) aj. Obdobně lze o frazeologické povaze uvažovat i u **onomastiky** v různém smyslu, srov.

(5) *labuť avonská* (= William Shakespeare), *Skočdopole*, *Hradec Králové*, *Achillova pata*, *kecá jak Palackej* aj.

Ze **syntaxe**, kterou má ve frazeologii smysl brát jako kritérium vždy až sekundárně, tj. při respektování pravidel sémantických a frazémů jako funkčních nedělitelných jednotek, je pak zajímavé si všítat jejich restrikcí, ať už jde o valenci, kterou pravidelná syntax nezná (takto synchronně v pravidelném jazyce neexistuje *ležet něčím*, ale v IF ano, srov. např. *ležet táborem*), či o kombinatoriku slovních druhů (srov. obtížnost, ne-li nemožnost pravidelné syntaxe tvořit kombinace substantivum + adverbium, a naproti tomu frazém *místo nahore*).

Dodejme, že do frazeologie zasahuje zde někdy i **fonologie** zvláště tím, že frazémy větné mají svou specifickou intonaci (suprasegmentální fonémy), srov. *To zrovna!*, které je vyslovitelné a použitelné pouze s jedinou specifickou intonací.

Je pak téměř zbytečné dodávat, že frazeologie je i významným prostředkem jazykové **stylistiky**, především tam, kde je jí většina, tj. ve frazémech mluvených, nespisovných; podobnou roli tu má z více hledisek i **rétorika**.

I fakt většinové příslušnosti české IF do mluveného (a nespisovného) jazyka ukazuje mimochodem na omezenost a nefunkčnost teorie spisovnosti. Vzato striktně by pak spisovná čeština neměla mít frazeologii (téměř) žádnou.

Z alternativních a dalších jazykových aspektů nacházíme frazémy také v **dialektologii** a historicky pak i v dosud málo studované **etymologii**, srov. např. znovu (6) *mateřídouška* jako prastarý kalk z lat. *matris animula* (matky dušička, odborně thymus), mající už tenkrát za sebou představu, resp. pověst, u nás zpopularizovanou K. J. Erbenem, že se do dané květiny přetělila duše zemřelé matky (podobně v dalších slovanských jazycích), či *číst někomu levity* (podle středověkého, snad pův. klášterního zvyku kdesi v Belgii připomínat provinilci pravidla správného chování čtením pasáží z biblického Levitica).

K jazykovým disciplínám majícím vztah k frazeologii můžeme konečně volně přiřadit ještě i lingvistickou **sémiotiku**. Zde si jen stručně připomeneme prudký kontrast mezi povahou denotátu komponentů frazému a denotátu frazému jako celku. Zatímco zřejmě většina denotátů prvního druhu má povahu výrazně konkrétní, denotáty frazémů bývají abstraktní povahy, srov. obojí pohled např. u *vrazit někomu kudlu do zad*, *jít i přes mrtvolu* aj. Nicméně zná sémioticky frazeologie jako jediná oblast jazyka i případy, kdy, přestože frazém jako celek denotát a význam má, jsou jeho komponenty bez něj, resp. denotát ani význam nemá, srov. *číst levity*, *vydat všanc* aj.

Dodejme ještě, že **mezijazykově**, zvláště v oblasti srovnávání a učení jazyků, se frazeologie a idiomatika spíše intuitivně odjakživa chápaly jako cosi zvlášť obtížného a nesnadno přenositelného z jednoho jazyka do druhého. Sám termín *idiom* (z řec. *idios* „vlastní, soukromý“, zvláště přes *idiōma* „soukromý majetek a/n výrazivo“) často označuje výraz danému jazyku vlastní a zvláštní, popř. jinde nenalezitelný. Především díky kalkům a příbuznosti jazyků ten poslední atribut tak úplně vždy neplatí, jeho síla je však obecně velká a platí v různé podobě pro každý jazyk. Frazémy a idiomy jsou jazykové universale, jejich vznik a distribuce jsou však velmi specifické.

## 2. IF a nejazyk: jiné obory a disciplíny

Výše jsme naznačili, že se s frazémy či kvazifrazémy setkáváme i v oblasti terminologie, tedy nejrůznějších věd, oborů a disciplín včetně techniky. I tady platí obecně, že čím je disciplína exaktnější, tím méně frazémů se v ní pro jejich vágní povahu použije; případ úzu astronomické *černé díry* je nepochybně dán i tím, že její podstata ještě zcela jasná není a záhadnost jejího jevu zvolený název vystihuje docela dobře.

Všimneme si tudíž stručně především vztahu frazeologie a nejazykových disciplín, u kterých jde o (1a) jejich dílčí odraz v samotné frazeologii (historie, etnografie), anebo naopak o (1b) využití frazémů v nich (literatura, umění), popřípadě o (2) zájem o frazémy určitého druhu, při němž se studuje a srovnává souvislost či dokonce paralelnost jevů studovaných danou disciplínou a vlastní frazeologií (logika, psychologie).

**Historie (1a)** a její události, postavy apod. do národní frazeologie vstupuje v podobě konkrétních frazémů běžně; je to jeden z jejích velkých zdrojů. Je jistě na historických, aby posoudili míru (subjektivní) důležitosti té či oné události a ne/oprávněnost jejího „zvěčnění“ v podobě frazémů, nicméně nějak se toto děje a dále vždy. Srov. české dějiny a frazémy *dopadnout jak sedláci u Chlumce*, *kecat jak Palackej*, *starej Procházka* (tj. císař František

Josef I.) či mezinárodní, zvl. antické *Sejdeme se u Filipp, potěmkinovské vesnice, Znáám své Pappenheimské* (F. Schiller, ústy Valdštejnovými) aj. Významný je tu i podíl **etnografie**, popř. **antropologie**, zvl. v reflexi starých zvyků, způsobů jednání, pověr, popř. řemesel, nástrojů apod., srov. *dát na modlení, svíčková bába, utopit červa, číst někomu levity, oko za oko...*, *vzít si někoho na paškál* (z řec. *paschalis* = velikonoční /svíce/), *světit modrý pondělek* (pův. snad ševcovské privilegium v pondělí nepracovat, později pak pít), *malovat čerta na zeď, jít do kanafasu* (*kanafas* = hrubší látka z konopí užívaná na lůžkoviny), *mluvit do někoho jako do dubu* (*dub* = magický strom Slovanů), *ani zblo* (< stéblo), *je to na jedno brdo* (*brdo* = hřeben tkalcovského stavu), *být dlouhý jak Lovosice* aj. Volně sem lze přiřadit i bájně neexistující krajiny naznačované frazémy jako *jít do Hajan/hajan* apod.

Ponecháme-li stranou samotné zásadní využití frazémů v textu a jeho způsoby, víme (1b) z krásné **literatury** dobře o známé oblíbě využívat potenciál frazémů, především metaforický, obrazný, k označení svých různých produktů, románů, dramat apod. Srov. např. hry Voskovce a Wericha *Nebe na zemi, Rub a líc, Pěst na oko* (vs. *mít nebe na zemi, to je rub a líc* těžé věci, *je to jako pěst na oko*), Řezáčův román *Větrná setba* (vs. *Kdo seje vítr, sklízí bouři*) či Pujmanové *Hra s ohněm* (vs. *hra/hrát si s ohněm*) aj. Jen stručně tu lze naznačit a označit literaturu i jako zdroj vzniku některých frazémů (srov. aspoň *boj s větrnými mlýny*, který proslavil Cervantesův *Don Quijote*, Vančurovské *Tento způsob léta zdá se mi být nešťastným* aj.). Oproti literatuře však **hudba** v této oblasti příliš významná není, srov. ale aspoň operu W. A. Mozarta *Don Juan/Don Giovanni*, který proslavil původní španělskou historickou postavu a dal následně i vznik frazému *donchuán*. Jiný příklad nabízí česká tvorba, kde antický hrdina Tantalos byl u nás napřed zdramatizován J. Vrchlickým pod názvem *Smír Tantalův* a následně zhudebněn Z. Fibichem; frazém *Tantalova muka* ovšem existoval v češtině nezávisle na Vrchlickém.

Frazeologie a její některé výrazy, resp. představy stojící za nimi, zvláště antické, popř. biblické, naopak poměrně hojně v minulosti využívalo **malířství**. Je mnoho obrazů znázorňujících např. *roh hojnosti, návrat ztraceného syna, babylonskou věž* apod. Unikátní a historicky pozoruhodný počín, zastíňující cokoliv jiného, tu však představuje holandský malíř Pieter Breughel ml., který r. 1559 svým svérázným a často nelichotivým způsobem geniálně znázornil a na jediné plátno vměstnal přes sto nizozemských dobových přísloví. O plátnu se píšou knihy, analýzy historické, kunsthistorické, ale i lingvistické (srov. jednu z posledních W. Mieder, ed., 2004, *The Netherlandish Proverbs. An International Symposium on the Pieter Breug(h)els*, Burlington). Originální a řídký český počín tohoto druhu lze však spatřovat i v posmrtně vydané knížce J. Lady (*Ilustrovaná frazeologie*, Praha 1971), který se pokouší svým typickým, obvykle žertovným způsobem převádět do jednoduché kresby několik desítek českých frazémů. Nacházet další podobné příklady, v nichž frazémy jsou inspiračním zdrojem, zvláště v karikatuře, není nijak těžké.

**Psychologie** (2) nabízí řadu způsobů, jak srovnávat pojmy své a frazeologicky podané. Ze srovnání je rychle zřejmé, že často je třeba se smířit s tím, že co do výstižnosti, zkratkovitosti a taky monopolnosti vyjádření a pojmenování psychologie a tedy pravidelný a exaktní jazyk za možnostmi frazeologie jasně pokulhává a leccos tak jednoduše jako ona vyjádřit neumí. Srov. výrazivo z oblasti intelektu, zvláště psychických a paměťových procesů (*lovit v paměti, sypat něco z rukávu, strílet bez míření/od boku, sedlat koně od ocasu*), inte-

lektuálních vlastností (*mít za ušima, nebýt dnešní, mít otevřenou hlavu*), postojů a chování (*přehazovat si něco jako horkej brambor – hodit to někomu na krk – bejt z obliga, sekat latinu, plakat nad rozlitym mlíkem, bejt hned na koni*), vlastností povahových (*neublížil by ani mouše, nechat na sobě dříví štípat, nasadit za někoho vlastní kůži/krk, dřít někomu kůži z těla*) nebo z oblasti citů a pocitů, odpovídajících reakcí apod. (*závidět někomu nos mezi očima, dštít oheň a síru na někoho, všechno se v něm vaří, ronit krokodýlí slzy, snesl by jí modrý z nebe*) aj. Příznačná tu je nominativní ekonomie frazémů tam, kde standardní vyjádření je buď příliš dlouhé a neohrabané, anebo se nenabízí žádné, srov. *bejt z obliga, bejt hned na koni, srazit někomu hřebínek*. V tomto smyslu je zřejmé, že **nominativně** frazeologie není až druhotnou sférou pojmenování, jak se běžně má za to; v řadě ohledů pravidelný jazyk naopak selhává a nevyhovuje a za primární, popř. jediné (prakticky) možné pojmenování je třeba považovat frazémy a idiomy. I zde jsou pak jasné meze a hranice pravidelného jazyka, jeho pravidel sémantických a celé gramatiky. Dodejme, že i v oblasti **logiky**, v níž v podobě své jedinečné knihy jejímu vztahu k přísloví věnoval pozornost O. Zich (1956), se ukazuje, že jak stavba pojmu, tak soudu se v příslovích dělá podstatně jinak, a někdy dokonce zdánlivě „nelogicky“, tj. na rozdíl od klasické logiky, srov. *Pomalů dál ujdeš, třepit si hubu, mít jazyk na vestě* aj.

Historie je ovšem otevřená a stejně tak i frazeologie z hlediska vývoje. Jedinečný případ, kdy naopak frazeologie ovlivnila realitu a vyvolala v život hromadné společensko-sportovní hnutí, představuje dnešní známý pochod Praha–Prčice (realizovaný, resp. zhmotněný a zdoslovněný podle frazému *poslat někoho do Prčic < prdele*, resp. *jít do Prčic < prdele*).

V širším kulturním smyslu je tu na místě si konečně připomenout to, co už bylo zřejmé F. L. Čelakovskému, totiž že ve frazeologii, především však v příslovích, se uchovává značná část tradiční národní zkušenosti a lidové životní moudrosti, a ne tedy pouze různě trefné, okrasné či jen zajímavé výrazy. Odtud pak název jeho jedinečného *Mudrosloví národu českého...* (1852) a zaznamenaných přísloví, jako jsou *Žádný učený z nebe nespad., Kdo mnoho mluví, lže, anebo se chlubí., Kam vítr, tam pláší* aj.

## Summary

### Introduction: Relationship of Idiomatics and Phraseology and Other Disciplines

*In language*, the phenomenon of idioms and phrasemes is limited to combinations of only those units that have a meaning. However, idioms and phrasemes (IP) made up in this way have both meaning and function that displays some kind of anomaly and defy standard semantical and grammatical rules. Being specific language nominations, IP form part of the lexicon, having a number of specific relations to other parts of the language, too.

It is just not enough to see IP in metaphorical combinations of words suggesting some kind of image only. This would exclude non-metaphorical, grammatical or sentence idioms, such as

(1a) *otevřít dokořán, dát najevo, být hin,*

(1b) *co do, na úkor někoho, v souvislosti s něčím, kór když, co kdyby,*

(1c) *Že by?, To zrovna!, Všechno má svůj čas., Blázni a děti mluví pravdu., Sedni si, at' ti kapsa nevisí!* etc.

It is, however, necessary to reconsider and include into IP anomalous combinations of morphemes below the level of word combination, too, which is, in some cases, signalled by a parallel existence as in

(2) *svéhlavý* vs. *mít svou hlavu*, *pecivál* vs. *válet se za pecí* or *měkkosrdcatý* vs. *mít měkké srdce*.

The morphology proper or word-formation offers numerous examples, such as

(3) *najít* (where *na*  $\equiv$  “above” + *jít*  $\equiv$  “move on foot, walk”  $\neq$  “find”, cf. Engl. *find*, Fr. *trouver*, Swed. *finna*, Fin. *löytää*, all of these being simple verbs without any prefix).

Then there are items that belong both IP and to terms, i.e. both single- and multi-word phraseological terms, such as

(4) *mateřídouška*, *zlatý déšť* (laburnum), etc.

One may discern idioms among proper names, too, such as

(5) *labuť avonská* (= William Shakespeare), *Hradec Králové*, *Achillova pata*, *kecá jak Palackej*, etc.

Semiotically, idioms display a number of special features, these including an almost universal abstract denotate (in contrast to the predominant concrete character of its constituents) or, having, oddly enough, constituents that have no denotate at all, such as in *číst levity*, *vydat všanc* etc.

*Outside* language, i.e. in various disciplines, one may find idioms too, or these disciplines may have a kind of relation to idioms in the language.

A notorious example is history, both domestic and international (classical), being reflected in many of its heroes, characters and events in idioms of various kind, cf.

*dopadnout jak sedláci u Chlumce*, *kecat jak Palackej*, *starej Procházka*, *Sejdeme se u Filipp*, *potěmkinovské vesnice*, *Znám své Pappenheimské*, etc.

A similar source of idioms is to be found in ethnography and anthropology, especially in the idioms' petrification of old customs, superstitions, instruments, domestic life, magic or crafts, such as in

*dát na modlení*, *svíčková bába*, *utopit červa*, *číst někomu levity*, *oko za oko ...*, *vzít si někoho na paš-kál* (from the Greek *paschalis*), *světit modrý pondělek*, *malovat čerta na zeď*, *jít do kanafasu*, *mluvit do někoho jako do dubu*.

In literature, a common habit to use idioms not only in texts, but also in the titles of books, play, etc. is well-known, cf.

*Nebe na zemi*, *Rub a líc*, *Pěst na oko* (vs. *mít nebe na zemi*, *to je rub a líc*, *je to jako pěst na oko*), *Větrná setba* (vs. *Kdo seje vítr, sklízí bouři*), *Hra s ohněm* (vs. *hra/hrát si s ohněm*) etc.

An interesting point of contact is to be found in art, especially in painting where, from Antiquity, a number of paintings are familiar bearing names such as *Return of the Prodigal Son* (*Návrat ztraceného syna*), *babylonská věž* etc. A record of its kind may be seen in P. Breughel's famous painting from 1559 depicting over a 100 Dutch proverbs (cf. one of many books on this, W. Mieder, ed., 2004, *The Netherlandish Proverbs. An International Symposium on Pieter Breug(h)els*, Burlington).

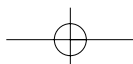
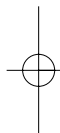
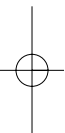
There are a number of idioms offering an alternative to standard names for emotions, intellectual processes, properties, attitudes, etc. all of them clearly showing that not all is comfortably or better expressed in the area covered by psychology, cf.

*lovit v paměti*, *sypat něco z rukávu*, *střílet bez míření/od boku*, *sedlat koně od ocasu*, *mít za ušima*, *nebýt dnešní*, *mít otevřenou hlavu*; *přehazovat si něco jako horkej brambor* – *hodit to někomu na krk* – *bejt z obliga*, *sekat latinu*, *plakat nad rozlitém mlékem*, *bejt hned na koni*; *neublížil by ani mouše*,

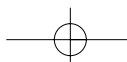
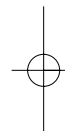
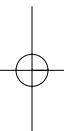


*nechat na sobě dříví štípat, nasadit za někoho vlastní kůži/krk, dřít někomu kůži z těla; závidět někomu nos mezi očima, dštít oheň a síru na někoho, všechno se v něm vaří, ronit krokodýlí slzy, snesl by jí modrý z nebe etc.*

It is important to note that idioms, notably proverbs, have been scrutinized from a logical point of view, too, by O. Zich in a book (1956).



# **A** **TEORIE a PODSTATY** **THEORY and SUBSTANCE**



# Základy a zásady frazeologie a idiomatiky

1. Úvod: Frazeologie a idiomatika v širším rámci .....	24
2. Principy a základy idiomatiky a frazeologie .....	25
2.1 Povaha a místo idiomatiky a frazeologie .....	25
2.2 Paradigmatika a syntagmatika .....	26
2.3 Anomálie ve frazeologii a idiomatice .....	30
2.4 Definice frazému a idiomu .....	31
3. Frazém a idiom: jejich komponenty a útvar .....	32
3.1 Povaha frazému a idiomu obecně .....	33
3.2 Forma frazémů a idiomů .....	35
3.21 Struktura frazému a idiomu .....	35
3.22 Komponenty idiomu a frazému .....	36
3.3 Význam a nominace .....	36
3.31 Povaha významu idiomu a frazému .....	36
3.32 Složky a aspekty významu frazému a idiomu .....	38
3.4 Varianty frazému a idiomu .....	40
3.5 Sémiotika frazému a idiomu .....	41
3.6 Kvazifrazémy a kvaziidiomy .....	42
4. Systém idiomatiky a frazeologie a klasifikace .....	42
4.1 Distribuce idiomů a frazémů v jazyce .....	42
4.2 Klasifikace idiomů a frazémů .....	44
4.3 Formální třídy idiomů a frazémů .....	44
4.31 Třídy kolokačních idiomů a frazémů .....	44
4.311 Autosémantika v úloze komponentů .....	44
4.312 Synsémantika v úloze komponentů .....	46
4.313 Autosémantika a synsémantika v úloze komponentů .....	47
4.314 Přirovnání .....	48
4.315 Verbonominální kvazifrazémy .....	50
4.316 Binomiály .....	51
4.32 Třídy propozičních frazémů a idiomů .....	52
4.321 Frazémy a idiomy propoziční a polypropoziční monosubjektové .....	53
4.3211 Formální struktura a sémantika propozičních frazémů a idiomů .....	53
4.3212 Funkční třídy propozičních frazémů a idiomů .....	55
4.3213 Pragmatická funkce propozičních frazémů a idiomů .....	55
4.3214 Některé tradiční typy propozičních frazémů a idiomů .....	57
4.322 Frazémy a idiomy polypropoziční intersubjektové .....	59
4.33 Frazémy lexikální .....	60
5. Funkce idiomů a frazémů .....	60
5.1 Strukturní složka funkce .....	61
5.2 Nominativní složka funkce .....	63
5.3 Komplexní funkční klasifikace .....	65
6. Transformace frazémů a idiomů .....	66
6.1 Podstata transformace ve frazeologii a idiomatice .....	66
6.2 Typy transformací .....	67
6.21 Mezirovinné transformace .....	67
6.22 Vnitrorovinné transformace mezistrukturní .....	68
6.23 Vnitrorovinné transformace vnitrostrukturní .....	68
6.3 Transformace aktualizací .....	69
7. Závěr: Místo frazeologie a idiomatiky v jazyce a textu .....	71

## 1. Úvod: Frazologie a idiomatika v širším rámci

Idiomatika a frazeologie (IF) je lingvistická disciplína poměrně mladá a až postupně se ukazuje, kde všude v jazyce její jednotky, idiomy a frazémy, můžeme hledat. V jazyce jsou obecně řečeno k nalezení všude tam, ve všech **jazykových disciplínách**, kde jde o kombinace jednotek nadané významem. IF má však i navenek vazby další **disciplíny nejazykové**. Málo studované tu jsou souvislosti s psychologií a logikou, zvl. v oblasti vyjádření pojmů a soudů, s filozofií, resp. etikou hodnocení a tradování zkušenosti, s historií a uměním (trópy, symboly ap.), avšak tradiční vazbu má především s kulturní antropologií a etnologií, zvláště v oblasti studia přísloví. Tam se pak celkem odůvodněně mluví už dlouho o pokladnici národní zkušenosti pod názvem mudrosloví, jak ho použil F. L. Čelakovský pro svou monumentální sbírku (Čelakovský 1852).

Jestliže etnografické, sémantické, historické, sociologické a další přístupy objasňují mj. národní specifika, včetně metaforiky a symboliky, kterou IF vyjadřuje, lze na první pohled říct, že široce a volně řečeno je co do své skladby česká IF svou významnou historickou částí rustikální, zemědělská, zatímco námořní tematika, běžná např. pro angličtinu, jí zcela chybí. Takovéto charakteristiky ale svou obecností mnoho neříkají a nemají valnou cenu. Na druhé straně se jen pomalu ujímá představa, že v IF nejde jen o pouhou sbírku nepravidelností a kuriozit kdesi na periferii jazyka a zájmu, že jde naopak o specifickou a málo studovanou oblast jazyka, kde s běžnými prostředky a postupy nevystačíme a která se nabízí i k aktivnímu studiu a využití v různých aplikacích v řadě komunikativních činností a oblastí, např. pedagogické ve škole nebo rétorické, resp. textové při výstavbě projevů různého typu. K základním deskriptivním aplikacím tu patří **frazografie**, shromažďující a popisující frazeologické výrazivo jazyka.

Volně a nejen už lingvisticky je možné frazémy a idiomy chápat jako nejvýznamnější součást stereotypů, především jazykových. **Stereotyp** je v jazyce (tj. vedle širšího pojetí např. v sociologii) lexikální kombinace vstupující do textu jako hotová, relativně frekvencovaná a známá, resp. tradovaná jednotka, která je ustálená a negenerovatelná, za níž stojí model s podobně malou či (téměř) žádnou produktivností. Komunikacně svou prefabrovanou povahou stereotyp budí dojem povědomosti, navozuje často jistý typ situace či chování a jazykově usnadňuje, urychluje i automatizuje tvorbu textu; ten v těchto částech ztrácí svou originalitu a jeho autor tak i svou kreativní možnost se vyjádřit originálně. Stereotypů je zvláště hodně v institucionalizovaném jazyce (takový jazyk je pak výrazně stereotypní), kde se používá řada formulí, které promluvu formalizují a dodávají jí uzuální a očekávaný (obvykle zdvořilostní) nebo i předepisovaný charakter, např. etiketou. Míra závazné platnosti může pak být tudíž sankcionována sociálně nebo zákonem. Typické jsou proto např. výrazy a formule pozdravové, děkovné, přací, rituální nebo formule přísežné, zatýkácí, svatební a další; jsou i formule užívané v některých specializovaných oborech.

Přestože se formule užívají v institucionalizovaném jazyce a dokonce v jazyce mluveném např. v různých konverzačních obrazech, nelze je zjednodušeně považovat za jev negativní. Sám výběr mezi řadou typů stereotypu může ukazovat na autorský přínos ke stavbě textu; téměř všechny typy textů stereotyp také vyžadují, protože lidé nemluví ani zdaleka jen kreativně a zcela originálně a určité uzuální a dohodnuté pojmy a situace je třeba vyjád-

řit uzuálním a povědomým způsobem. Původní typografický význam výrazu (pevné spojení více liter dohromady proti literám volným) tu do jisté míry je tedy stále přítomný. K nejběžnějším stereotypům patří propoziční a polypropoziční frazémy, zvláště přísloví, pořekadla, korelované pozdravy a přání apod. a řada dalších typů frazémů, ale i aforismy, různé situační průpovídky, standardní začátky či konce promluv a jiné kontaktové formule apod., hlavně však velké množství verbonominálních frazémů. O vztahu frazémů a stereotypu se diskutuje a problém je to otevřený. Z jiného hlediska jsou do značné míry identické a překrývají se, mají-li pevnou jazykovou formu. Řada stereotypů (obvykle negativně konotovaných), studovaných především sociologií, však takovou pevnou formu nemá a pak s frazeologií už přestává být shodná. Záleží to také na kritériích vymezení samotné frazeologie. Sociologické pojetí stereotypů pracuje často, zvláště v oblasti tradičních pohledů na jiné národy, s představami, které nemusejí být pravdivé ani lichotivé: např. Skotové jsou takto v češtině stereotypně vnímáni jako lakomí (srov. typické anekdoty o nich), Němci jako pořádkumilovní a přesní (*německá přesnost*), Francouzi jako kuchaři a milovníci, cikáni jako zloději nebo špinaví ap. Přesto se i některé tyto nejazykové stereotypní představy promítly do lingvistiky, resp. frazeologie v podobě tam existujících frazémů navozujících danou představu o Romech (*krade jako cikán*), Čechů coby hudebníků (*Co Čech, to muzikant*), Turců jako chaotických ekonomů (*turecké hospodářství*) aj., jakkoliv realitě nemusejí odpovídat. Pak se propojují oba zmíněné typy stereotypů dohromady.

IF ovšem není disciplínou studující **fráze**. Fráze jako ustálený konverzační výraz, chápáný zpravidla jako ořelý a sémanticky téměř prázdný, je pojem subjektivní a nedefinovatelný. Ani odborné pojetí fráze, známé zvl. z generativní gramatiky, není náplní studia této disciplíny. Zabývá se frazémy a idiomy a v dalším bude věnována pozornost jim.

## 2. Principy a základy idiomatiky a frazeologie

### 2.1 Povaha a místo idiomatiky a frazeologie

Disciplína idiomatiky a frazeologie (dále IF) se zabývá studiem a popisem frazémů a idiomů všech úrovní, kvůli jejich početnosti především však studiem lexémů víceslovných. Není to zvláštní jazyková rovina, jak se dříve mělo za to, je to ale oblast ustálených anomálních kombinací (syntagmatických a paradigmatických) forem vždy nižšího řádu, které jsou komplementární k oblasti pravidelných jazykových kombinací všeho druhu a na všech relevantních rovinách; rezultující frazém, uložený v systému, je pak členem vyššího řádu, než jsou jeho komponenty. Tou relevantní rovinou se míní každá rovina, jejíž jednotky jsou nadány významem (sémantikou).

**Pravidelnými** kombinacemi v obecně syntagmatickém smyslu se tu míní všechny kombinace řízené analogickými pravidly, tj. pravidly opírajícími se o analogii a tedy kombinace morfémů, lexémů, kolokací i vět. Těmito pravidly se tu obecně rozumějí jak (vždy primární) pravidla sémantická, založená na významové slučitelnosti kombinovaných prvků a smysluplnosti jejich výsledné kombinace (viz ale dále i kompatibilita, 2.2), tak (vždy sekundární) pravidla formální a gramatická (např. v syntaxi syntaktická apod.). IF je tedy v oblasti kombinatorických výsledků extenzí jazykových kombinatorických možností dále, za hranice pra-

videl do oblasti kombinací, které podle standardních pravidel v jazyce už nemůžou a nemají vznikat, a tedy kombinací **anomálních**. Už tím je pak dána i inherentní nepravidelnost a nemodelovost utvoření každého idiomu a frazému. Např. jednoduchý frazém *ležet polem* je formálně anomální a má jinde v jazyce, u obdobných komponentů (je tu statické sloveso a substantivum *pole*), analogicky nenalezitelnou další instrumentálovou konstrukci (a syntakticky cosi vzdáleně připomínající např. objekt); přitom sémanticky se v IF oba komponenty kombinují pouze zcela tím jediným odlišným způsobem, srov. např. tento frazém a pravidelnou kolokaci *ležet na poli*, tj. sloveso s adverbiale na otázku „kde?“.

Diachronně anomální frazeologické kombinace vznikaly různým způsobem, u kolokačních frazémů zvláště skrze přirovnání a individuální metafory, a můžou mimo jiné dnes odrážet i rezidua kombinatorických možností dříve běžných, dnes ale diachronních, které dnešní jazykový systém už opustil. Ve schopnostech jazykových jednotek kombinatorických, ležících v samém jádru každého jazyka (lze tu obecně mluvit o **kombinatorickém principu**), tkví i možnost vytváření stále nových pojmenování včetně anomálních, a tedy frazémů a idiomů. Existence frazeologie a idiomatiky je zákonitým jevem každého jazyka, stejně jako jejich užívání je pro jejich specifické vlastnosti a výrazové možnosti odstíněným prostředkem jazykové kultury.

Společná **nominativní funkce** frazémů a idiomů všech typů oblast IF řadí především do lexikologie jako její zvláštní část; v dalších ohledech ovšem má specifickou souvislost i s řadou disciplín dalších. Charakteristickou zvláštností a podstatou IF však je, že tu jsou platnost a fungování principů založených na pravidlech známých odjinud z jazyka zrušeny a nelze je pak ani pro analýzu systematicky využít: co tu vypadá jako pád, funkčně pád není, co se zdá být objektem verba, jím není, slovo tu není slovem se svým povědomým významem, tvar slova není tvarem lexému stojícího v pozadí, protože ten tu není relevantní a jako celek sem vůbec nevstupuje apod. Je to dáno všudypřítomnou **anomálností** různého druhu, která především neumožňuje uplatnění řádné analýzy frazému založené na kritériích pravidelného jazyka ani naopak jeho syntézy z výchozích komponentů. Přesto lze kritéria pravidelného jazyka využít v IF aspoň na úrovni komponentů (pouze) k vystižení jejich vnější charakteristiky, resp. vnějškové a povrchní podobnosti komponentům pravidelného jazyka, např. při formální klasifikaci frazémů založené na takovéto podobnosti se slovními druhy. Lze tudíž výhodně vystihnout strukturu a typ např. frazému *nosit kůži na trh* jako kombinaci **V-SprepS**, tj. verba, substantiva, prepozice a substantiva, jakkoliv tu o žádná substantiva ve skutečnosti nejde, popsaná struktura nebo případná analýza, která ji užívá, je jen vnějšková a o složení a povaze frazému nic (přesně) nevypovídá (o využití viz blíže zvl. v 4.).

## 2.2 Paradigmatika a syntagmatika

Paradigma je abstraktní třída prvků, jednotek libovolné roviny systému, které patří samo obvykle taky do systému a je založené na společném rysu těchto členských prvků (tj. formálním, sémantickém nebo funkčním), který je sjednocuje v paradigmata různého typu (srov. i Čermák 2000a). Členy paradigmatu, mezi nimiž je vzájemná ekvivalence, jsou v textu obecně vzájemně substituovatelné a pro jejich výběr v něm platí pravidlo *bud' – anebo*. Paradigma je obecně buď otevřené, tj. relativně snadno přístupné obohacování o další



členy a tedy zpravidla i rostoucí, anebo zavřené, dál už v tomto smyslu nerostoucí (srov. už Čermák 1978). Vedle např. intuitivně nepřehledného a velkého paradigmatu domácích kořenů, desubstantivních adjektiv či např. druhů jídel, zvířat apod. je naproti tomu pro frazeologii vždy důležité **zavřené paradigma**, které je obvykle velmi malé, popř. i jednočlenné, a jeho členy už zjevně dále početně nerostou. Podívejme se na tuto situaci poněkud podrobněji; pro další bude přitom mj. třeba rozlišit paradigma virtuální a kolokační.

Každé slovo, lexém, morfém atd. patří v systému do nějakého **paradigmatu** (jednoho nebo více); paradigmaty jsou tudíž různě rozsáhlá a v systému jazyka i různě abstraktní. Tak lze na jedné straně zjišťovat např. velmi širokou (početnou) třídu všech substantiv (tj. substantiv jako slovního druhu, čítajících v češtině několik desítek tisíc jednoslovných lexémů), jejíž povaha je především gramatická a jejíž ekvivalence a kombinatorika, projevující se v jejich širě pojímané vzájemné kontextové zaměnitelnosti (je shodná s obecnou pojmenovávací platností substantiv), je značně obecná; shoduje se v podstatě se syntaktickými rolmi substantiva. O jejich sémantické povaze se toho konkrétního však říct mnoho nedá. Naproti tomu ale např. třída posesivních adjektiv odvozených od názvů živočichů (např. *psí, kočičí*) je už výrazně užší, méně početná a její povaha je už výrazně sémantická. Projevuje se v široce ekvivalentní funkci v omezeném množství možných kontextů (adjektivum *psí* stojící mj. před substantivy látkovými, např. *chlupy, maso, kosti*, substantivy označujícími pohyb a zvukový projev, např. *štěkot*, ale nikoliv obvykle před názvy osob, nástrojů, výrobků, míst apod.).

Paradigmat je obecně ovšem mnoho, nejen co do počtu, ale především typu, a to zvláště v závislosti na různé výši jejich příslušnosti do jazykové hierarchie. Obecně a zhruba platí, že povaha paradigmatu souvisí s povahou jeho funkce a že je tím sémantičtější, čím výrazněji širší paradigmatu a počet jeho prvků klesá, tj. čím je užší (menší), a naopak.

**Virtuální paradigma** je takováto třída výše uvažovaná, tj. třída (široce) ekvivalentních prvků spjatých společnou funkcí a významem, která je dána abstraktně, je uložena v jazykovém systému a je vymezitelná svým místem v jeho hierarchii. Jeden druh virtuálního paradigmatu tvoří např. výše uváděná adjektiva typu *psí, kočičí, lví, koňský*, jiný např. substantiva látková typu *maso, sádlo, kůže, mléko* a ještě jiný např. názvy nástrojů typu *nůžky, kladivo, šroubovák, vrták* aj. První naznačené virtuální paradigma (*psí* aj.) se řadí s posesivními adjektivy vztahujícími se k člověku (*otcův, tetin*, popř. *otcovský, mateřský*...) do vyšší třídy posesivních adjektiv vůbec, ta pak ještě do vyšší třídy a tedy abstraktnějších adjektiv obecně apod. Virtuální paradigma vnímáme obvykle snadno i intuitivně, dokážeme si ho do jisté míry snadno vybavovat a buduje na něm, resp. na nich většina (sémantických) klasifikací lexémů, studium hyponymie aj. Každý člen virtuálního paradigmatu má přitom obecně stejnou funkci jako ostatní (tj. obecně vystupovat ve stejné vyšší jednotce, struktuře), individuálně se však vždy co do své kombinatoriky liší; na úrovni lexikální se mluví o **kolokabilitě**, která je pro každý z nich individuální.

**Kolokační paradigma** je naproti tomu taková třída prvků (zde stále slov), s nimiž se všemi je prvek jiného paradigmatu schopný se obvykle a smysluplně kombinovat; intuitivně se však taková třída jako celek vybavuje velmi obtížně a je-li rozsáhlá, není to zpravidla možné vůbec. Z příkladu s adjektivem *psí* to jsou všechna substantiva s tímto adjektivem

spojitelná (vzhledem k jejich heterogenosti půjde o více významově rozrůzněných paradigmát zároveň), např. o

- (1) typ představovaný substantivy zvukového projevu, jako je *štěkot*, *vrčení*, *kňučení* aj.,
- (2) typ částí těla, jako je *chlupy*, *maso*, *kosti* aj., popř.
- (3) další (pohyby, vlastnosti aj.).

Kolokační paradigma je tudíž vymežitelné jen skrze určitý **externí** prvek, k němuž se **jedinečně** vztahuje na základě společné kolokability (spojitelnosti) s ním, který však sám do něj nepatří (zde **psí**). Syntagmata, která prvky kolokačního paradigmatu se svým konkrétním externím („definujícím“) prvkem vytvářejí, jsou obvykle významově i funkčně různá. Svě (aspoň jediné) kolokační paradigma má každý prvek systému (tj. slovo aj.) a jeho explicitním výčtem lze vyjádřit možnosti i hranice užití (**kolokability**) každého prvku. Takto se např. běžně vytváří syntagma *psí štěkot*, ale ne už *\*psí mňoukání*, resp. *\*kočičí štěkot*, *\*psí bučení* atd. (kde brání sémantická pravidla). Na rozdíl od paradigmatu virtuálního se rozsah paradigmatu kolokačního v časovém vývoji může rychleji měnit a také často měnit; zde je mj. základ zrodu metaforiky, obecněji změny významu lexému a nejobecněji změny jazyka vůbec. Povaha obou typů paradigmát (virtuálního i kolokačního) se promítá do pojetí, vnímání i funkce každého frazému a idiomu.

**Syntagma**, minimální kombinace dvou obvykle heterogenních prvků, tj. pocházejících (zpravidla) ze dvou různých paradigmát virtuálních, je dána sémanticky povahou jeho prvků, obecně však i druhem jejich paradigmát a rovinnou příslušností; je tudíž podmiňována kombinatorickými pravidly řídícími chování prvků. Většina syntagmat je neustálená, textová, a jen některá, především frazémy, jsou ustálená a jsou součástí langue. Na syntagmatu a kombinaci, využijeme-li příkladu *psí štěkot* výše, která je neustálená, z hlediska frazeologie nic zvláštního není. Je zřejmě vždy znovu a znovu vytvářena ad hoc, a to podle pravidel sémantických (komponenty se tu nejen neodpuzejí, ale tradičně kombinují, je mezi nimi společný sémantický rys). Naproti tomu syntagma *psí počasí* frazém je, protože je anomální: neuplatňuje se tu žádné sémantické pravidlo (*psí* a *počasí* jsou normálně neslučitelné stejně jako *\*psí déšť*, *\*psí klima*, *\*psí podnebí* apod.), obě slova stejně jako virtuální paradigmata, kam patří, se spolu nekombinují; z hlediska virtuálního paradigmatu nepatří slovo *počasí*, popř. *čas* do žádného z výše naznačených kolokačních paradigmát vztahujících se k adjektivu *psí*. Jinými slovy slovo *počasí* je ve vztahu k adjektivu *psí* teoreticky i prakticky kolokačním paradigmatem s jediným členem (*čas* je tu vlastně jeho varianta), zcela zavřeným, které ale konkrétně, resp. sémanticky do okruhu slov kombinujících se uzuálně se slovem *psí* nepatří. Frazém *psí počasí* (resp. starší *psí čas*) je založen tudíž na jedinečné anomálii, bez opory v jakékoliv analogii.

Výše je uveden pro názornost jen jednoduchý příklad; jiné případy frazémů mohou být složitější. Nicméně už z něj je podstata frazému zřejmá. Protože frazémy a idiomy je třeba chápat jako kombinace prvků z určitého hlediska vždy anomální, je nutné jejich výskyt hledat většinou v neproduktivních a nevyužívaných oblastech kombinatoriky jazykových prvků. V oblasti pravidelného jazyka mají co do rovin největší produktivitu modely syntaktické, resp. kolokační, nejmenší slovotvorné, zatímco flektivní stojí mezi nimi. Jen obecně proto platí, že výskyt idiomů a frazémů i možnost (ne však nutnost) vzniku nových je těmto poměrům pravidelného jazyka zhruba úměrná. Proto také hlavní soustředění IF nacházíme v oblasti kolokací a vět, jakkoliv je i počet kombinací morfologických nemalý (ten však dosud systematicky zkoumaný není).

**Kolokabilita** je obecně individuální formální a významová spojitelnost, schopnost kombinace každého jazykového prvku s jiným nebo jinými, vytvářená jeho jedním nebo více kolokačními paradigmaty a (u pravidelných kombinací) podmíněná jeho kompatibilitou s nimi; spolu s valencí patří k hlavním složkám syntagmaticnosti každého jazykového prvku. Individuální realizací kolokability vzniká konkrétní **kolokace**. Je-li např. **valenci** verba *číst* akuzativ, je jeho kolokabilitou (kolokační) paradigma, třída konkrétních substantiv, která tento akuzativ konkretizuje, např. *číst (knihu, román, dopis...)*, a kolokace *číst knihu* je jedna jeho konkrétní možná kombinace. Do kolokačního paradigmatu slovesa *číst* však nepatří např. inkompatibilní *\*číst kámen, \*číst vodu, \*číst maso...*, i když to jsou také akuzativy. Obecně platí, že kolokabilita každého slova je úzce restringovaná a v zásadě daná jeho sémantikou a i když se časem i autorsky mění, nikdy není neomezená a libovolně široká: žádný lexém jazyka se nekombinuje se všemi dalšími. Kompatibilita je ve svém úhrnu zásadní složkou saussurovské hodnoty lexému a je pro libovolné dva lexémy, i synonymní lexémy, vždy odlišná, resp. jedinečná a nezaměnitelná. O kolokabilitě se na úrovni jiné než lexikální (autosémantické) zpravidla neuvažuje.

Kolokabilita lexémů, resp. slov (jejich tvarů) je různá jak co do typů, tak co do množství slov, lexémů aj., s nimiž se dané slovo kombinuje; závisí to na povaze jeho virtuálního paradigmatu, jeho místu v abstraktní hierarchii systému aj. (viz výše). Mluvit o kolokabilitě slov jako *velký* prakticky, bez vyčerpávajícího studia podepřeného korpusem, nelze (jde zřejmě o statisíce slov, především substantiv, a ty v hlavě nikdo mít nedokáže, nedokážeme si je ani vybavit; v korpusu SYN2005 jde o cca 200 000 kombinací); slovo *velký* takto tvoří, resp. má k sobě nepřehledně velká a otevřená kolokační paradigmata, která jeho kolokabilitu vytvářejí. Naproti tomu jsou i kolokační paradigmata krajně malá a zavřená, srov. omezené možnosti kolokability adjektiva *širý* se substantivy {*moře, pole, pláň, svět, obloha, krajina*}. Je zřejmé, že tu opět nacházíme jednu z nejběžnějších anomálií ve frazeologii, členství komponentu v krajně malém a uzavřeném paradigmatu. Mezi složkami kombinací lexémů a slov pravidelného jazyka, pokud jsou smysluplné, je sémanticky vždy vztah kompoziční. Z jiného hlediska tyto kombinace zásadně umožňuje vždy jejich princip významové slučitelnosti, **kompatibilita**, spočívající v přítomnosti společného sémantického rysu v obou kombinujících se slovech (zhruba: jak forma *číst*, tak *knihu* mají v sobě rys „čitelnosti“, řečeno jen náznakově).

Na rozdíl od pravidelného jazyka, jehož kombinace vycházejí pouze a nutně z kompatibility, frazémy a idiomy jsou jiné. Mezi jejich komponenty může být vedle kompatibility i její absence, vztah inkompatibility.

**Inkompatibilita**, významová neslučitelnost, je tedy sémantický syntagmatický vztah mezi dvěma vzájemně se vylučujícími lexémy nebo jinými prvky sémantické povahy, které se tudíž nemůžou běžně kombinovat ve smysluplném syntagmatu, srov. *\*ostrá voda, \*cítit kolmo* nebo vztah *pes a strom* (ty se běžně v jazyce nekombinují vůbec). V idiomech je ale inkompatibilita možná a častá, srov. *kamenný obličej, otevřená hlava, umět se narodit, ztratit hlavu, to tě bude stát krk* ap.

Od pojmu zavřeného paradigmatu se odvíjí pro frazeologii zásadní pojem a pojetí **monokolokability**. Je to taková kolokabilita lexikální formy, která je jasně omezená na možnost kombinovat se pouze s několika málo (2–7), resp. jedinou další formou v jazyce; ta (resp. ty)

ho chrání před jazykovým zánikem, protože bez kombinace, v úplné izolaci, v jazyce není možné nic. Tyto známé unikátní a **monokolokabilní** formy, jako jsou vedle výše uvedeného *širý* tvary jako *holičkách*, *duhu*, *pré*, *namále*, *vsříc*, *třetice*, *tratoliště*, *dokořán*, *jakž takž*, které nacházíme ve frazémech *nechat někoho na holíčkách*, *jít někomu k duhu*, *mít pré*, *mít namále*, *vyjít někomu vsříc*, *do třetice všeho dobrého*, *tratoliště krve*, *otevřít/být/nechat/zůstat dokořán*, *jakžtakž*, jsou v důsledku své drasticky omezené kolokability zcela jiné povahy než běžné lexémy: nejenže nemají svá lemmata (tvar *\*holičky* je nesmyslný a uměle rekonstruovaný), nemají ani svůj slovní druh ani další kategorie (otázka, zda *\*duhu*, *\*pré* je substantivum či v druhém případě objekt, nedává opět smysl, není tu žádná analogie pro úsudek, protože tu žádná syntax dosud nevznikla). Tyto prvky nemají dokonce ani vlastní význam, ba ani vlastní existenci, protože všechny tyto jen velmi nepřesně **odhadované** vlastnosti by byly možné pouze na základě dostatečné analogie ve více různých výskytech, ve kterých by daná forma musela existovat. Ta však neexistuje. Monokolokabilita je tudíž jen zdánlivě samostatná existence formy, která je ale ve skutečnosti bez svého frazému zrušená a fiktivní (daný frazém je její jediný kontext a tedy i jediná forma její jazykové existence); ve skutečnosti by se takové frazémy mohly psát i nediskrétně, dohromady (srov. poslední příklad), na jejich povaze by se však nic nezměnilo.

Diskrétnost formy je zavádějící; svádí k tomu, postulovat samostatnou existenci tomu, co se píše zvlášť. Specificky je to problém češtiny a podobných jazyků, obecněji pak latiniky; v jazycích užívajících např. jiná písmena tento problém do té míry nevystává, srov. čínštinu apod.

## 2.3 Anomálie ve frazeologii a idiomatice

Ve frazeologii se anomálie vždy vnímá v kontrastu k pozadí sémantických a formálních poměrů v pravidelném jazyce, tj. v jazyce řízeném pravidly, a je také umožněna pouze díky tomuto kontrastu se vším většinovým, co je v jazyce pravidelné. Je chápána jako mnohoaspektová anomální, nepravidelná povaha poměrů paradigmatických (viz výše) nebo realizace syntagmatického nebo transformačního vztahu (často v propojení), která je vnímána jako negenerovatelná odchylka od pravidel různého druhu, založených na analogii. Anomálie je v různém stupni a druhu ve frazeologii všudypřítomná a je jejím konstitutivním rysem. Zatímco pravidelné analogické kombinace jsou v jazyce opřeny o sémantická a formální pravidla a umožňují relativně neomezeně tvořit nové a nové kombinace, srov. *otevřený + dveře/okno/krabice/oko/...* (pravidlo sémantické: věci „otevratelné“, pravidlo gramatické: slovosled a rodová shoda), anomální frazeologické kombinace jsou vždy jedinečné a omezené (absolutně, někdy jen na několik málo jasných členů). Srov. frazém *otevřená hlava*, kde ale už nejde žádný z jeho komponentů variovat jinak a nahrazovat ho za analogický jiný ve stejné funkci či významu (a dostávat tak další frazém), což je základní rys členů téhož paradigmatu v pravidelném jazyce, např. *\*otevřená záda*, *\*otevřené břicho*, *\*otevřené oko* nebo *\*zavřená/přivřená hlava* aj. (pokud nejde o odlišný význam „rány“). Podobně platí, že lexémy zpravidla tvoří pravidelné třídy (např. *pes* patří mezi savce, *stůl* mezi nábytek), kdežto ve frazeologii takové analogické třídy komponentů nejsou, srov. definici frazému dál. Do řady druhů anomálie patří např. nemožnost určit slovní druh u komponentu *hin*, *duhu* (srov. *být hin*, *jít k duhu*), absence významu u těchto forem, ale i tvaru *otevřená hlava* apod., které se neshodují se žádným z běžných lexikálních významů, a nelze u nich tudíž

mluvit o lemmatech, jde jen o tvary, resp. prvky. K syntagmatickým anomáliím různého druhu patří např. *ležet ladem*, *malá domů*, ale i *hodit něco za hlavu* apod.

Nejzjevnějším anomálním rysem je v duchu vymezení výše anomálnost funkční, a tedy **nemožnost paradigmatické záměny** některého komponentu analogickým jiným ve stejné, resp. obdobné funkci (= **komutační test**), srov. už výše a dále např. *\*mít z něčeho lehkou/obtížnou hlavu*, *\*mít z něčeho těžké tělo/břicho/koleno*, *\*umět toto/tohle s někým*, *\*šesté/třetí kolo u vozu*, *\*co od/pro/k*, *\*kdo když*, *\*jestli/zda jo*, *\*Jak si někdo odestele, tak se položí* atd. (tj. vymyšlené nefrazémové variace k frazémům *mít z něčeho těžkou hlavu*, *umět to s někým*, *páté kolo u vozu*, *co do*, *co když*, *že jo*, *Jak si kdo ustele, tak si lehne*). Vzhledem k povaze frazémů se však anomálnost může projevit i v samotném statutu komponentů. Řada frazémových komponentů nemá totiž analogickou obdobu vůbec v žádném tvaru (či lemmatu) v pravidelném jazyce a vyskytuje se pouze ve frazeologii.

Velmi výrazně se anomálie projevuje v **gramatických transformacích**, srov. např. frazém *přirůst někomu k srdci* a nemožnost transformace a tedy jeho užití např. v otázce, pasivu, kondicionálu, imperativu či první osobě singuláru i plurálu. Tento druh anomálie je dobře vidět v omezeních textového úzu toho kterého frazému, která jsou vždy individuální a je třeba je pro každý frazém znát stejně tak, jako známe jeho význam (viz 6.3).

Protože se různé druhy anomálie u konkrétního frazému zpravidla sdružují v různém složení a počtu, lze podle míry jejich zapojení mluvit i o stupni **idiomatizace (frazologizace)** takového frazému; nejnižší mívají periferní kvazifrazémy a kvaziidiomy (viz dál, též 3.6). Nemá-li daná kombinace anomálií žádnou, není to frazém.

Anomálních rysů je mnoho druhů a lze je tedy využít pro funkční klasifikaci frazémů a idiomů (viz 5.3).

## 2.4 Definice frazému a idiomu

Frazém a idiom nelze zpravidla vymezit jedinou (anomální) vlastností, ať formální či sémantickou nebo kolokační. Ani jejich snad dosud nejrozšířenější pojetí jako *ustáleného a reprodukovatelného spojení prvků (zvláště slov)*, jehož význam je (zčásti nebo zcela) neodvoditelný z významu jeho složek nevyhovuje všem typům (např. frazémům s monokolokačním komponentem) a je třeba ho odmítnout. Výše uplatňovaný test analogické záměny komponentu (2.3, příklad *otevřená hlava*) lze vyjádřit pak operačně tak, že

*idiom a frazém je jedinečné spojení minimálně dvou prvků, z nichž některý (popř. žádný) nefunguje stejným způsobem v jiném spojení (resp. více spojeních), popř. se vyskytuje pouze ve výrazu jediném (resp. několika málo).*

I v tomto vymezení, třebaže implicitně již naznačuje více než pouhou sémantickou zvláštnost a lze ho užívat zvláště k praktické **identifikaci** frazému a idiomu, se však jednostranně zdůrazňuje stránka kolokační. Uvedený **test komutace** některého prvku v daném smyslu lze, po ověření primárního aspektu ustálenosti dané kombinace, použít v praxi dobře

pro identifikaci frazémů a idiomů při textové analýze a odlišit ho od případů pravidelných, kde platí běžná pravidla.

Je tedy stále zřejmé, že všechny zmiňované i další vlastnosti do jednoznačné konkrétní a jen pro frazém a idiom specifické definice vtěsnat nelze. Sám statut těchto vlastností, vymezovaných negativně jako anomálie, které jsou nazírány na pozadí pravidelného v jazyce, je relativní; je závislý na zpracování pojmu **pravidelný** (popř. pravidelně generovatelný) a na jeho explicitní náplni v jazykovém popisu podle určité koherentní teorie jazyka. V zobecněné podobě a s důrazem na to, že všechny vlastnosti jsou jen projevy vlastností jazykového systému a struktury, je však možné uvažované stránky shrnout a zobecnit dále. Takto je pak

*frazém a idiom takové nemodelové a ustálené syntagma prvků, z nichž (aspoň) jeden je z hlediska druhého členem extrémně omezeného a uzavřeného paradigmatu (formálně a většinou i sémanticky).*

Prvkem se přitom rozumí běžná jazyková jednotka libovolné roviny (kromě fonologické), ve skutečnosti však pouze určitá podoba, forma této jednotky, tj. např. určitý slovní tvar, morf apod. Výrazy jako „kombinace slov, vět“ apod. jsou proto v této souvislosti užívány, přísně řečeno, jen s jistou licencí. Dva, popř. více frazémů a idiomů se však samy jako prvky idiomatického syntagmatu vyššího řádu setkat zásadně nemůžou. Aby uvedená definice nezahrnovala také víceslovné termíny (srov. např. ptáka *vdovka královská* /*vidua regia*/, kde komponent *vdovka* se užívá ještě s několika dalšími adjektivy, např. *vdovka rajská*), je pak třeba přihlídnout také k dalším rysům; pouhým rysem omezené kolokability se z kombinace frazém ještě nestává (zde jde pouze o kvazifrazém, viz 3.6).

Od frazémů a idiomů je třeba lišit doslovná **homonyma**, srov. *otevřená hlava* ve smyslu chirurgickém „rozoperovaná“, ale i *chytat lelky*, *zlatý důl*, *házet flintu do žita* a další, s jejich doslovnými protějšky, které jsou zpravidla řidší a mají v textu zcela jinou distribuci.

### 3. Frazém a idiom: jejich komponenty a útvar

U IF je zásadně důležité rozlišovat úroveň vstupních forem, komponentů a úroveň výstupních hotových kombinatorických výsledků, idiomatických celků. Pro každou je třeba zvláštního a odděleného přístupu, navzájem nepřenosného. Zatímco pro vstupní formy platí možnost jejich analýzy, tj. ještě před vstupem do idiomatické kombinace celku, která je obdobná jako v pravidelném jazyce (tu využívá strukturní klasifikace), je analýza na úrovni celku už nemožná, protože vstupní komponenty tu ztratily svou původní vnější povahu a tvoří součásti celkového útvaru jen podmíněně. Naproti tomu se nabízí na této úrovni bohatá možnost jejich funkční analýzy, založené na textovém chování. Podstata výsledné idiomatické povahy frazému je v zásadě individuální a daná úhrnem jednotlivých anomálií toho kterého frazému a ty lze nazírat škálovitě: **čím více anomálií frazém má, tím je idiomatičtější a naopak**, tj. čím méně jich má, tím je blíže kombinacím pravidelného jazyka. Toto škálovité pojetí pak ve svém důsledku naznačuje možnost vymezení jakéhosi idioma-

tického centra celé oblasti (na každé kombinatorické úrovni zvláštní) stejně tak jako periferie s oblastmi ležícími mezi obojím.

Naznačuje však také možnost vystihnout i početné případy přechodné, jejichž míra idiomatičnosti je slabá, tj. kvazifrazémy a kvaziidiomy (3.6).

### 3.1 Povaha frazému a idiomu obecně

Jednotlivé složky frazému a idiomu (o jeho obecné povaze viz 2. výše) jsou minimálně tyto: prvky a jejich povaha, kombinace včetně sledu, ustálenost, počet prvků a funkce prvků, které jsou zpravidla vždy různě anomální povahy. Pro bližší vymezení frazému a idiomu je třeba posuzovat tyto aspekty z hlediska formálního, sémantického a kolokačního, resp. funkčního; u posledního aspektu pak i z hlediska jeho anomálních transformací (viz 5.3). Frazému a idiomu, který je základní jednotkou frazeologie a idiomatiky jazyka, se lidově a nespécificky, resp. v různém a kolísavém smyslu, občas dosud říkává **rčení** (popř. **úsloví**, **obrat** apod.) na rovině nevětných kombinací slovních, na rovině větné pak především **pří-sloví** či **pořekadlo**, popř. **pranostika**. Mezi větné (propoziční) frazémy však patří množství různých dalších typů, různé **formule** aj. Zvlášť stojí **přírovnání** (můžou být povahy kolokační i větné, viz 4.314) a **binomiály** (viz 4.316), které můžou zahrnovat i nefrazémy; done-dávna zcela ignorované jsou **frazémy synsémantické**, obecně většinou kolokační povahy (viz zvl. části 4.313 a 4.312). Značná část typů frazémů však žádné tradiční, jakkoliv nepřesné označení nemá, a odborný úzus je proto zpravidla vymezuje funkčně jako frazémy a idiomy té které roviny, popř. i podle formálního subtypu uvnitř každé z nich, např. *ukázat někomu paty* je takto kolokační frazém verbonominálního typu (podle struktury V-S).

**Prvky**, které do frazeologických kombinací jako **komponenty** formálně vstupují, jsou různé podle své členské roviny, kam patří (obecně viz dál, zvl. 3.22), zvl. morfémy (kořeny nebo afixy), lexémy, kolokace, klauze nebo propozice (věty), a zároveň se vyskytují zpravidla (kde je to relevantní) pouze v určitém tvaru, nikoliv podobě lemmatu. Výsledné kombinace, frazeologické celky, resp. frazémy jako **útvary** patří funkčně vždy na úroveň vyšší. Srov. např. **frazém lexikální přestřelit** a jeho komponenty *pře-střelit* (proti pravidelnému *pře-hodit* ap.), **frazém kolokační** a jeho komponentní lexikální tvary *mít strejčka* (proti pravidelnému *mít tetičku* ap.) nebo **frazém propoziční** a jeho komponenty *To je slovo do praxe* (proti pravidelnému *To je slovo do diskuze* ap.).

U **dichotomie frazém a idiom** (a také **frazeologie a idiomatika**) je vidět jak jasný rozdíl, tak určité překrývání. Obecně jde o rozdíl v hledisku; je proto vhodné ve snaze rozdílné hledisko zdůraznit udržet i rozdíl terminologický. V zásadě platí, že analyzujeme-li kombinatorický útvar formálně, z hlediska formálních rysů, mluvíme o **frazému**, zatímco analyza sémantická, z hlediska relevantních sémantických rysů (včetně nominativních), opravňuje užití názvu **idiom**. Tato jen zdánlivá dvoukolejnost není zbytečná: na jedné straně je metodologicky často nutné pro jejich nápadnou odlišnost si všímat zvlášť stránky sémantické nebo stránky formální, na druhou stranu nemají obě stránky stejný rozsah. Vzhledem ke klesajícímu stupni a počtu anomálií a při přechodu k periférii frazeologie a idiomatiky je tam sémantická stránka zpravidla silně oslabená, nejasná a pak převažuje, resp. prakticky zůstává pro analýzu pouze stránka formální a tedy i termín frazém. Taková je situace např.

u frazémů prepozičních typu *na rozdíl od*, **kvazifrazémů** verbonominálních typu *věnovat pozornost*, **termínů** jako je *zlatý déšť* aj. Z praktických důvodů se však poněkud zjednodušeně mluví často o všech jednotkách pouze jako o frazémech. Jejich formální typologie pochopitelně vychází z rovinné příslušnosti a typologie formálních aspektů komponentů.

Funkčně jsou kolokační frazémy ekvivalenty všech slovních druhů, přičemž početně nejbohatší jsou frazémy s funkcí verbální. Komponenty kolokačních frazémů čerpají ze všech slovních druhů (apelativ i proprií), které se ve své systémové platnosti však zpravidla nezapojují; jde pouze o jejich tvary. K nejnápadnějšímu, tradičnímu a zřejmě univerzálnímu typu (dá se zřejmě nalézt ve všech jazycích) patří z hlediska sémantického **frazémy somatické**, založené na názvech částí lidského těla, u kterých se výrazně projevuje **antropomorfický princip** jazyka, srov. *vzít si něco do hlavy*, *být někomu v patách* ap. (nejčastější komponent *ruka* se objevuje v několika stovkách frazémů). Řada somatických frazémů se užívá také k označení **gesta**, srov. *polkrčit nad něčím nos*, *mávnout nad něčím rukou* a ty mají pak i dvojí denotát a dva odlišné významy (u druhého frazému jde o „vlastní gesto“ vs. „nedbání, nezájem“). Těmto idiomům se někdy také říká **idiomy paralingvistické**, protože spojují jazyk (verbální) s nejazykem. Nicméně takový pohled na frazém skrze sémantický typ jeho komponentu je spíše impresionistický a nemá velkou hodnotu kromě studia metafory; obdobně se někdy mluví i o frazému **animálním, botanickém** ap.

Přes svou diskrétní formu je frazém a idiom funkčně i sémanticky integrálním **celkem**, který nelze synchronně dělit, analyzovat do komponentů; není jinými slovy prostou sumou významů svých komponentů, není ani založený na principu kompozičnosti významu. Všechny tradiční a opakující se pokusy o sémantickou analýzu (na rozdíl od analýzy diachronní) jsou nesmyslné, protože zůstávají vždy dílčí a opírají se o vnější kritéria, srov. frazém *leje jako z konve* (zhruba: „prší hodně“) a pokusy vyčleňovat z něj (aspoň) význam „prší“ pro komponent *leje*. V nedomyšlené „logice“ tohoto přístupu, který by měl být dováděný do konce, pak *jako z konve* by muselo znamenat „hodně“, což neznamená a jazyk toto spojení v tomto údajném smyslu neužívá (srov. *\*snědl toho jako z konve*, *\*spí jako z konve*). Takto je pak i lépe pochopitelné, proč jsou komponenty frazému pouze tvary formálně shodné, homonymní s tvary lexémů známých z lexikonu, a ne samotné lexémy; někdy ovšem v pravidelném jazyce takové předpokládané lemma neexistuje, jakkoliv se to povrchně může zdát, srov. (*nechat na*) *holičkách* a nesmyslné lemma *holička* (ani diachronně tu o „dámskou kadeřnici“ či „nedozrálé ovoce“ ap. přímo pochopitelně nikdy nešlo). Případy jako *leje jako z konve* je tedy třeba chápat jako zčásti homonymní a kvazimotivované (motivovanost). Většina frazémů však neumožňuje ani takto částečnou pseudoanalýzu. Tím je zároveň zřejmé i to, že frazém není svými komponenty přímo a měřitelně motivovaný.

IF je značně rozrůzněna i sociolingvisticky, v plném stylovém rozpětí dnešní češtiny od frazémů a idiomů argotických až po poetické. Protože není na rozdíl od zbytku lexikonu předmětem kodifikačního úsilí, které by frazeologii svým oblíbeným způsobem zjednodušeně deformovalo, vykazuje výraznou a přirozenou míru **variability** (variací) projevující se množstvím různě alternujících variant. Ty na rozdíl od zjednodušeného a umělého kodifikovaného jazyka odrážejí přirozený variantní, alternativní stav vývoje jazyka v této oblasti; pro svou výrazovou výstižnost je IF předmětem zvlášť nápadné snahy o tvorbu stále nových obměn, **variant autorských**, popř. nových frazémů a idiomů. Převážná většina IF je přiro-



ženě vytvářena a užívána však v jazyce mluveném, většinou nespisovném, protože právě v něm se nejvíce uplatňuje funkční potřeba druhu vyjádření, které frazémy a idiomy představují.

## 3.2 Forma frazémů a idiomů

### 3.21 Struktura frazému a idiomu

**Strukturu** frazému nelze interně z důvodů nemožnosti jeho objektivní analýzy a mnohostranné anomálie, jak jsme viděli, určit; místo toho se jí pak rozumí zpravidla jen vnější podoba frazému, založená na podobnosti jeho komponentových forem s tvary vnějšího, pravidelného jazyka. Takto se pro ni využívá analogicky, ale jen do dané míry, především slovnědruhová platnost komponentů, pokud je jasná. Takto stanovená struktura však o skutečné povaze daného frazému říká jen málo, má jen pomocnou úlohu. Tímto způsobem pojatá je struktura frazému pak vytvářena prototypickou a invariantní kombinací forem různých slovních druhů (včetně dalšího druhu monokolokabilního) v daném sledu (pořadí) spolu s explicitním relátorem (pokud je přítomný, tj. obvykle prepozicí nebo konjunkcí), který je spojuje (pokud není přítomný, je relace mezi nimi jen implicitní) a valencí. Srov. takto frazémy a zápis jejich struktury

*vzít nohy na rameny*

**V-S-prep-S**

*mít na někoho spadeno*

**V-na-V<sub>part</sub>**

*mít pré*

**V-mono**

*utíkat, {až se za ním prášilo}*

**V-Kl** (klauzi lze podle potřeby ovšem rozepsat).

Takto se konkrétněji např. kolokační frazém *mít z něčeho těžkou hlavu* (**V-A-Sprep**) formálně skládá ze tří komponentů (*mít, těžkou, hlavu*), jednoho **relačního prvku** (relátoru, z) vyjadřujícího explicitně vztah mezi nimi (který jindy může chybět) a **valenčního prvku** (valence, *něčeho*), realizovaného a konkretizovaného v kontextu; ten je v systému zastoupený jen obecně a potenciálně. Za komponenty lze považovat na této úrovni (kolokačních frazémů, největšího subtypu autosémantik) členy čtyř hlavních nominativních a autosémantických slovních druhů, tj. substantiva, adjektiva, verba a adverbia. V závislosti na své úrovni a subtypu se může proměňovat i povaha komponentů, srov. např. frazém kolokační složený z komponentů pouze synsémantických jako *co do* (**pron-prep**), *co když* (**pron-konj**), *že jo?* (**konj-int**), nebo kolokační z komponentů smíšených, jako je *od oka, pro tentokrát*, anebo frazém polypropoziční, složený z komponentů propozičních, resp. větných, srov. *Jak si kdo ustele, tak si lehne* (**PROP-PROP**) ap. (viz 4.31 a dál).

V některých případech je však vhodné uvažovat i o **substituci** autosémantických komponentů. Pokud frazémy tohoto subtypu obsahují pronomeny nebo numerale, považují se v nich podle své funkce za **substituty** substantiva či adjektiva, nikoliv za další komponenty, srov. *umět to s někým* (**V-pron-prep**), *páté kolo u vozu* (**num-SprepS**), viz 4.312 a 4.311. Substituty (zvl. substantiv a adjektiv) jsou číslovky a zájmena (num, pron), srov. *mít na to*, **V-prep-pron**, *dát si dvacet*, **V-num**. Poukaz na přítomnost substitutů je užitečný hlavně při posuzování obecné souvztažnosti kombinatorického typu, resp. struktury frazému s podtypy (substituční případ je pak podtypem) a při klasifikaci, srov. případ **V-prep-pron**, který je takto substitučním podtypem obecné struktury **V-prep-S**. Substi-

tut, relační či valenční prvek, může u frazému být přítomný nebo chybět (i když jsou obecně společné všem typům).

Příklady výše ilustrují všechny specifické idiomatické kombinace dvou až tří komponentů (nejčastější struktury), vytvářené různými tvary prvků, resp. slov slovníku; prvků, resp. komponentů idiomatické kombinace může být ovšem víc než dva, srov. *být někomu věrný až za hrob, neprodat svou kůži lacino* aj.

**Ustálenost** těchto kombinací, která umožňuje i jejich opakované užívání, tj. reprodukovatelnost, se projevuje v tom, že do kontextu vstupují jako pevný celek, syntakticky nečlenitelný, ale konkrétně i v tom, že mají zpravidla pevné pořadí, slovosled svých komponentů, srov. např. zpravidla obtížnou přijatelnost zkusmých obměn *\*těžkou hlavu z toho má, \*to s ním umí, \*u vozu kolo páte, \*do co, \*když co, \*jo že, \*Tak si lehne, jak si kdo ustele* atd. Nemožnost takovýchto slovnědruhových transformací je jednou z typických anomálií frazému.

## 3.2 Komponenty idiomu a frazému

V IF je **komponent** konstitutivním členem, složkou frazému v určitém tvaru a určité pozici ve struktuře jeho celku. Jakkoliv např. u kolokačního frazému složeného z autosémantik nemá vzhledem k celkové anomálii jeho struktury komponent jasnou slovnědruhovou funkci, navenek se tak zdánlivě jeví, srov. dva „autosémantické“ komponenty (*vzít, hůl*) a **relátor**, který je spojuje (zde předložka *na*, jindy spojka), ve frazému *vzít někoho na hůl* (přičemž *někoho* je **valenci**). U jiného subtypu může být komponent však i jiný „slovní druh“, např. předložka, interjekce či partikule nebo celá klauze, srov. *být pro, jemu je hej, až na, jen aby, utíkat, {až se za ním prášilo}*. Poslední příklad uvádí **interstratální frazém**, jehož komponenty pocházejí ze dvou rovin (strat) a u něhož neexistence neutrálního tvaru je naznačena ozávkováním {}. U **lexikálního frazému**, což jsou formálně jednoslovné lexémy, jsou vytvářeny komponenty z oblasti morfémů; v praxi v nich jde především o odvozeniny a kompozita (viz 4.33).

## 3.3 Význam a nominace

### 3.31 Povaha významu idiomu a frazému

Frazém je vždy specifické **pojmenování (nominace)**, který se výběrem a případnou kontextovou modifikací v textu jen aktualizuje, ale netvoří se jako pojmenování pravidelná. Jeho aktualizace záleží na aspektu zdůrazněném a potřebném v daném kontextu.

Nepřítomnost kompoziční funkce mezi významem komponentů a významem útvaru vede k **pojetí významu** idiomu jako významu integrálního, celistvého a nerozložitelného (viz i výše); odráží se v nemožnosti tento (pravidelně) nemotivovaný význam zpětně (a beze zbytku) segmentovat, dekomponovat na vstupní významy komponentů. Starší představa sémantického posunu zde situaci nevystihuje, případy typu *stará bela, růžové brýle, tlouct špačky* za posun považovat nelze; běžný posunutý význam má zřetelný a pravidelný vztah k tomu, z něhož vznikl, tady však ne. Ani případy typu *železná zásoba, mít těžkou hlavu*

nebo vyjít vstříc, vydat všanc nejsou založeny na nějakém posunu spočívajícím v metafoře a metonymii; liší se jak kolokabilitou, tak kompatibilitou komponentů aj. Stejně málo významný a výstižný je přístup přiblížit význam idiomů ve vágních pojmech přenesenosti, obraznosti, průhlednosti apod. (což jsou jen jiné názvy pro metaforu a metonymii).

Integrální význam idiomu je dán konjunkcí řady faktorů paradigmatické a syntagmatické povahy lišících se nejen proměnlivou mírou a způsobem své účasti, ale často i svou absencí. Třebaže anomální a nemotivovaná povaha významu (ve vztahu ke komponentům) znemožňuje jeho bezprostřední klasifikaci, povahu jeho komponentů vně frazému určovat lze. Své komponenty idiom čerpá z pravidelného jazyka, v němž platí objektivní a měřitelná pravidla, která se promítají i do klasifikace IF. Jen výjimečně je některý komponent idiomu sám idiomem, srov. *doprovodit někoho na poslední cestě* a *poslední cesta*; případy jako *sud prachu* jsou obv. sekundární transformací z širšího idiomu *sedět na sudu prachu*.

Ve srovnání s určitým jazykem se mnohý idiom jeví jako přímo nepřeložitelný, tj. nemající k sobě cizí ekvivalent. Tato kdysi zdůrazňovaná nepřeložitelnost idiomů má heuristickou cenu jen okrajovou a je tak relativní, že je obecně téměř bezcenná; ve vztahu ke všem cca 7–8000 existujících jazyků světa by pokaždé musela dávat, aspoň v teorii, jiný soubor nepřeložitelného.

Tam, kde pravidelný jazyk může vyjádřit např. radu ke střízlivosti a zdrženlivosti v úsudku až kombinací lexikálních prostředků (nespecifických) a prostředky suprasegmentálními, tam IF často nabízí již hotový a ekonomický výraz pro tuto potřebu dávno vytvořený a určený, srov. např. *Nevěř na duchy!* Z hlediska ustáleného **pojmenování** lze ve **vztahu idiomu k pravidelnému jazyku** obecně rozlišit tři případy. Idiomy

- (1) jsou monopolním a primárním pojmenováním (zpravidla i vysoce ekonomickým) a pravidelný jazyk je pro daný význam nevyužívá;
- (2) jsou pojmenováním paralelním, konkurujícím pravidelnému;
- (3) pro daný význam (a denotát) se nerealizují a existují tu jen pojmenování pravidelná.

Srov. příklady:

- (1) *vzít někoho za slovo, mít u někoho rozlitý ocet, mít máslo na hlavě, být hned na koni*,
- (2) *jít do sebe : zamyslet se nad sebou, mít nahnáno : bát se, dát si dvacet : zdřímnout si*;
- (3) *leptat, letovat, lyžovat* aj.

Typy významu zprostředkované idiomy v prvním případě nelze ani stručně a výstižně zachytit lexikograficky, natožpak jedním slovem. Ani druhý případ, obsahující zdánlivě ekvivalenty, není jednoduchý, vezmeme-li v úvahu všechny rysy pragmatické, kontextovou funkci apod. Paralelní existence pojmenování obou druhů je tu velmi volná a výrazy lze zaměňovat jen v některých (málo konkrétních a vágních) kontextech; nejde u nich tudíž o konkurenci, ale spíše o komplementární výskyt v textech téhož typu, podobně jako u některých synonym.

Třetí případ dobře ukazuje, že idiomatický výraz se zpravidla nekonstituuje v oblasti významu vázaného na konkrétní denotát. Hlavní doménou idiomatického je naopak oblast významů (co do denotátu) „nekonkrétních“ s výraznou pragmatickou složkou. Proto i s ohledem na (2) lze právě zde pro běžné významy uvedených typů obecně očekávat, nikoliv ovšem předurčovat, existenci idiomů.

Z takových příkladů vyplývá málo reflektovaná skutečnost, že ne vše v jazykové komunikaci nominativně obslouží standardní pravidelný jazyk monopolně a lépe než IF a že místo frazeologie v jazyce není vždy periferní. Integrace přístupů a metodologie užívané pro pravidelný jazyk a IF však dosud chybí a lingvistika pravidelného jazyka mívá někdy sklon v IF spatřovat především jen nežádoucí a dále nereflektovaný úhrn nepříjemných výjimek, pokud IF ovšem vůbec rozpozná; v tomto pohledu je jasné, že pokusy některých syntaxí o analýzu konkrétního frazému jsou absurdní a směšné.

S obecně vágní povahou sémantiky idiomu souvisí zákonitě i distribuce typů významu frazémů v jazyce. V pestrosti významových typů lze aspoň negativně říct toto: v těch sémantických a nominativních oblastech, které jsou pro svou určitost a exaktnost doménou pravidelného jazyka (nejen ve smyslu odborné terminologie), se IF v zásadě nevyskytuje. Pokud se tu s ní setkáme, jde opět o výrazné zvážnění a jen volnou paralelu významu pravidelného. Srov. zvláště vyjadřování určenosti, kvantifikace, lokálnosti a temporálnosti a např. idiomy *jeden každý, kdo měl nohy, ten tam utíkal; ani za mák; lítat po všech čertech; až naprší a uschne*.

### 3.32 Složky a aspekty významu frazému a idiomu

Význam idiomu je komplex vytvářený řadou rysů různé povahy. Proti významu pravidelného pojmenování téže roviny tu jsou důležité rozdíly. Vedle výše zmíněného rozdílu co do nemotivovanosti jde především o odlišný poměr zastoupení typů sémantických rysů v celku a o odlišný poměr mezi významem a smyslem. Z úhrnu sémantických rysů různého typu konstituujících obě základní složky významu, tj. význam **nocionální** (konceptuální) a **pragmatický** (konotativní, popř. asociativní), má obecně význam pragmatický, charakteristicky přítomný v celé IF, ve srovnání s poměry v pravidelném jazyce mnohem výraznější a častější zastoupení i v jednotlivých idiolech.

V rámci **pragmatických** aspektů, které jsou součástí sémantiky idiomu, pokud jsou stále a systémové, jsou nejnapadnější složky expresivní, symbolická a evaluativní (hodnotící), dané jednak vztahy k formě, jednak ke vstupním významům komponentů a jednak k obvyklé (pragmatické) funkci výrazu. Na rozdíl od významu nocionálního nelze pragmatickým typům významu obvykle přiřadit zřetelný sémantický rys jednoznačně, popř. vůbec.

**Expresivní** složka ustáleného významu, založená (1) **interně** na nápadné odchylce a kontrastu buď (části) formy nebo významu vůči formě nebo vůči významu základnímu, obvyklému, je projevem a signálem anomálního vztahu (mezi částmi formy nebo významem komponentů a celku apod.). Kromě toho i (2) **externě** samo užití idiomu v textu, kde idiom kontrastuje s jeho zbytkem, okolním kontextem, bývá expresivní; je obecné a častější. Srov. pro (1) expresivní motivaci formální v příkladech *být kulatoučký jako melounek* (zvukomalba a deminutiva), *přihřívát si svou polívčičku* (deminutivum), *mít popotahovačky s někým* (reduplikace prefixu), *dát si repete* (rytmičnost a prvek *e*), *mít rajc z něčeho* (nespisovnost komponentu) a sémantickou *mít pré* (monokolokabilita), *nechat na holičkách* (monokolokabilita a zdánlivá motivovanost komponentu), *šlápnout někomu na krk* (inkompatibilita a tabu), *chytat lelky* (absence sémantické participace) aj.

**Symbolickou** složku významu, opírající se o možnost asociace, si lze objasnit na rozdíl od dvou idiomů, jejichž nociónální význam je blízký, popř. synonymní. Umožňuje svou odlišnou asociací vhodné začlenění do zřetelně odlišného kontextu a tématu, srov. *byl hned na koni* (s významem „využít situační převahy“ a „jezdeckou“ složkou symbolickou, vedle evaluativní složky nadhledu, povýšenosti apod.) a *byla to voda na jeho mlejn* (s významem „situace se obrátila v jeho prospěch“ a symbolickou složkou „mlynářskou“, vedle nociónálního „vzchopení se“, hodnoceného jako prudké, popř. chtivé apod.).

**Evaluativní** složka významu, často doprovázející jeho složku kognitivní, je nejužší spjatá s jeho pragmatickou funkcí a zprostředkuje hodnotící stanovisko mluvčího na škále „dobrý-špatný“. Srov. trojici výrazů *letí jako vítr*, *žene se jako k ohni* a *utíká, jako by mu zadek zapálil*. I s odhlédnutím od poněkud odlišné intenzity a různého projevu významu komponentů lze pozorovat rozdíl v uзуálním hodnocení, které tyto výrazy zprostředkují. V prvním je evaluativní složka zřetelně kladná (mluvčí může takovým výrazem např. i chválit), v druhém je kolísavá a v třetím (mluvčí jím obvykle vyjadřuje vedle vlastního obsahu pojmenování i posměch a hanu) je záporná. V kontextu ovšem výraz běžně přebírá podle potřeby neustálená hodnocení další. Evaluace je zpravidla jen bipolární typu *dobrý-špatný* a lze ji členit podle protikladu *senzorický-intelektuální* (srov. *smrdí to jak bolavá noha* vs. *je chytřej jak vopice*) a *absolutní-relativní*. Zatímco evaluace absolutní je věcí obecně přijímané axiologie, je evaluace relativní nepřímá a daná, např. v přirovnáních, existencí momentálně přijímaného modelu a prototypu, srov. *má hlad jak vlk*. Jistý stupeň evaluativnosti má i řada funkčně gramatických frazémů, srov. *za přispění* (vždy pozitivní souvislost), *jen aby* (negativní pochybnost) aj.

Zvýrazněný podíl **pragmatických** složek významu, které všechny mívají vedle své ustálené a systémové části i část dodávanou kontextem (ta je již doménou smyslu a vlastní textové pragmatiky), mění i závislost idiomu na **kontextu**. Ta je zásadně větší než u pravidelných pojmenování. Idiom v uvedených a dalších složkách (zvl. v míře intenzity) z kontextu tedy výrazně čerpá. Zvlášť je tato závislost vysoká právě u idiomů propoziční povahy. Úzké sepětí výsledného významu (smyslu) idiomu s kontextem je však dáno ještě jinak, vysokou mírou **potenciálnosti a neurčitosti**, resp. **vágnosti** jeho významu. Idiomatický význam takto dovoluje poměrně snadno a často neobvyklou **aktualizaci**, tj. jeho posun nebo rozšíření. Srov. *dát někomu košem*, kde k ustáleným (nociónálním) složkám významu lze počítat odmítnutí nápadníka dívkou a nověji i odmítnutí nabídnutého pohoštění hostem; výraz však lze užít bez zvláštní invence a metafory i v řadě kontextů dalších, např. u nabízeného daru apod.

Na rozdíl od víceznačnosti je **vágnost** neurčitost, resp. neúplně specifikovaná významová určitost jazykové, textové formy, srov. *přijít/být jen tak*, *zaonačit to nějak*, *dát někomu co proto* aj. Tento denotativně neostrý dojem, projev jazykové formy, může být důsledkem úzu (data nemusejí být vždy ostře vymezená) nebo vědomého užití mluvčím, který buď nechce nebo se neumí vyjádřit přesněji, srov. *No von to Karel to... , von je takovej divnej, zatentovat, přitentovat*. Vágnost je v různé míře podstatnou vlastností většiny frazémů, ovšem už přímo v jejich systémové platnosti; tato jejich povaha, často z tohoto důvodu využívaná, vyvstává do popředí v kontrastu k přesnému, zvláště terminologickému vyjadřování.

Zvlášť výrazná je tato **potencialita** dalšího významu a úzu, daná neustáleností a splývavostí hranic významu ustáleného s neustáleným, u idiomů s monokolokabilním (a tedy nemotivujícím) komponentem, např. *být hin* aj.

Tento poslední rys významu má své důsledky i pro statut **polysémie** a **monosémie** idiomu; stanovit předěl mezi ustáleným významem, byť řídkým, a potenciálním významem neustáleným nelze vždy jednoznačně a není to dáno jen odlišným územ různými mluvčích. Avšak ani uvnitř oblasti významu ustáleného není určování jednotlivých významů (nociálních) vždy stejně spolehlivé. Oba výše zmiňované významy u

- (1) *dát někomu košem* tvoří polysémii jednoho typu (společný průnik „odmítnutí“), avšak významy
- (2) *být jak vosk* (materiál: „být měkký“; člověk: „být poddajný, povolný“) jsou jiné, odlišnější; výrazy typu
- (3) *dát někomu co proto, vzít to hopen* již ale za polysémii lze považovat jen obtížně.

Idiom tedy polysémní být může, ale míra jeho polysémie, která je sama více typů a závislá na povaze komponentů, je mnohem nižší než u jednoslovných pojmenování pravidelných. Závažná složka, kterou do významu podle konkrétních cílů komunikace vnáší kontext, i sama otevřenost významu takové kontextové specifikaci svědčí o výrazné a charakteristické vágnosti idiomu. Jí na vrub připadá i zdánlivá polysémie případů výše. Srov. např. zdánlivě samostatné významy „namlátit“ a „vynadat“ u *dát někomu co proto*, které jsou ve skutečnosti jen nejběžnějšími kontextovými konkretizacemi jakéhosi širokého „vytrestat“, které lze obohatit celou řadou běžných i méně běžných kontextových realizací dalších, např. „prohnat“ (v závodě), „udolat“ (např. množstvím jídla), „ztížit práci“ (např. nepromyšlenou úpravou) apod.

### 3.4 Varianty frazému a idiomu

Mimo jiné i díky tomu, že nejsou naštěstí preskriptivně kodifikovány, frazémy odrážejí proměnlivost svého vývoje věrně a nezkrusleně ve svých **variantách**. Lze odhadovat, že až každý čtvrtý frazém má nějakou formální variantu, srov. *být/stát na výši doby, nasadit někomu veš (do kožichu)*. Od variant je třeba lišit funkčně odlišné transformace frazémů (viz 6.). Varianta je taková ustálená či neustálená obdoba, tj. relativně malá formální nebo sémantická modifikace frazému, která nemění jeho funkci (viz transformace) a v zásadě ani podstatu významu, srov.

- (1) *do bílého rána/dne, prohnat si hlavu kulí/prohnat si kuli hlavou,*
- (2) *dostat (jich) nepočítanejch, nasadit (vlastní) krk.*

Od variant systémových a ustálených se liší **varianty neustálené, autorské**. Co do druhu je variantnost velmi pestrá a zahrnuje (často v kombinaci) jevy lexikální, morfologické i syntaktické, popř. i fonologické. V přehledu jde především o tyto typy variant (označené závorkou a lomítkem):

#### I syntagmatické

##### A *komponentní*

##### 1 kvantitativní

**a** aditivní: *přinést někomu něco (až) k nosu, nasadit (vlastní) krk*

**b** redukční: *dostat (jich) nepočítaných*

**2** kvalitativní

**a** kontaktní (zvl. inverze): *anděl strážný/strážný anděl, sypat si hlavu popelem/sypat si popel na hlavu*

**b** distanční (oddělitelnost): přechází do gramatických transformací (6.)

**B relační** (vnější, valenční)**1** explicitní

**a** synsémantické (zvl. předložkové): *ohrnovat nos (nad něčím)*

**b** flektivní (zvl. pádové): *měřit (někomu) dvojím metrem*

**2** implicitní: *být na místě (někde)***II paradigmatické**

**A komponentní** (obv. synonymní): *do bílého rána/dne, být s to/šton, slzavé údolí/údolí slz* (přechází již do transformací)

**B relační** (vnitřní): přecházejí již do transformací

Někdy je určování jedné varianty z více jako základní a reprezentativní, a tedy invariantu, obtížné. **Invariant** je v zásadě bezpříznakový tvar jak neutrální (z nejrozličnějších hledisek), tak i nejfrekventovanější a slouží mj. jako nejvhodnější východisko jednak k posouzení variant (druhých ustálených tvarů výrazu) a jednak k tvorbě transformací. Tyto požadavky splňuje v češtině běžně infinitiv, nominativ, popř. pozitiv. Nemá-li je frazém, pak je třeba za invariant považovat tvar nejčastější (např. 3. sg. prez. ind. apod.), srov. *Abys to nezakřik(l)!* (2. osoba sg. tu zastupuje všechny ostatní) nebo případ rozvinuté struktury **V-S** (tj. **V-KL**, verbum a klauze) *říct, co /mu/ slina na jazyk přinese*. Srov. ještě *být rád, že /je/ rád*. Klauzi kombinující se se slovesem v první pozici nelze ovšem zcela zobecnit (notace /mu/, /je/ je zástupná), přesto je zřejmé, že tu jde o výraz vždy jediný, a nikoliv množinu několika desítek ustálených vět lišících se jen jednotlivými morfologickými kategoriemi.

**3.5 Sémiotika frazému a idiomu**

Frazémy jsou sémioticky různé, jejich povaha, obvykle složená ze dvou složek, je však dána jejich tvarem a jeho funkcí, ne sémiotickou povahou jejich komponentů, která tu může být odlišná. Velmi rozšířené **pragmatické** aspekty frazému jako znaku jsou většinou dány jeho **evaluativností**; jeho denotativní stránka je rozložená mezi sekundární ikony, indexy a symboly dost nerovnoměrně. Zásadní rozdíl je mezi **denotáty** (symboly) konkrétní a abstraktní povahy, kudy vede i obvyklá hranice mezi standardními frazémy a kvazifrazémy (srov. *prásknout do bot* a *zout si boty*), daná i nízkou mírou anomálnosti v druhém případě; většina frazémů kolokačních má denotát takto abstraktní povahy. Výrazná sémiotická korelace mezi typy frazémů a jejich znakovou povahou se nenabízí, i když jsou k ní náležitě, srov. funkční podtyp relačních frazémů **prep-S-prep** (v *souvislosti s*) a blízkost indexálním frazémům. Srov. dále i sekundární frazeologický ikon *kamenný obličej*, index *nějaký ten* a většinový, resp. nejčastější typ mezi frazémy, totiž symbol *otevřená hlava, hodit něco za hlavu*. Podstatné je, že denotativně mohou komponenty těchto frazémů ještě před vstupem do nich být nejen jiné povahy (srov. indexální *ten* v symbolu *být ten tam*), ale nemusí mít denotát vůbec žádný (u monokolokabilních prvků, srov. *pré* v *mít pré*). Ani u propozič-

ních frazémů, jejichž sémiotická povaha je prozkoumána ještě méně, však přímočará klasifikace možná není. I Peirceova propozice se hodí na propoziční frazém (např. *To se říká kdoví jestli!*, *To by v tom byl čert, aby...!*), mj. pro její ustálenost, jen volně (sekundární propozice) ap.

### 3.6 Kvazifrazémy a kvaziidiomy

Každý frazém a idiom lze vždy komplexně vymezit více kritérii zároveň, především stupněm jeho anomálie. Neostrost hranic IF a pozvolný přechod do pravidelného jazyka vede k tomu, že u některých typů klesá i relevantní počet kritérií a aspektů anomálnosti výrazu. Periférii IF, tj. mj. pásmo přechodu do pravidelného jazyka, tvoří oblast kvazifrazémů a kvaziidiomů. Vedle zapojení jednotlivých kritérií je pro posouzení centrálnosti či perifernosti výrazu v tomto smyslu třeba brát v úvahu i míru jeho idiomatičnosti (viz 4.3, 5.3) a frekvenci.

Hlavní typy kvazifrazémů a kvaziidiomů naznačuje přehled:

- a) výrazy s konkrétním denotátem: *kočičí hlava*
- b) výrazy terminologické povahy (je tu exaktní definice významu): *mrtvý chod*
- c) některé výrazy se synsémantickými komponenty: *ve srovnání s někým* (mají specifickou jen svou kolokabilitu, význam a funkci)
- d) některé výrazy s desémantizovaným komponentem: *bejt napakovanej* („mít peníze“), *bejt potřebnej* („potřebovat peníze“) (adjektivum je možné jen v predikátové funkci)
- e) některé výrazy s omezenou kolokabilitou komponentů: *zarytě mlčet*, *otevřít dokořán*, *širé moře*, *poulit oči*
- f) některé výrazy verbonominální s abstraktním substantivem: *dávat pozor*, *věnovat pozornost* (omezená kolokabilita abstrakta) aj.

## 4. Systém idiomatiky a frazeologie a klasifikace

IF tvoří nominativně systém ustálených kombinací různého typu a různé funkce na více rovinách, který je komplementární k nominacím pravidelným, ale i neustáleným. Pokusíme-li se formálně struktury jednotlivých typů IF shrnout, vidíme i formálně složitý **systém**, který v pravidelném jazyce není.

### 4.1 Distribuce idiomů a frazémů v jazyce

IF příslušností svých jednotek, frazémů a idiomů, do více rovin do jisté míry naznačuje i svou vnitřní hierarchizaci. Skutečnost, že výsledný frazém patří zpravidla do roviny vyšší než komponenty, které do něj kombinatoricky vstupují, vystihuje schéma (+ označuje výskyt, kladné zastoupení):



<i>Rovina komponentů → výrazu ↓</i>	<i>A morfemická</i>	<i>B lexikální</i>	<i>C kolokační</i>	<i>D propoziční</i>
1 morfemická				
2 lexikální	+			
3 kolokační	+	+		
4 propoziční (větná)	+	+	+	
5 polypropoziční monosubjektová			+	+
6 polypropoziční intersubjektová				+

Podle této vnější povahy komponentů lze dospět k nejobecnějšímu rozvrstvení a tak i k **klasifikaci** frazémů a idiomů, u kterého se vychází z příslušnosti komponentů do stejné roviny (strata), srov.

**2A frazém a idiom lexikální** (*najít, předsevzetí*)

**3A frazém a idiom kolokační** v různém smyslu (*ke všemu, co do, co když*, se synsémantickými komponenty)

**3B** (*chytat lelky, jablko sváru*, s autosémantickými komponenty)

**3A–B** (*k polednímu, od oka, bez okolků*, s komponenty obojího druhu)

**4A frazém a idiom propoziční** (*To tak!, Že by?*, se synsémantickými komponenty)

**4B** (*Chcíp tam pes., Pozdě bycha honit*, s autosémantickými komponenty)

**4C** (*To je pro něj španělskou vesnicí*, s kolokačními komponenty)

**5C frazém a idiom polypropoziční monosubjektový** (*Mladý kurvy, starý modlářky*, s kolokačními komponenty)

**5D** (*Vono se řekne semenec, ale zobat ho z lahve!*, s propozičními komponenty)

**6D intersubjektový** (A: *Pozdrav pánbu!* – B: *Dejž to pánbu!*, s propozičními komponenty u různých mluvčích).

Tento přehled nezachycuje řídký případ **frazému a idiomu interstratálního**, ve kterém se kombinují komponenty různých rovin, srov. např. *dát někomu, co se do něj vejde* (lexém + propozice, resp. klauze), *tvářit se, jako by se nechumelilo* aj. Oblast IF tudíž zahrnuje všechny relevantní typy ustálených anomálních kombinací s pevnou funkcí a významem. Jestliže proto IF nemůže na jedné straně zahrnout kvůli absenci významu kombinace fonémových komponentů, není na druhé straně jasné, co vše lze sem zahrnout z úrovní vyšších ustálených kombinací, tj. nad frazémy polypropozičními, které už patří do literárních mikrožánrů (např. hádanky, bajky apod.). V mladé historii IF, tradičně se zprvu soustřeďující především na oblast kolokačních frazémů, se však zvláště z nedostatku poznání do IF celé rozsáhlé oblasti kombinací někdy nezačleňovaly, zvláště frazémy a idiomy propoziční i lexikální. V takových přístupech se pak u takto vágně vymežované IF ani neprojevovala snaha vystihnout tyto oblasti jinak a nechávaly se zcela bez povšimnutí, popř. vysvětlení.

Počet českých frazémů a idiomů kolokačních a propozičních lze odhadovat na třicet tisíc; není to tudíž oblast malá či zanedbatelná. Není to však ani oblast v jazyce okrajová

vá; vrství se podle frekvence užití na výrazy s nejvyšší frekvencí, patřící v češtině k samotnému *centru jazyka* (zvl. gramatické povahy jako *že jo, co když, to je jedno*), zatímco jiné jsou užívané středně často a konečně další jsou řídké a patří na jazykovou *periferii* (srov. *rozsvítit jednu lampu, nepadnout někomu do noty*). Právě zde lze najít jak početné výrazy zastarávající a klišé, tak některé neologismy a výrazy módní (srov. *To je o něčem jiném.*).

## 4.2 Klasifikace idiomů a frazémů

Hledisek, podle kterých lze frazémy členit, je více. Některá jsou téměř bezcenná pro nejasnost své podstaty (jako dříve přeceňovaná klasifikace stylistická), přílišnou obecnost (klasifikace sémiotická, znaková), nedostatek spolehlivých pramenů a informací (klasifikace diachronní a etymologická) nebo náročnost a neujasněnost kritérií (klasifikace sémantická, resp. onomaziologická, která neexistuje ani v lexikonu). Významné třídy členění frazémů lze však získat na základě klasifikace jejich funkcí (ta dosud ale chybí, srov. však 6.). Zbývá tedy klasifikace formálně-strukturní, založená na evidenci vstupních komponentů, které sice o frazému mnoho nevypovídají, navenek tu však sledované komponenty jsou většinou povědomé (s výjimkou monokolokabilních slov) a vnějškově shodné se slovy pravidelného jazyka.

## 4.3 Formální třídy idiomů a frazémů

Třídy frazémů a idiomů se nejjasněji vyčleňují na základě svých formálních struktur, jejichž úhrn nabízí pohled na jejich strukturní klasifikaci. V dalším se klasifikace IF omezí na frazémy kolokační (4.31), propoziční (4.32) a lexikální (4.33) z hlediska jejich struktur; ty občas doprovodí i poznámky o jejich funkci.

### 4.31 Třídy kolokačních idiomů a frazémů

Patří sem ty frazémy a idiomy (kolokace), jejichž komponenty jsou slova, resp. určité slovní formy, a které samy o sobě ještě netvoří promluvu, větu. Růzností struktur, jevů i počtem výrazů je to nejvýraznější oblast IF.

#### 4.311 Autosémantika v úloze komponentů

Substantiva, adjektiva, verba a adverbia dávají celkem šestnáct dvoumístných variací s opakováním (tj. 4<sup>2</sup>), které představují základní binární struktury tohoto typu frazému. IF tu, na rozdíl od přirozeného jazyka (který to neumožňuje), realizuje všechny teoreticky možné struktury, tedy všech 16. Některé z nich patří nebo přecházejí do propozic (jsou označeny ! a zde se uvádějí jen pro úplnost), jiné co do pořadí komponentů splývají (zvl. V-ADV s ADV-V, srov. však pouze *bohatě stačit* v tomto pořadí). K řídce realizovaným a početně málo obsazeným strukturám patří S-ADV a ADV-S, popř. i S-A a A-A, nejbohatší a uvnitř dále členěné jsou struktury V-S, A-S a S-S. Srov. stručný základní přehled

1 **S-V:** *pravda vítězí*

2 **V-S:** *číst někomu levity, tlouct špačky, chytat lelky*

3 **A-S:** *zlatý důl, černá hodinka, šťastnou cestu, nejvyšší měrou*

- 4 **S-A**: *anděl strážný, bohem zapomenutý, !panenko skákavá, !paličák paličatý*  
 5 **S-ADV**: *místo nahoře, !klobouk dolů*  
 6 **ADV-S**: *vzhůru nohama, !mnoho štěstí*  
 7 **V-A**: *kápnout božskou, oplatit někomu stejnou*  
 8 **!A-V**: *zmýlená neplatí*  
 9 **V-ADV**: *přijít zkrátka, dopadnout bledě, držet se stranou*  
 10 **ADV-V**: *bohatě stačit, trpce litovat, zarytě mlčet*  
 11 **A-ADV**: *černej vzadu, !celej nakřivo, (!) malá domů*  
 12 **ADV-A**: *neblaze proslulý, andělsky trpělivý, zbrusu nový*  
 13 **S-S**: *jablko sváru, roh hojnosti, den brannosti, palma vítězství, obratem ruky*  
 14 **V-V**: *umět se narodit, dát se poddat, zůstat někde viset*  
 15 **A-A**: *chudý příbuzný, neslaný nemastný, staří mladí*  
 16 **ADV-ADV**: *široko daleko, kolem dokola*

V tomto přístupu se sleduje slovnědruhová povaha prvků a nebere se v úvahu jejich sekundární a heterogenní povaha syntaktická; proto se i případy jako *božská, stejná* apod. chápou formálně za adjektiva. Syntaktický přístup by tu znamenal určitou interpretaci a rozklad na komponenty, který obecně u idiomu možný není.

**Základní kalkul** (binární) se dále člení více směry, především na případy účasti vlastních komponentů a **(a) substitutů** (zvl. zájmen a číslovek), vyjádření **relace mezi nimi** implicitně nebo **(b) explicitně** (předložkou a spojkou), na struktury binární nebo **(c) subtypy o více komponentech** (tj. druh strukturních transformací) aj. Po stránce sémantické je z řady možností zvláště výrazné odlišení substantiv abstraktních ve struktuře **V-S** (viz 4.315). Srov. např.

- 1 (a) **Pron-V**: *!ta píše, !to bodne*, **Num-S**: *dvaatřicet loupežníků*,  
 (c) **S-ADV-V**: *všeho moc škodí*, **Pron-S-V**: *co dům/týden dal*  
 2 (a) **V-Pron**: *zakřiknout to, být bez sebe*, **V-Num**: *olíznout všech deset, dostat pětadvacet*  
 (b) **VprepS**: *mít za ušima, lézt po čtyřech, být s to*, **Vprep(S)prepS**: *jít na něco od lesa*  
 (c) **V-A-S**: *mít krytá záda*, **VprepA-S**: *žít na vysoké noze*, **VprepS-S**: *mít z ostudy kabát*  
 3 (a) **Pron-S**: *moje maličkost, svého času, každým coulem*, **Num-S**: *jedním vrzem, jedna báseň, pár kroků*  
 (b) **AprepS**: *zralý pro šibenici*, **prepA-S**: *za velkou louží, v pravý čas*, **Pron-A-S**: *moje drahá polovička*, **prep-Pron-S**: *za každou cenu, za našich časů*  
 (c) **A-S-A**: *všemi mastmi mazaný*, **A-A-S**: *zlatá střední cesta*  
 4 (a) **Pron-A**: *nic kloudného*  
 (b) **prepS-A**: *za vlasy přitažený, po čertech chytrý*, **SprepA**: *obrat k lepšímu*  
 5 (b) **prepS-ADV**: *s hubou dokořán*  
 6 (a) **ADV-Pron**: *zhola nic*, **Num-S**: *pár kroků*  
 (b) **ADVprepS**: *jednou za čas*  
 7 (a) **V-Pron**: *být nsvůj, vědět své*  
 (b) **V-Aprep**: *být krátký na někoho*, **VprepA**: *tnout do živého*  
 (c) **V-S-A**: *mít to polepený*, **V-ADV-A**: *být ještě bosý*, **V-AprepA**: *plést páte přes deváté*  
 9 (c) **VprepADV**: *pít přes moc*, **Vprep(S)-ADV**: *jít od něčeho dál*, **V-ADV-ADV**: *nevědět kudy kam*  
 13 (a) **Pron-S**: *moje maličkost, něco na zub, kterej čert?*

- (b) **SprepS**: *pohled pro bohy, místo na slunci, zmatek nad zmatek*, **prepS-prep-S**: *ve dne v noci* **S-SprepS**: *boj muže proti muži*, **Pron-prep-S**: *ta s kosou, něco na zub*, **prepS-S-prep-S-S**: *od slunka východu do slunka západu*
- (c) **S-S-A**: *vlk v rouše beránčím*, **S-prep-A-S**: *agent s teplou vodou*
- 14 (c) **VprepS-V**: *umět v tom chodit*
- 15 (a) **Num-A**: *první poslední*
- (b) **AkonA**: *živ a zdrav*
- 16 (a) **ADV-ADV**: *tak nějak*
- (b) **ADVkonADV**: *tak či onak, zkrátka a dobře, dříve či později, pusto a prázdno*
- part-ADV-ADV**: *ani dost málo, jen tak tak*
- (c) **ADVkonADV-ADV**: *víc než kdy jindy*

Sémanticky a funkčně jsou kombinace autosémantik velmi různorodé (viz zvl. 5. a 4.321).

#### 4.312 *Synsémantika v úloze komponentů*

Pro jednoduchost jsou k vlastním synsémantikům (předložky, spojky, částice) připojeny slovní druhy, které jejich platnost mají jen v některých ohledech (zájmena, číslovky), nebo ji nemají vůbec (citoslovce). Formálně jsou sem dále zařazena i jinak zcela zvláštní **monokolokabilní slova**, u nichž je v různé míře oslabena nejen slovnědruhovná příslušnost, ale i sám statut slov; jsou to pouhé ekvivalenty slov, verboidy (a zde tedy vystupují jako jakýsi mimořádný jedenáctý slovní druh). Avšak ani klasifikace běžných synsémantik (zvláště partikulí) tu není vždy jednoznačně možná, mj. i vzhledem k jejich polyfunkčnosti v pravidelném jazyce, srov. např. částicovou povahu *ale* v *ale co!* O substituci (a) zde jednoznačně uvažovat nelze, některé případy na její možnost ale ukazují (srov. **pron-pron** *nějaký ten* a **A-A**), jiné však nikoliv (srov. *co do*). Podobně je zpravidla nemožné uvažovat o (b) explicitním vyjádření relace mezi komponenty; subtypy (c) jsou u některých typů (zvláště partikulí) však běžné. Z plného kalkulu  $7^2$  variací s opakováním (49) se realizují asi dvě třetiny případů. Komponenty jsou v přehledu uspořádány a zkracovány jako 1 **prep** (prepozice), 2 **kon** (konjunkce), 3 **pron** (pronomen), 4 **num** (numerale), 5 **int** (interjekce), 6 **part** (partikule), 7 **mono** (monokolokabilní prvek). Srov. např.

- 1A **prep-pron**: *mimo jiné, ke všemu, podle mého, mezi námi, beze všeho (!), podle toho!*
- (c) **prep-pron-pron**: *k tomu všemu, z ničeho nic, prep-pron-prep-pron*: *pro nic za nic*
- 1B **prep-num**: *do jednoho*
- 1C **prep-int**: *do alelujá*
- 1D **prep-part-mono**: *na to tata*
- 1E **prep-mono**: *na just, na blind, na ex*
- 1F **prep-prep**: (b) *(všechna) pro a proti*
- 2A **kon-pron**: *jako takový, a to*
- 2B **kon-part**: *či spíše, či co*
- (c) **kon-part-part-part**: *když už tak už (!)*
- 2B **kon-kon**: *byť i, a sice, i když*
- 3A **pron-prep**: *co do něčeho, co za někoho*
- (c) **pron-prep-pron**: *sám od sebe, co na tom!*
- 3B **pron-kon**: *co kdyby (!)*
- 3C **pron-part**: *jaképak copak!*

- 3D **pron-mono:** *tě pic/péro!*
- 3E **pron-pron:** *nějaký ten, ta jeho, ten který, ten samý, jaký – takový, co to?*  
(b) *ten a ten*  
(c) *mě druhé já*
- 4A **num-pron:** *jeden každý*  
(b) *jeden a týž*
- 4B **num-num:** *jedna dvě*  
(c) **prep-num-num:** *o sto šest; num-prep-num:* *jeden za druhým*
- 5A **int-pron:** *bác ho!, ale co!*
- 5B **int-part:** *ale ano!, ale kdepak!, no a?*
- 6B **part-kon:** *ještě že, kdoví jestli*  
(c) **part-kon-part-kon:** *a když už-tak*
- 6C **part-pron:** *no tohle!, no proto!, kdoví co/kdo*  
(c) **part-(S)-pron:** *copak (Karel), ten...!*
- 6D **part-int:** *no maucta!, a basta!, no jéje!*
- 6E **part-mono:** *ani muk (!)*
- 6F **part-part:** *což teprve, jen aby (!), jen jestli (!), to zrovna!, že by (?), že ano?, aby ne!, kdyby tak*  
(b) *ano i ne*  
(c) **part-part-part:** *taky že jo!, bode(j)t' by ne!, part-part-kon-part:* *tak tak že ne-*
- 7A **mono-kon:** *kór když*
- 7B **mono-pron:** *lautre nic*
- 7C **mono-mono:** *jakýs takýs, jakž takž, semo tamo, volky nevolky aj.*

Funkčně jednoznačné **číslovky** se jako komponenty frazémů účastní nejméně (význam jejich komponentů většinou přetrvává), ani ty však nestojí překvapivě zcela mimo; slabší je i účast citoslovců. Nejvíce jsou tu zastoupeny **částice**. Charakteristickým rysem některých frazémů je větší či menší míra oddělitelnosti jejich komponentů (distančnost), srov. *když už* (0/to udělá/...), *tak* (špatně aj.); *copak* (Karel), *ten* (všechno zná!); *kdo* (to četl/...), *ten* (ví aj.). Zvláštní případy představuje účast vázaného morfému (prefixu), srov. (každého) *hned tak ne-* (pozve aj.), *co možná nej-* (větší aj.). Řada případů funguje jen jako propoziční frazémy (označeny jsou pomocí !, popř. ?), srov. *taky že jo!*, *co to?*; některé jsou však homonymní s kolokačními frazémy, srov. *Jen jestli!* a *Jen jestli* (přijde včas)! a mohou se lišit i významově. Funkce a význam frazémů tohoto typu frazémů jsou velmi různé, zůstávají však většinou v rámci funkcí a významů tradičních synsémantik a slov jim sem přiřazených, popř. též adverbů. Pochopitelně tento rámec však výrazně rozšiřují a pozměňují.

#### 4.313 Autosémantika a synsémantika v úloze komponentů

Smíšené kombinace, resp. variace autosémantik a synsémantik (a dalších typů) v roli komponentů dávají teoreticky 2 x (4 x 7) struktur; v IF se z nich realizuje necelá polovina. Substituce se tu nevyskytují; i případy s explicitně vyjádřenou relací mezi komponenty a vícečlenné subtypy jsou řídké. Srov. např.

- 1A **prep-S:** (1) *bez okolků, od oka, s bídou, do omrzení, o chlup* (+ komparativ), *proti všemu očekávání, za boha ne-, ani za mák*

(2) *s ohledem na něco, v souvislosti s něčím, na úkor něčeho, v intencích něčeho, v intencích někoho*

(3) *s ohledem na to, že/co... v souvislosti s tím, že/co*

1B **prep-A**: *k polednímu, za mlada, při nejmenším, až na další*

1C **prep-ADV**: *pro jednou, pro tentokrát*

2A **kon-S**: *co chvíli*

2B **kon-A**: *co živ*

2C **kon-V**: *co nevidět*

3A **pron-ADV**: *co naplat!, nic moc (!), co a jak*

4A **int-S**: *ejhle člověk!*

4B **int-ADV**: *šup sem (šup tam) (!)*

5A **part-S**: *až hanba, ani nápad!, jen klid!, už ani slovo!*

5B **part-A**: *až černo*

5C **part-V**: *ani hnout!, jen říct!, ani nezabučet!, jen kvete!*

5D **part-ADV**: *až moc, taky dobře!, jak kdy/kde!*

6A **mono-S**: *baj vočko, kýho výra!, basama s fousama!*

6B **mono-A**: *fungl/zbrusu nový*

6C **mono-V**: *seč byl!*

7A **V-prep**: *být pro, stát mimo*

7B **V-pron**: (viz 4.311)

7C **V-num**: (viz 4.311)

7D **V-int**: *dělat na někoho bububu*

7E **V-part**: *mít něčeho ažaž, dostat co proto*

7F **V-mono**: *být tentam, být hin, být auf, být nasnadě, být zajedno s někým, být (moc) hr do něčeho*

8A **S-mono**: *mocí mermo, křížem krážem*

8B **A-mono**: *opilý namol*

9A **ADV-prep**: *až na někoho*

9B **ADV-pron**: *sotva kdo, jak kdo/co, zhola nic* (viz 4.311) *a kdesi cosi; dost na tom že*

9C **ADV-part**: *jak to?*

9D **ADV-mono**: *skrz naskrz*

Struktury s adjektivem nebo substantivem v první pozici se v IF často nerealizují; jde tu mj. o jejich předložkovou valenci. V důsledku funkčního překrývání jsou kombinace z celé této oblasti početně celkově okrajové; vymyká se z nich poměrně bohatá struktura **prep-S**. Dva její subtypy (2–3) slouží k vytváření **víceslovných předložek** (*s ohledem na něco*) a částí **víceslovných spojek** (*s ohledem na to, že/kdo/co...*). Řada struktur patří na rovinu propozic (srov. označení !).

#### 4.314 Přirovnání

Povaha přirovnání je v hierarchii jednotlivých oblastí IF výjimečná. Jako útvar má ustálené a formalizované přirovnání většinou funkci kolokace. Jeho komponentem (stojícím za komparátorem) však může být slovo, kolokace i klauze (věta), srov. např. II A(d) níže. Jako typ jde přirovnání tudíž do jisté míry napříč různými rovinami IF. Formalizované **přirovnání** je *idiom, který explicitně označuje vztah podobnosti mezi referentem, kontextově dopl-*

ňovaným, a modelem, který je součástí tohoto idiomu. Je tedy založeno na druhu sémantické iterace a blíží se některým typům binomiálů. Přirovnání je ovšem po stránce formální více typů; podobnost se v nich může vyjadřovat i implicitně, popř. i neustáleným způsobem, srov. lexikální idiom *sněhobílý* (tj. *jako sníh*), kolokaci *holubičí povaha* (*jako holubice*), *brčálově zelený* (*jako brčál*) a např. neustálenou kombinaci *mluví nepřátelsky* (*jako nepřítel*) apod. (Čermák in SČFI1, 1984).

Sémantická struktura přirovnání je **/Kd/-R-(Tk)-k-Kt**,  
kde označuje

**/Kd/ comparandum**, tj. referent (obv. subjektivní valence) určený konkrétním kontextem, který je celým výrazem ožřejmován, např. */Karel/ je zarostlý jako Ezau*;

**R relátor**, tj. komponent relační povahy (obvykle verbum), který idiomu dodává povahu predikátu, např. */Karel/ je jako Ezau, /kluk/ poslouchá jako hodiny*;

**(Tk) tertium comparationis**, tj. znak, rys, vlastnost aj., který Kd a Kt mají společný; jen výběrově označuje podobnostní vztah explicitně, např. */Karel/ je 0 jako slepý*; někdy splývá s R, např. */Karel/ řve jako lev*;

**k komparátor**, tj. formální znak (obv. *jak/jako*) signalizující obecně podobnostní vztah, který je v Tk konkretizovaný;

**Kt comparatum**, tj. zobecněný a obvykle známý model, prototyp, k němuž je kontextem daný referent (Kd) co do své podobnosti vztažen a takto ožřejmován; může mít povahu entity, akce, situace aj., např. */hrnec/ je děravý jako cedník*, */Karel/ byl červený, jako by ho krví polil*; *to je pěšky jako za vozem* aj.

**Komparandum** (Kd) je zhruba v 80 % případů člověk, popř. živočich (přirovnání specifická jen pro živočichy čeština však nemá), zbytek připadá na věci, abstraktní pojmy, např. místa, srov. *je toho jako nastláno*, *být kyselý jako šťovík*. **Komparatem** (Kt), zdrojem podobnosti, jsou naopak asi z 50 % konkrétní předměty, z 20 % živočichové aj., srov. *být studený jako led*, *mít hlad jako vlk* apod. Přirovnání je tudíž na pravé straně (Kt) výrazně konkrétní povahy, orientace na člověka na jeho levé straně (Kd) prozrazuje silnou antropomorfizaci celé jeho oblasti; ta je charakteristická však pro celou IF. Vyjadřovaná podobnost je téměř vždy doprovázena různě silným pragmatickým vztahem evaluativním, srov. *je to kyselý jako šťovík* (tj. nepřijemně), a často i důrazem, jehož vyjádření může odsunout do pozadí i vlastní podobnost, srov. *uteklo to jako voda* (tj. velmi rychle).

Typy formálních struktur přirovnání jsou do přehledu zařazeny dva, (I) kolokace a (II) propozice; okrajové struktury, jejichž vysoká variabilita je daná velkým počtem základních kombinatorických prvků, ho však rozšiřují na několik desítek možností. Vedle (a) základních struktur tu jsou naznačeny i (b) některé substituce, (c) explicitní vyjádření relace mezi komponenty a (d) vícekomponentové subtypy. Konvenční zkratky označují čtyři zapojená autosémantika, **k** označuje komparátor (*jak/jako*); klauze se tu pro jednoduchost zvlášť neoznačuje.

- |                  |  |
|------------------|--|
| <b>I 1A Vkv:</b> | (a) <i>lže, jako když tiskne</i> , (d) <i>řve, jako by ho na nože brali</i>  |
| <b>1B Vka:</b>   | (a) <i>být jako slepý</i> , (c) <i>jednat s někým jako rovný s rovným</i> , (d) <i>byl červený, jako by ho krví polil</i>  |
| <b>1C Vks:</b>   | (a) <i>poslouchat jako hodiny</i> , <i>pít jako duha</i> , (b) <i>dýchat jako čtyři</i> ,<br>(c) <i>přihnát se jako na koni</i> (d) <i>být jako pěst na oko</i> , <i>být jako otevřená kniha</i> |

- 1A **VkADV:** (a) *mít se jako doma*, (b–d) *dělat něco jako nikdy v životě*  
 2A **V-AkV:** (a) 0, (d) *je hubený, jako by cvrčky louskal*  
 2B **V-AkS:** (a) *být lehký jako pírko*, (c–d) *mají se rádi jako kočka s myší*  
 3A **V-SkS:** (a) *mít krk jako holínku*, (c) *mít nohy jako z olova*, (d) *mít hlavu jako vrabčí hnízdo*

- II 4A **S-VkV:** (a) 0, (b) *je jich jak naseto*, (d) *to je, jako když dá hluchému dobrýtro, to je prašť jako uhod'*  
 4B **S-VkA:** (a) *den je jako malovaný*  
 4C **S-VkS:** (a) *je to jako broky, čas utíká jako voda*, (b) *uteklo to jako nic*, (c) *je to jako na potvoru* (d) *leze to z něj jak z chlupatý deky, je toho jako z apatyky*  
 4D **S-VkADV:** (a) *to je jako naschvál*  
 5A **(S)-V-ADVkV:** (a) *je tam jako nastláno*  
 5B **(S)-V-ADVkS:** (a) *je to tam zatažené jako pytel*, (c) *to je pěšky jako za vozem*  
 6A **S-V-AkV:** (a) *voda je teplá jako kafe aj.*

Z přehledu je zřejmé, že většina přirovnání je budována ve své levé straně na verbu (popř. adjektivu) a v pravé straně na substantivu (popř. deverbativním adjektivu). Typickým přirovnáním je proto v tomto smyslu formálně např. *sedět jako pecka*, *jít jako svázaný*. Výrazy typu *chlap jako hora*, *den jako malovaný*, *lehký jako pírko* lze chápat za různé ustálené transformace z bázi obsahujících sloveso (zde všude *být*).

#### 4.315 Verbonominální kvazifrazémy

Početný subtyp struktury **V-S** je vytvářen účastí abstraktního substantiva (tj. bez konkrétního denotátu a označujícího zhruba vlastnost, stav, situaci, děj, událost, činnost, instanci, jev, pocit, konstrukt ap.), vhodněji tedy **V-S<sub>abstr.</sub>**. Např.

*věnovat pozornost*, *klást důraz*, *uvést do pohybu*, *tvořit skupinu*, *způsobit škodu*, *mít většinu*, *doznat změny*, *nabýt paměti* aj.

Tyto frazémy, které patří už zpravidla mezi kvazifrazémy a kvaziidiomy, jsou vysoce příznačné specifickou distribucí (kolokabilitou) slovesného komponentu, který je plně podřízen povaze substantiva (Čermák 1974). Sloves svým substantivem takto predeterminovaných však bývá zpravidla několik málo (v průměru 6–8). Protože jde o slovesnou skupinu pevně strukturovanou, která se u téhož typu substantiv opakuje pravidelně, lze ji chápat za soubor transformací (agenciální a statutární, viz 6.3). Hlavní rozdíly jsou ve vyjádření tří aspektů **fáze**, resp. **stavu** (*inchoativní*, *durativní* a *terminativní*, zkratky INCH, DUR, TERM) a přítomností či nepřítomností **kauzativy** (KAUZ). Plná struktura, resp. makrostruktura sdružující více těchto frazémů dohromady, je tedy **INCH-DUR-TERM**, která může být zdvojená kauzativitou, naznačující přítomnost druhého agentu, srov. např. typy

- 1 **INCH** *nabýt paměť*, **DUR** *mít paměť*, **TERM** *pozbyt/ztratit paměť* (a podobně i *klid*, *síla*, *zájem* aj.)  
 2 **INCH** *dostat hlad*, **DUR** *mít hlad*, **TERM** *ztratit hlad*, **TERM-KAUZ** *utišit/zahnat hlad* (a podobně i *chuť*, *žízeň* aj.)



3 **INCH** *projevit/nabýt zájem*, **DUR** *mít zájem*, **TERM** *ztratit/pozbýt zájem* aj.  
**INCH-KAUZ** *vyvolat zájem*, **DUR-KAUZ** *udržet/udržovat zájem*, **TERM-KAUZ** *odradit/uspokojit zájem*

V ideálním případě se dospívá k šestimístné kombinaci a makrostruktuře; ani trojmístná konfigurace však nebývá vždy realizována plně (což je dáno povahou významu substantiva), popř. je některá její fáze řídká, srov. *mít paměť*. Některé případy se větví do subtypů, a jsou-li ustálené, posouvají se tak k centru IF, srov. *mít krátkou paměť*, *mít dobré základy* aj. Do těchto vztahů vstupuje naprostá většina abstraktních substantiv a tyto makrostruktury představují jejich nutné a základní chování v textu. Abstraktní substantiva uvedená výše v závorkách (tj. *klid*, *síla*, *zájem*, *žízeň*, *chuť* aj.) naznačují, že se chovají analogicky a přijímají stejnou trojdomou makrostrukturu také. Podle stupně jemnosti rozlišení lze takových základních makrostruktur pro všechna česká abstrakta najít asi dvacet. Je třeba zdůraznit, že každá jednotlivá ustálená kombinace (např. *mít hlad*, *vyvolat zájem*) je samostatným kvazifrazémem, který je však třeba skrze jeho některou a zde jen naznačenou konfiguraci vnímat v jasné souvislosti s ostatními.

U tohoto typu frazému, resp. kvazifrazému převládá zřetelně stránka kolokační, charakteristicky omezená jak co do rozsahu, tak kvality; jiné stránky jsou tu okrajové, srov. např. sémantika substantivního komponentu či monokolokabilní komponenty v *kout pikle* aj.

#### 4.316 Binomiály

Binomiály jsou velmi staré struktury, rozšířené v mnoha jazycích, srov. české *ze dne na den*, *den co den*, *z ručky do ručky*, *sakum prásk*, *jednou jedenkrát*, *tak či onak*, *Spejbl a Hurvínek*, srov. ale např. i ekvivalentní (čes., něm., fr., angl., švéd., rus.) **pes a kočka**, *Hund und Katze*, *chien et chat*, (*fight like*) *cat and dog*, *hund och katt*, *кошка и собака*.

**Binomiál** lze charakterizovat především jako kolokační frazém či idiom, který je složený ze sledu dvou komponentů téhož slovního druhu, někdy i identické formy, má obvykle nominální povahu a je založený na iteraci třídné nebo individuální (zde pak je někdy dokonce založený na stejném základu). Svou povahou se opírá o koordinaci či juxtapozici a má obvykle neměnný slovosled. Součástí binomiálů můžou být a bývají i propria. Na druhou stranu je však síla iterativního modelu natolik silná, že se uplatňuje i v tvorbě kombinací, které jdou už za rámec frazémů, srov. *bratři a sestry*, *doba a místo*, *jídlo a pití*. Jakkoliv je binární a původně nominální struktura největší, její síla vede k tomu, že se analogické binární struktury tvoří do jisté míry i dále, s jinou slovnědruhovou funkcí, srov. **formální** struktury

##### 1 – Struktury substantivní

**S-S:** *hlava nehlava* (mono), *kniha knih*, *letem světem*, *sliby chyby*, *škoda přeškoda*

**S-prep-S:** *den za dnem*, *děšť se sněhem*, *marnost nad marnost*, *půjčka na/za oplátku ruku v ruce*

**S-kon-S:** *alfa i omega*, *cukr a bič*, *den co den*, *děs a hrůza*, *krev a mlíko*, *muž a žena*, *nebe a dudy*, *Petr nebo Pavel*, *Romeo a Julie* aj.

##### 2 – Struktury adjektivní

**A-A:** *neslaný nemastný*, *zdravý nemocný*

**A-kon-A:** *nečesaný a nemytý, širší než delší, takový nebo makový, vážení a milí*

### 3 – *Struktury verbální*

**V-V:** *chtě nechť, vstávaje lehaje*

**V-kon-V:** *být či nebýt, čekat a doufat, číst a psát, jíst a pít, rozdělit a panuj*

### 4 – *Struktury adverbální*

**ADV-ADV:** *skrz naskrz, široko daleko, zprava zleva*

**ADV-kon-ADV:** *pusto a prázdno, pomalu ale jistě, snadno a rychle, tady a teď, tu a tam, zas a zas*

Vedle substantivní je toto druhá největší struktura, zatímco struktury založené na adjektivu a verbu jsou mnohem menší. Existují však řídce i jiné, srov.

### 5 – *Struktury další*

**prep-Pron-prep-Pron:** *pro nic za nic, pro mne za mne*

**Pron-prep-Pron:** *ten či onen, kdo s koho*

**prep –Pron-Pron:** *z ničeho nic*

**Num-prep-Num:** *páté přes deváté*

**Prep-kon-Prep:** *pro a proti*

**kon-kon:** *bud' anebo*

**Int-Int:** *tik tak*

**mono-mono:** *jakýs takýs, jakž takž, hógo fógo, láry fáry, saky paky, semo tamo, šupito presto, techtle mechtle*

U binomiálů se uplatňují mezi oběma iterovanými komponenty různé **druhy sémantických vztahů** (zvl. statické jako je opozice, synonymie, komplementárnost, srov. *kočka a myš, děs a hrůza, déšť se sněhem*, ale i dynamické jako konverze, implikace či alternativnost aj., srov. *vzestup a pád, rozdělit a panuj!, peníze nebo život*), které se opírají o vztahy formální (zvl. heteronymie, homonymie nebo aliterace a rýmu, srov. *chléb a hry, ze dne na den, kniha knih, kost a kůže, jakýs takýs*) aj.

**Sémanticky** se tu však na úrovni významu celku vyskytuje hlavně tento poměrně sevrěný okruh významů:

**A – Identita:** *jeden jediný, šalba a klam*

**B – Totalita:** *od kolébky do hromu, široko daleko, chléb a hry, Jeníček a Mařenka, tělem i duší, v zimě v létě*

**C – Diference, alternativa, volba:** *peníze nebo život, Kain a Ábel, kdo s koho, oko za oko*

**D – Indiferentnost, popř. nutnost:** *pán nepán, prašit jako uhod'*

**E – Intenzifikace a důraz:** *ze dne na den, poprvé a naposled, dál a dál, běda přeběda, kniha knih, hlava na hlavě, smůla na smůlu, stát ve státě*

**F – Bezprostřednost, těsnost:** *tváří v tvář, bok po boku, ruku v ruce*

## 4.32 Třídy propozičních frazémů a idiomů

Základní rozdíl tu je v množství účastněných mluvčích a tedy promluv; frazémy s jedním subjektem (mluvčím) lze pracovně seskupit dohromady, i když se podle své složenosti mohou dělit dále; jde o frazémy a idiomy propoziční a polypropoziční monosubjektové (4.321). Proti nim stojí frazémy polypropoziční intersubjektové a tedy s dvěma mluvčími (4.322).

Propozičním frazémům, které tvoří uzavřenou a ustálenou větu, se v minulosti pozornost žádná nevěnovala. Určitou výjimkou byla přísloví, která ovšem tradičně studovala etnografie, a ne lingvistika. Některé typy těchto frazémů se časem začaly různě nazývat, ale názvy jako *adagium*, *aforismus*, *anekdota*, *apothegma*, *axióm*, *citát*, *floskule*, *formule*, *fráze*, *gnóma*, *hádaná*, *heslo*, *hlavolam*, *hříčka*, *kalambúr*, *loci communes*, *maxim(a)*, *motto*, *obrat*, *okřídlená slova*, *parémie*, *pořekadlo*, *poučka*, *pozdrav*, *pranostika*, *průpovídka*, *příslloví*, *rčení*, *rozpočítávací*, *říkadlo*, *říkánka*, *sentence*, *slogan*, *spojení slov*, *truismus*, *úřloví*, *wellerismus*, *zaklínadlo*, *zařikávací*, *zásada* se překrývají, byly a jsou co do svých hranic i podstaty často nejasné, nedefinované; proto je třeba je v zásadě většinou odmítnout.

#### 4.321 Frazémy a idiomy propoziční a polypropoziční monosubjektové

**Frazémy a idiomy propoziční**, které jsou formálně již hotovou propozicí (promluvou), jsou vytvářeny kombinací komponentů z roviny lexikální nebo kolokační (A), zatímco **frazémy a idiomy polypropoziční** jsou vytvářeny z komponentů z roviny propoziční (B, přítomnost aspoň dvou propozic). Oba typy mají společného **jediného mluvčího**, resp. pevný subjekt (což nemusí být totéž). Řada případů však splývá, srov.

**A:** *Ani kuře zadarmo nehrabe., Do třetice všeho dobrého., Na Nový rok o slepičí krok., Cti otce svého i matku svou., Leze mu tlustý do tenkého., Za zeptání nic nedáš., Chcíp tu pes., Zlatý voči!, Dobré pořizení!, Čelem vzad!, To si piš!, To zrovna!, Až po vás!, Ke všem čertům!, Jak to tak vypadá... aj.*

**B:** *Jak si kdo ustele, tak si lehne., Vrána k vráně sedá, rovný rovného si hledá., Komu dal pámbu úřad, tomu dal i rozum., Když to nejde po dobrém, půjde to po zlém., Ať je to ten nebo ten, vyženem ho z kola ven., Kdyby blbost kvetla, budeš samý poupě., Prodej písničku a kup si notu., Povídali, že mu hráli!, Nech to koňovi, ten má větší hlavu., Koukej, abys viděl!, Kde se vzal, tu se vzal... aj.*

#### 4.3211 Formální struktura a sémantika propozičních frazémů a idiomů

Tyto frazémy jsou stavěny z různých **komponentů** pocházejících ze všech slovních druhů včetně proprií a slov monokolokabilních, srov. např. ... *fuk*, *fujtajbl*, *fušku*, *fáry*, *glanc*, *go*, *grešli*, *grog*, *habaděj*, *hapruje*, *hecu*, *hej*, *hejsa*, *hlásku*, *hoj*, *hokus*, *hop*, *hopsa*, *hot*, *housti*, *hráchy*, *hrana*, *hrklo*, *hromský*, *huj*, *hýn*, *chtění*, *ježkovy*, *jímá*, *jobova*, *jonáci*, *just*, *kalné*, *kalup*, *kápě*, *kaput*, *karbid*, *kázáníčka*, *kejhák*, *klapačku*, *klempíruje*, *klíďo*, *kloudnýho*, *knap*... Součástí komponentů však zde mohou být i klauze (v asi 9 %), srov. *Jak si kdo ustele, tak si lehne*. Ve velké míře se těchto frazémů účastní i verbální komponenty (celkem asi 85 %) a protože je podoba věty zde pevná, je možné ji do určité míry i charakterizovat: dané sloveso tu v téměř polovině případů stojí v prezentu v 3. osobě singuláru. Významné je zjištění, že většina těchto frazémů má podobu věty oznamovací (přes 70 %), zatímco necelá třetina je oznamovací či zvolací a jen malá část (cca 5 %) má podobu otázky.

Struktura propozičních frazémů buduje především na transformacích základního syntagmatu **S-V**, které však je samo přímo využíváno jen málo, srov. přísloví *Samochvála smrdí*. Pro tyto frazémy je charakteristická i typická **intonace**, zvláště v podobě koncové kadence (v sedmi obecných typech), srov. např. *Co tě nemá!* s kadencí zvolací klesavě-stoupavou od

posledního slova. K **základním formálním strukturám** lze počítat věty různě dlouhé, v počtu slov od 2–14; nejčastější struktury mají 4 nebo 5 slov (cca 19 % a 16 %). Základních struktur je tu deset; mají však různou frekvenci:

- 1 To je S/A/ADV/mono/prep-S:** *To je vyloučeno., To je k smíchu.*  
a var. *Tím to hasne., V tom je jádro pudla., To se říká kdoví jestli.*
- 2 S-V:** *Věc se obrátila., Budou křtiny., Pravda vítězí., Čas běží., Než by řek švec., Není pochyb.*  
a var. *Jen kvete., Není zač., Ani hnout!, Ani co by se za nehet vešlo., Zdař bůh!, Ba co dím...*
- 3 S-V-S:** *Úsilí bylo korunováno úspěchem., Pálí ho dobré bydlo., To/všechno chce svý., Něco je ve vzduchu.*  
a var. *Kápní božskou!, Kdo by si pomyslel, že...!*
- 4 S-V-A:** *Vzduch je čistej., Je to s ním všelijaký., Přišlo mu to blbý.*
- 5 S-V-ADV:** *To hodně/něco pamatuje.* a var. *de mu to jedním uchem tam a druhým ven.*
- 6 ADV-V/mono-V:** *Pozdě bycha honit.*
- 7 S-S:** *Učedník mučedník., Mladost radost.*  
(okrajové) a var. *Důvěra za důvěru., Pro dobrotu na žebroutu., Můj dům můj hrad., A eště tu vo Červený Karkulce., Dobrou chut!, Chyba lávky!, Panna nebo vorel?*
- 8 prepS-A-S/S-A/S-ADV/A-ADV aj.:** *Po účinku zlá rada., Samá voda., Malá domů., Dobrýho pomálu.*  
a var. *Starý zlatý časy!, Husa jedna pitomá!, Klobouk dolů!, Hlavu vzhůru!, Starou Blažkovou!, Houby s voctem!, Čelem vzad!, Čerta starýho!, Všechna čest!, Ve jménu zákona!, S dovolením!, Hanba mluvit!*
- 9 ADV-ADV:** *Zkrátka a dobře., Jak kdy., Jako doma!*
- 10 syn-syn:** *To tak!, Ale co! Kdo s koho., No a?, No ne!, Že by?, Ani za nic!, To zrovna!, Láry fáry!, Hej rup!, To jistě!, Na ex!, No těbůh!*

**Sémanticky** jsou i tyto frazémy specifickým typem pojmenování. Označují především (1) lidské projevy, aktivity, postoje a situace, často různě kvalifikované (*Blázen, kdo dává, ještě věšší, kdo nebere., Když Pánbůh dopustí, i motyka spustí.*), (2) lidské fyzické stavy, vlastnosti a schopnosti (*Člověk na ní může oči nechat*), (3) předměty a věci, výrobky ap. (*To něco pamatuje.*), (4) obecné pojmy, události a situace v oblasti času, zvyků, zákona ap. (*Jak se do lesa volá, tak se z lesa ozývá.*) a (5) vnější přírodu a prostřední včetně počasí (*To je psí čas!*). Jen řídce se týkají něčeho jiného, např. zvířat. Častá je, zvláště u přísloví, složka významu gnómičského (*Člověk je sám sobě špatným rádcem.*). Z dalších rysů je nápadný aspekt hravosti či žertovnosti (*Když je něčeho moc, tak je toho příliš*), někdy spojený s rýmem (*Sedni si, ať ti kapsa nevisí*) nebo jazykolamem (*V kapli plakal pan kaplan*) či jazykovou hříčkou.

Oba typy propozičních frazémů jsou charakteristické relativně nízkým, popř. nulovým stupněm paradigmatické obměnitelnosti (transformovatelnosti) a **neexistencí neutrálního tvaru**; ten splývá s jejich (nejčastější) textovou podobou. Pozitivně jsou dány existencí pevného subjektu (pokud ho takový frazém má), který nepodléhá transformacím. Možnost paradigmatické záměny subjektu ovšem řadí danou kombinaci mezi **kolokace** (verbální), srov. *bylo to* (ale i: udělat to/strčit to do kapsy/...) *dílem okamžiku*, tj. tedy *být dílem okamžiku*, proti **propozičním frazémům** *nedá/nedalo se svítit, pálí/pálilo ho/jí... dobré bydlo* aj. Ana-

logicky i případy typu *neví, kde mu hlava stojí* je pak nutno považovat za kolokace, tj. *nevě-dět, kde /mu/ hlava stojí* (pro jejich formu tu je třeba užít metajazykového zápisu). Množství relací mezi komponenty (chápaných zčásti i logicky, zvl. u přísloví) je vyšší než u kolokací a v důsledku toho je tu i iterace řidší, srov. *já pán, ty pán; vem, kde vem; rozděl a panuj, zab se a uteč* (srov. binomiály) aj. **Monokolokabilní** prvky jsou tu zastoupeny stejně jako v celé IF, srov. *je po něm veta, do třetice všeho dobrého, na viděnou/slyšenou! trefil ho šlak, nic naplat, a basta fidli, ani muk!, šťastné pořizení!* aj.

#### 4.3212 Funkční třídy propozičních frazémů a idiomů

V závislosti na své strukturní funkci lze obojí frazémy formálně lišit na **funkční** komunikativní **třídy**

- 1 – **vokativní** vyjadřující reakce a postoje mluvčího a volní vliv na posluchače:  
*Vida ho!, To mi může být ukradený!, Koukej mazat!, Mě vynech!, Trhni si nohou!*
- 2 – **kontakto**vý vyjadřující vztah mezi mluvčím a posluchačem, kdy vlastní kontakt má často podobu pozdravů:  
*Zdař bůh!, Tak se měj!, Dobrou chuť!, S tebou sem domluvil!, To ti teda povím...*
- 3 – **tematický** vyjadřující referenční funkci mluvčího ve vztahu k tématu (tomu, čeho se komunikace týká), zvláště ve smyslu podtypů:
  - a – **deskriptivní** (jednoduché tvrzení nebo výrok a popis): *Vzduch je čistý., Zatmělo se mu před očima.*
  - b – **postojový** (atitudinální, jako výraz evaluace a stanoviska): *To zrovna!, Opatrnosti nikdy nezbývá.*
  - c – **konektivní** (vazby mezi částmi a fázemi tématu, řidké): *Bylo, nebylo..., Jak se věci mají...*
  - d – **kategorizační** (ztotožnění se známým typem či prototypem): *To je slovo do praxe.*
- 4 – **metajazykový** vyjadřující občas aktivovaný vztah mluvčího ke kódu, prostředku komunikace, a tedy jazyku:  
*Mám to na jazyku., Na tom si zlomíš jazyk., Co jde šeptem, to jde s čertem.*

Do kontaktového typu funkce patří i frazeologické vyjadřování **zdvořilosti**, které obvykle jako doprovodný rys prostupuje mnoha typy propozičních frazémů. Srov. však i specifitější případy *Budte tak hodnej a...!, Nechte si chutnat!, Není zač., Chovejte se jako doma!* a odpovídající funkce zdvořilého souhlasu, žádosti, přání, odpovědi na děk, pozvání a vybidnutí. Zdvořilost přitom může být i přímým obsahem třeba konkrétního přísloví, srov. *S kloboukem v ruce projdeš celý svět.*

Korelace typu významu, resp. funkce a formy frazému, je většinou nejednoznačná a jen u některých případů lze pozorovat jasnější vazby. Takový je například vztah deskriptivních výrazů zvláště v tematickém typu (a ve faktuální třídě) a indikativní formy. Podobně jde o souvztažnost zvolání, která patří do typu vokativního, a imperativní formy apod.

#### 4.3213 Pragmatická funkce propozičních frazémů a idiomů

Z řady ohledů je však u propozičních frazémů často významnější **pragmatická funkce**, která může mít především pět základních podob. Srov.

**1 Faktuální třída frazémů (F):** mluvčí s ohledem na posluchače se tu vztahuje, popř. fakticky, věcně zavazuje k pravdivosti propozice (faktu) a jejím různým stupňům a odstínům, popř. modifikacím. V praxi sem patří to, čemu se běžně říká **informace**, jakkoliv je velmi různého druhu. K řadě podtříd, které ji vytvářejí, patří frazémy vyjadřující (a) tvrzení, možnost/nutnost, ne/určitost, existenci, (b) ne/jistotu, ne/pravdivost, odhad, (c) ne/pravděpodobnost, předpověď, (d) ne/závislost, ne/podmíněnost, (e) omezení/rozšíření/změnu platnosti, všeobecnost, (f) kategorizaci, zařazení do třídy, typologii.

Např. *To je stará písnička.* (Ff), *Hlad je nejlepší kuchař.* (Fa), *V hlavě mu už svítá.* (Fa), *To je vod tebe sice hezký, ale...* (Ff), *Straší mu to v hlavě.* (Fa/Exi), *Hlava ho pro to nebolí.* (Fc/Exg), ... *a je po ftákách,* (Fa-e/Emc), *To je v háji.* (Fa/Emc), *To dá fušku.* (Fa/Emf), *Bůh ví.* (Fb), *Věřte si tomu nebo ne.* (Fc), *Člověk je to, co jí.* (Fd).

Extenzí této třídy je **metajazykový** subtyp, srov. *Mám to na jazyku* (Fa) (srov. metajazykovou funkci výše). K charakteristickému přechodu mezi typy patří tradiční **příslíví**, která pro svou gnómičnost, normativnost a nabádavost patří především k voluntativnímu typu, která však můžou být užita zároveň i jako faktuální, srov. *Hlad je nejlepší kuchař.* (Fa).

**2 Voluntativní (direktivní) třída frazémů (V):** mluvčí jejich užitím chce vyvolat u posluchače nějaký účinek, popř. ho k něčemu přimět; role mluvčího tu bývá aktivní, autoritativní. Zahrnuje řadu možných podtříd, jako jsou (a) rozkaz, příkaz, žádost, (b) ne/souhlas, přijetí, spokojenost, nechut, dovolení, odmítnutí, připomínka, upozornění, (c) hrozba, varování, napomenutí, snaha odradit, (d) ujištění, slib, uklidnění, konejšení, povzbuzení, (e) lhostejnost, rezignovanost, (f) rada, výzva, (g) přání, (h) přesvědčování, doporučení, poučení, (i) podpora, pobídnutí, uklidnění.

Např. *Abys to nezakřik!* (Vc), *Kolik hlav, tolik rozumu.* (Vf), *Buď rád, že seš rád.* (Vd), *Nevěš hlavu!* (Vd), *Di do háje!* (Vb/Emc), *To je hloupý.* (Vd/Fa), *Když je hlava blbá, trpí celý tělo.* (Fa/Vh/Emh), *Ještě to tak!* (Vb/Emc), *Dej si facku, nikdo se nekouká!* (Vb/Emc), *Pro mně za mně!* (Ve), *Táhni k čertům!* (Vab), *Dej bůh!/Bože dej!* (Vg), *Jen si dej, seš bledej!* (Vi).

Užití frazémů je často funkčně, resp. pragmaticky vícere i proto, že jinou roli ve vzájemném vztahu zaujímá mluvčí a jinou posluchač. Méně zřetelnou pragmatickou povahu mají široce pojaté **pozdravy**, které lze chápat za extenzi především voluntativního typu a elipticky za přání, srov. *Poroučím se., Tě pic!*, pozdravy ritualizované pro určité situace a specializovaný úzus, srov. *Čest práci!, K počtě zbraň!* U **rituálů** a jejich výrazů je zesílená složka respektu k rituálu, organizaci, církvi apod. a závaznost tradice. Zvláště v této oblasti se mj. překrývá s podtypy zde zastoupenými vyjadřování zdvořilosti (viz výše).

**3 Expresivní třída frazémů (Ex):** mluvčí jimi sděluje posluchači především intelektuálním, racionálním postojem své zaujetí situací a výraz tohoto zaujetí apod; role posluchače je tu někdy pasivní. K jejím podtřídám lze počítat zvláště (a) gratulace, (b) díky, (c) omluvu, (d) pochybnost, (e) sympatie, uznání, nelibost, odpor, znechucení, pohrdání, (f) důvěru/nedůvěru, (g) obvinění, výčitku, odsudek, (h) prominutí, prosbu o prominutí, (i) žertování, (j) bezradnost (vyvedenost z míry), vyhybavost.

Např. *Vsadím na to hlavu, že...* (Exf/V), *Jen aby!* (Exd), *Klobouk dolů!* (Exa), *Hádej, můžeš třikrát.* (Va/Exi-j), *To je dlouhá historie.* (Fa/Exj), *To se říká kdoví jestli!* (Exd), *Vloudila se chybička.* (Exc), *Máš recht!* (Exe), *Sem ti dobrej jen k tomu, abych...* (Exg).

Jistou extenzi tohoto funkčního typu lze spatřovat například v **odpovědi** na některé typy výrazů, zvláště na **děk**, srov. *Není zač*. Další extenzi představuje vyjádření **zdvořilého přání**, obvykle v podobě konvenčního pozdravu, srov. *Dobrou chuť!*, *Všechno nejlepsi!*, *Tisíceré díky!* (Exb), *Když dovolíte...* (Exh).

**4 Emocionální třída frazémů (Em):** mluvčí frazémem sděluje svou emocionální reakci vůči posluchači, která bývá výrazně evaluativně polarizovaná na škále *dobrý-špatný*; její podtřídy zahrnují zvláště (a) překvapení, údiv, (b) obdiv, chválu, pohrdání, (c) potěšení, radost, odpor, nechuť, (d) závist, (e) uspokojení, zklamání, lítost, (f) obavu, strach, klid, důvěru, (g) zlost, pomstychtivost, (h) výsměch, posměch, sarkasmus.

Např. *Krev mu ztuhla v žilách.* (Emf/F), *Srdce se na to směje.* (Ema-d), *Dobře, že deš.* (Eme), *To je ale kus!* (Emb), *Krve by se v něm nedořezal.* (Emf/Va), *Jen houšť a větší kapky!* (Emc), *Ten má ale časy!* (Emd), *To jsem si dal!* (Eme), *To sme šlápli do hovna.* (Eme), *Kat aby to spral!* (Emg), *Boha jeho!* (Emg), *Neví, co chce, ale nedá pokoj, dokud to nedostane.* (Emh).

Do této třídy patří i frazeologické **nadávky a kletby**, které se výrazně liší podle osoby; zpravidla se nejsilněji nadává vůči přítomnému posluchači, zatímco kletby (dnes už většinou bez religiozního zabarvení či obsahu) bývají neosobní nebo orientované na třetí osobu apod., srov. *Ty kluku pitomá!* (Emc), *Kat aby to spral!* (Emg). Vlastních nadávek je však ve frazeologii poměrně málo. Z pohledu teorie mluvních aktů patří některé ze subtypů do aktu perlokučního.

**5 Deklarativní třída frazémů (D):** mluvčí užitím frazému nastoluje, konstituuje novou situaci, stav, který je takto deklarovaný. Je to třída okrajová a její podtřídy nejsou zcela jasné, srov. však aspoň (a) propuštění, exkomunikaci, (b) jmenování, ustanovení, (c) odsouzení, (d) realizaci, nastolení.

Např. *Hej rup!* (Va/Dd), *To je hanebnost!* (Dc), *Dělají-li dva totéž, není to totéž.* (Dd), *At je po tvým!* (Db).

Lze jistě uvažovat i o **(6) intenzifikační třídě (I)** ... až zrak přechází (I), ... co to dá (I), kde se však frazémy v čisté podobě samostatně vyskytují zřídka. Naproti tomu je častá **(7) evaluativní třída (Ev)**, která ale také sama obvykle v čisté podobě nevystupuje, jen se kombinuje s jinou. Takto frazém *To je věc!* (Em/Ev) vyjadřuje jak obdiv, tak pozitivní evaluaci apod.

#### 4.3214 Některé tradiční typy propozičních frazémů a idiomů

Některé z typů propozičních frazémů jsou známy pod **tradičními nálepkami či jmény**. Nejznámější je starý, tradiční a většinou etnografický pojem **příslaví** (dříve též v Evropě **parémie** a později i **proverbia**, Erasmus je však řadil pod **adagia**), patřící mezi větné, tj. propoziční frazémy; disciplína, která se jimi tradičně zabývá, se nazývá **paremiologie**. Jejich třídění bylo v minulosti jen orientační a dnes ho nelze brát vážně. V celém rozsahu propozičních frazémů nejsou nejdůležitější, zaujímají jen jeho malou část (necelých 10 %), jakkoliv jsou i historicky zřejmě nejznámější. Proti dobám Čelakovského, největší sbírce českých přísloví, je cca 40 % přísloví nových, pozdějších, většina přísloví nasbíraných Čelakovským se však už dnes neužívá. Starší a obecná povaha přísloví pokrývající dříve všechny běžné oblasti života a zkušenosti a pravdy z nich odvozené (srov. přísloví „medicinské“ *Kam nechodí slunce, tam chodí lékař.* „obchodní“ *Zadarmo ani kuře nehra-*

be., „vojenské“ *Nejlepší obranou je útok.* anebo přísloví „právnícké“ *Oko za oko, zub za zub.*) se nověji vyvinula a vyhranila tak, že se tematicky už nechápou. Jediným zbytkem přísloví o počasí (meteorologická přísloví) jsou zvlášť nazývané **pranostiky**, kterých se stále několik desítek ještě užívá a lze je chápat, stejně jako vlastní přísloví, jako podtyp s faktuální funkcí (srov. *V máji nemá pastýřovi oschnout hůl.*, *Sv. Martin přijel na bílém koni.*). Jejich dodnes „předpovídací“ funkci napovídá sám název, který je v češtině už zakrytý (srov. přes latinské *prognostica* původní řecké základové *prognosis*, obsažené ve slově *prognóza*). Základní autoritativní povaha pravdy a zkušenosti, o které se nepochybuje, u přísloví však zůstává i po ztrátě legální normativnosti; odtud jejich rys, který se někdy nazývá nadosobní platnost, popř. didaktická funkce apod. Sémioticky a funkčně je u přísloví vždy třeba počítat s dvěma aktanty přímými (zhruba: „poučující/nabádající“ a „poučovaný/nabádaný“ apod.), v pozadí je však vždy přítomný anonymní aktant třetí, autor pronášeného výroku, pravdy, poučení apod., jehož porušení vede ke společenské sankci. Podobnou povahu však má často i leckterý citát (viz dál), tam je ovšem onen třetí aktant známý.

**Hesla** mají voluntativní funkci přesvědčování a přes svou typickou ad hoc povahu, podmiňnou dobovým územ a (především politickými) preferencemi, jich několik, zvláště jako kulturní relikty, v ustálené podobě přežívá, srov. *Zpěvem k srdci, srdcem k vlasti.*, *Jsmem s vámi, buďte s námi.*, *Rovnost, volnost, bratrství.*

Frazémy se často užívají i jako výrazy pro **radu**, srov. *Dvakrát dává, kdo rychle dává.*, *Co tam, to tam, o to hlavy nelam!* Podobně jako rada má voluntativní funkci i **přání**, které je svou podstatou založené, podobně jako pozdrav, na elipse typu *já ti přeju*, srov. *Kéž by!*, *Čest jeho památce!*, *Chléb náš vezdejší dejž nám dnes.*, časté jsou však jeho extenze a komentáře nebo repliky na něj, srov. aspoň *Starý lidi jsou jako malý děti.*

**Kontakťová funkce** frazémů se obvykle a nejčastěji explicitně zprostředkovává jako **pozdrav** (viz výše), který je v širším smyslu hojný svým územ, ale i počtem jednotek ve frazeologickém systému. **Kontaktu mezi lidmi** v jiném smyslu, než jaký zprostředkovává pozdrav, se týkají frazémy jako *Chová se, jako by se nic nestalo.*, *S kým mám tu čest?* apod. (výrazně fatická funkce se tu dá izolovat jen obtížně). Poměrně časté jsou v kontakťové oblasti ještě výrazy **díku**, srov. *Díky bohu!*, *Zaplat' pámbu!*, *Jako by se stalo!* Forma **kletby a nadávky**, které mají také kontakťovou povahu, bývá obvykle zvolací, srov. *A sakra!*, *Aby do toho (už)!*, *Di k čertu!*, *Ty seš korunovanej vů!* Někdy se šíře mluví o **kontakťových formulích** existujících a užívaných pro různé druhy sociálního kontaktu. Sem patří i všechny ostatní frazémy vyjadřující, resp. doprovázející nebo naznačující různé **sociální rituály**. Většina rituálů, resp. jejich občasné verbální doprovod, **rutinní formule**, ovšem pro svou specializovanou a kvaziterminologickou povahu mezi frazémy nepatří. Srov. však aspoň frazeologické případy pozitivní vztahující se k různým tradičním a dodržovaným **zvykům** *Dejž to pámbu!*, *S dovolením!*, *Srdečné blahopřání!*, *Čest jeho památce!*, *Sláva vítězům, čest poraženým!*, *Zlom vaz!*, *Pozdrav pámbůh!* – *Až na věky!*, *Jménem zákona!*, *Přísahám na svou čest a svědomí.* *K budování a obraně vlasti buď připraven!*, *Arbeit macht frei.* aj. Řada dalších typů (srov. už výše), jakkoliv jsou známé, však svá jazyková jména nedostala; sem patří např. výrazy označující **pověry** (*Musím to zaklepat*) a jiné.



**Citát** (ve frazeologii však jen citát ustálený) je přesné znění vysloveného názoru nebo hodnocení obvykle významné osobnosti minulosti, který zobecňuje a aktualizuje, ale i zasazuje do rámce doby. Cituje se mnohé, ale jen málo z toho je známé a běžně užívané, popř. přesné, srov. aspoň masarykovské *Nebát se a nekrást!* či vančurovské *Tento způsob léta zdá se mi poněkud nešťastným*. Pokud se u citátu ztratí povědomí o jeho autorovi, stává se zpravidla příslovím. I v češtině existuje málo užívaný výraz **okřídlená slova** (podle Homéra, *epea pteroenta*) stojící uprostřed mezi citátem a příslovím (proslavil je svou knihou G. Büchmann, *Geflügelte Worte*, 1864), který se dnes však už užívá méně, srov. Descartovo *cogito, ergo sum*, *cherchez la femme*, hamletovské *to be or not to be* aj. Častěji se užívají v překladu, srov. *být či nebýt, znám(e) své Pappenheimské (Daran erkenn' ich meine Pappenheimer)*, a pak jde o idiomatické kalky, v celé IF velmi běžné, běžně však mluvčími již takto nepocitované.

Řada jiných výrazů, ať už pojmově či terminologicky, se v češtině (téměř) neužívá, jakoliv můžou být v jiných jazycích běžné, ba oblíbené, jako např. v angličtině **limerick** (pětiřádkový humorný verš) či **wellerismus** (podle Samuela Wellera, vtipného sluhy pana Pickwicka ze slavné Dickensovy knihy, který je obvykle tříčlenný a končí směšným paradoxem), nebo haiku, oblíbený japonský básňový útvar (17 slabik), je-li ustálený, srov. aspoň ilustraci původního Wellerovského limericku *Every beggining is hard said the thief when he began stealing an anvil.*, a do jisté míry i české *Uvidíme, pravil slepej.* (Srov. mj. Permjakov 1970, 1979, 2000, Mukařovský 1971, Taylor 1931, 1975, Zich 1956, Kuusi 1957, 1972.)

#### 4.322 Frazémy a idiomy polypropoziční intersubjektové

Frazémy a idiomy **polypropoziční intersubjektové**, které jsou vytvářené kombinací alespoň dvou propozic a vztahují se v komunikaci také alespoň k dvěma mluvčím, tvoří již přechod do drobných útvarů literárních a folklorních různého typu. Jejich užití je zvlášť silně normováno ve vztahu k jazykové a společenské kultuře při mezilidském styku (odtud také někdy název *kulturní idiomy*). Zároveň tyto výrazy nepřímou naznačují, které stránky tohoto styku jsou nejčastější a nejvýraznější; spadají tudíž i do oblasti sociolingvistiky a antropolingvistiky. K nejznámějším typům této především dialogické oblasti IF patří

- (1) korelované pozdravy,
  - (2) zdvořilostní fráze,
  - (3) kontaktní formule,
  - (4) žertovné (a jiné) reakce a repliky,
  - (5) intersubjektové anekdoty, hádanky aj. Srov.
- (1) A: *Na shledanou!* – B: *Na shledanou!*; A (v reakci na kýchnutí): *Pozdrav pámbu!* – B: *Dejž to pámbu!*
- (2) A: *Děkuji.* – B: *Není zač/Rádo se stalo* aj.; A: *Dobré chutnání!* – B: *Nápodobně;* A: *Pro-miňte.* – B: *Nic se nestalo;*
- (3) A: *Jak se vede?* – B: *Děkuji, ujde to* aj.; A: *To je pan XY.* – B: *Těší mě.;*
- (4) A: *Proč?* – B: *Pro slepičí kvoč.;* A: *Já mám hlad.* – B: *Tak ho hlad' (a říkej mu malej!)* apod.

Význam těchto frazémů je vázaný na druh funkce, především na ose mluvčí-posluchač. I u tohoto typu frazémů patří k nejspeciřičtější jejich (především vnitřní) syntagmatická stránka; konkrétní odpověď např. u pozdravu je zcela či téměř zcela daná pozdravem prv-

ním a ten je zase předurčen kontextem, situací a dalšími sociolingvistickými ohledy (viz např. (1) výše).

Není zcela jasné, kam až lze v identifikaci frazémů jít, pohybujeme-li se rozsahově nahoru; např. některé anekdoty anebo hádanky (angl. riddle) oblíbené zvláště ve folkloru africkém, můžou být značně rozsáhlé. Je pak otázkou, (1) nakolik jsou ještě ustálené a (2) které rysy idiomu si uchovávají, především z aspektů anomálie. Jiné takové folklorně-literární mikrožánry, jako jsou různé lidové stereotypy, nemívají však nutně pevnou formu.

### 4.33 Frazémy lexikální

**Frazémy lexikální** patří svými komponenty do roviny morfémové (slovotvorné a flektivní) a mají jedinou syntetickou, nediskrétní formu; funkčně jsou jednoslovným lexémem. Patří tradičně k nejspornější a také nejzanedbávanější oblasti frazeologie. Srov. kompozita *budižkničemu*, *kratochvíle*, *tlučhuba*, *tvrdošijný*, *každopádně*, deriváty *najít* (kde *na* = „povrch“ + *jít* = „pohybovat se“ ≠ „nalézt, find, trouver etc.“), *neplecha*, *zákeřný*, *našup* a řídce i flektivní formy *Nevídáno!*, *budiž*, *nač*, o jejichž povaze se však diskutuje. Přechod pro svou diskretní povahu tvoří případy s reflexivitou a explicitní valencí *obejít se bez*, *zašít se*, *zapomenout se*, *bodnout někomu* ap. Platí pro ně stejné vymezení jako pro frazémy ostatní a jsou mj. tedy stejně anomální jako frazémy kolokační apod. Pokud si jich teorie tvoření slov vůbec všímá (většinou však ne), považuje je za jakési výjimky, které jsou však dost početné (pro SSČ se jich odhaduje kolem 5 % hesel). Například všech 21 prefixálních derivátů slovesa *být/bývat*, které existují, lze takto považovat za lexikální idiomy (typu *nabýt*, *dobýt*, *zbývat* aj.). Plný popis této oblasti však dosud chybí.

## 5. Funkce idiomů a frazémů

Funkce frazému a idiomu je v textu primárně dvojí, formální a sémantická. V obou případech se jeho vnější syntagmatická stránka projevuje především v jeho kolokabilitě a valenci. Obecně je funkce frazému a idiomu daná potřebou pojmenovat (**nominativní** funkce) a plnit úlohu ve vyšší jednotce (**strukturní** funkce). K nim se řadí, zvláště u frazémů propozičních, ještě výrazná funkce **pragmatická** (viz zvl. 3.5 a 4.3213).

Do kontextu v užším smyslu, který je daný úhrnem všech dalších pojmenování, která s frazémem mají společnou referenci a vytvářejí s ním (bezprostředně) vyšší strukturní jednotku, vstupuje tento frazém různým způsobem. **Kolokabilita** frazému, tj. jeho schopnost spojovat se v textu s určitými typy pojmenování, má výrazně povahu sémantickou. **Valence** frazému určuje konkrétně způsob jeho formálního zapojení do kontextu, ať větného, nebo nadvětného. Toto zapojení si může v souladu s povahou výrazu a jeho transformačními možnostmi vynucovat i formální obměny výrazu. V kontextu se idiom a frazém aktualizuje i po stránce sémantické, tj. ruší se na jedné straně jak jeho polysémie, tak rozsah jeho významu, popř. se na druhé straně do jeho významu vnáší v souladu se záměrem mluvčího i některé složky nové. Výraz tak nabývá konkrétního smyslu. Zachovány však zůstávají další charakteristické složky, jako je jeho vágnost, evaluativní povaha a další.

Jako každé pojmenování slouží i idiom a frazém komunikativním potřebám. Úzus IF je závislý jak na tématu a situaci (a zvoleném stylu), tak na osobnosti mluvčího. Charakteristicky nejvlastnější doménou užití IF je neformální mluvený jazyk, popř. umělecká próza ho odrážející a publicistika; v odborném jazyce se naproti tomu, s výjimkou některých kvazifrazémů, IF nevyskytuje. V důsledku povahy významu IF, dané zvl. napětím mezi významem komponentů a celku, je míra kontextové aktualizace průměrného idiomu větší než u slova. To je pak i jeden z hlavních faktorů vedoucích ke slabší normovanosti takového výrazu. Skutečný úzus, který jeho význam a funkci podmiňuje, se na druhé straně sám poměrně rychle proměňuje, popř. kolísá, zvl. u výrazů slabě motivovaných. I místo IF ve slovníku a v idiolektu jednotlivce je velmi proměnlivé. Na jedné straně lze vyšší výskyt IF u mluvčího i spisovatele považovat za znak jeho vytříbeného a odstíněného jazyka; na druhé straně však slabý výskyt IF nemusí signalizovat výrazovou chudost a autor může volit místo IF tvorbu vlastních metafor (příkladem tu může být do určité míry Vladislav Vančura).

Sémantická povaha idiomu, ostře kontrastující s povahou okolních pravidelných pojmenování, klade i zřetelné meze přemíře užití IF v textu. Za text na výskyt IF dnes bohatý lze považovat již takový, v němž stojí jeden idiom v průměru na 80–120 slov (Čermák 1982), (psaný) text v tomto smyslu bohatší (tj. takový, ve kterém idiom připadne na menší počet slov) je už vzácnější a delší souvislý psaný text s jedním idiomem na deset nebo méně slov (především umělecký) pravděpodobně neexistuje.

Funkcí, které IF v textu plní, je ovšem více. Volněji se někdy mluví i o dalších funkcích, funkci **estetické** (řada frazémů vyjadřuje např. zvláště podařenou či vtipnou metaforu, obraz), **ekonomické** (nominativně je frazém obvykle velmi ekonomický a „krátký“ a doslovný opis, pokud ho vůbec má, je vždy delší), popř. i **evaluativní** (většina frazémů je prostředky pro výrazné hodnocení vůbec v systému nejbohatší), kterou je však třeba uvažovat v rámci pragmatiky. Patří sem i funkce **metajazyková**, založená na sledování a zasahování do vlastní linie sdělení vloženým komentářem nebo opravou, srov. funkci a smysl užití *co jsem to chtěl vlastně říct, mám to na jazyku, na tom si zlomíš jazyk, jinými slovy, lépe řečeno* aj.

## 5.1 Strukturní složka funkce

Každý frazém tvoří uzavřenou strukturu a rozkládat ho, ať již sémanticky, nebo syntakticko-formálně, není možné, stejně jako je i naopak nemožné libovolný, popř. nový frazém generovat. Strukturní funkce jeho komponentů se s jeho domnělou syntaktickou stavbou nekryje (protože ji nemá), srov. např. nemožnost rozkladu na vlastní predikát a objekt u výrazu *Čte mu levity* nebo podobně *Je z toho hin* aj. Jako celek však frazém do syntaktických vztahů vstupuje. I syntaktické vztahy ke kontextu zprostředkované realizovanou valencí a (zvláště aktualizacími) transformacemi jsou však v závislosti na druhu anomálie výrazu v různé míře vždy oslabeny. Nepřítomností pravidelné kompoziční funkce ve stránce sémantické (viz 3.31), která je obecně v pozadí každého pravidelného vztahu, se z tohoto vztahu i u frazému správně utvořeného (a analogického tudíž pravidelné kombinaci) zachovává jen vnější stránka formální, resp. formální prostředky takového vztahu, ne však plný vztah sám. Např. nominativní singulárová forma nominálního frazému nemusí automaticky znamenat, jako u substantiv, že frazém lze užít v roli subjektu (*otevřená hlava*) nebo neshodného atributu (*růžové brýle*) apod.

**Frazémy kolokační** jsou funkčně dvou druhů, frazémy (A) adkolokační a (B) adpropoziční, a to podle toho, zda se kombinují s pravidelnými kolokacemi (popř. slovy) nebo větami. Uvnitř se dělí dále, zčásti v závislosti na strukturní klasifikaci (uváděny jsou jen některé hlavní typy):

## A adkolokační

### 1 verbální:

- a) verbo-substantivní (V-S)
- b) verbo-adjektivní (V-A)
- c) verbo-adverbální (V-ADV, ADV-V)
- d) verbo-verbální (V-V)

### 2 nominální:

- a) adjektivně-substantivní (A-S, S-A)
- b) substantivně-adverbální (S-ADV, ADV-S)
- c) substantivně-substantivní (S-S)

### 3 modifikační:

- a) adverbálně-adjektivní (ADV-A)
- b) adjektivně-adjektivní (A-A)
- c) adverbálně-adverbální (ADV-ADV)

### 4 relační

### 5 substituční aj.

## B adpropoziční

### 1 modifikační

### 2 relační aj.

Napříč 1–3 jsou frazémy binomiální (A1d, 2c, 3b–c); přirovnání tvoří většinou zvláštní subtyp A1a. Příklady:

A1a: *číst někomu levity*, b: *kápnout božskou*, c: *přijít zkrátka*, *bohatě stačit*, d: *umět se narodit*;

A2a: *zlatý důl*, *anděl strážný*, b: *místo nahoře*, *vzhůru nohama*, c: *jablko sváru*;

A3a: *andělsky trpělivý*, b: *neslaný nemastný*, c: *široko daleko*;

A4a: *na úkor někoho*, *ve srovnání s někým*, *bez ohledu na něco*, b: *dřív než*, *hned jak*, *na jedné straně – na druhé straně*, *bez ohledu na to že /kdo/ co...*;

A5: *Co (je to) za (člověka)?*, *čertví kdo?* aj.

B1: *pro jednou*, *mezi námi*, *při čerstvé paměti*;

B2a: *Že by*, *Kdyby tak*, *Co když*, *Jen ať...*, *Co V/S – to V/S* (Co věta, to chyba), b: *byť i*, *až když*, *proto aby*, *ten který*, *kdo – ten*;

Určitý strukturní typ někdy odpovídá více funkčním typům a podobně i týž funkční typ může využívat více typů strukturních. Zvlášť je to zřetelné u komponentů sysémantických.

Klasifikace strukturní a (strukturně-) funkční u sysémantických komponentů do značné míry splývá a je méně vyhraněná. Výjimkou tu jsou početné viceslovné (idiomatické) prepozice (A4a), konjunkce (A4b, B2b), partikule (B2a) a pronomina (A5), které jsou funkčními ekvivalenty příslušných gramatických slovních druhů. Nejvýraznější strukturu z nich mají prepozice: **prep-S(-prep)** (užité substantivum je abstraktní).

Hlavní strukturní **funkční** typy frazémů **propozičních** jsou v podstatě dány mody věty, srov.

(1) indikativní,

(2) imperativně vokativní a

(3) interogativní,

každý s (a) pozitivní a (b) negativní varietou, srov.

(1a) *to je stará písnička, to je vrchol!*;

(1b) *ani pes by to nežral*;

(2a) *ruku na srdce!, trhni si nohou!*;

(2b) *abys to nezakřik(l)!*;

(3a) *a já jsem vosk?* (negativní otázka 3b je vzácná).

Pro tyto frazémy je navíc specifická i jejich formálně komunikativní funkce a funkce sémiotická (viz 4.321); více viz 4.32.

## 5.2 Nominativní složka funkce

**Pojmenovávací, nominativní** funkce idiomu, která je dána především jeho sémantickou povahou, má dvě hlavní složky, funkci deznativní a pragmatickou. Poměr zastoupení obou složek se u různých výrazů značně liší. Základní je funkce **deznativní**, založená na vztahu frazému jako znaku k jeho denotátu a nociónální (konceptuální) složce jeho významu; je společná pojmenováním všech druhů. **Pragmatická** funkce, založená na (subjektivním) vztahu účastníků promluvy k frazému, je naproti tomu zvláště výrazná v IF. Zprostředkovává obecně různé subjektivní aspekty mluvčím při komunikaci zdůrazňované (zvláště emocionální a postojoyé) a inventář IF je jejím nejdůležitějším výrazovým prostředkem. Existence IF svědčí mimo jiné i o tom, že i pro individuum jedinečná a subjektivní zkušenost je společensky a generačně opakovatelná a má své charakteristické prostředky, jimiž mluvčí chválí, zlobí se, žasne, zdůrazňuje apod. Zvláště výrazné uplatnění pragmatické funkce umožňuje denotativní vágnost frazému a dualita jeho skutečného významu kontrastující s povrchovou informací zprostředkovanou jeho komponenty.

Primárními nositeli deznativní funkce u frazémů kolokačních jsou deznátory, primárními nositeli funkce pragmatické jsou především různé operátory (Weinreich 1966, Greenbaum 1969 aj.); každý frazém má však v sobě zastoupeny v různé míře složky obě. Mezi operátory je třeba lišit aspoň prostředky vyjadřující různý subjektivní vztah mluvčího k obsahu promluvy (komentory), vztah deixe (deiktika), určení způsobu a míry platnosti obsahu promluvy (kvantifikátory) a široce chápané formální prostředky (syntaktické) organizace promluvy (relátory). Hlavní typy:

### A Deznátory

#### 1 primární:

a argumenty (A-S, S-S, S-ADV)

b predikátory (V-S, V-A, V-ADV, V-V, přirovnání)

#### 2 kvalifikační (sekundární)

a kvalifikátory (A-S, S-ADV, ADV-A, A-A, přirovnání)

b cirkumstativa (tj. okolnostní určení, ADV-ADV)

### B Operátory

#### 1 komentory:

a evaluativní (co do úspěšnosti, příjemnosti, správnosti aj.)

b generalizační (shrnující a zobecňující)

- c expresivní (vyjadřující zaujetí, upřímnost mluvčího aj.)
- d veritativní (vyjadřující postoj k pravdivosti promluvy nebo situace, nebo z nich dedukující)
- e intenzifikační (zdůrazňující)

## 2 deiktika

### 3 kvantifikátory:

- a existenční (výlučná platnost, rozdíl typu *každý* – *žádný*)
- b restriktivní (omezená platnost, rozdíl typu *zvláště* – *výlučně* – *hlavně*)

### 4 relátory

Příklady:

- A1a: *Ta hospoda na rozcestí je **zlatej důl**. Zachránila ho **zbožná lež**. Stalo se z toho **jablko sváru**.*
- A1b: *Táta mu pak proto **četl levity**. Po chvíli **kápnul božskou**. Při tomhle podniku **zkrátka nepřijdeš**. Je jí šedesát, ale pořád **se umí otáčet**. Taková ranka a ty **řveš jako tur**.*
- A2a: *Nakreslil to **jedním tahem**. **Andělsky trpělivej** učitel. **Neslaná nemastná** odpověď. **Voda studená jako led**.*
- A2b: ***Široko daleko** není žádný stín.*
- B1a: ***S bídou** stačí jen půlku. Je to **k zbláznění** dlouhé. **Z ničeho nic** vykřikl. **Ještě že** mám rezervní klíč.*
- B1b: *Je to **celkem vzato** docela slušné. **Zkrátka a dobře** jsem se spletl.*
- B1c: *Moc se mi to, **popravdě řečeno**, nelíbí. Nemusíš mu, **mezi námi**, říkat všechno.*
- B1d: *Tak Karel o tom **podle tebe** vůbec neví? **Podle všeho** už asi nepřijde.*
- B1e: ***Co chvíli** mu to připomínám. Lidí tam je **až černo**. **Za nic na světě** už tam nepůjdu.*
- B2: ***Nějakou tu** korunu to stát bude.*
- B3a: *Udělat to musí **jeden každý** z vás. **Odpadávali jeden po druhým**.*
- B3b: *Je to kreslír **nad jiné** dovedný. Zbývá mu už **jedna jediná** možnost. **Na prvním místě** bych se zeptal jeho.*
- B4: ***Ve srovnání s loňskem** byla letos úroda ovoce lepší. **S ohledem na nedostatek času** se to musí zkrátit. Byli tam tři, **a to** Novák, Holub a Skalka. Zlepšil se, **byť i stále ne dost**. **Ve srovnání s tím, jak** to dělají jinde, to vy děláte lépe. **Co slovo, to** nadávka.*

Nominativní funkce u **frazémů propozičních** závisí na druhu komunikativního vztahu a funkce (mluvčí – posluchač – kontext – situace – téma – médium). Hlavní případy:

#### A vokativní typ (vztah mluvčí – posluchač):

- 1 interjekční (vyjadřuje emocionální reakce, postoje apod.)
- 2 imperativní (vyjadřuje volní vliv na posluchače)
- 3 kontaktové (vyjadřuje fázi kontaktu)
- 4 kategorizační (přechod do B4)

#### B tematický typ (vztah mluvčí – téma):

- 1 deskriptivní (prostý popis)
- 2 postojový (hodnocení, výraz stanoviska apod.)
- 3 konektivní (naznačuje vazbu mezi úseky, fázemi tématu apod.)
- 4 kategorizační (identifikace tématu se známým typem)

**C mediální** typ (vztah mluvčí – prostředek komunikace)

**D situační** typ (vztah mluvčí – situace) aj.

Vokativní typ zahrnuje tradiční funkci (Bühler 1934, Jakobson 1960, Leech 1974 aj.) jak expresivní a apelovou (resp. emotivní i konativní), tak i fatickou (zvl. A3); funkce sdělovací (informační) je zahrnuta hlavně v typu tematickém a funkce metajazyková v typu mediálním. Podíl obou složek nominativní funkce, deznativní a pragmatické, je u propozičních frazémů mnohem vyrovnanější a žádná z nich zpravidla výrazně nepřevažuje.

Příklady:

A1: *Vida ho! To mi může být ukradený!*

A2: *Koukej mazat! Mě vynech!*

A3: *Zdař bůh! Tak se měj!*

B1: *Vzduch je čistý. Zatmělo se mu před očima.*

B2: *To zrovna! Chraň pámbu! To je fuk! Aby ho nebylo! Opatrnosti nikdy nezbyvá.*

B3: *Bylo nebylo... Jak se věci mají...*

B4: *To je slovo do pranice. Lehko nabyl, lehko pozbyl.*

C: *Na tom si zlomíš jazyk* (srov. metajazykovou funkci v 5.2)

D: *A teď babo radí, A co teď?*

Jednotlivé funkční typy a podtypy se různě modifikují nebo vzájemně kombinují, některé mají i své konvenční názvy, např. přípitky, pozdravy, díky apod. Za užití dalších kritérií, jako je platnost osobní – nadosobní, odbornost – neodbornost apod., lze takto klasifikovat i další tradiční a pojmově dosti vágní útvary, např. říkánky, rozpočítávadla, hesla, průpovědky, aforismy, hádanky aj. (viz i 4.3214).

### 5.3 Komplexní funkční klasifikace

Anomálie různého druhu a stupně, kterými jsou frazémy a idiomy charakterizovány a klasifikovány po stránce formální, sémantické i koloakační, se zpravidla odrážejí i v textovém chování těchto výrazů a jejich úzu. **Idiomatičnost** výrazu lze tudíž chápat jako míru jeho anomálnosti v některém ze sledovaných smyslů; vzhledem k tomu, že textové chování výrazu se odráží ve skladbě a poměru jeho transformací možných k nemožným, je jeho idiomatičnost daná a objektivně také testovatelná (popř. i kvantifikovatelná) restrikcí jeho transformací, zvláště aktualizáčních. Jsou-li všechny relevantní transformace možné, je to výraz pravidelný: stupeň idiomatičnosti výrazu je nepřímě úměrný počtu a povaze jeho transformací realizovaných, resp. realizovatelných.

Uvažujeme-li např. pouze **transformace aktualizáční** u verbálního typu idiomů a vycházíme-li z jeho 17 položek, tj.

(a) pozitivní – negativní, (b) aktivum – pasivum, (c) tvrzení – otázka – imperativ, (d) indikativ – kondicionál, (e) minulost – přezens – futurum, (f) singulár – plurál, (g) (osoba) první – druhá – třetí),

pak jen při pravidelné realizaci jedné alternativy z každé naznačené skupiny (a–g), tj. kombinace o sedmi prvcích, má výraz 244 možných realizací. Frazémy však zpravidla realizo-

vány všechny konfigurace svých transformací nemají; počet teoreticky možných transformačních kombinací takto „výběrových“ pro celý typ pak nesmírně roste (v daném případě zhruba  $10^7$  možností). V praxi lze však skutečný počet transformačních anomálních kombinací odhadovat jako mnohem nižší, avšak stále dostatečně vysoký na to, aby se staly bází nejpřesnější klasifikace v IF vůbec.

Počet a složení jednotlivých transformací možných a nemožných, pro každý frazém vysoce specifický, může být tedy i podkladem pro **komplexní funkční klasifikaci** vždy téhož strukturního typu výrazů. Srov. např. frazémy a řadu neužívaných transformací (uvozených 0):

<i>nechodit pro slovo daleko</i>	0: <i>ot, poz, pas, kond, imp, imp neg, 1sg a pl</i>
<i>dělat někomu hláskou troubu</i>	0: <i>ot, neg, pas, imp, 1sg a pl</i>
<i>vědět svoje</i>	0: <i>ot, neg, pas, kond, imp, imp neg, fut</i>

Uvedené verbální frazémy se liší řadou anomálních rysů (zde jen ve třech konfiguracích). Je jasné, že každou takovou konfiguraci anomálních aspektů, kterou má společnou víc frazémů, které jsou i svým textovým chováním podobné, lze využít pro stanovení zvláštní třídy či podtřídy a tedy i pro **funkční klasifikaci**. Každá taková konfigurace vyjadřuje komplexně textovou funkci daného frazému.

Takovou bázi je třeba modifikovat pak přihlédnutím ještě ke všem transformacím dalším (tj. jiným než aktualizacím, viz 6.3), ke stránce sémantické i strukturní (viz Čermák 1982).

## 6. Transformace frazémů a idiomů

### 6.1 Podstata transformace ve frazeologii a idiomatice

**Transformace** v IF je takový proces (strukturní nebo gramatické) změny, jehož rezultat (nazývaný také **transformace**) se od své báze liší **funkcí** a tedy postavením v textu, avšak je své bázi formálně a sémanticky v různém stupni podobný; odvozenost tu je zřejmá. To platí pro transformace **strukturní**, tj. vztahující se k abstraktní struktuře, která je společná určité třídě individuálních frazémů. Transformace **individuální** se objevují takovou proměnou formy vždy jen u jednotlivých frazémů. Vztah obou druhů transformací však nelze považovat za vztah třídy a jejího prvku. Strukturně totiž určitá transformace realizována být může, ale nikdy jí nepodléhají všechny jednotlivé frazémy, které tuto strukturu obecně mají. Navíc některé typy transformací jsou pouze obecně typu strukturního (zvl. vnitrostrukturní, viz 6.23), a individuální transformace chybějí. Tím je dán i hlavní rozdíl transformací od variant (viz 3.4), které jsou bezkontextové (pokud jsou ustálené), je pro ně závazná **identita sémantická i funkční**, a jsou tudíž omezeny pouze na obměnu části formy individuálního frazému. **Varianta** je tudíž vždy obměna téže jednotky, která si zachovává svou funkci, transformací však může vzniknout jednotka další. **Báze** transformace je základní podoba obecné struktury nebo individuálního frazému, z níž proces transformace vychází; v praxi je obvykle shodná s invariantem. Vzhledem k mnohostranné inherentní anomálnosti celé IF a absenci přesných pravidel (gramatických i sémantických) je však vhodnější transformace v IF považovat za pouhé **kvazitransformace**.



Transformace můžou být **ustálené**, např. (transformace:) *jehla v kupě sena* (← báze: *hledat jehlu v kupě sena*), i **neustálené**, které jsou nejčastější v nominalizacích, srov. *postavení něčeho na hlavu* (← *postavit něco na hlavu*). Transformace se liší nejen stupněm obecnosti, povahou frazémů, jeho strukturou a rovinnou příslušností, ale i tím, že jejich rezultat již do IF patří nemusí, srov. *ostřížít pohled* (← *mít pohled jako ostříž*). Transformací z jediné báze může být několik, srov. *hledání jehly v kupě sena* a *jehla v kupě sena* (← *hledat jehlu v kupce sena*).

## 6.2 Typy transformací

Transformace strukturní a individuální, tj. transformace (různě rozsáhlé) třídy a (individuálního) frazému, tvoří protiklad ve všech oblastech, kde se vyskytují. Zvláště výrazný se stává v rámci transformací vnitrostrukturních (viz 6.23). V přehledu hlavní typy představují transformace

- I Mezirovinné
- II Vnitrovinné
  - A mezistrukturní
  - B vnitrostrukturní
    - 1 syntagmatické
    - 2 paradigmatické
    - 3 smíšené

Tyto transformace bývají zpravidla ustálené a změna jejich funkce je obecné povahy. Při vstupu do konkrétního kontextu však každý individuální frazém podstupuje ještě **transformace aktualizační** (méně přesně gramatické), jimiž se do něho konkrétně začleňuje. Jsou podobné jako transformace vnitrostrukturní trojího druhu: syntagmatické, paradigmatické a smíšené (viz 6.3).

### 6.21 Mezirovinné transformace

Při mezirovinných (interstratálních) transformacích se liší báze od výsledné transformace příslušností do odlišné roviny; zpravidla však báze patří do roviny vyšší. Nejběžnějším typem takových strukturních transformací jsou

- (1) lexikalizace
- (2) transformace kolokační (kolokalizace)
- (3) propozicionalizace.

Uvnitř obou lze dále rozlišovat

- (A) nominalizace
- (B) verbalizace
- (C) adjektivizace
- (D) adverbializace aj.

Obecně lze uvažovat i o dalších typech (partikularizaci, prepozicionalizaci, konjunkcionizaci, komparativizaci aj.). Individuální transformace tu jsou jejich výběrovými realizacemi na některé konkrétní případy. Např.

1A: *tlučhuba* (z V-S), *květomluva* (z S-S, V-S); 1B: *lelkovat* (z V-S), *špačkovat* (z přirovnání); 1C: *sněhobílý* (z přirovnání); 1D: *tvrdšíjně* (přes A z V-A-S);

- 2A: *tichá voda, devatero řemesel* (z propozic); 2B: *sejít z mysli* (z propozice) aj.  
 3B: *Stalo se to v hodině dvanácté* (z kolokace *udělat něco v hodině dvanácté*).

## 6.22 Vnitrorovinné transformace mezistrukturní

Charakteristické jsou především na rovině kolokačních frazémů. Hlavní přechody mezi jednotlivými strukturami pokrývají transformace:

- 1 nominalizace** (rezultát: S-S, A-S, A-S-A aj.; z různýchází)  
**2 adjektivizace** (rezultát: ADV-A, popř. A-S, S-A aj.; z různýchází)  
**3 verbalizace** (rezultát: V-S, V-ADV, V-V aj.; z různýchází)  
**4 komparativizace** (rezultát: přirovnání; z různýchází)

Srov. individuální transformace, které výše uvedené transformace strukturní výběrově ilustrují:

- 1 (*překročit Rubikon* → *překročení Rubikonu*; *zhroutit se jako domeček z karet* → *domeček z karet*;  
 2 (*postavit něco na hlavu* → *postavený na hlavu*; *(křivě někoho nařknout)* → *křivě nařčený*;  
 3 (*poslední cesta* → *doprovodit někoho na poslední cestě*; *(být silný jako lev)* → *mít sílu lva*;  
 4 (*prašivá ovce* → *být jako prašivá ovce*; *(polít někoho studenou vodou)* → *byl, jako by ho studenou vodou polil* aj.

Jednotlivé transformace strukturní mají někdy několik zdrojů (bázových struktur) a naopak jednotlivé struktury se mohou státází pro více druhů transformací. Vnitrorovinné (intrastruktální) transformace tudíž tvoří mezi strukturami bohatou síť křížících se vztahů. Pro řadu frazémů je příznačné, že se vyvíjejí z původních přirovnání, které je takto často první instancí a vstupní branou do IF, srov. transformační sekvenci: přirovnání → verbonominální frazém → nominální frazém (existující dosud ve všech fázích zároveň), tj. *mít zrak jako ostrůž* → *mít ostrůž* → *ostrůž zrak*. Jednou z možných a důležitých cest dalšího vývoje je vznik polysémie jednoslovného pojmenování již mimo rámec IF, srov. *dupe jako kůň* → (*to je ale*) *kůň*.

## 6.23 Vnitrorovinné transformace vnitrostrukturní

Jsou to transformace, jimiž se základní kombinační struktura (obvykle binární) větví ve vícečlennou. Je jich řada typů (viz též jednotlivé subtypy v přehledech v 4.311–6, včetně substitucí; PROP = propozice). Hlavní typy tvoří transformace:

### 1 Syntagmatické

A *kvantitativní* a aditivní: A-S → A-A-S, A-S → ADV-A-S, V-S → V-A-S aj.

b redukční: PROP → úsek PROP, V-A-S → V-S aj.

B *kvalitativní* (inverzní): A-S → S-A aj.

### 2 Paradigmatické

A *kvantitativní* (zvl. substituce číslovkou, tj. numeralizace): V-S → V-Num, V-A-S → V-Pron-Num aj.

B *kvalitativní* (zvl. substituce zájmen, tj. pronominalizace): A-S → Pron-S, V-S → V-Pron aj.

- 3 Smíšené**, tj. zvl. explicitní vyjádření relace mezi komponenty (tj. předložkou a spojkou): V-S → VprepS, S-S → SkonS aj.

Individuální příklady, na kterých se ilustrují výše uvedené transformace strukturní, zpravidla samy o sobě transformacemi z určité konkrétní báze nejsou; vznikly již jako hotové frazémy přímo z transformované obecné struktury, transformační vztah však přesto naplňuje. Frekvence jednotlivých typů je značně rozdílná. Srov.

1Aa *staré zlaté časy, jednou za uherský měsíc, mít šťastnou ruku*

1Ab (*Chytrému napověz, hloupému dolož* →) *Chytrému napověz..., (hledat si postranní cestičky →) hledat si cestičky*

1B *anděl strážný*

2A *dát si dvacet, olíznout všech deset*

2B *moje maličkost, hodit sebou, jet v tom*

3 *jít pod nůž, rub a líc*

Za redukční transformace lze považovat mj. tradiční elipsy (1Ab, první příklad).

### 6.3 Transformace aktualizační

Tyto vždy individuální transformace jsou závislé na kontextu frazému a má je, pokud mu to jeho anomálie dovolí, každý frazém s proměnlivou morfologií některého komponentu. Co do množství jich má každý frazém téhož typu a struktury teoreticky stejný počet. V praxi však jich má vzhledem ke své individuální anomálii vždy méně a v různém množství; zároveň se liší od frazému k frazému i kvalitou. Soubor realizovatelných transformací frazému (vycházíme-li z jeho invariantního tvaru) je vysoce příznakový a individualizující u každého frazému nejen z hlediska jeho funkce, ale i míry jeho idiomatičnosti. Kromě dalších závislostí (na typu anomálie formální, kolokační a sémantické) je míra idiomatičnosti frazému (idiomu) obvykle komplementární a nepřímo úměrná především úhrnu jeho možných transformací (zvl. aktualizačních). V souboru realizovatelných transformací, který u jednotlivých i velmi podobných frazémů zpravidla není zaměnitelný, se tudíž projevuje individuální anomálie výrazu po stránce funkční. Aktualizační transformace, resp. poměr realizovatelného a nerealizovatelného v nich, jsou proto základem komplexní funkční klasifikace frazémů (5.3). Hlavní typy (v 1 značené závorkou) jsou:

**1 Syntagmatické** A **kontaktní** a aditivní: (*zcela*) *zahladil stopy, (neúspěšný) pokusný balónek*

b redukční: *chytrému napověz (a hloupému dolož)*

c inverzní: *hory doly sliboval* (invariant: *slibovat hory doly*)

**B distanční:** *má srdce (již) zadané, měl (konečně) čistý stůl*

**2 Paradigmatické** A **statutární**

**a fázové** – inchoativnost:

(*dostat někoho na krk*)      *vzít si něco do hlavy*

– durativnost:

*mít někoho na krku*      *mít něco v hlavě*

– terminativnost:

*mít někoho z krku*      *pustit něco z hlavy*

**b agenciální** (zvl. kauzativní)

**agens:**

– inchoativnost:

*hodit někomu někoho na krk*    *nasadit někomu něco do hlavy*

– durativnost:

–

– terminativnost:

–

*vyhnat někomu něco z hlavy*

**patiens:**

– inchoativnost:

*(dostat někoho na krk)*

–

– durativnost:

*mít někoho na krku*

*mít něco v hlavě*

– terminativnost:

*mít někoho z krku*

*(pustit něco z hlavy)*

**B morfologické** (gramatické)

**a** deiktické: osoba, číslo, čas, popř. determinace a komparace

**b** relační: pád

**c** modulační: otázka, negace, modus

### 3 Smíšené: pasivum

Nemožnost podstoupit určitou transformaci je u silněji idiomatických frazémů charakteristická. Srov. takovéto negativní možnosti např. u 1Aa *číst někomu levity* (tj. nelze např. \**zcela mu četl levity*), 1Ab *nasadit někomu psí hlavu* (tj. nic nelze redukovat), 1Ac *chytat lelky* (tj. pevnost slovosledu); B *mít v hlavě*, *zlatý důl* aj.

Transformace typu *byly to (úplně/hotově/...) lukulské hody, dostal za to (příslověčnou) jidášskou odměnu* aj., jimiž mluvčí **komentářovým způsobem** vstupuje do základní komunikativní linie, jsou zvláštním případem **metajazykové transformace** (vyskytují se zde jen určitá slova) a jsou založené na omezeném inventáři prostředků metajazykové funkce (viz 4.3212).

Běžné transformace aditivní jsou založené na možnosti kvalifikace výrazu podle běžných syntaktických pravidel (rozvíjení); řídká redukce (v podstatě elipsa) je příznačná zvláště pro výrazy delší a frekventované. Distanční transformace, spočívající v možnosti frazému rozdělit se a přijmout mezi komponenty své struktury vnější kontextový prvek, je také řídká; charakteristická pro IF je tudíž její nerealizovatelnost.

Paradigmatické transformace statutární struktury **V-S** (viz též frazémy se substantivem abstraktním v 4.315), založené na v různé míře realizovatelných fázích prostého nebo kauzativního děje, činnosti a stavu, jsou charakteristické jak svou specifickou distribucí pouze u některých frazémů, tak i souborem verb, která do nich vstupují. Morfologické (gramatické) transformace patří k nejcharakterističtějším vůbec, srov. jejich nemožnost např. pro 1. sg. verba (*nechat na hlavě*), singulár substantiva (*pálit si prsty*), futurum (*mít za ušima*), komparativ (*šedá eminence*), libovolný další pád (*každým coulem, otevřená hlava*), otázku (*mít oči z Kašparovy krávy*), pozitiv (*nemoci od něčeho odtrhnout oči*), imperativ (*mít dlouhou chvíli*) aj. Soubor teoreticky možných morfologických transformací je pro každý strukturní typ nutně jiný (zde je ilustrovaný hlavně na V-S a A-S).

Pro svou povahu je zvlášť vydělena silně příznaková **pasivizace** (3 výše), srov. opět její nemožnost např. u frazému *hodit flintu do žita, vzít do zaječích*. Realizované pasivum

u prvního příkladu (*flinta byla hozena do žita*) svědčí o existenci doslovného homonyma (viz 2.4).

Možnost či nemožnost transformace je však ještě širší. Idiomatická povaha určitého frazému může vést k tomu, že zabraňuje i zcela běžné transformaci v pravidelném jazyce, v němž vystupuje jako komponent širšího výrazu, přičemž je mu taková vněfrazologická transformace i strukturně zcela cizí. Srov. např. vztah pasivizace k nominálnímu frazému *pokusný králík* a k pravidelnému výrazu *divoký králík*: *Ty škody byly způsobeny divokými králíky* x *\*Ty škody byly způsobeny pokusnými králíky*. Podobně nelze frazém *pokusný králík* běžně užít ani v subjektu věty aj.

## 7. Závěr: místo frazeologie a idiomatiky v jazyce a textu

Každý přirozený systém včetně jazykového je kvůli své neustálé interakci s okolím, jeho vlivy i vývojem otevřený. Jedním z důsledků této otevřenosti je jeho nízká uspořádanost, resp. vysoká míra **entropie**. Projevuje se vysokou nepravidelností, nepravidelnou distribucí jednotek systému (včetně redundance), a tedy i neprediktabilností jejich výskytu i vzniku. Zdá se, že v jazyce je dobrým příkladem vysoké entropie právě frazeologie a idiomatika, projevující se při analýze frazémů a idiomů v procesu komunikace: jejich informační zpracování, zvl. při jejich dekódování, je vždy složitější a vrstevnatější (děti jim do jistého věku nerozumějí, protože je chápou doslovně). Přestože je míra informace při jejich postupném dekódování nutně velmi malá, je naproti tomu míra informace přenášaná jimi jako známými a naučenými celky (ustálenými pojmenováními) ve srovnání s pravidelnými kombinacemi jazyka vyšší. Významným projevem entropie je i nesystematická existence uzavřených paradigmát v kontrastu k otevřeným. Nepravidelnost bude obecně v jazyce v důsledku této obecné povahy jeho systému vždy a vždy bude i jednou z nejvíce nápadných vlastností tam, kde se objevuje masově, tj. ve frazeologii a idiomatice. Jedním z důsledků tohoto je i **uni-verzálnost** frazeologie ve všech jazycích. Není proto náhodou, že i nejstarší a stále užívaný **umělý jazyk** esperanto začíná pro své lepší fungování nabývat vlastností přirozeného jazyka a frazémy a idiomy se začaly objevovat i v něm.

IF je oblastí, stručně řečeno, **anomálních hotových nominací** komplementárních k většinovým nominacím pravidelným, které bývají většinou neustálené. V důsledku své anomálnosti však frazémy a idiomy mohou „přeskočit či ignorovat“ pravidla známá odjinud do té míry a takovým způsobem, že nabývají charakteristické vlastnosti jiné, které pravidelný jazyk systematicky nevykazuje (např. motivovanost delšího pravidelného výrazu může být nevýhodou a překážkou pro jeho možné zkrácení). Frazémy a idiomy jsou pak jazykové prostředky významně **ekonomické**, které na malou plochu své formy vážou značně komplexní význam a funkci, a to v jasném protikladu k prostředkům pravidelného jazyka. Je obtížné najít frazém, který by byl rozvláčný, příliš dlouhý a neohrabaný. IF je tedy významným prostředkem vyhovujícím potřebě jazykové ekonomie. Z hlediska slovnědruhových funkcí například stojí v ostrém kontrastu k (absolutním i frekvenčním) proporcím základního pravidelného lexikonu nápadně rozsáhlý výskyt i míra užití frazémů verbální povahy, které jako by kompenzovaly omezený počet pravidelných verb: IF je zde velmi významným tezaurem hotových víceslovných predikátů.

Na druhou stranu je IF nejrozsáhlejší oblastí lexikonu, která je specializovaná mj. na pragmatické funkce vyjadřování **evaluativního**, tj. subjektivní a často i emocionální evaluace v postojích a názorech svých uživatelů. Je zřejmé, že každý mluvčí vše kolem sebe v různé míře neustále hodnotí, upřednostňuje nebo odmítá, často v doprovodu aspektů emocionálních. Vedle potřeby sdělovat fakta je v lidské komunikaci potřeba hodnotit zcela základní. Této evaluativní oblasti se pravidelné prostředky nepropůjčují snadno ani ne vždycky, a proto i pozornost jazykové evaluaci věnovaná tradiční lingvistikou je dosud mizivá. Naproti tomu si nápadná evaluativní většinová povaha frazémů a idomů o to přímo říká: je k tomu jakoby stvořena.

Sémanticky je valná část IF, především v souvislosti s entropií, také významně **vágní**, a to už v oblasti svých denotátů. Frazémy jsou takto obecně neurčité, neexaktní (a tedy opakem exaktních termínů) s nejasnými hranicemi. V textu se pak často takový význam do značné míry snadno dotváří a modifikuje. Lze tedy mluvit u IF o významné míře **potenciality** dalšího, nestandardního úzu, který se tu komunikativně nabízí. Standardním územ IF, který našťastí nemá cenu kodifikovat a ani ho kodifikovat nelze, se ovšem míní jen momentálně a statisticky platný úzus většinový a typický, který mluvčí snadno přizpůsobuje svým potřebám (což např. s běžným významem slov lze dělat jen do určité míry a někdy). Frazeologie a idiomatika se tudíž i pružně a rychle **vyvíjí**.

---

A translation of this paper is called *Foundations and Principles of Phraseology and Idiomatics*, see also elsewhere in this volume. For this, my thanks go to A. Klégr.

# Foundations and Principles of Phraseology and Idiomatics

1. Introduction: Phraseology and Idiomatics in a Wider Perspective .....	74
2. Principles and Foundations of Phraseology and Idiomatics .....	76
2.1 The Nature and Position of Phraseology and Idiomatics .....	76
2.2 Paradigmatic and Syntagmatic Relations .....	77
2.3 Anomaly in Phraseology and Idiomatics .....	81
2.4 The Definition of the Phraseme and Idiom .....	83
3. The Phraseme and Idiom: Their Components and Structure .....	84
3.1 The Character of the Phraseme and Idiom in General .....	84
3.2 The Form of Phrasemes and Idioms .....	87
3.21 The structure of Phrasemes and Idioms .....	87
3.22 Components of the Idiom and Phraseme .....	88
3.3 Meaning and Denomination .....	89
3.31 The Nature of Meaning in the Idiom and Phraseme .....	89
3.32 The Elements and Aspects of the Meaning of the Phraseme and Idiom .....	91
3.4 Variants of the Phraseme and Idiom .....	93
3.5 Semiotics of the Phraseme and Idiom .....	94
3.6 Quasiphrasemes and Quasiidioms .....	95
4. The System of Idiomatics and Phraseology and Classification .....	95
4.1 Distribution of PI in Language .....	95
4.2 Classification of Idioms and Phrasemes .....	96
4.3 Formal Classes of Idioms and Phrasemes .....	96
4.31 Classes of Collocational Idioms and Phrasemes .....	96
4.311 Lexical Words in the Roles of Components .....	96
4.312 Synsemantic or Function Words in the Role of Components .....	99
4.313 Lexical words and Function Words in the Role of Components .....	101
4.314 Similes .....	102
4.315 Verbonominal Quasiphrasemes .....	104
4.316 Binomials .....	105
4.32 Classes of Propositional Phrasemes and Idioms .....	106
4.321 Propositional and polypropositional monosubject phrasemes and idioms .....	107
4.3211 The Formal Structure and Semantics of Propositional Phrasemes and Idioms ...	107
4.3212 Functional Classes of Propositional Phrasemes and Idioms .....	109
4.3213 Pragmatic Function of Propositional Phrasemes and Idioms .....	109
4.3214 Some Traditional Types of Propositional Phrasemes and Idioms .....	111
4.322 Polypropositional Intersubject Phrasemes and Idioms .....	113
4.33 Lexical Phrasemes .....	114
5. The Function of Idioms and Phrasemes .....	115
5.1 The Structural Component of the Function .....	116
5.2 The denominative Component of the Function .....	118
5.3 Complex Functional Classification .....	120
6. Transformation of Phrasemes and Idioms .....	121
6.1 The Essence of Transformation in Phraseology and Idiomatics .....	121
6.2 Types of Transformation .....	122
6.21 Interstratal Transformations .....	122
6.22 Intrastratal Interstructural Transformations .....	123
6.23 Intrastratal Intrastructural Transformations .....	123
6.3 Actualizing Transformations .....	124
7. Conclusion: The Place of Phraseology and Idiomatics in Language and Text .....	126

## 1. Introduction: Phraseology and Idiomatics in a Wider Perspective

Phraseology and idiomatics (PI henceforth) is a relatively young branch of linguistics, and it only gradually comes to light where in language we may look for its units, idioms and phrasemes. In language they are, generally speaking, found in all **branches of linguistics** wherever combinations of units endowed with meaning are involved. Moreover, PI has links to other **non-linguistic disciplines**, outside of language. Little attention has been paid to the study of connections with psychology and logic, esp. in the area of expressing concepts and propositions, with philosophy or ethics of assessing and conveying experience, with history and arts (tropes, symbols, etc.). However, there is a traditional tie particularly with cultural anthropology and ethnology, especially in the study of proverbs. This sphere has been long described, and with good reason, as a treasure house of the nation's experience under the name 'wisdom-lore', a term used by the 19<sup>th</sup>-century Czech revivalist poet, F. L. Čelakovský, in his monumental collection of proverbs (Čelakovský 1852).

If ethnographic, semantic, historical, sociological and other approaches cast light, among other things, on national specific features, including the use of metaphors and symbolism which PI expresses, it might be said, at first glance, that a significant historical part of Czech PI is, broadly and loosely described, rustic or agricultural in its structure, whilst sea-related themes (so common to English) are entirely missing. Because of their generality, such characteristics do not say very much and have little value. On the other hand, the view is only slowly gaining ground that PI is not a mere collection of irregularities and curiosities, somewhere on the periphery of language and attention, but, on the contrary, that it is a specific and little explored area of language in which standard ways and means will not get us very far, although it is an area open to active study and application in a number of communicative activities and spheres, such as pedagogical at school or rhetorical, i.e. textual, in the creation of different types of text. Its basic descriptive application is **phraseography**, which collects and describes phraseological means of expression of a given language.

Phrasemes and idioms may, loosely and not only linguistically, be seen as the most important part of stereotypes, especially verbal. A **stereotype** in language (i.e. apart from the much wider conception in sociology and elsewhere) is a lexical combination which enters a text as a ready-made, relatively frequent and well-known, or established, unit which is fixed and non-generable, a unit with an underlying model with similarly small or (almost) non-existent productivity. From a communicative point of view, the prefabricated nature of a stereotyped expression creates the impression of familiarity; the expression often induces a certain type of situation or behaviour and in terms of language-production it facilitates, speeds up and automates the generation of a text. Such a text loses its originality in these parts and its author forfeits his or her claim to an original creative expression. Stereotyped expressions abound especially in institutionalized language (such a language is distinctly stereotyped) which uses a large amount of formulae which tend to make the utterance formalized and thereby endow it with customary and predictable (usually polite) or even prescribed character, for instance in keeping with etiquette. The



degree of obligatoriness may be then sanctioned socially or legally. Typical examples are therefore expressions and formulae of greeting, thanking, wishing, rituals or formulae of swearing, arresting, marriage rituals and others. There are also formulae restricted to certain specialized fields.

Although formulae are used in institutionalized language and even in spoken language, for example in various parts of conversation, it would be a simplification to regard them as something negative. The very act of selecting between different types of stereotyped expression may be seen as an authorial contribution to text building. Also, almost all types of text require that stereotyped expressions be used inasmuch as people are by no means only creative and wholly original speakers; certain customary and conventional notions and situations need to be expressed in a customary and familiar way. To some extent, the original typographic meaning of the term stereotype (solid combination of more letters compared to free letters) is still present in it. The commonest stereotyped expressions include propositional and polypropositional phrasemes, esp. proverbs, sayings, correlated greetings and wishes, etc., and a number of other types of phrasemes, but also aphorisms, various kinds of situational clichés, standard beginnings or endings of utterances and other contact formulae, and finally, a great number of verbonominal phrasemes. The relationship between phrasemes and stereotyped responses is a much debated issue; in fact, it is an open one indeed. On one level, the two are largely identical and overlapping if they have a fixed form. However, a number of stereotyped responses (usually with negative connotations) which are studied especially in sociology do not have such fixed form and there the correspondence with phraseology ends. It is also a matter of the criteria whereby phraseology itself is defined. The sociological conception of stereotypes often operates with images (especially in the sphere of traditional views of other ethnic or national groups) which can be neither true nor flattering: for instance, the Scots are in Czech stereotyped as being miserly (a source of frequent anecdotes), the Germans as orderly and accurate (*německá přesnost* – ‘German precision’), the French as gourmets and lovers, the Gypsies as thieves or slatterns. Some of these non-linguistic stereotyped images have entered language, i.e. its phraseology, in the form of phrasemes which reflect a given stereotype, such as those about the Romany (*krade jako cikán* – ‘he is given to stealing like a Gypsy’), the Czechs as being musical (*Co Čech, to muzikant* – ‘Every Czech is a musician’), the Turks as being poor managers (*turecké hospodářství* – ‘Turkish, i.e. poorly organized, economy’), etc., however much removed they may be from reality. When this happens we have a conjunction of the two types of stereotyped responses.

It should be stressed that PI in Czech is not the study of **phrases** or ‘turns of phrase’ (clichés or hackneyed phrases). A ‘turn of phrase’ is a habitual expression used in conversation, usually felt to be unoriginal and semantically almost empty; it is a subjective concept, difficult to define. Nor is the ‘phrase’, as known especially in generative grammar, its subject of study. PI, as conceptualized here, deals with phrasemes and idioms and in the following attention will concentrate on these two.

## 2. Principles and Foundations of Phraseology and Idiomatics

### 2.1 The Nature and Position of Phraseology and Idiomatics

Phraseology and idiomatics (PI) is concerned with the study and description of phrases and idioms at all levels (strata), with special focus – on account of their frequency – on multiword lexemes. PI does not represent a special level of language, as was previously thought, but it is an area of fixed anomalous combinations (of syntagmatic and paradigmatic forms always of a lower order) which is complementary to the area of regular language combinations of all kinds and at all relevant levels. The resultant phraseme, stored in the system, is then a unit of a higher order than its constituents are. By a relevant level is meant any plane whose units are endowed with meaning (semantics).

By **regular combinations** in a generally syntagmatic sense we mean all combinations governed by analogous rules, i.e. rules based on analogy, that is combinations of morphemes, lexemes, collocations and even clauses. Generally speaking, by these rules we mean both semantic rules (always primary), based on semantic compatibility of the combined elements and on the meaningfulness of their resultant combinations (see also compatibility, below in 2.2), and (always secondary) formal and grammatical rules (e.g., syntactic ones in syntax). Thus, PI goes a step further in the area of combinatory outcomes of the extensions of combinatory possibilities in language, crossing the boundaries and operating in an area of combinations that, according to the standard rules of language, cannot or should not take place, i.e. **anomalous combinations**. This predetermines the inherent irregularity and non-model-based formedness of every idiom and phraseme. For instance, the simple Czech phraseme *ležet polem* (to pitch a camp) is formally anomalous and no other such instrumental construction (and syntactically something remotely reminiscent of, e.g., object) can be found anywhere else in the language with comparable components (static verb and the noun *pole*). In PI both components combine only in this unique way, cf. this phraseme and the regular collocation *ležet na poli* (to lie in a field), consisting of a verb and an adverbial answering the question ‘where?’.

Diachronically anomalous phraseological combinations came into being in various ways, in collocational phrasemes especially through a simile and individual metaphors, and may, among other things, reflect the residuum of combinatory possibilities previously common, but currently diachronic, that are no longer found in the present-day language system. Within the combinatory properties of linguistic units, i.e. properties which form the very core of every language (we may generally speak of the **combinatory principle**), lies the potential to form ever new naming units, including anomalous ones, and so even phrasemes and idioms. The existence of phraseology and idiomatics is a natural phenomenon in every language, just as the use of phrasemes and idioms is a multi-layered means of conveying language culture, due to their specific features and expressive possibilities.

The common **denominative function** of all types of phrasemes and idioms places PI straight within lexicology as its special branch. In other respects, though, it has specific relationships with a number of other disciplines. PI’s uniqueness and essence consists in that it is an area in which the validity and operation of principles based on rules known to be

functional anywhere else in language are suspended and therefore cannot be systematically applied in analysis: what looks like a morphological case here is functionally not a case, what appears to be an object of a verb is not an object, a word is no longer a word here with its familiar meaning, a word form is not a form of the underlying lexeme because the lexeme has become irrelevant and does not figure here as a whole at all, etc. All this is due to the omnipresent **anomaly** of different kinds which above all prevents the proper analysis of a phraseme in terms of criteria of the regular language and, conversely, its synthesis from the initial components. Yet it is possible to make use of the criteria of the regular language in PI at least at the level of components to describe (at least) their external characteristics, or their outward and superficial similarity with components of the regular language, for example in the formal classification of phrasemes on the basis of their similarity with the word-classes. Thus it is possible to describe the structure and type of, for example, the phraseme *nosit kůži na trh* (to carry one's skin to the market, i.e., to expose oneself to danger) as the combination **V-S-prep-S**, i.e. that of verb, noun, prepositions and noun, although in actual fact there are no nouns involved, and the structure described or the analysis made of it is only external and says nothing (specific) about the composition and nature of the phraseme (on the use of it see more especially in 4).

## 2.2 Paradigmatic and Syntagmatic Relations

A paradigm is an abstract class of elements, units at whatever level of the system, and it is itself usually part of the system. It is based on a feature (formal, semantic or functional) which is common to all these elements and unites them into a paradigm of whatever type it may be (cf. also Čermák 2000a). Members of a paradigm are mutually equivalent and in a text generally also mutually substitutable, their choice being governed by the *either-or* rule. A paradigm is either open, i.e. relatively accessible to being enriched by other members and, as a consequence, usually expanding, or closed, i.e. non-expandable in this sense (cf. Čermák 1978). Compared to the intuitively opaque and extensive paradigms of, say, indigenous roots, denominal adjective or kinds of dishes, animals, etc., of greater importance for phraseology is always a **closed paradigm**, which is typically very small, if not single-member, whose membership will evidently not grow. Let us have a more detailed look at this situation. In the following discussion, it will be necessary to distinguish, among other things, between 'virtual' and 'collocational' paradigms.

Within the language system every word, lexeme, morpheme, etc., belongs to a certain paradigm (one or more). Accordingly, paradigms within the system of language are of different sizes and different levels of abstraction. On the one hand, there is, for instance, a huge class of all nouns (i.e., nouns as a word-class amounting to several tens of thousands of one-word lexemes), mostly grammatical by nature, whose equivalence and combinatorics, manifest in the width of their mutual contextual substitutability (it corresponds to the general denominative function of nouns), are quite extensive. Basically, it coincides with the syntactic roles of the noun. However, there is not much to say specifically about their semantic nature. On the other hand, there is a class of, say, possessive adjectives derived from the names of animals (e.g. *psí* – dog's, *kočičí* – cat's), which is considerably smaller, less numerous, but semantically markedly defined. This is shown by its broadly equivalent function in a limited amount of possible contexts (the adjective *psí* appears, for example,

with mass nouns, such as *chlupy* – hairs, *maso* – meat, *kosti* – bones, nouns denoting movement and acoustic manifestations, e.g. *štěkot* – barking, but usually not with the names of persons, instruments, products, places, etc.).

Of course, paradigms are generally numerous and of different types, depending on which level of the language hierarchy they belong to. On the whole, the nature of a paradigm is related to the nature of its function: the more restricted (narrower) the paradigm and the smaller the number of its elements, the more semantic it tends to be, and vice versa.

A **virtual paradigm** is such a class as described, i.e. a class of (broadly) equivalent elements related by a common function and meaning which is abstract in nature, found in the language system, i.e. existing in mind only, and definable by its place in the system's hierarchy. One such type of virtual paradigm consists of the above adjectives of the type *psí* – dog's, *kočičí* – cat's, *lví* – lion's, *koňský* – horse's, another of mass nouns of the type *maso* – meat, *sádlo* – fat, *mléko* – milk, and yet another of the names of tools of the type *nůžky* – scissors, *kladivo* – hammer, *šroubovák* – screwdriver, *vrták* – borer, etc. The first type of virtual paradigm (*psí*, etc.), together with possessive adjectives related to man (*otcův* or *otcovský* – father's or fatherly, *matčin* or *mateřský* – mother's or motherly, *tetín* – aunt's, etc.), belongs to a higher class of possessive adjective as such, these belonging to yet a higher class and thereby to highly abstract adjectives in general, and so on. A virtual paradigm is usually easy to grasp even intuitively, and to some extent we can recall it quite easily and use it in mental operations; as a matter of fact most (semantic) classifications of lexemes, the study of hyponymy, etc., are based on this type of paradigm. Every member of a virtual paradigm shares the same general function with the others (i.e., it can participate in the same higher unit, structure), though individually each member differs from the others in its combinatorics. On the lexical level, we say that the **collocability, collocational potential** of every one of them is different.

A **collocational paradigm**, on the other hand, is a class of elements (still meaning words here), with every one of which a member of a different paradigm may habitually and meaningfully combine. However, this type of class is difficult to recall intuitively as a whole, and if extensive, the recall is as a rule not possible at all. To use the example of the adjective *psí* (dog's), its nominal collocational paradigm will include all nouns combinable with this adjective (because of their heterogeneity they will represent several semantically differentiated paradigms at once), for example

- (1) the type represented by nouns of acoustic manifestation such as *štěkot* – barking, *vrčení* – growling, *kňučení* – whining, etc.
- (2) the type represented by body parts, such as *chlupy* – hairs, *maso* – meat, *kosti* – bones, etc., or
- (3) other types (motions, properties, etc.)

Hence, a collocational paradigm is definable only through a particular **external** element, to which it **uniquely** relates on the basis of their mutual collocability (combinability); this external element (here the adjective *psí*) does not belong to this paradigm, though. Syntagmas which are formed by the elements of the collocational paradigm with this specific external ('defining') element usually differ in meaning and function. Every element of the system (every word, etc.) has its own (at least one) collocational paradigm and by

explicitly listing its members it is possible to delineate the possibilities and the boundaries of using the elements of the system (their **collocability**). Thus, it is usual to form the syntagma *psí štěkot* – dog's barking, but not *\*psí mňoukání* – dog's miaowing, or *\*kočičí štěkot* – cat's barking, *\*psí bučení* – dog's mooing, etc. (which are forbidden by semantic rules). In contrast to the virtual paradigm, the extent of the collocational paradigm may change more quickly over a time and it also often does; this may actually be seen as the source of metaphoric shifts or, in more general terms, of the semantic shift of a lexeme and, on the most general level, of the change in language as such. The nature of both types of paradigm (virtual and collocational) is reflected in the conception, perception and function of each phraseme and idiom.

The **syntagma**, a minimal combination of two usually heterogeneous elements, i.e. those originating (as a rule) from two different virtual paradigms, is semantically conditioned by the nature of its elements, generally also by the type of their paradigms and their hierarchical level. Hence it is conditioned by the combinatory rules governing the behaviour of these elements. Most syntagmas are unfixed, textual, and only some, primarily phrasemes, are fixed and form part of langue. If we take the above example *psí štěkot* – a dog's barking, there is nothing special about a syntagma and combination which is unfixed from the point of view of phraseology. Apparently it is formed ad hoc again and again according to semantic rules (not only are the components mutually non-exclusive, but they customarily combine and appear to share a common semantic feature). On the other hand, the syntagma *psí počasí* (foul weather) is a phraseme because it is anomalous: no semantic rule applies here (*psí* – dog's and *počasí* – weather are normally incompatible just as *\*psí déšť* – dog's rain, *\*psí klima* – dog's climate, etc.), both words and the virtual paradigms they belong to do not combine. From the point of view of the virtual paradigm the word *počasí* – weather, or *čas* – time for that matter, does not belong to any of the collocational paradigms mentioned which relate to the adjective *psí*. In other words, with regard to the adjective *psí* the word *počasí* forms both theoretically and practically a completely closed collocational paradigm with only one member (*čas* in *psí čas* is in fact only its variant), a paradigm, though, which concretely or semantically does not belong to the range of words habitually combining with the word *psí*. The phraseme *psí počasí* (or its older variant *psí čas*) is therefore based on a unique anomaly without support in whatever analogy.

The above example was chosen for its simplicity; other phrasemes may present a much more complicated picture. However, it makes the essence of a phraseme sufficiently clear. Inasmuch as phrasemes and idioms must be seen as combinations of elements which in some respect are always anomalous, it is necessary to look for them mainly in the non-productive and little used areas of combinatorics of language elements. In the sphere of regular language, among the different levels the most productive are syntactic, or collocational, models, the least productive are word-formation models, while the inflectional ones stand somewhere in between them. However, it is only very generally true that the incidence of idioms and phrasemes and the possibility (though not necessity) of the appearance of new ones is roughly proportionate to these productivity rates. Still, we do find the main concentration of PI phenomena in the sphere of collocations and sentences, although the number of morphological combinations is far from small (no systematic study has been done here so far, though).

**Collocability** is, generally speaking, individual, formal and semantic combinability, i.e., the ability of each language element to combine with another or others. It results from one or more of the element's collocational paradigms and (in regular combinations) is conditioned by its compatibility with them. Together with valency it is the main component of syntagmaticity of any language element. The individual realization of collocability produces a **collocation**. If, for instance, the **valency** of the verb *číst* (to read) requires a noun in the accusative, its collocability involves a (collocational) paradigm, i.e., a class of concrete nouns which provide the potential accusative forms, such as *číst (knihu – book, román – novel, dopis – letter, etc.)* and the collocation *číst knihu* is one of the common potential combinations. The collocational paradigm of the verb *číst* does not include, however, the incompatible *\*číst kámen* (to read a stone), *\*číst vodu* (to read water), *\*číst maso* (to read meat) ..., even though these are accusative forms too. In general, the collocability of each word is narrowly restricted and in principle follows from its semantics. Although it changes with time and users, it is never unlimited and never has an arbitrary scope: no lexeme of a language combines with all others. In its sum, compatibility is the essential component of the Saussurean value of the lexeme. For any two lexemes, even synonymous lexemes, it is always different, i.e. it is unique and noninterchangeable. As a rule, collocability is not thought of in terms of any other level but lexical.

The collocability of lexemes, i.e. words (or rather their forms), differs with respect to the types and the number of words, lexemes, etc., with which each lexeme combines. It depends on the nature of the lexeme's virtual paradigm, its place in the abstract hierarchy of the system, etc. (see above). To speak of the collocability of words like *velký* (large) is virtually impossible without an exhaustive investigation based on a corpus (their collocability probably involves hundreds of thousands of words, mostly nouns, which no speaker can carry in his head, in fact we cannot even recall them; in the Czech National Corpus SYN2005 they involve some 200 000 combinations). Thus the word *velký* forms or has at its disposal overwhelmingly large and open collocational paradigms which establish its collocability. By contrast, there are collocational paradigms which are extremely small and closed, such as that of the nouns defining the collocability of the adjective *širý* (wide, spacious), including {*moře – sea, pole – field, pláň – plain, svět – world, obloha – skies, krajina – landscape*}. Quite clearly, this is one of commonest types of anomaly in phraseology, the component being a member of an extremely limited and restricted paradigm. The relationship between the components of combinations of lexemes and words of regular language, if meaningful, is semantically always that of compositionality. Viewed from a different angle, these combinations are always made possible thanks to the principle of semantic **compatibility** which consists in the presence of common semantic features found in either of the two combining words (roughly: both the form *číst – read* and the form *knihu – book* imply the feature 'readability', to use a simplification).

In contrast to regular language, whose combinations are based solely and necessarily on compatibility, phrasemes and idioms are different. Their components may display compatibility as well as its absence, the relationship of incompatibility. **Incompatibility**, semantic disharmony, is a semantic syntagmatic relationship between two mutually exclusive lexemes or other elements of a semantic nature, which as a result cannot normally combine in a meaningful syntagma, e.g. *\*ostrá voda – sharp water, \*cítit kolmo – to feel*

perpendicularly, or the relationship between *pes* – dog and *strom* – tree (which normally never combine in language). In idioms, though, incompatibility is possible and frequent, as in *kamenný obličej* – stony face, *otevřená hlava* – open, i.e. clever, head, *umět se narodit* – know how to be born, i.e., to be lucky, *ztratit hlavu* – lose one's head, *to tě bude stát krk* – this will cost you your neck/head, etc.

The concept of a closed paradigm is important for another concept, crucial for phraseology, namely that of **monocollocability**. It is the collocability of a lexical form which is clearly defined by its being combinable with only a few (2–7), or strictly speaking, only one other form in a language. This single form (or forms) prevents the given lexeme from its linguistic extinction, because without combining, in a state of isolation, nothing is possible in language. These well-known unique and monocollocable forms, such as the above *širý*, or such forms as *holičkách*, *duhu*, *pré*, *namále*, *vstříc*, *třetice*, *tratoliště*, *dokořán*, *jakž*, *takž*, found in the phrases *nechat někoho na holičkách* (leave sb in the lurch), *jít někomu k duhu* (do sb good), *mít pré* (be free to), *mít namále* (be close to dying, have a narrow escape), *vyjít někomu vstříc* (be helpful), *do třetice všeho dobrého* (third time lucky), *tratoliště krve* (a pool of blood), *otevřít/být/nechat/zůstat dokořán* (have/be/leave/stay wide-open), *cítit se jakžtakž* (feel so-so), are on account of their severely restricted collocability entirely different from normal lexemes. Not only do they lack their own lexemes (unlike the locative form *holičkách* the nominative form *\*holičky* is nonsensical and artificially reconstructed), but also their word-class and any other category. Once again, the questions of whether *\*duhu*, *\*pré* are nouns and, in the latter case, an object do not make sense: there is no analogy to base our judgement on because no syntax has yet evolved here. These elements do not have their own meaning or even existence, because all of those only imprecisely guessed properties could be derived only on the basis of a sufficient analogy deducible from several different occurrences in which the form would have to exist. It does not exist though. Monocollocability therefore suggests only a seemingly independent existence of a form which, however, is non-existent in reality without its phrase and therefore fictitious (a given phraseme is its only context and so even the only form of its linguistic existence). In actual fact, such phrasemes could be written in a non-discrete way, as a solid sequence (see the last example), as it would make no difference to their character.

Discreteness of form is a misleading feature: it tempts us to postulate an independent existence to what is written separately. It is a problem specific to Czech and similar languages, and more generally the Latin script. In languages which use other types of script, this problem is less prominent, for example in Chinese and others.

## 2.3 Anomaly in Phraseology and Idiomatics

In phraseology anomaly is always perceived in contrast to the background of the semantic and formal situation in regular language, i.e., a sphere of language governed by rules. It is also made possible only thanks to this contrast with all phenomena which are both prevalent and regular in language. Anomaly is seen as the multiaspectual exceptional, irregular nature of the paradigmatic situation (see above) or as the realization of the syntagmatic or transformational relationship (often combined) which is viewed as a nongenerable deviation from rules of various types based on analogy. Anomaly is

omnipresent in phraseology in different degrees and shades, it is its constitutive feature. While regular analogous combinations in language are based on semantic and formal rules permitting relatively without limits the formation of new and new combinations (cf. *otevřený + dveře/okno/krabice/oko* ..., etc., i.e. 'open + door/window/box/eye' combinations in keeping with the semantic rule: 'openable' things and the grammatical rule: prescribing word order and gender agreement), anomalous phraseological combinations are always unique and restricted (totally, sometimes to just a few clearly delimited members). For instance, neither of the components of the phraseme *otevřená hlava* (open, i.e. clever, head) can be varied and replaced with another, analogous one in the same function or meaning (to create another phraseme), which is the basic feature of members of the same paradigm in regular language, as documented by *\*otevřená záda* (open back), *\*otevřené břicho* (open stomach), *\*otevřené oko* (open eye) or *\*zavřená/přivřená hlava* (closed/semiclosed head) (unless, of course, the opening refers to a different sense, say 'wound'). Similarly, whereas lexemes usually form regular sets (e.g. *pes* belongs among mammals, *stůl* among furniture), no such analogous sets of components exist in phraseology, cf. the definition of the phraseme below. The different types of anomaly include, for example, the impossibility to identify the word-class in the components *hin*, *duhu* (*být hin* – be exhausted/destroyed, *jít k duhu* – do sb good), the absence of meaning in these forms, noticeable even in the word forms *otevřená*, *hlava*, etc., which do not correspond to any of its standard lexical senses. It is therefore impossible to speak of lemmas, but only of forms, or elements. Syntagmatic anomalies of different types are exemplified by expressions such as by *ležet ladem* (lie fallow/waste), *malá domů* (passing the ball to one's goalie) or *hodit něco za hlavu* (throw st behind your head, i.e. not worry about it).

The most obvious sign of anomaly is – in keeping with the way functional anomaly was defined above – the **impossibility of paradigmatic substitution** of one component by another, analogous component in the same or similar function (= **commutation test**), as exemplified above. Other examples are *\*mít z něčeho lehkou/obtížnou hlavu*, *\*mít z něčeho těžké tělo/břicho/koleno*, *\*umět toto/tohle s někým*, *\*šesté/třetí kolo u vozu*, *\*co od/pro/k*, *\*kdo když*, *\*jestli/zda jo*, *\*Jak si někdo odestele, tak se položí*, etc. (which are made up non-phrasemic variations of the phrasemes *mít těžkou hlavu* z – feel bad about st, *umět to s někým* – can wind sb round one's little finger, *páté kolo u vozu* – fifth wheel of car, i.e. redundant, *co do* – as regards, *co když* – what if, *že jo* – isn't it, *Jak si kdo ustele, tak si lehne* – You've made your bed and you must lie in it). Because of the nature of phrasemes, anomaly may appear in the very status of the components. A number of phraseme components do not actually have an analogous counterpart in any lexical form (or lemma) in regular language and occur only in phraseology.

Anomaly is quite conspicuous in **grammatical transformations** as shown, for example, in the phraseme *přirůst někomu k srdci* (grow onto sb's heart, i.e. become dear to sb) which does not allow transformations and so to be used in questions, the passive, conditional, imperative or in the first person singular or plural. This type of anomaly is particularly noticeable in the restrictions of textual uses of the phrasemes, restrictions which are always highly individual and the user must know them just as he has to know the meaning of the phraseme (see 6.3).



Since different kinds of anomaly tend to associate in different constellations and numbers with particular phrasemes, the extent of their presence may be used as an indicator of the degree of idiomatization (phraseologization) of such phrasemes. The lowest is found in peripheral quasiphrasemes and quasiidioms (see below, also 3.6). If a given combination displays no anomaly, it is not a phraseme.

There are many kinds of anomalous features and it is therefore possible to use them in the functional classification of phrasemes and idioms (see 5.3).

## 2.4 The Definition of the Phraseme and Idiom

The phraseme and idiom cannot as a rule be defined by a single (anomalous) property, whether formal, semantic or collocational. Nor does the current, and perhaps, most widespread conception of the phraseme and idiom – as *a fixed and reproducible combination of elements (especially words) whose meaning is (partly or completely) not deducible from the meaning of the components* – meet all their types (such as phrasemes with a monocollocable component) and therefore it has to be rejected. The test of analogous substitution of components mentioned above (2.3, the example *otevřená hlava*) may be formulated operationally as follows

*the idiom and phraseme is a unique combination of minimally two elements, one (or more) of which does not function in the same way in another combination (combinations), or it occurs in just one expression (or a severely limited number of such expressions).*

Even this definition, although it implicitly suggests the presence of more than semantic peculiarity and can be used especially for the practical **identification** of a phraseme and idiom, one-sidedly stresses the collocational aspect. The **test of commutation** of some of the elements may in this sense be usefully applied in practice (after testing the primary aspect of fixedness of a given combination) for the identification of phrasemes and idioms in textual analysis and for excluding regular cases which follow the standard rules.

Still it is evident that all the features discussed, and others that could be mentioned, cannot be squeezed into an unambiguous specific definition applying only to the phraseme and idiom. The very status of these features, defined negatively as anomalies, which are perceived against the background of the regularities of language, is relative. It depends on the elaboration of the concept of regularity (or regular generability) and its explicit meaning in the linguistic description according to a specific coherent theory of language. However, in a generalized form, stressing the fact that all these features are only manifestations of the properties of the language system and structure, it is possible to sum up these features and render them more abstract. This leads to the following formulation:

*the phraseme and idiom is such a non-model and fixed syntagma of elements of which (at least) one is with respect to the other a member of an extremely limited (both formally and, mostly, even semantically) and closed paradigm.*

By an element is meant a standard linguistic unit from whatever level (except phonological), although in actual fact it is only a certain shape, form of this unit, i.e. a particular word form,

morph, etc. Expressions like a ‘combination of words, sentences’, etc., are therefore used in this connection with a certain amount of licence. Two, or more, phrasemes and idioms, however, can by no means come together as elements of an idiomatic syntagma of a higher order. In order to preclude the definition from covering also multiword technical terms (such as the bird’s name *vdovka královská* – *vidua regia* – where the component *vdovka* can be used with only several other adjectives, e.g. *vdovka rajská*), further features have to be taken into account. The feature of mere restricted collocability does not turn a combination into a phraseme (such cases are only quasiphrasemes, see 3.6).

Phrasemes and idioms also need to be distinguished from literal homonyms, cf. *otevřená hlava* (open head) in the sense of being subject to a ‘surgical’ operation), and others, *chytat lelky* (to waste time, dawdle, literally to catch fern-owls), *zlatý důl* (gold mine), *házet flintu do žita* (to give up prematurely, literally to throw one’s gun into a rye field), whose literal counterparts are usually rarer and have a completely different distribution in text.

### 3. The Phraseme and Idiom: Their Components and Structure

In phraseology and idiomatics (PI) it is imperative to make a distinction between the level of the input forms, components, and the level of the final combinatory outputs, the idiomatic wholes. Each one requires a special and separate approach, not mutually transferable. While the input forms can be analyzed before they enter the idiomatic combinatory whole which is not unlike that found in regular language (at this point structural classification may still be used), analysis at the level the resultant whole is impossible inasmuch as the input components have lost their original external character and form components of the total combination only conditionally. On the other hand, the output level offers a rich array of possible kinds of functional analysis, based on textual behaviour. The nature of the resultant idiomatic phraseme is essentially individual and arises from the sum of the individual anomalies it displays. These anomalies can be seen as a scale: **the more anomalies a phraseme displays, the more idiomatic it is and vice versa**, the fewer it has, the closer it is to regular language. As a consequence, this scaled approach suggests the possibility to identify an idiomatic centre of sorts for this whole area (separately for each combinatory level) and, conversely, its periphery with zones situated in between this centre and the regular language.

It also suggests the possibility of defining the frequent transitional cases whose degree of idiomacity is weak, i.e. quasiphrasemes and quasiidioms (3.6).

#### 3.1 The Character of the Phraseme and Idiom in General

The components of the phraseme and idiom (for their overall description see 2. above) include minimally the following ones: elements and their character, a combination including the sequence, fixedness, the number of elements and their functions which always display some kind of anomaly. In order to define the phraseme and idiom in more detail, these components have to be considered from the formal, semantic and collocational, or functional, aspects. In the last one also its anomalous transformations have to be taken into account (see

5.3). Phrasemes and idioms, which are the basic units of the PI of a language, are popularly and nonspecifically given various names with different meanings, such as **rčení** (or **úsloví**, **obrat**, etc.; saying, saw or phrase) at the level of non-sentential word combinations, and, at the sentential level, especially **příslví** or **pořekadlo**, or **pranostika** (proverb, adage, maxim, etc.). At the same time, sentential (propositional) phrasemes include a number of other types, such as **formulae**. Special cases are **similes** (they can be both collocational and sentential, see 4.314) and **binomials** (see 4.316), which may include non-phrasemes as well. Completely ignored until recently were **synsemantic phrasemes**, mainly of a collocational nature (see esp. sections 4.313 and 4.312). A considerable proportion of phrasemes, though, have no traditional, however imprecise, name, and in technical usage they are as a rule described functionally as phrasemes or idioms of this or that level, or according to a formal subtype within each group. For example *ukázat někomu paty* (show sb one's heels, i.e. take to one's heels) is classified as a collocational phraseme of the verbonominal type (according to its structure **V-N**).

**Elements** which formally enter phraseological combinations as **components** are of different kinds depending on their status (for a general description see below, especially 3.22); they are mostly morphemes (roots or affixes), lexemes, collocations, clauses or propositions (sentences). At the same time they typically occur only in a certain form (where relevant), not as a lemma. The resultant combinations, phraseological wholes, i.e. phrasemes, always – as **formations** – functionally belong to a higher level. To give some examples: the **lexical phraseme** *přestřelit* (to overshoot, i.e., overdo) and its components *pře-střelit* (compared with the regular *pře-hodit* – to throw over, etc.), the **collocational phraseme** and its component lexical forms *mít strejčka* (literally to have an uncle, meaning 'to have something in store' compared to the regular *to have an aunt*, etc.) or the **propositional phraseme** and its components *To je slovo do pranice* (That's a word to have a fight about, i.e. a matter of contention, compared with the regular *To je slovo do diskuse*, That's a matter for discussion).

The **dichotomy phraseme and idiom** (and **phraseology and idiomatics** for that matter) shows both a clear difference and also a certain overlap. By and large, it is a difference of perspective. In order to emphasize the divergent points of view, it is suitable to preserve the terminological difference. Basically, if a combinatory formation is analyzed in terms of formal features, we use the term **phraseme**, while semantic analysis, in terms of relevant semantic features (including denominative) justifies the use of the term **idiom**. This seemingly parallel approach is not redundant: on the one hand, it is often methodologically necessary to pay separate attention to the semantic aspect and the formal aspect because of their marked difference; on the other hand, the two aspects are not commensurate. With a decreasing degree and quantity of anomaly and with transition to the periphery of phraseology and idiomatics the semantic aspect tends to become markedly weaker, less distinct, which leaves the formal aspect dominant and in fact the only aspect to be analyzed and thus the term phraseme is the logical choice. Such a situation occurs, for instance, in phrasemes of the prepositional type, such as *na rozdíl od* (in contrast to), **quasiphrasemes** of the verbonominal type, such as *věnovat pozornost* (to pay attention to), technical **terms** such as *zlatý déšť* (golden rain or laburnum). For practical reasons, though, it is common to simplify things and speak of all units as phrasemes. Their formal typology

is clearly enough based on their hierarchical status and the typology of the formal aspects of the components.

Collocational phrasemes are functional equivalents of all types of word-class, the most varied ones being phrasemes with verbal function. The components of collocational phrasemes come from all kinds of word-class (appellatives and proper names), but do not usually conform to their syntactic functions; they are mere word-forms. From a semantic point of view, the most conspicuous, traditional and apparently universal type (probably found in all languages) is **somatic phrasemes** exploiting the names of different parts of the human body, in which the *anthropomorphic principle* becomes very prominent, cf. *vzít si něco do hlavy* (take st, i.e., get an idea, into one's head), *být někomu v patách* (be on the heels of), etc. The most frequent component *ruka* (hand/arm) appears in several hundred phrasemes in Czech. A number of somatic phrasemes are also used to describe gestures, e.g. *po/krčit nad něčím nos* (turn up one's nose at st), *mávnout nad něčím rukou* (to sweep one's hand over st, i.e. to show one's lack of interest in st), which then have two kinds of referent and two distinct meanings (in the latter phraseme it is the 'actual gesture' versus 'absence of care, interest'). These idioms are sometimes called **paralinguistic** since they combine (verbal) language and nonverbal communication. Nevertheless, such an approach to the phraseme based on the semantic type of its components tends to be rather impressionistic and has little value apart for the study of metaphor. Similarly, authors speak of **animal phrasemes**, **botanic phrasemes**, etc.

In spite of its discrete form, the phraseme and idiom functionally and semantically represents an integral **whole** which cannot be synchronically divided and analyzed into components. In other words, it is not the sum of the meanings of its components, nor is it based on the principle of compositionality of meaning. All traditional and recurrent attempts at semantic analysis (in contrast to diachronic analysis) are futile for they are bound to be incomplete and reliant on external criteria, cf. the phraseme *leje jako z konve* (it rains cats and dogs, i.e. it rains heavily) and the attempts to extract from it (at least) the meaning 'prší' (rains) for the component *leje*. If the dubious logic of this approach were to be followed through, then *jako z konve* would have to mean 'a lot', which it does not and the users never use this combination in this hypothetical meaning (cf. *\*snědl toho jako z konve*, he ate like from a watering can, *\*spí jako z konve*, he sleeps like from a watering can). This makes it easier to understand why components of a phraseme are only word-forms formally that are identical, homonymous with the forms of lexemes one knows from the dictionary, but not the lexemes themselves. Sometimes, however, such assumed lemmas are missing even in regular language, no matter how it may look on the surface, e.g. (*nechat na*) *holičkách* (to leave on 'little sticks', i.e. in the lurch) with the phantom lemma *holička*. Cases such as *leje jako z konve* must therefore be understood as being partly homonymous and quasimotivated (presumably having some motivation). Still, most phrasemes do not allow even such a partial pseudoanalysis. This also makes it clear why the phraseme is not directly and in a measurable way motivated by its components.

PI is considerably differentiated even in sociolinguistic terms, subsuming a full stylistic range of Modern Czech from argotic phrasemes and idioms to poetic ones. As it

is not – in contrast to the rest of the lexicon – subject to standardizing efforts that would deform phraseology in its own inimitable way, phraseology and idiomatics displays a distinct and natural degree of **variability** (variation) reflected in the multitude of alternating variants. These, unlike the simplified and artificially codified language, exhibit the natural variable, alternative state of development of language in this particular sphere. On account of its expressive appositeness, PI is the subject of particularly conspicuous efforts to introduce new and new innovations, **authorial variants**, or even new phrasemes and idioms. An overwhelming part of PI is naturally coined and used in spoken language, mostly non-standard, for it is here that the functional need for the kind of expression which phrasemes and idioms represent is most urgently felt.

## 3.2 The Form of Phrasemes and Idioms

### 3.21 The Structure of Phrasemes and Idioms

The **structure** of the phraseme cannot be internally determined, as we have seen, for reasons of the impossibility of it being objectively analyzed and of its multiple anomalousness. Instead, what is meant by structure is usually only the external form of the phraseme, based on the similarity between its component forms and the forms of external, regular language. Likewise, especially the word-class status of the components is analogously invoked, but only to the extent to which it is clear. In reality, its structure thus determined says very little about the nature of the given phraseme, and plays only an auxiliary role. The phraseme structure conceived in this way then consists of the prototypical and invariant combination of the forms of various word-classes (including another, so-called monocollocable, class) in a given sequence (order), together with the explicit relator (if present, i.e. usually a preposition or conjunction) which connects them (if not present, the relation between them is only implicit), and valency. Let us consider phrasemes and the notation of their structure based on this approach

<i>vzít nohy na ramena</i>	<b>V-N-prep-N</b>
<i>mít na někoho spadeno</i>	<b>V-na-V</b> <sub>participle</sub>
<i>mít pré</i>	<b>V-mono</b>
<i>utíkat {až se za ním prášilo}</i>	<b>V-Cl</b> (the clause may be developed as necessary)

In this way, for example, the collocational phraseme *mít z něčeho těžkou hlavu* ('have a heavy head about st', i.e., be worried about st) (**V-A-Nprep**) formally consists of **three components** (*mít*, *těžkou*, *hlavu*), one **relational element** (relator, *z*) which explicitly specifies the relationship between them (though it may be missing in other phrasemes) and the **valency element** (valency, *něčeho*), which is realized and specified in context. It is represented in the system only as a general and potential feature. At this level (i.e., that of the collocational phraseme, the commonest lexical subtype), we may regard as the components members of the four main denominative and lexical word-classes, i.e., nouns, adjective, verbs, and adverbs. Depending on the level and subtype, the nature of the components may vary. We may find, for example, a collocational phraseme composed of only function words, such as *co do* (**pron-prep**), *co když* (**pron-conj**), *že jo?* (**conj-interj**), a collocational phrasemes made up of mixed components, such as *od oko*, *pro tentokrát*, or a polypropositional

phraseme made up of propositional components, that is clauses, e.g. *Jak si kdo ustele, tak se lehne* (**PROP-PROP**), etc. (see 4.31 and the following).

In some cases it is useful to consider the possibility of **substituting** lexical components. If phrasemes of this type contain a pronoun or a numeral, these are regarded as **substitutes** of the noun or adjective depending on their function, not as further components, cf. *umět to s někým* (**V-pron-prep**), *páté kolo u vozu* (**num-N-Prep-N**), see 4.312 and 4.311. Substitutes (especially of nouns and adjectives) are numerals (**num**) and pronouns (**pron**), cf. *mít na to* (**V-prep-pron**), *dát si dvacet* (**V-num**). Reference to the presence of substitutes is useful mainly in the assessment of the general correlation between a combinatory type (the phrasemes structure) and its subtypes (an instance with substitution is then a subtype) and in classification, cf. **V-prep-pron** viewed as a substitution type of the general structure **V-prep-N**. A substitute, relational or valency element, may but need not be present in the phraseme (although generally they are common to all types).

Examples of this may be found in all specific combinations of two to three components (these are the commonest structures), made up of various forms of the elements, i.e. words of the lexicon. Of course, an idiomatic combination may have more than two elements, or components, cf. *být někomu věrný za hrob* (be faithful to sb 'beyond his/her grave'), *neprodát svou kůži lacino* (not to sell one's skin cheaply, i.e. not to give in easily), etc.

**Fixedness** of these combinations, which makes their repeated use – easy reproducibility – possible, is manifested by the fact that they enter the context as stable wholes, syntactically indivisible, but also by having fixed word order of the components, which usually hinders the acceptability of experimental variations, cf. *\*těžkou hlavu z toho má*, *\*to s ním umí*, *\*u vozu páté kolo*, *\*do co*, *\*když co*, *\*Tak si lehne, jak si kdo ustele*, etc. The impossibility of such transformations is one of the typical anomalies found in the phraseme.

### 3.22 Components of the Idiom and Phraseme

In PI, a **component** is a constitutive element, a constituent of the phraseme, characterized by a certain form and a certain position in the structure of the whole. Although, for example, components in the collocational phraseme composed of lexical words do not have clearly identifiable word-class functions on account of the overall anomaly of the phraseme structure, outwardly they seem to have it – consider the two 'lexical' components (*vzít*, *hůl*) and the **relator** which connects them (in this case the prepositions *na*, elsewhere it may a conjunction) in the phraseme *vzít někoho na hůl* – 'to take sb on a stick', i.e. to rip sb off (with *někoho* indicating **valency**). A different subtype may, however, have another 'word-class', such as a preposition, interjection or particle, or even a clause, cf. *být pro* (be for), *je mu hej* (he feels hurray, i.e. fine), *až na, jen aby, utíkat {až se za ním prášilo}*. The last example illustrates an **interstratal phraseme** whose components come from two levels (strata) and the non-existence of a neutral form is indicated by braces {}. In the case of a **lexical phraseme** which is formally a one-word lexeme, the components come from the level of morphemes; in practice it means that lexical phrasemes are mostly derived words and compounds (see 4.33).

### 3.3 Meaning and Denomination

#### 3.31 The Nature of Meaning in the Idiom and Phraseme

The phraseme is always a specific **naming unit (denomination)** which becomes specific in text through selection and possibly by contextual modification (modulation), but it is not formed in the same way as regular naming units. Its contextual specification depends on the relevant aspect emphasized and required in a given context.

The absence of the compositional relation between the meaning of the components and the meaning of the whole structure results in the **view of the meaning** of an idiom as integral, holistic and indivisible (see also above). It is reflected in the impossibility of this (typically) unmotivated meaning being (completely) re-segmented, decomposed in the respective input meanings of the components. The previously held idea of a semantic shift does not explain the situation, for cases of the type *stará bela*, *růžové brýle*, *tlouct špačky* can hardly be regarded as a shift. Standard cases of shifted meaning exhibit a clear and regular relationship to what they have developed from, which does not apply here. Nor are the cases of the type *železná zásoba*, *mít těžkou hlavu*, or *vyjít stříc*, *vydat všanc* attributable to any shift based on metaphor or metonymy. They differ in both the collocability and the compatibility of the components, and other things. Similarly irrelevant and inappropriate is the approach attempting to account for the meaning of idioms in vague terms of figurativeness, transparency, etc. (which are merely different terms for metaphor and metonymy).

The integral meaning of an idiom is the outcome of a number of factors of a paradigmatic and syntagmatic nature, differing not only in their variability and the manner of their presence, but often even their absence. Although the anomalous and non-motivated nature of meaning (with relation to the components) prevents its immediate classification, it is nonetheless possible to determine the character of the components outside the phraseme. An idiom draws its components from regular language in which objective and measurable rules hold and these are projected into the classification of PI. Only exceptionally does a component of an idiom happen itself to be an idiom, cf. *doprovodit někoho na poslední cestě* (to accompany sb on his last journey to the grave) a *poslední cesta* (last journey); cases like *sud prachu* (powder keg) are usually a secondary transformation of a broader idiom (*sedět na sudu prachu* – to sit on a powder keg).

Comparison with another language makes many idioms look as if they are untranslatable, i.e. as if they had no equivalent. This untranslatability of idioms, once so emphasized, is of only a marginal heuristic value and generally proves to be so relative as to be almost irrelevant. Given the ca. 7–8000 existing languages in the world, each time this would, at least in theory, have to yield a different set of untranslatables.

Where regular language may express, for example, a recommendation and restraint to make a sober and rational judgment only by a combination of (non-specific) lexical means and suprasegmental means, PI can often supply a ready-made economical expression for

just this occasion, created a long time ago, e.g. *Nevěř na duchy* (Don't believe in ghosts)! As regards a fixed **denomination**, it is possible to distinguish three cases of the **relation of the idiom to regular language**. The idiom

- (1) is a primary and monopoly means of expression (usually also highly economical) and regular language is not used for a given meaning;
- (2) is a parallel expression competing with the regular one;
- (3) does not obtain for a given meaning (and referent) and there is only a regular expression.

Examples to illustrate this are:

- (1) *vzít někoho za slovo, mít u někoho rozlitý ocet, mít máslo na hlavě, být hned na koni*,
- (2) *jít do sebe : zamyslet se nad sebou, mít nahnáno : bát se, dát si dvacet : zděmnout si*;
- (3) *leptat, letovat, lyžovat*, etc.

The types of meaning mediated by idioms in the first case cannot be captured lexicographically either briefly or fittingly, let alone by a single word. Nor is the second case, seemingly offering an equivalent, a simple one, if such things as pragmatic features, contextual function, etc., are taken into account. The equivalence of parallel expressions, fixed and regular, is very loose and the expressions are substitutable in only some (vague and not very specific) contexts. Therefore their distribution is not so much competing as complementary in texts of the same type, similarly as with some synonyms.

The third case is a good illustration of the fact that an idiomatic expression does not usually arise in an area of meaning tied up with a concrete referent. On the contrary, the main domain of the idiomatic is the sphere of meanings (whose referents are) 'non-concrete' with a distinct pragmatic component. That is why (also with respect to (2)) it is here that we may generally expect, though not predict, the existence of idioms for standard meanings of these types.

Such examples lead to a rarely mentioned conclusion that in verbal communication standard language does not have a monopoly on everything, nor can it cover it better than PI, and that the place of phraseology in language is not always on the periphery. The integration of the approaches and methodology used for regular language and PI, however, is still missing and the linguistics of regular language tends to view PI as something undesirable, as an aggregate of unpleasant exceptions not worthy of further consideration, if it recognizes PI at all. This makes it clear that the attempts made by some types of syntax to analyze a particular phraseme are preposterous and ridiculous.

The generally vague nature of the semantics of idioms is inevitably connected with the distribution of the types of meaning of phrasemes in language. At least the following negative comment can be made on the range of their meaning types: PI is strictly absent from those semantic and denominative spheres which, because of their concreteness and exactness, are the domain of regular language (not only in the sense of technical terminology). If phraseology is encountered there, it again involves distinct fuzziness and only a very loose parallel to regular meaning. It is instructive to consider especially the expressions of definiteness, quantification, locality and temporality and their treatment in idioms, e.g., *jeden každý; kdo měl nohy, ten tam utíkal; ani za mák; lítat po všech čertech; až naprší a uschne*.



### 3.32 The Elements and Aspects of the Meaning of the Phraseme and Idiom

The meaning of an idiom is a complex configuration composed of a number of features of a different nature. Compared to the meaning of a regular expression from the same level there are some important dissimilarities. In addition to the difference in motivation, the primary distinction is in the different distribution of the types of semantic features *in toto* and in the different relation between meaning and sense. In the sum of the semantic features of different types that constitute the two fundamental components of meaning, i.e., **notional** (conceptual) and **pragmatic** (connotative, or associative), it is the pragmatic meaning, characteristically permeating the whole of PI, that is far more noticeably and frequently represented in PI, and even in individual idioms, compared to the circumstances in regular language.

Among the **pragmatic** aspects which are part of the semantics of an idiom, if permanent and systemic, the most conspicuous are the expressive, symbolic and evaluative components, determined by the relations to the form, the input meanings of the constituents and by the relation to the usual (pragmatic) function of the expression. Unlike with the notional meaning, the pragmatic types of meaning cannot usually be assigned a distinct semantic feature unambiguously, or even at all.

The **expressive** component of the fixed meaning, based (1) **internally** on a marked deviation and contrast of either (a part of) the form or the meaning to the form, or to the basic, usual meaning, is a manifestation and signal of an anomalous relationship (between parts of the form or between the meaning of the components and the whole, etc.). Besides, even (2) **externally**, the very use of an idiom in text, where the idiom contrasts with the rest of the text, the surrounding context, tends to be expressive; it is general and more frequent. To illustrate (1), we may consider the expressive formal motivation in following expressions, *být kulatoučký jakou melounek* (onomatopoeia and diminutives), *přihřívát si svou polívčičku* (diminutive), *mít popotahovačky* (reduplication of the prefix), *dát si repete* (rhythmicity and the element 'e'), *mít rajc z něčeho* (the substandard character the component), and the semantic motivation in the expressions *mít pré* (monocollocability), *nechat na holičkách* (monocollocability and the apparent motivation of the component), *šlápnout někomu na krk* (incompatibility and taboo), *chytat lelky* (the absence of semantic participation), etc.

The **symbolic** component of the meaning, based on the possibility of association, can be clarified by the difference of two idioms whose notional meaning is close, if not synonymous. Their distinct associations make it possible for them to be used appropriately in distinctly different contexts and with different subject-matter, cf. *byl hned na koni* ('to be immediately on the horse', grab one's chance and triumph and exploiting the 'horse-riding' symbolic components, including the evaluative component of superiority, superciliousness, etc.) and *byla to voda na jeho mlejn* ('it was water on his mill' meaning 'the situation has changed to his advantage' and the symbolic 'mill' element, besides the notional 'pulling oneself together', viewed as swift or even rapacious, etc.).

The **evaluative** component of meaning, often accompanying its cognitive component, is tied up with its pragmatic component most closely and mediates the evaluative attitudes of the

speaker on the scale 'good-bad'. This can be shown by the following three expressions, *letí jako vítr*, *žene se jako k ohni* and *utíká, jako by mu zadek zapálil*. Disregarding the somewhat different intensity and the different images introduced by the components, it is possible to observe the different default evaluative judgments conveyed by the expressions. In the first, the evaluative component is distinctly positive (the speaker may even use the expression to praise), in the second the tenor is undecided and in the third the judgment is negative (the speaker typically uses it to express not only the actual denominative content but also derision and deprecation). However, depending on the demands of the context, the expressions frequently assume other, non-fixed evaluative meanings. Evaluation is usually only of the bipolar type **good-bad**, and can be subdivided according to the **sensory-intellectual** contrast (cf. *smrdí to jako bolavá noha* versus *je chytrý jako vopice*) and the **absolute-relative** contrast. While absolute evaluation is a matter of generally accepted axiology, relative evaluation is indirect and arises, e.g., in similes, from the existence of a currently accepted model and prototype, cf. *má hlad jako vlk*. A certain degree of evaluativeness is found in many functionally grammatical phrasemes too, cf. *za přispění* (always used in positive connection), *jen aby* (negative doubtfulness), and others.

The increased participation in the meaning of the idiom of **pragmatic** components, all of which tend to have, besides their fixed and systemic ingredients, also an ingredient part supplied by context (this latter ingredient is the domain of sense and textual pragmatics proper), influences the dependence of the idiom on **context**. In principle, this dependence is higher than in regular expressions. Thus in the above three and other semantic components (especially in the degree of intensity) the idiom draws heavily on context. The dependence is particularly strong in propositional idioms. However, the close connection of the resultant meaning (sense) of the idiom with the context is due to yet another factor, a high degree of the **potentiality and uncertainty**, i.e. **vagueness** of its meaning. Hence idiomatic meaning is relatively easily and frequently susceptible to **modification** and updating, i.e. to being shifted and expanded. Consider *dát někomu košem*, whose fixed (notional) components of meaning include refusal of a suitor by a girl, but recently also refusal of refreshments by a guest, and the expression may be used in many other types of context, such as refusal of a gift, etc., without employing special invention or metaphor.

In contrast to polysemy, **vagueness** implies indefiniteness, i.e., an incompletely specified semantic scope of the linguistic, textual form, cf. *přijít/být jen tak*, *zaonačit to nějak*, *dát někomu co proto*, and others. This denotatively unfocused impression or application of a linguistic form may be the result of usage (data need not be always sharply delimited) or of the speaker's conscious use as when he either does not want to or cannot express himself precisely, cf. *No von to Karel to ..., von je takovej divnej*, or *zatentovat, přitentovat*. In varying degrees, vagueness is an essential property of most phrasemes, a property which, however, is inbuilt in their systematic function. This characteristic property of phrasemes, which accounts for their frequent use, comes to the fore in contrast with precise forms of expression, especially terminological.

This **potentiality** of another sense and usage, due to the instability and merging character of the semantic borders between the fixed and unfixed meanings, is particularly strong in idioms with a monocollocable (and thus unmotivating) component, e.g., *být hin*.

This last feature of meaning has its consequences even for the status of **polysemy** and **monosemy** in idioms. It is not always possible to determine the dividing-line between the fixed meaning, however rare, and the potential unfixed meaning unambiguously, and the reason is not just the different usage of different speakers. Moreover, not even within the sphere of the fixed meaning is the determination of individual (notional) senses always equally reliable. Both of the above senses in

- (1) *dát někomu košem* form polysemy of one type (with the semantic intersection 'refusal'), but the senses of
- (2) *být jak vosk* (material: 'be soft'; man: 'be yielding, submissive') are different, more apart; while expressions of the type
- (3) *dát někomu co proto, vzít to hopen* can be considered polysemous with only a difficulty.

Clearly, an idiom may be polysemous, but the degree of its polysemy which in itself involves several types and depends on the nature of its components, is much lower than that of regular single-word expressions. The importance of the contribution made to the idiom's meaning by the context according to the specific goals of the communication and the very openness of meaning to such contextual modulation attests to the strong and characteristic vagueness of the idiom. This vagueness is responsible even for the seeming polysemy of the above cases. We may see this in the apparently separate meanings 'to beat sb' and 'to tell sb off' in *dát někomu co proto* which in reality are the commonest type of contextual concretization of the very broad 'to punish' which can be expanded by a number of other frequent and less frequent realizations, such as 'to give a good chase' (in a race), 'to exhaust' (e.g., by a large amount of food) or 'to make sb's work difficult' (e.g., by using a bad format).

### 3.4 Variants of the Phraseme and Idiom

Thanks to, among other things, the fact that phrasemes have luckily not been prescriptively standardized in Czech, the changeability of their development is reflected in their **variants** in a true and undistorted manner. It is possible to estimate that as many as every fourth phraseme has a variant of one sort or another, cf. *být/stát na výši doby, nasadit někomu veš (do kočichu)*. Variants have to be distinguished from functionally distinct transformations of phrasemes (see 6.). A variant is a fixed or unfixed, relatively small formal or semantic modification of a phraseme which does not alter its function (see transformation) and essentially not even its meaning, cf.

- (1) *do bílého rána/dne, prohnat si hlavu kulí/prohnat si kuli hlavou,*
- (2) *dostat (jich) nepočítaných, nasadit (vlastní) krk.*

A different type from systemic and fixed variants is the group of **unfixed, authorial variants**. Variants come in quite a range of types and include (often in combination) lexical, morphological and syntactic, or even phonological, phenomena. The overview below presents especially these types of variant (denoted by brackets and a slash):

#### I syntagmatic

##### A component

##### 1 quantitative:

**a** additive: *přinést někomu něco (až) k nosu, nasadit (vlastní) krk*

**b** reductional: *dostat (jich) nepočítaných*

**2** qualitative

**a** contact (esp. inversion): *anděl strážný/strážný anděl, sypat si hlavu popelem/sypat si popel na hlavu*

**b** distant (separability): merges into grammatical transformations (6.)

**B relational** (external, valency)

**1** explicit: **a** synsemantic (esp. prepositional): *ohrnovat nos (**nad něčím**)*

**b** inflectional (esp. case ones): *měřit (**někomu**) dvojím metrem*

**2** implicit: *být na místě (**někde**)*

**II paradigmatic**

**A component** (usu. synonymous): *do bílého rána/dne, být s to/šton, slzavé údolí/údolí slz* (merges into transformations)

**B relational** (internal): merges into transformations

Sometimes it is difficult to determine which of the several variants is the basic and representative one, i.e., an invariant. An **invariant** is in principle an unmarked form which is both neutral (from various points of view) and the most frequent, and which serves as the most suitable starting-point for assessing variants (the rest of the fixed forms of the given expression) and for making transformations. In Czech these requirements are routinely met by the infinitive, the nominative, or the positive. If the phraseme does not have any of these, then the most frequent form should be viewed as the invariant (for example, 3<sup>rd</sup> person sg, present indicative, etc.), cf. *Abys to nezakřik(l)!* (2<sup>nd</sup> person sg represents here all others), or the case of a complex structure **V-S** (i.e., **V-CL**, verb and clause) *říct, co /mu/ slina na jazyk přinese*, or *být rád, že /je/ rád*. A clause combining with a verb in the first position cannot, however, be completely generalized (notation */mu/ /je/* is substitutional); still, it is obvious that we are always dealing with a single expression, not a set of several dozens of fixed sentences differing in only particular morphological categories.

**3.5 Semiotics of the Phraseme and Idiom**

Phrasemes represent a semiotically diverse group. However, their character, usually made up of two components, is conditioned by their formation and function, not the semiotic nature of the constituents which may be dissimilar. The pervasive **pragmatic** aspects of a phraseme as a sign are mostly due to its **evaluativeness**. Its denotative side is distributed among secondary icons, indexes and symbols rather unevenly. There is a fundamental difference between **referents** (denotata, symbols) of a concrete and those of an abstract nature which is where the borderline between standard phrasemes and quasiphrasemes often runs, owing to the typically low level of anomaly in the latter ones (cf. *prásknout do bot* and *zout si boty*). Most collocational phrasemes have referents of such an abstract nature. There is no pronounced semiotic correlation between types of phrasemes and their sign character, although there are traces of it, cf. the functional subtype of relational phrasemes **prep-N-prep** (*v souvislosti s*) and the closeness to indexal phrasemes. Also, it is interesting to compare the secondary phraseological icon *kamenný obličej*, the index *nějaký ten* and the predominant, or most frequent, type among phrasemes, i.e., symbols such as *otevřená hlava, hodit něco za hlavu*. The important thing is that denotatively components of these phrasemes, prior to becoming part of them, may not only be of a different nature (cf. the indexal *ten* in the

symbol *být ten tam*), but may even have no referent at all (as in monocollocable elements, e.g., *pré in mít pré*). Similarly in propositional phrasemes, whose semiotic properties are even less explored, it is impossible to provide a straightforward classification. Even Peirce's proposition can be applied only very loosely (secondary proposition) to a propositional phraseme (e.g. *To se říká kdoví jestli!*, *To by v tom byl čert, aby ...!*), among other things, because of its fixedness.

### 3.6 Quasiphrasemes and Quasiidioms

Every phraseme and idiom can be always defined in a complex way by several criteria at once, most of all by the degree of its anomaly. The fuzziness of the borders of PI and the gradual transition to regular language are responsible for the fact that in some types of phraseme the number of relevant criteria and manifestations of the anomaly is declining. The PI periphery, including the zone of transition to regular language, is formed by quasiphrasemes and quasiidioms. Besides the application of individual criteria the assessment of the centrality or peripheral position of an expression in this sense must also take into account the extent of its idiomaticity (see 4.3, 5.3) and its frequency.

The principal types of quasiphrasemes and quasiidioms are indicated by the following survey:

- a) expressions with a concrete referent: *kočičí hlava*
- b) expressions of a terminological nature (there is an exact definition of their meaning): *mrtvý chod*
- c) some expressions with synsemantic components: *ve srovnání s někým* (what is specifically their own is just the collocability, meaning and function)
- d) some expressions with a desemanticized component: *bejt napakovanéj* (to have money, to be loaded), *bejt potřebnej* (to need money) (the adjective can appear only in the predicative position)
- e) some expressions with limited collocability of the components: *zarytě mlčet, otevřít dokořán, širé moře, poulit oči*
- f) some verbonominal expressions with an abstract noun: *dávat pozor, věnovat pozornost* (limited collocability of abstract nouns), etc.

## 4. The System of Idiomatics and Phraseology and Classification

Denominatively, PI forms a system of fixed combinations of various types and functions at several levels and the system is complementary to expressions which are both regular and unfixed. An attempt to produce a formal outline of the structure of the individual PI types indicates the presence of a formally complex **system** which is not found in regular language.

### 4.1 Distribution of PI in Language

The occurrence of the PI units, phrasemes and idioms, at several levels of language to some extent also indicates the internal hierarchization of PI. The fact that the phraseme

typically belongs to a level higher than that of its components which combine to make it is described by the following diagram (+ denotes occurrence, positive representation):

<i>Level of components → of expression ↓</i>	<i>A morphematic</i>	<i>B lexical</i>	<i>C collocational</i>	<i>D propositional</i>
1 morphematic				
2 lexical	+			
3 collocational	+	+		
4 propositional (sentential)	+	+	+	
5 polypropositional monosubject			+	+
6 polypropositional intersubject				+

On the basis of this external character of components it is possible to arrive at the most general hierarchization and so at the **classification** of phrasemes and idioms which uses the membership of the components in the same level (stratum) as its starting-point, cf.

2A **lexical phraseme and idiom** (*najít, předsevzetí*)

3A **collocational phraseme and idiom** in different senses (*ke všemu, co do, co když*, with synsemantic components)

3B (*chytat lelky, jablko sváru*, with lexical components)

3A–B (*k polednímu, od oka, bez okolků*, with components of both types)

4A **propositional phraseme and idiom** (*To tak! Že by?*, with synsemantic components)

4B (*Chcíp tam pes, Pozdě bycha honit*, with lexical components)

4C (*To je pro něj španělskou vesnicí*, with collocational components)

5C **polypropositional monosubject phraseme and idiom** (*Mladý kurvy, starý modlářky*, with collocational components)

5D (*Vono se řekne semenec, ale zobat ho z lahve!*, with propositional components)

6D **polypropositional intersubject phraseme and idiom** (A: *Pozdrav pámbu!* – B: *Dejž to pámbu!*, with propositional components in different speakers)

This survey does not include the rare cases of an interstratal phraseme and idiom, in which components from different levels combine, such as *dát někomu, co se do něj vejde* (lexeme + proposition, or clause), *tvářit se, jako by se nechumelilo*, and others. The area of PI thus includes all relevant types of fixed anomalous combinations with stable functions and meaning. While, on the one hand, combinations of phonemic components cannot be incorporated in PI for the absence of meaning, it is not at all clear, on the other hand, what everything should be included under PI from the higher-level fixed combinations, i.e. those above polypropositional phrasemes which belong to literary microgenres (e.g., riddles, fables, etc.). In the short history of PI, which at first traditionally concentrated on collocational phrasemes, there were periods in which large areas of combinations, especially propositional and lexical phrasemes and idioms, were sometimes excluded from PI for lack of knowledge. In these approaches, with PI being as vaguely defined as it was,

there was no attempt to tackle these areas in a different way. Instead they were completely bypassed and left unexplained.

The number of Czech collocational and propositional phrasemes and idioms can be estimated at some thirty thousand, which is clearly not a small and negligible number. And it is not a peripheral area of language either. In terms of frequency, these expressions range from those with the highest frequency and thus belonging to the *language core* in Czech (esp. expressions of a grammatical nature, such as *že jo, co když, to je jedno*), those used with medium frequency and finally those which are rare and on the *periphery of language* (cf. *rozsvítit jednu lampu, nejít někomu do noty*). It is in the last group that we find a great many obsolescent expressions and clichés, also some neologisms and fashionable buzzwords (cf. *To je o něčem jiném*).

## 4.2 Classification of Idioms and Phrasemes

There are more aspects than one from which phrasemes can be classified. Some are almost worthless because of the vagueness of their premise (such as the previously overrated stylistic classification), excessive generality (semiotic or sign-based classification), lack of reliable sources and information (diachronic and etymological classification) or the difficulty of application and unclear criteria (semantic or onomasiological classification, which does not even exist in the lexicon). Important categories of phrasemes could be arrived at on the basis of the classification of their functions (such a classification has not been developed yet, however, see 6.). Hence we are left with a formal-structural classification based on the evidence provided by the input components. Although these components do not tell us much about a phraseme, superficially most of them are familiar enough to speakers (with the exception of monocollocable words) and externally identical with words of regular language.

## 4.3 Formal Classes of Idioms and Phrasemes

Classes of phrasemes and idioms are best singled out on the basis of their formal structures whose survey offers a glimpse of their structural classification. In the following the PI classification will be confined to collocational phrasemes (4.31), propositional phrasemes (4.32) and lexical phrasemes (4.33) and their structures. Occasionally, a brief comment on their function will be provided as well.

### 4.31 Classes of Collocational Idioms and Phrasemes

The set includes phrasemes and idioms (collocations) whose components are words, or rather specific word-forms, and which do not form an utterance, a sentence, on their own. In terms of the variety of structures, phenomena involved and the number of expressions it is the most prominent area of PI.

#### 4.311 Lexical (Autosemantic) Words in the Roles of Components

Nouns, adjectives, verbs and adverbs allow the total of sixteen double-figure variations with repetitions (i.e., 4<sup>2</sup>), which represent the basic binary structures of this type of

phraseme. PI realizes here the full potential of all theoretically possible structures, i.e. all 16 of them, unlike regular language (which does not). Some of them belong to or are on the way to become propositions (they are marked by ! and are given here for the sake of completeness), others fuse as regards the order of their components (especially V-ADV and ADV-V, though there are cases, such as *bohatě stačit*, where the order is invariant). The rare and numerically little encountered structures include N-ADV and ADV-N, or N-A and A-A, while the richest structures, with internal subdivisions, are V-N, A-N and N-N. The following is a brief survey

- 1 **!N-V**: *pravda vítězí*
- 2 **V-N**: *číst někomu levity, tlouct špačky, chytat lelky*
- 3 **A-N**: *zlatý důl, černá hodinka, !šťastnou cestu, nejvyšší měrou*
- 4 **N-A**: *anděl strážný, bohem zapomenutý, !panenko skákavá, !paličák paličatý*
- 5 **N-ADV**: *místo nahoře, !klobouk dolů*
- 6 **ADV-N**: *vzhůru nohama, !mnoho štěstí*
- 7 **V-A**: *kápnout božskou, oplatit někomu stejnou*
- 8 **!A-V**: *zmýlená neplatí*
- 9 **V-ADV**: *přijít zkrátka, dopadnout bledě, držet se stranou*
- 10 **ADV-V**: *bohatě stačit, trpce litovat, zarytě mlčet*
- 11 **A-ADV**: *černež vzadu, !celej nakřivo, !malá domů*
- 12 **ADV-A**: *neblaze proslulý, andělsky trpělivý, zbrusu nový*
- 13 **N-N**: *jablko sváru, roh hojnosti, den brannosti, palma vítězství, obratem ruky*
- 14 **V-V**: *umět se narodit, dát se poddat, zůstat někde viset*
- 15 **A-A**: *chudý příbuzný, neslaný nemastný, staří mladí*
- 16 **ADV-ADV**: *široko daleko, kolem dokola*

This approach focuses on the word-class of the components and disregards their secondary and heterogeneous syntactic character. That is why even cases such as *božská, stejná*, etc., are formally analyzed as adjectives. A syntactic approach would entail a certain interpretation and breakdown into components, which is not in principle possible with idioms.

The **basic calculus** (binary) further subdivides in many directions, especially into cases involving components proper and (a) **substitutes** (esp. pronouns and numerals), cases in which the **relation between them** is implicit or (b) **explicit** (by means of a preposition or conjunction), cases with binary structures or (c) **subtypes of more components** (i.e. involving some kind of structural transformations), etc. From a semantic point of view, of the many possibilities the presence of abstract nouns in the structure **V-N** is particularly prominent (see 4.315). To give some examples

- 1 (a) **Pron-V**: *!ta píše, !to bodne*, **Num-N**: *dvaatřicet loupežníků*, (c) **N-ADV-V!**: *všeho moc škodí*, **Pron-N-V**: *co dům/týden dal*
- 2 (a) **V-Pron**: *zakřiknout to, být bez sebe*, **V-Num**: *olíznout všech deset, dostat pětadvacet*  
 (b) **VprepN**: *mít za ušima, lézt po čtyřech, být s to*, **Vprep(N)prepN**: *jít na něco od lesa*  
 (c) **V-A-N**: *mít krytá záda*, **VprepA-N**: *žít na vysoké noze*, **VprepN-N**: *mít z ostudy kabát*
- 3 (a) **Pron-N**: *moje maličkost, svého času, každým coulem*, **Num-N**: *jedním vrzem, jedna báseň, pár kroků*  
 (b) **AprepN**: *zralý pro šibenici*, **prepA-N**: *za velkou louží, v pravý čas*, **Pron-A-N**: *moje drahá polovička*, **prep-Pron-N**: *za každou cenu, za našich časů*



- (c) **A-N-A**: všemi mastmi mazaný, **A-A-N**: zlatá střední cesta
- 4 (a) **Pron-A**: nic kloudného  
 (b) **prepN-A**: za vlasy přitažený, po čertech chytrý, **NprepA**: obrat k lepšímu
- 5 (b) **prepN-ADV**: s hubou dokořán
- 6 (a) **ADV-Pron**: zhola nic, **Num-N**: pár kroků  
 (b) **ADVprepN**: jednou za čas
- 7 (a) **V-Pron**: být nesvůj, vědět své  
 (b) **V-Aprep**: být krátký na někoho, **VprepA**: tnout do živého  
 (c) **V-N-A**: mít to polepený, **V-ADV-A**: být ještě bosý, **V-AprepA**: plést páte přes deváté
- 9 (c) **VprepADV**: pít přes moc, **Vprep(N)-ADV**: jít od něčeho dál, **V-ADV-ADV**: nevědět kudy kam
- 13 (a) **Pron-N**: moje maličkost, něco na zub, kterej čert?  
 (b) **NprepN**: pohled pro bohy, místo na slunci, zmatek nad zmatek, **prepN-prep-N**: ve dne v noci **N-NprepN**: boj muže proti muži, **Pron-prep-N**: ta s kosou, něco na zub, **prepN-N-prep-N-N**: od slunka východu do slunka západu  
 (c) **N-N-A**: vlk v rouše beránčím, **N-prep-A-N**: agent s teplou vodou
- 14 (c) **VprepN-V**: umět v tom chodit
- 15 (a) **Num-A**: první poslední  
 (b) **AconA**: živ a zdrav
- 16 (a) **ADV-ADV**: tak nějak  
 (b) **ADVconADV**: tak či onak, zkrátka a dobře, dříve či později, pusto a prázdno  
**part-ADV-ADV**: ani dost málo, jen tak tak  
 (c) **ADVconADV-ADV**: víc než kdy jindy

The combinations of lexical (autosemantic) words are semantically and functionally extremely varied (see esp. 5. and 4.321).

#### 4.312 Synsemantic or Function Words in the Role of Components

For the sake of simplicity, function words proper (prepositions, conjunctions, particles) were supplemented with word-classes which are used as function words only in some respects (pronouns, numerals) or not at all (interjections). Formally, another category was added here as well, i.e., **monocollocable words**, quite special in other respects, in which not only the word-class status, but also the very status of a word, have been weakened in varying degrees. They are merely formal equivalents of words, verboids (they function as an eleventh word-class of an exceptional type in phrasemes). However, not even the classification of ordinary function words (especially particles) is always unproblematic, among other things, owing to their multifunctionality in regular language, cf. for example the particle-like nature of *ale* in *ale co!*. Although this is not a clear case of (a) substitution, some cases admit of this possibility (cf. **pron-pron** *nějaký ten* a **A-A**), while others do not (cf. *co do*). Similarly it is not usually possible to speak of (b) an explicit expression of the relation between components; in some types (especially particles), however, (c) subtypes are quite common. Out of the full calculus of 7<sup>2</sup> variations with repetition (49), only about two thirds are realized. In the survey, components are arranged and shortened as 1 **prep** (preposition), 2 **con** (conjunction), 3 **pron** (pronoun), 4 **num** (numeral), 5 **int** (interjection), 6 **part** (particle), 7 **mono** (monocollocable element). Here are some examples:

- 1A **prep-pron:** *mimo jiné, ke všemu, podle mého, mezi námi, beze všeho (!), podle toho!*  
 (c) **prep-pron-pron:** *k tomu všemu, z ničeho nic, prep-pron-prep-pron:* *pro nic za nic*
- 1B **prep-num:** *do jednoho*
- 1C **prep-int:** *do alelujá*
- 1D **prep-part-mono:** *na to tata*
- 1E **prep-mono:** *na just, na blind, na ex*
- 1F **prep-prep:** (b) *(všechna) pro a proti*
- 2A **con-pron:** *jako takový, a to*
- 2B **con-part:** *či spíše, či co*  
 (c) **con-part-part-part:** *když už tak už (!)*
- 2B **con-con:** *byť i, a sice, i když*
- 3A **pron-prep:** *co do něčeho, co za někoho*  
 (c) **pron-prep-pron:** *sám od sebe, co na tom!*
- 3B **pron-con:** *co kdyby (!)*
- 3C **pron-part:** *jaképak copak!*
- 3D **pron-mono:** *tě pic/péro!*
- 3E **pron-pron:** *nějaký ten, ta jeho, ten který, ten samý, jaký-takový, co to?*  
 (b) *ten a ten*  
 (c) *mé druhé já*
- 4A **num-pron:** *jeden každý*  
 (b) *jeden a týž*
- 4B **num-num:** *jedna dvě*  
 (c) **prep-num-num:** *o sto šest; num-prep-num:* *jeden za druhým*
- 5A **int-pron:** *bác ho!, ale co!*
- 5B **int-part:** *ale ano!, ale kdepak!, no a?*
- 6B **part-kon:** *ještě že, kdoví jestli*  
 (c) **part-kon-part-kon:** *a když už-tak*
- 6C **part-pron:** *no tohle!, no proto!, kdoví co/kdo*  
 (c) **part-(N)-pron:** *copak (Karel)-ten...!*
- 6D **part-int:** *no maucta!, a basta!, no jéje!*
- 6E **part-mono:** *ani muk (!)*
- 6F **part-part:** *což teprve, jen aby (!), jen jestli (!), to zrovna!, že by (?), že ano?, aby ne!,*  
*kdyby tak*  
 (b) *ano i ne*  
 (c) **part-part-part:** *taky že jo!, bode(j)t'by ne!, part-part-kon-part:* *tak tak že ne-*
- 7A **mono-kon:** *kór když*
- 7B **mono-pron:** *lautre nic*
- 7C **mono-mono:** *jakýs takýs, jakž takž, semo tamo, volky nevolky aj.*

Functionally unambiguous **numerals** participate in phrasemes as components least often (the meaning of their components usually perseveres). Surprisingly, though, not even they are completely excluded. Also **interjections** are not heavily represented. The most frequent type of component is **particles**. A characteristic feature of some phrasemes is a varying degree of separability of their components (distancing), e.g. *když už* (0/to udělá/...), *tak* (špatně, etc.); *copak* (Karel), *ten* (všechno zná!); *kdo* (to četl/...), *ten* (ví, etc.). A special category involves cases with a participating bound morpheme (prefix), e.g. (každého) *hned*

*tak ne-*(pozve, etc.), *co možná nej-*(větší, etc.). Many expressions function only as propositional phrasemes (they are marked by !, or ?), cf. *taky že jo!*, *co to?*. Some of them, however, are homonymous with collocational phrasemes, e.g., *Jen jestli!* and *Jen jestli* (přijde včas)!, and may differ even semantically. The function and meaning of phrasemes of this type are very diverse, although they mostly stay within the function and meaning of the traditional function words and words added to them, or also adverbs. Clearly enough, they markedly broaden and change this range.

#### 4.313 Lexical (Autosemantic) Words and Function Words in the Role of Components

Mixed combinations, or variations, of lexical (autosemantic) words and function words (and other types) in the role of components, theoretically give 2 x (4 x 7) structures. In PI, only less than half of them are actually realized. Substitutions do not occur, and also cases with an explicitly expressed relation between components and multiple-element subtypes are rare. Consider the following examples

- 1A **prep-N**: (1) *bez okolků, od oka, s bídou, do omrzení, o chlup* (+ Comparative), *proti všemu očekávání, za boha ne-, ani za mák*  
 (2) *s ohledem na něco, v souvislosti s něčím, na úkor něčeho, v intencích něčeho, v intencích někoho*  
 (3) *s ohledem na to, že/co... v souvislosti s tím, že/co*
- 1B **prep-A**: *k polednímu, za mlada, při nejmenším, až na další*
- 1C **prep-ADV**: *pro jednou, pro tentokrát*
- 2A **con-N**: *co chvíli*
- 2B **con-A**: *co živ*
- 2C **con-V**: *co nevidět*
- 3A **pron-ADV**: *co naplat!, nic moc (!), co a jak*
- 4A **int-N**: *ejhle člověk!*
- 4B **int-ADV**: *šup sem (šup tam) (!)*
- 5A **part-N**: *až hanba, ani nápad!, jen klid!, už ani slovo!*
- 5B **part-A**: *až černo*
- 5C **part-V**: *ani hnout!, jen říct!, ani nezabučet!, jen kvete!*
- 5D **part-ADV**: *až moc, taky dobře!, jak kdy/kde!*
- 6A **mono-N**: *baj vočko, kýho výra!, basama s fousama!*
- 6B **mono-A**: *fungl/zbrusu nový*
- 6C **mono-V**: *seč byl!*
- 7A **V-prep**: *být pro, stát mimo*
- 7B **V-pron**: (see 4.311)
- 7C **V-num**: (see 4.311)
- 7D **V-int**: *dělat na někoho bububu*
- 7E **V-part**: *mít něčeho ažaž, dostat co proto*
- 7F **V-mono**: *být tentam, být hin, být auf, být nasnadě, být zajedno s někým, být (moc) hr do něčeho*
- 8A **N-mono**: *mocí mermo, křížem krážem*
- 8B **A-mono**: *opilý namol*
- 9A **ADV-prep**: *až na někoho*

9B **ADV-pron**: *sotva kdo, jak kdo/co, zhola nic* (see 4.311) *a kdesi cosi; dost na tom že*

9C **ADV-part**: *jak to?*

9D **ADV-mono**: *skrz naskrz*

Structures with an adjective or a noun in the first position are not very frequent in PI, the reason is, among others, their prepositional valency. Because of a functional overlap, combinations from all of this area are numerically insignificant on the whole. The only exception is the relatively rich structure **prep-N**. Two of its subtypes (2–3) are a frequent source of **multiword prepositions** (*s ohledem na něco*) and some **multiword conjunctions** (*s ohledem na to, že/kdo/co/...*). A number of the structures belong to the level of propositions (see the mark !).

#### 4.314 Similes

The character of similes (comparisons) in the hierarchy of the PI spheres is exceptional. As a formation the fixed and formalized simile assumes mainly the form of a collocation. As a component (following the comparator or comparison marker) it may have a word, a collocation or a clause (sentence), cf. I1A (d) below. Hence the simile as a type, in a way, traverses different levels of PI. A formalized **simile** is an idiom which explicitly denotes the relation of similarity between the referent supplied by the context and a model which is part of this idiom. It is therefore based on a kind of semantic iteration and is close to some types of binomial. Of course, from a formal point of view similes come in different types. The similarity may be expressed in them implicitly, or in a non-fixed manner, cf. the lexical idiom *sněhobílý* (i.e. *jako sněh*, like snow), by means of a collocation, e.g. *holubičí povaha (jako holubice)*, *brčálově zelený (jako brčál)* or by a non-fixed combination *mluví nepřátelsky (jako nepřítel)*, etc. (Čermák in SČF1, 1984).

The semantic structure of a simile is /**Cd**–**R**–(**Tc**)–**c**–**Ct** where the abbreviations stand for /**Cd**/ **comparandum**, i.e. the referent (usually subject valency) provided by the context whereby the whole expression is made clear, e.g. */Karel/ je zarostlý jako Ezau*;

**R** **relator**, i.e. a component of a relational nature (usually a verb) which gives the idiom the character of predicate, e.g. */Karel/ je jako Ezau*, */kluk/ poslouchá jako hodiny*;

(**Tc**) **tertium comparationis**, i.e. a sign, feature, property, etc., which Cd and Ct share; it only optionally denotes the relation of similarity explicitly, e.g. */Karel/ je 0 jako slepý*; sometimes it coincides with R, e.g. */Karel/ řve jako lev*;

**c** **comparator** (comparison marker), i.e. a formal sign (usually *jak/jako*) which signals the relation of general similarity, specified in Tc;

**Ct** **comparatum**, i.e. a generalized and usually well-known model, prototype, to which the referent (Cd) supplied by the context is related by reason of similarity and thus explicated; it can be an entity, action, situation, etc., for example, */hrnec/ je děravý jako cedník*, */Karel/ byl červený, jako by ho krví polil; to je pěšky jako za vozem*, etc.

In 80 per cent of cases, **comparandum** (Cd) is man, or an animal (Czech, however, does not have similes specific only to animals), the remaining cases involve things, abstract concepts, e.g., places such as *je toho jako nastláno*, *být kyselý jako šťovík*. By contrast, **comparatum** (Ct), the source of similarity, is in 50 per cent represented by concrete objects,

in 20 per cent by animals, etc., for example *být studený jako led*, *mít hlad jako vlk*. Thus a simile is markedly concrete on the right (Ct), while the orientation towards man on its left (Cd) suggests strong anthropomorphization of this area. This, however, is typical of the whole of PI. The similarity expressed is almost always accompanied by a strong pragmatically evaluative relationship, cf. *je to kyselé jako šťovík* (i.e. unpleasant), and often by emphasis which may eclipse the actual similarity, cf. *uteklo to jako voda* (i.e. very quickly).

Two types of formal structures of the simile are included in the overview, (I) collocations and (II) propositions. Peripheral structures, whose high variability is due to the large number of the basic combinatory elements, expand the number of possibilities to several dozen structures. Besides (a) basic structures the overview outlines some (b) substitutions, (c) an explicit expression of the relation between components and (d) multiple-component subtypes. Conventional abbreviations mark the four lexical (autosemantic) word-classes used, **k** denotes the comparator (*jak/jako*); clauses are not specifically indicated to keep the overview simple.

- I 1A V<sub>k</sub>V:** (a) *lže, jako když tiskne*, (d) *řve, jako by ho na nože brali*  
**1B V<sub>k</sub>A:** (a) *být jako slepý*, (c) *jednat s někým jako rovný s rovným*, (d) *byl červený, jako by ho krví polil*  
**1C V<sub>k</sub>N:** (a) *poslouchat jako hodiny*, *pít jako duha*, (b) *dýchat jako čtyři*, (c) *přihnát se jako na koni* (d) *být jako pěst na oko*, *být jako otevřená kniha*  
**1A V<sub>k</sub>ADV:** (a) *mít se jako doma*, (b–d) *dělat něco jako nikdy v životě*  
**2A V–A<sub>k</sub>V:** (a) 0, (d) *je hubený, jako by cvrčky louskal*  
**2B V–A<sub>k</sub>N:** (a) *být lehký jako pírko*, (c–d) *mají se rádi jako kočka s myší*  
**3A V–N<sub>k</sub>N:** (a) *mít krk jako holínku*, (c) *mít nohy jako z olova*, (d) *mít hlavu jako vrabčí hnízdo*
- II 4A N–V<sub>k</sub>V:** (a) 0, (b) *je jich jak naseto*, (d) *to je, jako když dá hluchýmu dobrýtro, to je prašť jako uhod'*  
**4B N–V<sub>k</sub>A:** (a) *den je jako malovaný*  
**4C N–V<sub>k</sub>N:** (a) *je to jako broky*, *čas utíká jako voda*, (b) *uteklo to jako nic*, (c) *je to jako na potvoru*, (d) *leze to z něj jak z chlupatý deky, je toho jako z apatyky*  
**4D N–V<sub>k</sub>ADV:** (a) *to je jako naschvál*  
**5A (N)–V–ADV<sub>k</sub>V:** (a) *je tam jako nastláno*  
**5B (N)–V–ADV<sub>k</sub>N:** (a) *je to tam zatažené jako pytel*, (c) *to je pěšky jako za vozem*  
**6A N–V–A<sub>k</sub>V:** (a) *voda je teplá jako kafe aj.*

The overview makes it clear that most similes are based on the verb (or adjective) on the left and the noun (or deverbal adjective) on the right. In this sense, a typical simile has a formal structure as in *sedět jako pecka*, *jít jako svázaný*. Expressions of the type *chlap jako hora*, *den jako malovaný*, *lehký jako pírko* may be seen as different kinds of fixed transformation from input structures containing a verb (in these cases the verb *být*, to be).

#### 4.315 Verbonominal Quasiphrasemes

This frequent subtype with the structure **V-N** is formed using an abstract noun (i.e., one without a concrete referent and denoting especially quality, state, situation, action, event, activity, institution, phenomenon, feeling, construct, etc.), and should be therefore described as **V-N<sub>abstr</sub>**. Here are some examples

*věnovat pozornost, klást důraz, uvést do pohybu, tvořit skupinu, způsobit škodu, mít většinu, doznat změny, nabýt paměti* etc.

These phrasemes, which are usually classified as quasiphrasemes and quasiidioms, are very much characterized by the specific distribution (collocability) of the verbal component which is fully subordinated to the character of the noun (Čermák 1974). However, verbs which are predetermined by their nouns in this way are very few (6–8 on average). As it is a clearly structured group of verbs which appear with the same type of noun regularly, it can be regarded as a set of transformations (agential and statal, see 6.3). The main differences are in the expression of three kinds of **aspect of phase**, or state (**inchoative**, **durative** and **terminative**, abbreviated as **INCH**, **DUR**, **TERM**), and the presence or absence of **causativeness** (**CAUS**). The full structure, or macrostructure, which brings more of these phrasemes together, is then **INCH-DUR-TERM** and it can be doubled by causativity, indicating the presence of another agent. Compare the following examples of types

- 1 **INCH** *nabýt paměť*, **DUR** *mít paměť*, **TERM** *pozbyt/ztratit paměť* (and similarly *klid, síla, zájem*, etc.)
- 2 **INCH** *dostat hlad*, **DUR** *mít hlad*, (**TERM** *ztratit hlad*) **TERM-CAUS** *utišit/zahnat hlad* (and similarly *chut', žízeň*, etc.)
- 3 **INCH** *projevit/nabýt zájem*, **DUR** *mít zájem*, **TERM** *ztratit/pozbyt zájem*, etc. **INCH-CAUS** *vyvolat zájem*, **DUR-CAUS** *udržet/udržovat zájem*, **TERM-CAUS** *odradit /uspokojit zájem*

In an ideal case it might be possible to arrive at a six-figure combination and macrostructure. However, not even a three-figure configuration is always fully realized (owing to the type of meaning of the noun), or some of its phases tend to be rare, cf. *mít paměť*. Some cases subdivide into subtypes and, if fixed, move towards the centre of PI, cf. *mít krátkou paměť*, *mít dobré základy*, etc. The absolute majority of nouns which enter these relations are abstracts and these macrostructures represent their obligatory and fundamental mode of existence in text. Abstract nouns mentioned above in brackets (i.e., *klid, síla, zájem, žízeň, chuť*, and others) suggest that they behave analogously and enter the same three-member structure as well. Using degrees of delicacy of differentiation it is possible to arrive at about twenty such basic macrostructures for Czech abstracts. It must be stressed that each fixed combination (e.g., *mít hlad*, *vyvolat zájem*) is an independent quasiphraseme which must be seen as being clearly linked to others through some of its configurations, here only briefly hinted at.

In this type of phrase, or quasiphraseme, the collocational aspect is distinctly prevalent, while being characteristically limited in scope and quality. Other aspects are peripheral, such as the semantics of the noun component or the monocollapsible components in *kout pikle* and others.

### 4.316 Binomials

Binomials represent a very old type of structure, found in many languages, cf. the Czech *ze dne na den*, *den co den*, *z ručky do ručky*, *sakum prásk, jednou jedenkrát, tak či onak*, *Spejbl a Hurvínek*, or, for example, the equivalent (Czech, German, French, English, Swedish, and Russian) structures ***pes a kočka***, *Hund und Katze*, *chien et chat*, (*fight like*) *cat and dog*, *hund och katt*, *kouka u cobáka*.

The **binomial** can be characterized primarily as a collocational phraseme or idiom which is composed of a sequence of two constituents of the same word-class, sometimes even identical in form. It is usually of a nominal nature and is based on class or individual iteration (sometimes even founded on the same base). Its essential feature is coordination or juxtaposition and it typically exhibits fixed word order. Binomials may and often do comprise proper names. On the other hand, the strength of the iterative model is such as to be used in the formation of combinations which go beyond the framework of phrasemes, cf. *bratři a sestry*, *doba a místo*, *jídlo a pítí*. Although the binary and originally nominal structure is the most dominant, its strength to some extent provides a further impetus for the formation of analogous binary structures with different word-class functions, cf. the **formal** structures

#### 1 – Nominal structures

**N-N:** *hlava nehlava* (mono), *kniha knih*, *letem světem*, *sliby chyby*, *škoda přeškoda*

**N-prep-N:** *den za dnem*, *děšť se sněhem*, *marnost nad marnost*, *půjčka na/za oplátku ruku v ruce*

**N-con-N:** *alfa i omega*, *cukr a bič*, *den co den*, *děs a hrůza*, *krev a mlíko*, *muž a žena*, *nebe a dudy*, *Petr nebo Pavel*, *Romeo a Julie* aj.

#### 2 – Adjectival structures

**A-A:** *neslaný nemastný*, *zdravý nemocný*

**A-con-A:** *nečesaný a nemytý*, *širší než delší*, *takový nebo makový*, *vážení a milí*

#### 3 – Verbal structures

**V-V:** *chtě nechtě*, *vstáváje lehaje*

**V-con-V:** *být či nebýt*, *čekat a doufat*, *číst a psát*, *jíst a pít*, *rozděl a panuj*

#### 4 – Adverb structures

**ADV-ADV:** *skrz naskrz*, *široko daleko*, *zprava zleva*

**ADV-con-ADV:** *pusto a prázdno*, *pomalů ale jistě*, *snadno a rychle*, *tady a ted'*, *tu a tam*, *zas a zas*

Together with nominal structures this is the second largest group of structures while those based on the adjective and verb are much smaller. There are also others, though rare, cf.

#### 5 – Other structures

**prep-Pron-prep-Pron:** *pro nic za nic*, *pro mne za mne*

**Pron-prep-Pron:** *ten či onen*, *kdo s koho*

**prep –Pron-Pron:** *z něčeho nic*

**Num-prep-Num:** *páté přes deváté*

**Prep-kon-Prep:** *pro a proti*

**con-con:** *bud' anebo*

**Int-Int:** *tik tak*

**mono-mono:** *jakýs takýs, jakž takž, hógo fógo, láry fáry, saky paky, semo tamo, šupito presto, techtle mechtle*

In binomials different **kinds of semantic relations** obtain between the two iterated constituents, especially static ones such as opposition, synonymy, complementarity (cf. *kočka a myš, děs a hrůza, déšť se sněhem*), but also dynamic ones such as conversion, implication or alternativeness, and others (cf. *vzestup a pád, rozděl a panuj!, peníze nebo život*). They are supported by formal relations, especially heteronymy, homonymy, or alliteration and rhyme (cf. *chléb a hry, ze dne na den, kniha knih, kost a kůže, jakýs takýs*, etc.)

**Semantically**, the following, relatively narrow, range of senses is found with binomials at the level of the meaning of the whole:

**A – Identity:** *jeden jediný, šalba a klam*

**B – Totality:** *od kolébky do hrobu, široko daleko, chléb a hry, Jeníček a Mařenka, tělem i duší, v zimě v létě*

**C – Difference, alternative, choice:** *peníze nebo život, Kain a Ábel, kdo s koho, oko za oko*

**D – Indifference, or necessity:** *pán nepán, prašť jako uhod'*

**E – Intensification and emphasis:** *ze dne na den, poprvé a naposled, dál a dál, běda přeběda, kniha knih, hlava na hlavě, smůla na smůlu, stát ve státě*

**F – Immediacy, closeness:** *tváří v tvář, bok po boku, ruku v ruce*

#### 4.32 Classes of Propositional Phrasemes and Idioms

The basic difference from other types of phraseme is in the number of participating speakers and thus also utterances. Phrasemes with a single subject (speaker) can be tentatively brought together, although they can be subdivided according to their composition. They are propositional and polypropositional monosubject phrasemes and idioms (4.321). Standing in contrast to them are polypropositional intersubject phrasemes involving two speakers (4.322).

Propositional phrasemes which form a closed and fixed sentence were paid no attention in the past. A certain exception was proverbs which, however, were traditionally studied by ethnographers, not by linguistics. With time, some types of these phrasemes have been given various names, but terms like *adagium, aforismus, anekdota, apothegma, axióm, citát, floskule, formule, fráze, gnóma, hádanka, heslo, hlavolam, hříčka, kalambúr, loci communes, maxim(a), motto, obrat, okřídlená slova, parémie, pořekadlo, poučka, pozdrav, pranostika, průpovídka, přísloví, rčení, rozpočítávadlo, říkadlo, říkánka, sentence, slogan, spojení slov, truismus, úsloví, wellerismus, zaklínadlo, zařikávadlo, zásada (adage, anecdote, aphorisms, apothegm or apophthegm, axiom, battle-cry, bromide, byword, calembour, catch-phrase, catchword, clichés, conundrum, counting-out rhyme, dictum, epigram, formula, gnome, greeting, guide, incantation, loci communes, magic formula, maxim, moral, mot, motto, parable, phrase, platitude, precept, principle, prognostic, proverb, proverbial saying, pun, riddle, quotation, rhyme, saw, saying, sentence, slogan, theorem, truism, turn of phrase, Wellerism, winged words, witticism, word combination, etc.), overlap and were and still are unclear as far as their boundaries are concerned. Therefore in most cases they must be basically rejected.*



#### 4.321 Propositional and Polypropositional Monosubject Phrasemes and Idioms

**Propositional phrasemes and idioms**, which are already formally complete propositions (utterances), are made up of components from the lexical and collocational levels (A), while **polypropositional phrasemes and idioms** consist of components from the propositional level (B; characterized by the presence of at least two propositions). Both types agree in having **one speaker** in common, or a fixed subject (which may not be the same thing). However, quite a few cases converge, cf.

**A:** *Ani kuře zadarmo nehrabe., Do třetice všeho dobrého., Na Nový rok o slepičí krok., Cti otce svého i matku svou., Leze mu tlustý do tenkýho., Za zeptání nic nedáš., Chcíp tu pes., Zlatý voči!, Dobré pořízení!, Čelem vzad!, To si piš!, To zrovna!, Až po vás!, Ke všem čertům!, Jak to tak vypadá... etc.*

**B:** *Jak si kdo ustele, tak si lehne., Vrána k vráně sedá, rovný rovného si hledá., Komu dal pámbu úřad, tomu dal i rozum., Když to nejde po dobrém, půjde to po zlém., Ať je to ten nebo ten, vyženem ho z kola ven., Kdyby blbost kvetla, budeš samý poupě., Prodej písničku a kup si notu., Povídali, že mu hráli!, Nech to koňovi, ten má větší hlavu., Koukej, abys viděl!, Kde se vzal, tu se vzal... etc.*

##### 4.3211 The Formal Structure and Semantics of Propositional Phrasemes and Idioms

These phrasemes are built from different components which come from all types of word-classes including proper names and monocollapsible words, cf., for example, ... *fuk, fujtajbl, fušku, fáry, glanc, go, grešli, grogy, habaděj, hapruje, hecu, hej, hejsa, hlásku, hoj, hokus, hop, hopsa, hot, housti, hrana, hrklo, hromský, hráchy, huj, hýn, chťení, ježkovy, jímá, jobova, jonáci, just, kalné, kalup, kápě, kaput, karbid, kázáníčka, kejhák, klapačku, klempíruje, klíďo, kloudnýho, knap...* The components may even include clauses as their parts (in about 90 per cent), as in *Jak si kdo ustele, tak si lehne.* To a large extent, these phrasemes contain verbal components (about 85 per cent in all) and, as the form of the sentence is fixed, they can in some measure even be characterized: the verb is in almost half of the cases in the present tense, in 3<sup>rd</sup> person singular. An important finding is that most of these phrasemes have the form of the declarative sentence (over 70 per cent), while less than one third are imperative or exclamatory sentences and only a fraction (about 5 per cent) are interrogative sentences.

The structure of propositional phrasemes derives especially from transformations of the basic syntagma **N-V**, which itself, however, is directly employed only rarely, cf. the proverb, *Samochvála smrdí.* These phrasemes are also characterized by a typical **intonation**, especially in the form of a conclusive tonal pattern (yielding seven general types), e.g. *Co tě nemá!*, with the exclamatory falling-rising tone beginning with the last word. The basic formal structures subsume sentences of varying length, 2 to 14 words long. The **most frequent formal structures** have 4 and 5 words (about 19 per cent and 16 per cent respectively). There are ten basic structures, each having a different frequency:

- 1 To je N/A/ADV/mono/prep-N:** *To je vyloučeno., To je k smíchu.*  
and var. *Tím to hasne., V tom je jádro pudla., To se říká kdoví jestli.*
- 2 N-V:** *Věc se obrátila., Budou křtiny., Pravda vítězí., Čas běží., Než by řek švec., Není pochyb.*  
and var. *Jen kvete., Není zač., Ani hnout!, Ani co by se za nehet vešlo., Zdař bůh!, Ba co dím...*
- 3 N-V-N:** *Úsilí bylo korunováno úspěchem., Pálí ho dobré bydlo., To/všechno chce svý., Něco je ve vzduchu.* and var. *Kápní božskou!, Kdo by si pomyslel, že...!*
- 4 N-V-A:** *Vzduch je čistej., Je to s ním všelijaký., Přišlo mu to blbý.*
- 5 N-V-ADV:** *To hodně/něco pamatuje.* and var. *De mu to jedním uchem tam a druhým ven.*
- 6 ADV-V/mono-V:** *Pozdě bycha honit.*
- 7 N-N:** *Učedník mučedník., Mladost radost.* (marginal) and var. *Důvěra za důvěru., Pro dobrotu na žebrotu., Můj dům můj hrad., A ešě tu vo Červený Karkulce., Dobrou chut!, Chyba lávky!, Panna nebo vorel?*
- 8 prepN-A/N-N-A/N-ADV/A-ADV,** etc.: *Po účinku zlá rada., Samá voda., Malá domů., Dobrýho pomálu.* and var. *Starý zlatý časy!, Husa jedna pitomá!, Klobouk dolů!, Hlavu vzhůru!, Starou Blažkovou!, Houby s voctem!, Čelem vzad!, Čerta starýho!, Všechna čest!, Ve jménu zákona!, S dovolením!, Hanba mluvit!*
- 9 ADV-ADV:** *Zkrátka a dobře., Jak kdy., Jako doma!*
- 10 syn-syn:** *To tak!, Ale co! Kdo s koho., No a?, No ne!, Že by?, Ani za nic!, To zrovna!, Láry fáry!, Hej rup!, To jistě!, Na ex!, No těbůh!*

**Semantically** even these phrasemes are a specific type of naming unit. They denote especially (1) human manifestations, activities, attitudes and situations, often variously qualified (*Blázen, kdo dává, ještě větší, kdo nebere., Když Pánbůh dopustí, i motyka spustí.*), (2) human physical states, qualities and abilities (*Člověk na ní může oči nechat.*), (3) objects and artefacts, products, etc. (*To něco pamatuje.*), (4) general concepts, events and situations in the sphere of time, customs, law, etc. (*Jak se do lesa volá, tak se z lesa ozývá.*), and (5) external nature and the environment including the weather (*To je psí čas!*). Only rarely do they concern anything else, such as animals. A frequent feature, especially in proverbs, is the gnomic component of the meaning (*Člověk je sám sobě špatným rádcem.*). Other striking features include playfulness and facetiousness (*Když je něčeho moc, tak je toho příliš.*), sometimes connected with rhyme (*Sedni si, ať ti kapsa nevisí.*), a tongue-twister (*V kapli plakal pan kaplan.*) or a pun.

Both types of propositional phrasemes are characterized by a relatively low, or even zero degree of paradigmatic variability (transformability), and the **non-existence of a neutral form** instead of which the (most frequent) textual form is used. They are positively delimited by the existence of a fixed subject (if the phraseme has one) which is not subject to transformations. The possibility of a paradigmatic substitution of the subject, however, places the given combination among (verbal) **collocations**, cf. *bylo to* (but also: *udělat to/strčit to do kapsy...*) *dílem okamžiku*, i.e. *být dílem okamžiku*, compared with the **propositional phrasemes** *nedál/nedalo se svítit, páli/pálilo ho/jí!... dobré bydlo*, and others. Analogously also cases of the type *neví, kde mu hlava stojí* must be regarded as collocations, i.e. *nevědět, kde /mu/ hlava stojí* (their form requires the use of a metalinguistic notation). The number of relations among the components (partly viewed as logical, especially in proverbs) is higher than in collocations and,

as a result, even the iteration is less frequent here, cf. *já pán, ty pán; vem, kde vem; rozděl a panuj, zab se a uteč* (compare with binomials), etc. **Monocollocable** elements are here represented in the same measure as in all of PI, cf. *je po něm veta, do třetice všeho dobrého, na viděnou/slyšenou! trefil ho šlak, nic naplat, a basta fidli, ani muk!, šťastné pořizení!, etc.*

#### 4.3212 Functional Classes of Propositional Phrasemes and Idioms

Depending on their structural function both types of phraseme may be formally divided into the following **functionally** communicative **classes**

- 1 – vocative** expressing the reactions and attitudes of the speaker and the volitional influence on the listener: *Vida ho!, To mi může být ukradený!, Koukej mazat!, Mě vynech!, Trhni si nohou!*
- 2 – contact** expressing the relationship between the speaker and the listener where the actual contact often assumes the form of greetings: *Zdař bůh!, Tak se měj!, Dobrou chut!, S tebou sem domluvil!, To ti teda povím...*
- 3 – thematic** expressing the referential function of the speaker with regard to the theme (that what the communication is about), especially in the following subtypes:
  - a – **descriptive** (simple assertion or a statement and description): *Vzduch je čistý., Zatmělo se mu před očima.*
  - b – **attitudinal** (as an expression of evaluation and a stance): *To zrovna!, Opatrnosti nikdy nezbývá.*
  - c – **connective** (the link between parts and phases of the theme, rare): *Bylo, nebylo..., Jak se věci mají...*
  - d – **categorizing** (identification with a known type or prototype): *To je slovo do pranice.*
- 4 – metalinguistic** expressing the occasional activated relationship of the speaker to the code, the means of communication, i.e. language: *Mám to na jazyku., Na tom si zlomíš jazyk., Co jde šeptem, to jde s čertem.*

The contact type of function includes even the phraseological expression of **politeness**, which usually as an accompanying feature, permeates all types of propositional phrasemes. There are, however, more specific cases, such as *Budte tak hodnej a...!, Nechte si chutnat!, Není zač., Chovejte se jako doma!* and the corresponding functions of polite agreement, request, wish, response to thanks, and invitation. Politeness may directly form the contents of a concrete proverb, as in *S kloboukem v ruce projdeš celý svět.*

The correlation between the type of meaning, or function, and the form of the phraseme is usually inconclusive and in only some cases a clear connection may be detected. This, for example, is the case of the descriptive expression, especially in the thematic type (and in the factual class) and the indicative form. Similarly, in exclamations which belong to the vocative type a correlation can be found with the imperative form, etc.

#### 4.3213 Pragmatic Function of Propositional Phrasemes and Idioms

In many respects the **pragmatic function** is often more important in propositional phrasemes than the above ones. It comes particularly in the following five basic forms:

**1 Factual class of phrasemes (F):** the speaker – with respect to the listener – relates, or factually, commits himself to the truthfulness of the proposition (fact) and its different

levels and shades, or modifications. In practice, this means the introduction of what is commonly called **information**, although it is of many different kinds. The subclasses which make up this type include phrasemes expressing (a) assertion, possibility/necessity, in/definiteness, existence, (b) un/certainty, un/truthfulness, estimate, (c) im/probability, forecast, (d) in/dependence, un/conditionality, (e) restriction/extension/change of validity, generality, (f) categorization, assignment to a class, typology.

E.g., *To je stará písnička.* (Ff), *Hlad je nejlepší kuchař.* (Fa), *V hlavě mu už svítá.* (Fa), *To je vod tebe sice hezký, ale...* (Ff), *Straší mu to v hlavě.* (Fa/Exi), *Hlava ho pro to nebolí.* (Fc/Exg), *...a je po ftákách.* (Fa-e/Emc), *To je v háji.* (Fa/Emc), *To dá fušku.* (Fa/Emf), *Bůh ví.* (Fb), *Věřte si tomu nebo ne.* (Fc), *Člověk je to, co jí.* (Fd).

An extension of this class is represented by the **metalinguistic** subtype, cf. *Mám to na jazyku* (Fa) (compare with the metalinguistic function mentioned above). A characteristic transition between the types is the traditional category of **proverbs**, which on account of their gnomic nature, normativeness and exhortativeness belong primarily to the volitional type, although they may be used even factually, cf. *Hlad je nejlepší kuchař.* (Fa)

**2 Volitional (directive) class of phrasemes (V):** through their use the speaker wants to achieve some kind of effect in the listener, or make them do something. The role of a speaker is usually active and authoritative. The class includes a number of possible subclasses, such as (a) command, order, request, (b) dis/agreement, acceptance, satisfaction, unwillingness, permission, rejection, reminding, alerting, (c) threat, warning, admonishment, discouragement, (d) reassurance, promise, soothing, comforting, encouragement, (e) indifference, resignation, (f) advice, appeal, (g) wishing, (h) persuading, recommendation, instruction, (i) support, prompting, calming.

E.g., *Abys to nezakřik!* (Vc), *Kolik hlav, tolik rozumu.* (Vf), *Bud' rád, že seš rád.* (Vd), *Nevěš hlavu!* (Vd), *Di do háje!* (Vb/Emc), *To je hloupý.* (Vd/Fa), *Když je hlava blbá, trpí celý tělo.* (Fa/Vh/Emh), *Ještě to tak!* (Vb/Emc), *Dej si facku, nikdo se nekouká!* (Vb/Emc), *Pro mně za mně!* (Ve), *Táhni k čertům!* (Vab), *Dej bůh!/Bože dej!* (Vg), *Jen si dej, seš bledej!* (Vi).

The use of phrasemes often serves functionally, or pragmatically, a multiple purpose because both the speaker and the listener play different roles in their mutual relationship. A less distinct pragmatic nature is displayed by the broadly conceived group of **greetings**, which can be seen as the extension of particularly the volitional type and elliptically as wishes, cf. *Poroučím se., Tě pic!*, greetings ritualized for certain situations and specialized usage, cf. *Čest práci!, K počtě zbraň!*. The element of respect for rituals, organizations, churches, etc., and the binding nature of tradition are emphasized in **rituals** and their expressions. Especially in this area it overlaps, among others, with subtypes represented here by expressions of politeness (see above).

**3 Expressive class of phrasemes (Ex):** the speaker uses them to inform the listener of his interest in the situation, to express this interest, etc., mainly through his intellectual, rational attitudes. The role of the listener is sometimes passive here. The subclasses of the Ex class include especially (a) congratulations, (b) thanks, (c) apologies, (d) doubts, (e) sympathies, recognition, dislike, repulsion, disgust, disdain, (f) trust/mistrust, (g) accusation, reproach, condemnation, (f) forgiveness, plea for forgiveness, (i) joking, (j) helplessness (being at a loss), evasiveness.

E.g. *Vsadím na to hlavu, že...* (Exf/V), *Jen aby!* (Exd), *Klobouk dolů!* (Exa), *Hádej, můžeš třikrát.* (Va/Exi-j), *To je dlouhá historie.* (Fa/Exj), *To se říká kdoví jestli!* (Exd), *Vloudila se chybička.* (Exc), *Máš recht!* (Exe), *Sem ti dobrej jen k tomu, abych...* (Exg).

A certain extension of this functional type can be seen, for example, in **responses** to some types of expression, especially to **thanks**, cf. *Není zač.* Another extension is the expression of a **polite wish**, usually in the form of a conventional greeting, cf. *Dobrou chut!*, *Všechno nejllepší!*, *Tisícere díky!* (Exb), *Když dovolíte...* (Exh).

**4 Emotional class of phrasemes (Em):** by this type of phraseme the speaker conveys his emotional reaction to the listener, which is usually distinctly polarized evaluatively on the *good-bad* scale. Its subclasses include especially (a) surprise, astonishment, (b) admiration, praise, contempt, (c) pleasure, joy, refusal, displeasure, (d) envy, (e) satisfaction, disappointment, regret, (f) worry, fear, calm, confidence, (g) anger, vindictiveness, (h) derision, ridicule, sarcasm.

E.g., *Krev mu ztuhla v žilách.* (Emf/F), *Srdce se na to směje.* (Ema-d), *Dobře, že deš.* (Eme), *To je ale kus!* (Emb), *Krve by se v něm nedořezal.* (Emf/Va), *Jen houšť a větší kapky!* (Emc), *Ten má ale časy!* (Emd), *To jsem si dal!* (Eme), *To sme šlápli do hovna.* (Eme), *Kat aby to spral!* (Emg), *Boha jeho!* (Emg), *Neví, co chce, ale nedá pokoj, dokud to nedostane.* (Emh).

This class also subsumes phraseological **expletives and abuses**, which markedly differ from person to person. The strongest are usually used about the listener who is present, while oaths (today mostly without religious allusions and content) are typically impersonal or aimed at a third person, cf. *Ty kluku pitomá* (Emc), *Kat aby to spral* (Emg). However, actual oaths are relatively infrequent in phraseology. In terms of speech act theory some of these subtypes are perlocutionary in nature.

**5 Declarative class of phrasemes (D):** by using the phraseme the speaker establishes, constitutes a new situation, state, which is thus declared. It is a marginal class and its subclasses are not always easy to identify, but it is possible to give at least (a) dismissal, excommunication, (b) appointment, assignment, (c) condemnation, (d) implementation, establishment.

E.g., *Hej rup!* (Va/Dd), *To je hanebnost!* (Dc), *Dělají-li dva totéž, není to totéž.* (Dd), *Ať je po tvým!* (Db).

It is also possible to consider another group, **6 intensification class (I)** ... *až zrak přechází* (I), ... *co to dá* (I), but phrasemes of this class rarely occur in their pure form. On the other hand, another group, **7 evaluative class (Ev)**, is quite frequent, but even this one does not usually appear on its own in a pure form and combines with another class. So, the phraseme *To je věc* (Em/Ev) expresses both admiration and positive evaluation, etc.

#### 4.3214 Some Traditional Types of Propositional Phrasemes and Idioms

Some of the types of propositional phrasemes are known under their **traditional labels or names**. The best-known type is the old, familiar, and mostly ethnographic concept, **proverb** or **byword** (previously known in Europe as **paroemia**, and only later as **proverbia**, though Erasmus placed it under **adages**), which clearly belongs with sentential, i.e., propositional, phrasemes. The branch which deals with proverbs is called **paroemiology** (a term common in English in the mid-19<sup>th</sup> century). The classification of proverbs in the

past was largely tentative and cannot be taken seriously any more today. Also, proverbs are not the most important part of the propositional phraseme range, as they take up only a small section of it (some 10 per cent), although historically they have probably been the most prominent. Since the time of Čelakovský's work, the largest collection of Czech proverbs, the number of new proverbs has increased by some 40 per cent, while most of the proverbs gathered by Čelakovský are no longer in use today. The general character of proverbs which used to cover all areas of everyday life and the experience and truths derived thereof (cf. the "medicinal" proverb *Kam nechodí slunce, tam chodí lékař*, "the commercial" proverb *Zadarmo ani kuře nehrabe*, the "military" *Nejlepší obranou je útok*, or the "legal" proverb *Oko za oko zub za zub*) have newly developed and become so generalized that they are not thematically interpreted any more. The only remnant of weather proverbs is sometimes called **prognostics**. There are still several dozens of them, they are still used and can be regarded, alongside the proverbs proper, as a subtype with factual function (cf. *V máji nemá pastýřovi oschnout hůl.*, *Sv. Martin přijel na bílém koni.*). Their "forecasting" function, still applied, is betrayed by their opaque name in Czech (**pranostika**), which has been partly obscured (cf. Latin *prognostica* and Greek *prognosis*). The basic authoritative nature of the conveyed truth and experience which is not to be doubted has been preserved in proverbs even after they lost their legal normativeness. This is where their distinct quality, sometimes called transpersonal validity, or didactic function, etc., springs from. Semiotically and functionally, proverbs can be always counted on to have two direct participants (roughly: the "instructor/exhorter" and the "instructee/exhortee", etc.). In the background, though, there is always a third, anonymous participant present, the author of the presented statement, truth, instruction, etc., which cannot be disregarded without this entailing a social sanction. A similar character is displayed by many a quotation (see below) but there the third participant is known.

**Slogans** have the volitional function of persuasion and in spite of their ad hoc nature, conditioned by the usage and (especially political) preferences of the times, there are still several of them in Czech surviving, especially as cultural relics with a fixed form, cf. *Zpěvem k srdci, srdcem k vlasti.*, *Jsmé s vámi, buďte s námi.*, *Rovnost, volnost, bratrství.*

Phrasemes are also frequently used as expressions of **advice**, cf. *Dvakrát dává, kdo rychle dává.*, *Co tam, to tam, o to hlavy nelam!*. Like advice even **wishes** have volitional function. Basically, they are just as greetings based on ellipsis of the type *I wish ...*, cf. *Kéž by!*, *Čest jeho památce!*, *Chléb náš vezdejší dejž nám dnes*. Connected to wishes, however, are their frequent extensions or commentaries and replies to them, cf. at least *Starý lidi jsou jako malý děti*.

The **contact function** of phrasemes is usually and most frequently explicitly expressed through a **greeting** (see above), which in its broad sense is quite an abundant unit both in the variety of uses and incidence in the system of phraseology. **Contact between people** in a different sense than that mediated by greetings is described by phrase such as *Chová se, jako by se nic nestalo.*, *S kým mám tu čest?*, etc. (the distinctly phatic function is difficult to isolate though). Relatively frequent in the area of contact are also expressions of **thanks**, cf. *Díky bohu!*, *Zaplať pámbu!*, *Jako by se stalo!*. The form of **expletives and abuses**, which also have contact function, is usually exclamatory, cf. *A sakra!*, *Aby do toho (už)!*, *Di*

*k čertu!*, *Ty seš korunovanej vů!*. The contact area also subsumes all the other phrasemes that express, accompany or indicate various **social rituals**. However, most rituals, i.e. their occasional verbal accompaniment, **routine formulae**, do not belong among phrasemes because of their specialized and quasiterminological nature. Still, there are some positive phraseological cases relating to various traditional and observed **customs**, cf. *Dejž to pámbu!*, *S dovolením!*, *Srdečné blahopřání!*, *Čest jeho památce!*, *Sláva vítězům*, *čest poraženým!*, *Zlom vaz!*, *Pozdrav pánbůh!* – *Až na věky!*, *Jménem zákona!*, *Přísahám na svou čest a svědomí.*, *K budování a obraně vlasti buď připraven!*, *Arbeit macht frei.*, etc. A number of other types (see above), however well-known they are, have not been given any particular names. They include, for example, expressions denoting **superstitions** (*Musím to zaklepat* – Touch wood), and others.

A **quotation** (in phraseology only a fixed quotation) is the accurate repetition of an opinion or evaluation expressed usually by an outstanding personality of the past which generalizes and provides a topical comment, but also puts something in the context of the times. There are many quotations but only few of them are well-known and in common use, or remembered accurately, in the Czech milieu it is possible to mention at least President Masaryk's *Nebát se a nekrást!*, or the Czech writer Vančura's *Tento způsob léta zdá se mi poněkud nešťastným*. If the knowledge of the authorship of the quotation becomes lost, it usually turns into a proverb. Even Czech knows the infrequent expression **winged words** (according to Homer's *epea pteroenta*) whose position is somewhere between a quotation and a proverb. The expression was made famous by G. Büchmann's book *Geflügelte Worte* (1864), though it is less commonly used nowadays, cf. Descart's *cogito, ergo sum*, the French *cherchez la femme*, or Hamlet's *to be or not to be*, etc. Winged words are more often used in their translated form, cf. *být či nebýt, znám(e) své pappenheimské* (*Daran erkenn' ich meine Pappenheimer*), then we speak of idiomatic calques, extremely common in the whole of PI, though speakers usually no longer feel about them in this way.

A number of other expressions, whether conceptual or terminological, are (almost) never used in Czech, however common or popular they may be in other languages, such as the English **limerick** (a humorous five-line poem) or a **Wellerism** (a form of humorous comparison named after Samuel Weller, Mr. Pickwick's witty servant from Dickens' famous book, typically in three parts and concluded by a funny paradox), or **haiku**, a popular short Japanese poem (of 17 syllables), provided they are fixed. As an illustration, it is possible to give the original Wellerian limerick *Every beginning is hard said the thief when he began stealing an anvil.*, and to some extent even the Czech *Uvidíme, pravil slepej* (*We'll see, said the blind man*). (Cf., among others, Permjakov 1970, 1979, 2000, Mukařovský 1971, Taylor 1931, 1975, Zich 1956, Kuusi 1957, 1972).

#### 4.322 Polypropositional Intersubject Phrasemes and Idioms

**Polypropositinal intersubject phrasemes and idioms** which are formed by the combination of at least two propositions and which also relate to at least two communicating speakers, form a transition to small literary and folklore genres of various types. Their use is particularly strongly standardized with regard to the accepted linguistic and social culture in interpersonal communication (hence sometimes the term *cultural*

idioms). At the same time these expressions indirectly suggest which aspects of this communication are the most frequent and dominant. Accordingly, they belong just as much to the sphere of sociolinguistics and anthropolinguistics. The best known types of this primarily dialogue-oriented area of PI are

- (1) correlated greetings
  - (2) polite phrases,
  - (3) contact formulae,
  - (4) facetious (and other) reactions and responses,
  - (5) intersubject anecdotes, riddles, etc. Cf.
- (1) A: *Na shledanou!* – B: *Na shledanou!*; A (in reaction to sneezing): *Pozdrav pámbu!* – B: *Dejž to pámbu!*
- (2) A: *Děkuji.* – B: *Není zač/Rádo se stalo* etc.; A: *Dobré chutnání!* – B: *Nápodobně; A: Promiňte.* – B: *Nic se nestalo;*
- (3) A: *Jak se vede?* – B: *Děkuji, ujde to* etc.; A: *To je pan XY.* – B: *Těší mě;*
- (4) A: *Proč?* – B: *Pro slepičí kvoč.*; A: *Já mám hlad.* – B: *Tak ho hlad' (a říkej mu malej)! etc.*

The meaning of these phrasemes is tied up with a particular kind of function, especially on the speaker-listener axis. In this type of phraseme, the most specific feature is their (mainly internal) syntagmatic aspect. The concrete answer, e.g., in greetings is entirely or almost entirely conditioned by the first part of greeting and the latter is in turn predetermined by the context, situation and other sociolinguistic respects (see, e.g., type (1) above).

It is not quite clear how far we can go in the identification of phrasemes, if we move upwards in terms of size. For instance, some anecdotes or riddles, especially popular in African folklore, can be very long. It is a question then, (1) to what extent they can still be fixed, and (2) which features of an idiom they still preserve, especially as regards anomaly. Other such folklore-literary microgenres, as various popular stereotypes, need not have necessarily a fixed form.

### 4.33 Lexical Phrasemes

Because of their components, **lexical phrasemes** belong to the level of morphology (inflectional and word-formation) and display a single synthetic, non-discrete form. Functionally, they are one-word lexemes. Traditionally they are the most debated and neglected area of phraseology. They may be exemplified by compounds such as *budižkničemu*, *kratochvíle*, *tlučhuba*, *tvrdošijný*, *každopádně*, the derivatives *najít* (where *na* = “surface” + *jít* = “go, move” ≠ “nalézt, find, trouver, etc.”), *neplecha*, *zákeřný*, *našup*, and rarely even by inflected form, such as *Nevídáno!*, *budiž, nač*, whose nature is a hotly contested issue. On account of their discrete form, cases involving reflexivity and explicit valency, such as *obejít se bez*, *zašít se*, *zapomenout se*, *bodnout někomu*, etc., represent a transition. The same criteria apply to them as to other phrasemes and they are just as anomalous as collocational phrasemes, and other types. If word-formation theory takes notice of them at all (mostly it does not), it regards them as exceptions of sorts which, however, are rather numerous (the Dictionary of Standard Czech is estimated to include them in about 5 per cent of entries). For example, all 21 prefixed derivatives of the verb *být/bývat* (to be) existing in Czech can be considered



lexical idioms of this type (e.g., *nabýt*, *dobýt*, *zbývat*, etc.). An exhaustive description of this area is still lacking.

## 5. The Function of Idioms and Phrasemes

The function of the phraseme and idiom in text is primarily twofold, formal and semantic. In both cases, its external syntagmatic aspect is manifested especially in its collocability and valency. Generally speaking, the function of the phraseme and idiom follows from the need to denominate (**denominative** function) and to fulfill its role in the higher unit (**structural** function). These two functions are, especially in propositional phrasemes, complemented by a distinct **pragmatic** function (see esp. 3.5 and 4.3113).

There are various ways in which this type of phraseme is put in the context in the narrow sense which consists of the sum of all other naming units which share reference with the phraseme and which form an (immediate) higher structural unit with it. The **collocability** of the phraseme, i.e. its ability to combine with a certain type of expressions in text, is of a distinctly semantic nature. The **valency** of the phraseme determines the specific way in which it will be formally integrated in the context, whether sentential or hypersentential. This integration may, according to the nature of the expression and its transformation potential, entail a formal variation of the expression. In the context, the idiom and phraseme is even activated/specified semantically, i.e., on the one hand both its polysemy and the range of its meaning are suspended, or on the other hand its meaning will incorporate some new features in keeping with the speaker's intention. The expression thus assumes a concrete meaning. Its other characteristic components, such as vagueness, evaluation character, etc., are preserved.

Like any other naming unit also the idiom and phraseme serve communicative needs. The use of PI depends both on the subject-matter, the situation (and the style chosen) and the speaker's personality. The central domain in which the use of PI is characteristically focused is informal spoken language, or fiction and journalistic style reflecting the informal spoken language. In technical language, on the other hand, PI does not occur with the exception of several quasiphrasemes. Due to the nature of meaning in PI, marked especially by the tension between the meaning of the components and that of the whole, the level of contextual semantic activation/specification of the average idiom is much higher than in a word. This is also one of the main factors leading to idioms being less standardized. Their actual usage, determined by their meaning and function, is subject to relatively fast change, or vacillation, especially in expressions with a weak motivation. Further, the place of PI in the lexicon and in the idiolect of individuals is very much variable. On the one hand, the higher incidence of PI in the speaker and writer may be regarded as a sign of their polished and finely shaded language; on the other hand, the small occurrence of PI need not signalize the poverty of expression since instead of using PI, the authors may choose to create their own metaphors (an example of this may, to some extent, be the Czech writer Vladislav Vančura).

The semantic nature of the idiom, sharply contrasting with the character of the surrounding regular expressions, sets clear limits on any excessive use of PI in text. At present, it is possible

to consider a text to be rich in PI if it contains one idiom per 80–120 words on average (Čermák 1982). Written text which is even richer in this sense (i.e., one in which there are fewer words per idiom) is rare and an unbroken written text (esp. literary) of some length that would contain one idiom per ten or fewer words probably does not exist.

Nevertheless, there are more functions which PI serves in text. The other, loosely defined functions sometimes mentioned include the **aesthetic** function (many phrasemes express, particularly good or witty metaphor, image), the **economic** function (a standard phraseme is usually a highly economic means of expression and its “short” or literal paraphrase, if there is any, is always longer), or the **evaluative** function (most phrasemes are distinct means of expressing assessment, actually they are the richest source of it in the system) which has to be thought of in terms of pragmatics. The list includes even the **metalinguistic** function, the purpose of which is to follow and comment on or correct the actual line of the message, cf. the function and purpose of *co jsem to chtěl vlastně říct, mám to na jazyku, na tom si zlomíš jazyk, jinými slovy, lépe řečeno*.

## 5.1 The Structural Component of the Function

Every phraseme forms a closed structure which cannot be decomposed either semantically or in a syntactic-formal way, just as it is impossible to generate an arbitrary or new phraseme. The structural function of its constituent does not coincide with the seeming syntactic make-up of the phraseme (because it does not have it), cf. the impossibility of analyzing into predicate and object the expressions *Čte mu levity* or *Je z toho hin*, and others. As a whole, the phraseme does enter into syntactic relations, though. But the syntactic relations to the context, mediated by the valency and (especially the actualizing) transformations, are always weakened in varying degree depending on the type of anomaly displayed by the expression. Due to the absence of a regular compositional function in the semantic aspect (see 3.31), which is generally underlying any regular syntactic relation, only the external formal aspect, i.e. the formal means of such a relation, are preserved in this relation in a well-formed phraseme (and thus analogous to a regular combination), but not the full relation itself. For example, the nominative singular form of a nominal phraseme does not have to mean automatically, as it does in a regular noun, that the phraseme can be used as subject (*otevřená hlava*) or as a grammatically incongruous modifier (*růžové brýle*), etc.

**Collocational phrasemes** are functionally of two kinds, (A) adcollocational and (B) adpropositional, depending on whether they combine with regular collocations (or words) or with sentences. Each kind further subdivides, partly in keeping with the structural classification (only some of the main types are given):

### A adcollocational

- 1 *verbal*
  - a) verbonominal (V-N)
  - b) verbo-adjectival (V-A)
  - c) verbo-adverbial (V-ADV, ADV-V)
  - d) verboverbal (V-V)

- 2 nominal** a) adjectival-nominal (A-N, N-A)  
 b) nominal-adverbial (N-ADV, ADV-N)  
 c) nominal-nominal (N-N)  
**3 modifying** a) adverbial-adjectival (ADV-A, A-ADV)  
 b) adjectival-adjectival (A-A)  
 c) adverbial-adverbial (ADV-ADV)  
**4 relational**  
**5 substitutive**, etc.

## **B adpositional,**

### **1 modifying**

### **2 relational**, etc.

Binomial phrasemes are dispersed across types 1–3 (A1d, 2c, 3b–c); similes or comparisons mostly form a special subtype of A1a. Examples:

A1a: *číst někomu levity*, b: *kápnout božskou*, c: *přijít zkrátka*, *bohatě stačit*, d: *umět se narodit*;

A2a: *zlatý důl*, *anděl strážný*, b: *místo nahoře*, *vzhůru nohama*, c: *jablko sváru*;

A3a: *andělsky trpělivý*, b: *neslaný nemastný*, c: *široko daleko*;

A4a: *na úkor někoho*, *ve srovnání s někým*, *bez ohledu na něco*, b: *dřív než*, *hned jak*, *na jedné straně – na druhé straně*, *bez ohledu na to že /kdo/ co...*;

A5: *Co (je to) za (člověka)?*, *čertví kdo?* etc.

B1: *pro jednou*, *mezi námi*, *při čerstvé paměti*;

B2a: *Že by*, *Kdyby tak*, *Co když*, *Jen ať...*, *Co V/N – to V/N* (Co věta, to chyba), b: *byť i*, *až když*, *proto aby*, *ten který*, *kdo – ten*;

A particular structural type sometimes corresponds to several functional types and similarly the same functional type may be associated with several structural types. This is especially noticeable in synsemantic (function) components.

The structural classification and the (structurally) functional classification in synsemantic components largely overlap and are less clear-cut. One exception is the group of numerous multi-word (idiomatic) prepositions (A4a), conjunctions (A4b, B2b), particles (B2a) and pronouns (A5) which are functional equivalents of the respective grammatical word-classes. The most prominent structure is found in prepositions: **prep-N(-prep)** (with an abstract noun).

The main structural **functional** types of **propositional** phrasemes basically follow from the communicative sentence types, cf.

(1) declarative

(2) imperative, exclamatory, and

(3) interrogative,

each having a (a) positive, and (b) negative variant, cf.

(1a) *to je stará písnička*; *to je vrchol*!;

(1b) *ani pes by to nežral*;

(2a) *ruku na srdce*!; *trhni si nohou*!;

(2b) *abys to nezakřik(l)*!;

(3a) *a já jsem vosk?* (negative question 3b is rare).

The specific feature of these phrasemes is also their formally communicative function and the semiotic function (see 4.321); for more see 4.32.

## 5.2 The Denominative Component of the Function

The **denominative function** of the idiom which follows especially from its semantic nature has two main components, designative and pragmatic. The ratio of representation of each component considerably differs from one expression to another. Of primary importance is the **designative** function, based on the relation of the phraseme as a sign to its referent and to the notional (conceptual) element of its meaning. It is common to expressions of all kinds. By contrast, the **pragmatic** function, based on the (subjective) relation of the participants of the discourse to the phraseme is particularly manifest in PI. In general, it mediates the speaker's various subjective aspects stressed by him in communication (especially emotional and attitudinal) the PI inventory being its most important source of means of expression. The existence of PI shows, among other things, that for an individual the unique and subjective experience is socially and generationally repeatable and that it has its characteristic means whereby the speaker praises, expresses anger, wonders, emphasizes, etc. The pragmatic function can be made use of especially thanks to the denotative vagueness of the phraseme and the duality of its actual meaning contrasting with the surface information conveyed by its components.

The primary bearers of the designative function in collocational phrasemes are designators; the primary bearers of the pragmatic function are especially various operators (Weinreich 1966, Greenbaum 1969, and others). Each phraseme contains both components in varying degrees, though. In operators, it is necessary to distinguish at least those which are means of expressing different subjective relations of the speaker to the content of the utterance (commentors), those expressing the relation of deixis (deictics), the manner and measure to which the content of the utterance applies (quantifiers) and the broadly conceived formal means of (syntactic) organization of the utterance (relators). The main types are as follows:

### A Designators

#### 1 *primary*

- a arguments (A-N, N-N, N-ADV)
- b predicators (V-N, V-A, V-ADV, V-V, simile)

#### 2 *qualifying (secondary)*

- a qualifiers (A-N, N-ADV, ADV-A, A-A, simile)
- b circumstantials (i.e. adverbials of circumstances, ADV-ADV)

### B Operators

#### 1 *commentors:*

- a evaluative (as to successfulness, pleasantness, correctness, etc.)
- b generalizing (summarizing and generalizing)
- c expressive (expressing interest, frankness of the speaker, etc.)
- d truth-assessing (expressing the speaker's attitude to the truth-value of the utterance or situation or making deductions from them)
- c intensifying (making the statement more intense or acute)

## 2 *deictics*

### 3 *quantifiers*

**a** existential (with exclusive validity, making distinction of the type *každý-žádný*)

**b** restrictive (with limited validity, differences of the type *zvlášť-výlučně-hlavně*)

### 4 *relators*

Examples:

A1a: *Ta hospoda na rozcestí je **zlatej důl**. Zachránila ho **zbožná lež**. Stalo se z toho **jablko sváru**.*

A1b: *Táta mu pak proto **četl levity**. Po chvíli **kápnul božskou**. Při tomhle podniku **zkrátka nepřijdeš**. Je jí šedesát, ale pořád **se umí otáčet**. Taková ranka a ty **řveš jako tur**.*

A2a: *Nakreslil to **jedním tahem**. **Andělsky trpělivej** učitel. **Neslaná nemastná** odpověď. **Voda studená jako led**.*

A2b: ***Široko daleko** není žádný stín.*

B1a: ***S bídou** stačí jen půlku. Je to **k zbláznění** dlouhé. **Z ničeho nic** vykřikl. **Ještě že** mám rezervní klíč.*

B1b: *Je to **celkem vzato** docela slušné. **Zkrátka a dobře** jsem se spletl.*

B1c: *Moc se mi to, **popravdě řečeno**, nelíbí. Nemusíš mu, **mezi námi**, říkat všechno.*

B1d: *Tak Karel o tom **podle tebe** vůbec neví? **Podle všeho** už asi nepřijde.*

B1e: ***Co chvíli** mu to připomínám. Lidí tam je **až černo**. **Za nic na světě** už tam nepůjdu.*

B2: ***Nějakou tu** korunu to stát bude.*

B3a: *Udělat to musí **jeden každý** z vás. **Odpadávali jeden po druhým**.*

B3b: *Je to kreslíř **nad jiné** dovedný. Zbývá mu už **jedna jediná** možnost. **Na prvním místě** bych se zeptal jeho.*

B4: ***Ve srovnání s** loňskem byla letos úroda ovoce lepší. **S ohledem na** nedostatek času se to musí zkrátit. Byli tam tři, **a to** Novák, Holub a Skalka. Zlepšil se, **byť** i stále ne dost. **Ve srovnání s tím, jak** to dělají jinde, to vy děláte lépe. **Co** slovo, **to** nadávka.*

The denominative function in **propositional phrasemes** depends on the kind of communicative relation and function realized (speaker-listener-context-situation-theme-medium). These are the main cases:

#### A **vocative type** (the speaker-listener relation):

1 interjectional (expresses emotional reactions, attitudes, etc.)

2 imperative (expresses volitional influence on the listener)

3 contact (expresses the phase of contact)

4 categorizing (transition to B4)

#### B **thematic type** (the speaker-theme relation):

1 descriptive (simple description)

2 attitudinal (evaluation, expression of a viewpoint, etc.)

3 connective (suggest a connection between segments, phrases of the theme, etc.)

4 categorizing (identification of the theme with a known type)

#### C **medial type** (the speaker-means of communication):

#### D **situational type** (the speaker-situation relation), etc.

The vocative type includes three of the traditional functions (Bühler 1934, Jakobson 1960, Leech 1974, and others), both expressive and that of Appell (or emotive and conative), and phatic (esp. A3). The informational function is included mainly in the thematic type and the metalinguistic function in the medial type. The proportion of both components of the denominative function, i.e. designative and pragmatic, is much more balanced in propositional phrasemes and neither of them markedly prevails.

Examples:

A1: *Vida ho! To mi může být ukradený!*

A2: *Koukej mazat! Mě vynech!*

A3: *Zdař bůh! Tak se měj!*

B1: *Vzduch je čistý. Zatmělo se mu před očima.*

B2: *To zrovna! Chraň pámbu! To je fuk! Aby ho nebylo! Opatrnosti nikdy nezbyvá.*

B3: *Bylo nebylo... Jak se věci mají...*

B4: *To je slovo do pranice. Lehko nabyl, lehko pozbyl.*

C: *Na tom si zlomíš jazyk* (cf. metalinguistic function in 5.2)

D: *A ted' babo radí, A co ted'?*

The individual functional types and subtypes are subject to various modifications and combine with one another; sometimes they are known by their conventional names, such as toast, greeting, thanks, etc. With the aid of other criteria, such as the personal-impersonal dimension, technical-nontechnical use, etc., it is possible to extend the classification to other traditional and conceptually rather vague locutions, e.g. children's rhymes, counting-out rhymes, slogans, witticisms, aphorisms, riddles, and so on (see also 4.3214).

### 5.3 Complex Functional Classification

Anomalies of different kinds and degrees whereby phrasemes and idioms are characterized and classified from the formal, semantic and collocational points of view usually make their presence felt in the textual behaviour of these expressions and their usage. So the **idiomaticity** of an expression can be seen as the measure of its anomalousness in some of the senses discussed. Inasmuch as the textual behaviour of an expression is reflected in the range and ratio of its possible and impossible transformations, its idiomaticity is determined and objectively testable (or quantifiable) by the restrictions placed on its transformations, especially actualizing (context-specific) ones. If all relevant transformations are possible, the expression is regular: the degree of the idiomaticity of the expression is in indirect proportion to the number and character of its realized, or realizable, transformations.

If, for instance, only **actualizing transformations** are taken into account in the verbal type of idioms and its 17 features are considered, i.e.,

(a) positive-negative, (b) active-passive, (c) statement-question-directive, (d) indicative-conditional, (e) past-present-future, (f) singular-plural, (g) first-second-third (person),

then with just the regular realization of one alternative of each group (a–g), i.e. with the combination of seven features, the expression has 244 possible realizations. Phrasemes, however, usually do not realize all configurations of their transformations. The number of

theoretically possible transformation combinations that are “selective” in this way for the whole type thus grows considerably (in the given case there are roughly  $10^7$  possibilities). In practice, though, the actual number of anomalous transformation possibilities can be estimated to be much lower, but it still is sufficiently high to be the basis of the most precise classification in PI ever.

The number and range of possible and impossible transformations is highly specific for each phraseme. So much so that it can serve as the basis for a **complex functional classification** of the same structural type of expressions. Compare, for example, the following expressions and a list of transformations which do not obtain (headed by 0):

<i>nechodit pro slovo daleko</i>	<b>0: quest, pos, pass, cond, direct, direct neg, 1sg and pl</b>
<i>dělat někomu hláskou troubu</i>	<b>0: quest, neg, pass, direct, 1sg and pl</b>
<i>vědět svoje</i>	<b>0: quest, neg, pass, cond, direct, direct neg, fut</b>

The above verbal phrasemes differ in a number of anomalous features (here only in three configurations). It is clear that each such configuration of anomalous aspects shared by more phrasemes whose textual behaviour is likewise similar can be used for the determination of a special class or subclass and so even for the establishment of a **functional classification**. Each such configuration provides a complex indication of the textual function of a given phraseme.

This classificatory basis must be further modified by taking into consideration all other transformations (i.e., other than the actualizing ones; see 6.3) and the semantic and structural aspects (see Čermák 1982).

## 6. Transformations of Phrasemes and Idioms

### 6.1 The Essence of Transformation in Phraseology and Idiomatics

**Transformation** in PI is such a process of (structural and grammatical) change whose result (which is also called transformation) differs from the input structure by its **function** and therefore its position in the text while preserving a formal and semantic similarity of a varying degree to this input structure. The fact that one is derived from the other is clear. This is true of **structural** transformations, i.e., those relating to the abstract structure which is common to a certain class of individual phrasemes. **Individual** transformations always produce such change of form only in individual phrasemes. The relation between two kinds of transformation cannot, however, be regarded as a relation between a class and its element. Although it is possible for a given transformation to be realized structurally, it can never be realized in all individual phrasemes which share this structure in general. Moreover, some types of transformation are of only generally structural type (especially the intrastructural ones, see 6.23), while individual transformations are absent. This also constitutes the principal difference between transformations and variants (see 3.4) which are context-independent (if fixed) and therefore must obligatorily display both **semantic and functional identity** and, accordingly, are confined to variations of only parts of the form of an individual phraseme. A **variant** is thus always a modification of the same unit which

preserves its function, while transformation may produce another unit. The **input structure (base)** is the basic form of a general structure or individual phrases from which the transformation proceeds. In practice it is typically identical with the invariant. Owing to the many-sided inherent anomalousness of all PI and the absence of precise (grammatical and semantic rules), it seems suitable to regard transformations in PI as mere **quasitransformations**.

Transformation can be **fixed**, such as (transformation:) *jehla v kupě sena* (← input structure: *hledat jehlu v kupě sena*), or **unfixed**, the latter being most frequent in nominalizations, cf. *postavení něčeho na hlavu* (← *postavit něco na hlavu*). Transformations differ in their degree of generality, the nature of the phraseme, its structure and the level to which it belong, but also in that their result may no longer be part of PI, cf. *ostříží pohled* (← *mít pohled jako ostříž*). There may be several transformations from a single input structure, cf. *hledání jehly v kupě sena* and *jehla v kupě sena* (← *hledat jehlu v kupce sena*).

## 6.2 Types of Transformation

Structural and individual transformations, i.e. transformations belonging to a (variously) extensive class and an (individual) phraseme, may contrast in all areas in which they occur. Such contrast becomes especially noticeable in intrastructural transformations (see 6.23). The following survey presents the main types of transformation

- I Interstratal
- II Intrastratal
  - A interstructural
  - B intrastructural
    - 1 syntagmatic
    - 2 paradigmatic
    - 3 mixed

These transformations are usually fixed and the change of their function is of a general nature. When placed in a particular context, each individual phraseme undergoes **actualizing** (or, less accurately, grammatical) **transformations** whereby it becomes integrated in it. Like intrastructural transformations they are of three kinds: syntagmatic, paradigmatic and mixed (see 6.3).

### 6.21 Interstratal Transformations

With interstratal transformations the input structure differs from the resultant transformation by being from a different stratum. Usually the input structure belongs to a higher stratum. The commonest types of such structural transformations are

- (1) lexicalization
- (2) collocational transformation (collocalization)
- (3) propositionalization.

Within the first two it is possible to distinguish



- (A) nominalization,
- (B) verbalization,
- (C) adjectivization,
- (D) adverbialization, etc.

Generally it is possible to envisage other types (particularization, prepositionalization, conjunctionalization, comparativization, etc.). Their optional realizations with regard to particular cases are implemented by individual transformations. For example

- 1A: *tlučhuba* (from V-N), *květomluva* (z N-N, V-N); 1B: *lelkovat* (from V-N), *špačkovat* (from a simile); 1C: *sněhobílý* (from a simile); 1D: *tvrdošíjně* (through A from V-A-N); 2A: *tichá voda, devatero řemesel* (from propositions); 2B: *sejít z mysli* (from a proposition) aj. 3B: *Stalo se to v hodině dvanácté* (from the collocation *udělat něco v hodině dvanácté*).

## 6.22 Intrastratal Interstructural Transformations

They are typically encountered at the level of collocational phrasemes. The main transitions between respective structures are realized by the following transformations:

- 1 nominalization** (resultant structure: N-N, A-N, A-N-A, etc; from various input structures)
- 2 adjectivization** (resultant structure: ADV-A. or A-N, N-A, etc; from various input structures)
- 3 verbalization** (resultant structure: V-N, V-ADV, V-V, etc; from various input structures)
- 4 comparativization** (resultant structure: simile; from various input structures)

We may compare individual transformations which selectively illustrate the above structural transformations:

- 1 (*překročit Rubikon* → *překročení Rubikonu*; *zhroutit se jako domeček z karet* → *domeček z karet*;
- 2 (*postavit něco na hlavu* → *postavený na hlavu*; *křivě někoho nařknout* → *křivě nařčený*;
- 3 (*poslední cesta* → *doprovodit někoho na poslední cestě*; *být silný jako lev* → *mít sílu lva*;
- 4 (*prašivá ovce* → *být jako prašivá ovce*; *polít někoho studenou vodou* → *byl, jako by ho studenou vodou polil*, etc.

Particular structural transformations sometimes have several sources (input structures) and, conversely, individual structures may become an input structure for more than one kind of transformation. Hence intrastratal transformations yield a rich network of overlapping relations between structures. It is typical of a number of phrasemes that they develop from an original simile which thus often serve as the first base and the gateway to PI, cf. the transformational sequence: simile → verbonominal phraseme → nominal phraseme (still existing in all of these phases at once), i.e. *mít zrak jako ostrůž* → *mít ostrůž* → *ostrůž* → *ostrůž zrak*. One of the possible and important ways leading to further development is the emergence of polysemy in a one-word expression already outside the scope of PI, cf. *dupe jako kůň* → (*to je ale*) *kůň*.

## 6.23 Intrastratal Intrastructural Transformations

These are transformation whose basic combinatory structure (typically binary) branches out into a multi-component one. There are many types of these (see also the individual subtypes in the overviews in 4.311–6, including substitutions; PROP = proposition). The main types are represented by the following transformations:

**1 Syntagmatic**

A **quantitative** a additive:  $A-N \rightarrow A-A-N$ ,  $A-N \rightarrow ADV-A-N$ ,  $V-N \rightarrow V-A-N$ , etc.

b reductive:  $PROP \rightarrow \text{úsek } PROP$ ,  $V-A-N \rightarrow V-N$ , etc.

B **qualitative** (inversive) :  $A-N \rightarrow N-A$ , etc.

**2 Paradigmatic**

A **quantitative** (esp. substitution by a numeral, i.e., numeralization):  $V-N \rightarrow V-Num$ ,  $V-A-N \rightarrow V-Pron-Num$ , etc.

B **qualitative** (esp. substitution by a pronoun, i.e., pronominalization):  $A-N \rightarrow Pron-N$ ,  $V-N \rightarrow V-Pron$ , etc.

**3 Mixed**, esp. the explicit expression of the relation between the components (i.e., by a preposition and conjunction):  $V-N \rightarrow VprepN$ ,  $N-N \rightarrow NconN$ , etc.

Particular examples which are used here to illustrate the above structural transformations are not, as a rule, themselves transformations from a concrete input structure. They came into existence as ready-made phrasemes directly from a transformed general structure; still the transformation relation is fulfilled. The frequency of the individual types differs considerably. Cf.

1Aa *staré zlaté časy, jednou za uherský měsíc, mít šťastnou ruku*

1Ab (*Chytrému napověz, hloupému dolož* → *Chytrému napověz...*, (*hledat si postranní cestičky* →) *hledat si cestičky*)

1B *anděl strážný*

2A *dát si dvacet, olíznout všech deset*

2B *moje maličkost, hodit sebou, jet v tom*

3 *jít pod nůž, rub a líc*

Incidentally, traditional ellipsis (1Ab, the first example) may be regarded as a reductive transformation.

**6.3 Actualizing Transformations**

These invariably individual transformations are dependent on the context of the phraseme and, if their anomalies permit it, they appear in every type of phraseme with a variable morphology of some of the components. As regards the number of these transformations, each phraseme of the same type and structure theoretically has the same number. In practice, a concrete phraseme will always have fewer of them and in a varying number due to its individual anomaly. At the same time, one phraseme differs from another in terms of quality. The set of realizable transformations of a phraseme (if the starting-point is its invariant form) is highly symptomatic and individuating for each phraseme not only from the point of view of its function, but also in the degree of its idiomaticity. In addition to other dependences (on the type of formal, collocational and semantic anomaly), the degree of the idiomaticity of a phraseme (idiom) is usually complementary and inversely proportional above all to the sum of its possible transformations (esp. actualizing). Within the set of realizable transformations which is usually not interchangeable even in very similar phrasemes, the individual anomaly of an expression is manifested from a functional point of view. Actualizing transformations – or rather the ratio of what is realizable and what is non-realizable in them – are therefore the basis for a complex functional classification of phrasemes (5.3). The main types (in 1 designated by brackets) are:

- 1 Syntagmatic** **A contact** a additive: (*zcela*) *zahladil stopy, (neúspěšný) pokusný balónek*  
 b reductive: *chytrému napověz (a hloupému dolož)*  
 c inversive: *hory doly sliboval* (invariant: *slibovat hory doly*)

**B distancing:** *má srdce (již) zadané, měl (konečně) čistý stůl*

**2 Paradigmatic** **A statal**

**a phasal** – inchoativeness:

(*dostat někoho na krk*)

*vzít si něco do hlavy*

– durativeness:

*mít někoho na krku*

*mít něco v hlavě*

– terminativeness:

*mít někoho z krku*

*pustit něco z hlavy*

**b agential** (esp. causative)

**agens:**

– inchoativeness:

*hodit někomu někoho na krk*

*nasadit někomu něco do hlavy*

– durativeness:

–

–

– terminativeness:

–

*vyhnat někomu něco z hlavy*

**patiens:**

– inchoativeness:

(*dostat někoho na krk*)

–

– durativeness:

*mít někoho na krku*

*mít něco v hlavě*

– terminativeness:

*mít někoho z krku*

(*pustit něco z hlavy*)

**B morphological** (grammatical)

**a** deictic: person, number, tense, or in/definiteness and comparison

**b** relational: case

**c** modulational: question, negation, mood

**3 Mixed:** passive

The impossibility to undergo a certain transformation is typical of more markedly idiomatized phrasemes. Examples of these negative possibilities are found in 1Aa *číst někomu levity* (i.e., the impossibility of, for example, *\*zcela mu četl levity*), 1Ab *nasadit někomu psí hlavu*, (i.e., the impossibility of reducing anything), 1Ac *chytat lelky* (i.e., fixed word-order); B *mít v hlavě, zlatý důl*, etc.

Transformations of the type *byly to (úplně/hotové/...) lukulské hody, dostal za to (přísloušnou) jidášskou odměnu*, etc., by which the speaker enters the basic communicative line in a **commentary-like way**, are special cases of **metalinguistic transformation** (only specific words are found here) and they are based on a limited inventory of means of metalinguistic function (see 4.3212).

Ordinary additive transformations are based on the possibility to qualify the expression according to standard syntactic rules (modification). The rarely occurring reduction (in fact

ellipsis) is typical especially of long and frequent expressions. Distancing transformations, consisting in the possibility for the phrase to divide itself and incorporate some external contextual element among components of its structure, are also scarce; characteristic of PI is therefore their nonrealizability.

Paradigmatic transformations of the statal structure **V-N** (see also phrasemes with an abstract noun in 4.315), based in varying degree on the realizable phases of simple or causative action, activity and state, are characterized both by its specific distribution in only some phrasemes and by the set of verbs which participate in them. Morphological (grammatical) transformations are the most typical of all, see in the following examples the impossibility of their occurrence with the 1<sup>st</sup> person singular verb (*nechat na hlavě*), a singular noun (*pálit si prsty*), the future (*mít za ušima*), the comparative (*šedá eminence*), any other case (*každým coulem, otevřená hlava*), a question (*mít oči z Kašparovy krávy*), the verb in the positive (*nemoct od něčeho odtrhnout oči*), the imperative (*mít dlouhou chvíli*), etc. The set of theoretically possible morphological transformations is for each structural type inevitably different (here illustrated mainly by V-N and A-N).

The special nature, i.e. strong markedness, of **passivization** sets this type of transformation apart (type 3 above). The impossibility to passivize can be documented by phrasemes such as *hodit flintu do žita, vzít do zaječích*. The realization of the passive in the first case (*flinta byla hozena do žita*) attests to the existence of a literal, homonymous structure (see 2.4).

However, the possibility or impossibility of a transformation has much wider implications. The idiomatic nature of a particular phraseme may lead to its preventing even quite a common transformation in regular language where the phraseme operates as a component of a larger expression; in fact such an extraphraseological transformation is structurally quite alien to it. This can be demonstrated by comparing the relation of passivization to the nominal phraseme *pokusný králík* and to the regular expression *divoký králík*: *Ty škody byly způsobeny divokými králíky* x *\*Ty škody byly způsobeny pokusnými králíky*. Similarly, the phraseme *pokusný králík* is, under normal circumstances, impossible to use as the subject of a sentence, etc.

## 7. Conclusion: The Place of Phraseology and Idiomatics in Language and Text

Every natural system, including the language system, is kept open on account of its constant interaction with the environment, its influences and development. One of the consequences of this openness is a low measure of its orderedness, i.e., a high measure of its **entropy**. This is manifested by the irregularity, irregular distribution of the units of the system (including redundancy), and so even by the unpredictability of their occurrence and emergence. It appears that in language we find a good example of the presence of high entropy precisely in phraseology and idiomatics, which comes to light in the analysis of phrasemes and idioms in the process of communication: their information processing, especially their decoding, is always more complicated and multilayered than in other units

(up to a certain age children find them difficult to understand, because they tend to interpret them literally). Although the measure of information required for their gradual decoding is necessarily very low, the measure of information conveyed by them as known and learned wholes (fixed naming units) is much higher whea, compared to regular combinations in language. A significant manifestation of the entropy is also the unsystematic existence of closed paradigms in contrast to the open ones. All in all, irregularity in language, due to this general property of its system, will always be present here and it will always be one of its most striking features wherever it maintains a massive presence, i.e., in phraseology and idiomatics. One of the consequences of this is also the **universality** of phraseology in all languages. It is no accident, therefore, that the oldest and still used **artificial language**, Esperanto, begins to acquire properties of a natural language as part of improving its function with phrasemes and idioms beginning to emerge even in its system.

In short, PI is an area of **anomalous ready-made naming units** complementary to the majority set of regular naming units which are mostly unfixed. In consequence of its anomalousness, however, phrasemes and idioms may “skip or ignore” the rules known from other areas of language to such an extent and in such a way that this leads to the appearance of other typical properties which regular language does not systematically exhibit (for example, the motivatedness of a long regular expression may be a disadvantage and an obstacle to its possible shortening). Phrasemes and idioms are linguistic means which are very economical, possessed of considerable **compactness**; their small-scale form may carry an enormously complex meaning and function, which is in marked opposition to the means of regular language. It is difficult to find a phraseme that would be too wordy, over-long and clumsy. Hence PI is an important vehicle suited to the needs of **language economy**. From the point of view of word-class function, for example, the strikingly large incidence and extensive use of verbal phrasemes stand in sharp contrast to the (absolute and frequency) proportions of verbs in the fundamental regular lexicon. It is as if the verbal phrasemes compensated for the limited amount of regular verbs: in this respect PI is a highly important store-house of ready-made multiword predicates.

One the other hand, PI is the largest area of the lexicon which, among others, specializes in the pragmatic functions expressing evaluative, i.e. subjective and often also emotional, **evaluation** in the attitudes and opinions of their users. It is obvious that every speaker is in varying degree engaged in constant evaluation of everything around, in expressing preferences and rejections, often with accompanying emotional features. Besides the need to convey facts, the need to evaluate is absolutely essential in human communication. Regular means lend themselves to the requirements of this evaluation area neither easily, nor always, and that is why even attention paid to linguistic evaluation in traditional linguistics is still very meagre. By contrast, the conspicuous predominantly evaluative nature of phrasemes and idioms begs to be examined from this aspect: it fits this purpose like a glove.

Semantically, especially in connection with entropy, most of PI is also considerably **vague**, starting with the denotata. Phrasemes are generally indeterminate, inexact (and so the very opposites of exact terms), with fuzzy borderlines. This kind of meaning is often largely easy to complete and modify in text. Consequently we can speak in PI about a high

degree of the **potentiality** of another, non-standard, type of usage which is on offer here communicatively. What is meant by the standard PI usage, which luckily is not worth standardizing, is the present and statistically valid usage, prevalent and typical, which the speaker easily adapts to his needs (which, by the way, can be done with the standard meanings of words only sometimes and to some extent). It follows from this that phraseology and idiomatics is a very flexible and **fast-developing field**.

---

Tento příspěvek je zde díky A. Klégrovi přeložený v úplnosti, viz česky *Základy a zásady frazeologie a idiomatiky*.

# Linguistic Units and Text Entities: Theory and Practice

(In: *Actes Euralex '98 Proceedings*, Vol. I–II, 1988, p. 281–290)

## Abstract

In this paper, both the linguistic and lexicographic aspects of the system and text units are examined and the question of their mutual mapping is taken up. Against the background of F. de Saussure's ideas about the language entity and unit, viewed semiotically, three levels of units and entities to be examined and included in the lexicographer's description are suggested. It is shown that the current preoccupation with typical collocations may be too narrow and one-sided, both in that it omits other units and rare syntagmas which have not been paid much attention, so far. Alternative views on some common collocations are examined and a need for other criteria is voiced. A detailed example of a functional classification of collocational idioms is offered for the Czech language followed by a survey of basic combinatory phenomena found in the text and system.

## 0. Introduction

There has always been a paradox. While in linguistic theory, usually, the notion of various familiar **system units**, such as **lexeme** is taken for granted, what is being made material basis used and presented in dictionaries is far from clear, both theoretically and practically. Hence, so many "practical" ad hoc solutions. Moreover, in some lexicographic approaches, no attempt or pretension is made to veil their refusal to come to grips with this mapping of system units into text and vice versa. What should this relation be, if any? To put it differently and ask more general questions:

- (1) are there two **different sets of entities**, for the language system and the text, respectively, or
- (2) is there some **overlapping** between them only, or
- (3) is there a more or less strong **mapping** of the former onto the latter?

Ever since Chomskyan beginnings, it has become fashionable, in the profusion of theories and countertheories, to stop using in linguistics such general Saussurean terms as **unit** or **entity**, as if they did not have any substance, although more specific terms, such as **word**, **lexeme**, **idiom** etc., are used here; this has been accompanied by a strong dislike to do any exhaustive material mapping (in the sense referred to above). Hence, there is a multitude of transformational grammars or rather theories and proposals of various denominations, but not a single transformational dictionary. This disregard for the units of the system has had a broad impact, unfortunately. On the other hand, corpus linguistics, true to its meticulous search for

any type of form or configuration of forms in text which might seem significant (Sinclair etc.), has shifted the focus from the system units to their textual forms and collocations and has always stressed the exhaustive type of approach. Hence the corpus-based dictionaries approach is advocated here, since, obviously, a considerable amount of data has to be made available first and only then a useful selection can take place.

Now, the corpus linguistics is basically right, of course, in its search of significant units, but what exactly is to be understood by **significant** (and meaningful), i.e. by significant forms and their combinations? Should it mean **typical** only, as is often the case nowadays, then a lot of what is **rare** in the corpus, and still represents some kind of the system unit, becomes a drop-out, automatically. Delving a bit deeper, one has to admit that the reason for this is to be seen in the insufficient size of the corpus used which is to be partly blamed.

In general, if we disregard any hasty planning of a dictionary, lack of expertise and/or theoretical background, then primarily, two basic sorts of problems relating to units have to be singled out:

- (1) *system-text mapping* and
- (2) *quality and quantity of data* available to us.

Clearly, solutions to these problems should be on a scale of various degrees of compromise, where some compromises would be better, some worse, and some hardly acceptable. In what follows, some comments on both problems will be made.

## 1. Units and Entities

De Saussure regarded **units** of *la langue* as being identical with language **entities** (de Saussure, 145), which are defined in his *Cours* as being based on oppositions (145) and differences (168) within the sum total of their mutual associations in the system (*la langue*, 189). To put it differently, these entities exist only through their meaning (“sens”) and function (191, 149). However, what he did not mention so explicitly, but what is recorded in the manuscript edition of *Cours* by Engler, is his emphasis on the necessary condition that any unit, for it to be a unit, must be felt as such (Engler 2152 B). This, clearly, implies its shared stability and fixedness in the system. De Saussure’s usage of the term **entity** is, unfortunately, somewhat misleading. It seems that by the **concrete entity** he generally meant both what one would, today, call system units as against text ones (cf. 145), i.e., for example, **lexeme** and its **manifestations** (if any), while the term **abstract entities** was used, primarily, for such notions as genitive case, word class, word order etc. (190), which he viewed as being always based, in a final analysis, on the concrete entities (190). Thus, having distinguished between **class** (category) and its **members** (units), he has not made any further systematic attempt to distinguish between the **type-token** levels of the unit’s existence. Semiotically, an entity may exist, in his well-known view, only thanks to its association of the signified (*signifié*) and the signifier (*signifiant*, 144).

What these statements imply, among other things, is that any case, where the link between an alleged entity, viewed as a form (signifier), and its function, meaning (or its signified) is not clear, stable or documented at all, must be viewed with some suspicion and



may be discarded, if a further analysis does not help in clearing it up. This is, in a nutshell, what linguists and specifically lexicographers do while trying to single out, from the text continuum, meaningful entities. However, on the basis of this Saussurean conclusion some uncomfortable questions can also be formulated. Given the obvious and patently true hypothesis that words do not collocate, combine mutually in the same degree and quality anywhere in the lexicon, a continuum of combination or collocation types must be presumed, then, along a scale starting with very obvious and frequent combinations through less frequent combinations, ending in highly infrequent and hypothetical (potential) ones. Now, if the notion of **collocation** (e.g., Sinclair, for a survey see, e.g., Heid) is to be taken to mean *any* type of meaningful combination, should the term not be discarded, since the term **lexical syntagmatics** would do the job? Should it, on the other hand, be taken to designate some section of the scale only, and if so, how does one define it? By valency, syntactic, semantic, pragmatic criteria, or any combination of these? Hence the widespread confusion about the term **collocation** and its ill-defined nature. Although there is no doubt about the relevance of **functional and semantic criteria** on which any sensible lexical **syntagma** (collocational and other) is based, there are still precious few statistical tools, such as MI-Score and t-score, that might help us to shed some light into the vast combinatory realms of lexical units. It has to be borne in mind that all of the approaches so far have dealt exclusively with combinatory phenomena which are common and typical. This clearly means that only a **minor** part of the field has been paid some attention.

While I do not propose to go into this matter any further here, I want to point to an obvious dividing line in the combinatory continuum which may be found in de Saussure's argumentation: where a stable and clear link between both form and function/meaning may be repeatedly found, these might be seen as fixed units, **lexemes**, made up of both single and multiple word forms and forming part of the language system, while the rest is different in its not being stable and fixed. There is no doubt about that the solution of the **fixity or stableness** problem is to be sought, at least in part, just here. But is the boundary between both parts of the combinatorial continuum really clear? Often it is not, and where it is, are there any specific areas to be delimited in what lies behind "the border line"? The obvious first candidates for fixed units, stable entities, found both in the system and the text, are **idioms** and (**technical**) **terms** of all sorts. But what about the rest?

## 2. Levels and Entities

### 2.1 Above the Word

The rest, largely unexplored, seems to be a tangle of problems. Turning over to the clearer and better-explored part of the combinatory continuum, how should we view such combinatory cases as *black coffee* and *white coffee*? Provided that we admit the combination *green coffee* as acceptable, too, even though it is not quite the same thing, and if, perhaps, *pale coffee* might be passed for an expression of a criticism, how acceptable are such combinations, i.e. of a colour adjective + the noun *coffee*, as *\*blue coffee*, *\*violet coffee*, *\*scarlet coffee*, *\*yellow coffee*...? There is no excluding these from our consideration on any ground whatsoever, but their probability of occurrence seems to be quite low and the present

size of corpora is simply not sufficiently large to be of much help here. Declaring *black/white coffee* to be collocations does not solve much as, in doing so, only the **frequency** (and, perhaps, **habituality**) aspect would be stressed, leaving the other aspects unattended, and we do not know whether that alone can be sufficient as a criterion. Should the size of corpora be made tenfold, would not the scope and horizon of frequent combinations be larger and a new selection and reclassification would become feasible? Even if we accept this frequency criterion with some reservation, we still have no way of knowing where to stop and draw some sort of line (lines?) in the realm of less frequent and rare combinations.

Hence, some of the problems to be dealt with here include the following:

- (A) Do we, by calling these combinations **collocations**, solve the problem of classifying them as system entities, **multiword lexemes**, or as something else? The oversimplification is rather obvious.
- (B) Should we, following the Prague school scholars, see here **complex denominations** (Mathesius, Filipec – Čermák)? Neither that, due to the psychological, i.e. external starting-point, solves anything in this context. It may be objected that rare combinations, dwindling gradually into mere **potentialities**, may not be of immediate concern for many practical lexicographers (yet), but they are very much the concern of theoretical linguists. To make matters more complicated, let us notice that neither *black coffee* nor *white coffee* is really black or white, respectively.
- (C) Should this aspect, then, be used for declaring these combinations to be **idioms**, since the familiar definition of idioms based on non-compositionality still seems to prevail?
- (D) Or, taking into account, at least in culinary circles, existence of some sort of definition (cooking instruction, recipe) existing for them, should *black/white coffee* be declared to be **terms**?

A corpus analysis (Cobuild 1995) of the collocations of *coffee* has, in fact, revealed that there is only one combination of it with a colour adjective, namely *black coffee* (occurring twice in 5 millions). The rest, which is different, including *fresh coffee*, *freshly ground coffee*, *instant coffee*, *strong coffee* etc. is hardly acceptable as stable and fixed combinations, perhaps with the exception of *instant coffee* which is a term. A check in BNC shows that while there are 97 instances of *black coffee* to be found, there is hardly any other colour name going with the noun (altogether, there are 4 instances of *white coffee* there).

It is clear that moving from what is (proto)typical and easily recognizable to the less typical brings one into the realm of the **vague**, non-determinate with growing scarcity of criteria. I am afraid that neither of the above criteria used offers any generally acceptable approach. This being so, all sorts of ad hoc approaches and partial classifications emerge. Yet, the “*coffee*” example used is a very simple one indicating its clear limitation due to the closed set of colour terms; but vast realms of words and their combinations are not like that at all.

There is something counterintuitive in the concept of **multiword units**, i.e. in viewing a discrete combination as a **whole**. Accordingly, this term is not very much employed,

except in the most obvious cases, such as typical idioms and terms. It is quite difficult, for instance, to find a grammar of English which would offer an **exhaustive list** of multiword prepositions, such as *as to*, *as for*, *in connection with*. In fact, they are not found in grammars of most languages; yet, for example, the latest count for the Czech language has come up with no less than some 400 of them. While they usually do appear in dictionaries, they may not be labeled there as prepositions at all; they are often presented in some other way (in boldface etc.) as some further unclassified combinations. This approach is specifically hard to accept in grammars, which almost always pretend to be so very much exhaustive, but neither is the solution solution to be found in dictionaries ideal and satisfactory.

## 2.2 Word Forms

But there are also entities of the opposite order, lying below the level of single word lexemes, namely **word forms**. Traditionally, they have been used to demonstrate that there is a kind of difference between the system entities and text ones, although perhaps in degree only, not principally. It does not automatically follow that between such forms as, for example, *go* and *went* there is the **type-token** difference. On the other hand, it is a familiar experience that some word forms, to the exclusion of other, have a different meaning (sense) not only contextually but repeatedly, a fact very much stressed nowadays by corpus linguistics. Beside irregular plural forms (such as the English *mice*) and suppletive forms (see *go-went* above), such (personal inflectional) forms as *I wonder (why...)* contrasted with other forms of the verb, such as *he wondered*, *I wondered (why...)* etc., may appear to be a typical case, especially if a translation into another language is attempted: while, in the Czech language, *I wonder* is best translated as *rád bych věděl* (i.e., with a Conditional Mood, literally “I would like to know”), the other forms of this verb correspond to different equivalents, such as *být zvědavý*, *divit se* and there is no possibility for a mutual substitution. However, our awareness and attention paid to the existence and number of such cases has been only marginal and not systematic, so far.

## 2.3 Below the Word

In some languages where there is typologically no inflection, this word form type of entity may assume a different appearance and include only a **part** of the word, or, rather, of a single **morpheme (root, affix)**. This is the case, primarily but not exclusively, of verbal prefixes in Slavonic languages, of some compositional roots in Germanic, Uralic languages or Sino-Tibetan type of languages and of other phenomena, such as those that are related to the phenomenon of incorporation and polysynthesis in many other types of languages.

Thus, for example, such Slavonic (mainly Czech, Slovak, Polish) **prefixes** as *vy-/wy-* (out of), *z-* (used for perfectivization) are specimens of a boundary phenomenon between the **affix** with a limited distribution (which belongs to a limited set of words, each of them being recorded in dictionary) and a free element. It enables an independent formation of new words, which are quite acceptable, even though they are never written in isolation, some of them not being found in dictionaries. To list them in a grammar only would greatly limit uses of a good dictionary. Such **suffixes** as the Indonesian *-kah* or not quite identical Finnish *-ko/-kö*, used for signalling questions, might represent a similar case. Some elements in

Chinese compounds, such as *wù* in *ren wù* “person/man (< man + thing)” are hardly ever used outside their respective compounds, and, for a number of reasons, a dictionary should record them, too.

Some **compound elements** may, too, appear as a surprisingly useful element in a bilingual dictionary between a language which abounds in compounds and one which does not use them much, cf. *house-*, *huis-*, *hus-*, *koti-*, *ház-* in English, Dutch, Swedish, Finnish and Hungarian and the corresponding Czech and Polish full-fledged adjectives *domácí*, *domowy* etc. The dictionary compiler creates here, for the benefit of the user, an entry artificially based on a decomposition in such languages as these five.

From the point of view of lexical language **nomination**, there are basically four possible ways how to express some of the common (simple) notions in various languages. These are to be seen in the equivalent expressions for the same thing, namely *RAILWAY*, in, for example, Czech, Finnish and French, cf.

- (1) a derivative *železnice*,
- (2) a compound *rautatie*, and
- (3) a collocation *chemin de fer*;

the fourth case, a simple root, is not available here. To return back to Chinese and other polytonic languages, one must take into account that it is the syllable that is the dominant unit here, overlapping largely with the morpheme (tantamount to root) and words; due to the Chinese way of writing (to be found in other languages, too) placing characters next to each other, it is difficult to distinguish all of the boundaries here. Thus, is

the Chinese *tiělù* (= iron + way) a compound or a collocation?

Accordingly, the distinction between the **compound and collocation**, that one is so confidently accustomed to recognize in European languages, is extremely blurred and problematic to draw here. It does cast some shadow on the wisdom of a one-sided preoccupation with collocations only. The same principle of writing, which may not be so apparent from Chinese due to its use of characters, is much more clear in Vietnamese which did use Chinese characters a short time ago, too.

Here, still another phenomenon to be mentioned is a number of dependent **function words** which serve purely grammatical and pragmatic functions, i.e. they are never used independently as words, such as *à* (a question particle) or *các* (non-definite noun plural particle). It could be argued that some of these words (Vietnamese and other) merit their inclusion into a dictionary (a bilingual one, primarily), since they do not seem to have any obvious counterpart, but that is only one of the reasons. Rather, they are specimens of the **non-existence of any sharp boundary between lexis and grammar**. Many languages do have very common and extremely frequent words used for grammar purposes solely which a dictionary must include, see yet another example of such words for signalling questions in the initial Polish *czy* or alternating Malay/Indonesian *b(id)ak* and *(bu)kan*. In all of these cases, we have to deal with rather specific syntagmatic combinations, not typical collocations.

It is evident that there are, then, at least **three levels of entities** of some interest for both the linguist and the lexicographer. Note that due to the existence of both types of entities, which are either larger or smaller than the word, in those languages where the term *dictionary* is based on the root used for the **WORD**, as in the Slavonic or Germanic languages, it may no longer be so fitting and in its pointing to words only it may become somewhat misleading, cf., for example, Czech *slovník*, Finnish *sanakirja*, or Swedish *ordbok* etc. (where *slovo*, *sana* and *ord* stand for “word”).

### 3. Multi-Word Entities and Units

While it is difficult to see in the diversity of **sub-word entities** any system or at least classes and languages differ here very much, it is the other, opposite type consisting of **multi-word units** that is different. To show some of the regularities and systematic aspects of this area of entities, examples from a single language will be, briefly, used (SČFI 1–3, Filipec – Čermák, Čermák 1994b).

In many respects, the criterion of the formal class membership (e.g. word classes) seems to be useful for a formal classification of word combinations and idioms; it has long been used in the Czech linguistic practice, too. Yet, on a closer look, it is a mere **external** type of criterion covering the type of component parts used in their construction, which may be particularly suitable for **formal** statistical and combinatory-based approaches, but less so for other. Should we want to use, on the other hand, meaning and (nominative) **function** as the criterion here, a different and more revealing approach would be implied. An important aspect of **idioms** on the level of collocations is that **functionally** they may be viewed as **extensions** of all standard word classes. Since we are dealing with units on and around the word level, sentence type idioms seem to be excluded, but that is not quite true. There are, traditionally, ten word classes recognized in the Czech language and there are as many functional classes corresponding to these in the region of the Czech collocational (i.e. sub-sentential) idioms; thus, every single word class has a kind of extension in functionally equivalent idioms. These may, then, be illustrated by such simple cases as:

**1 NOUNS:** *zlatý důl* (“a gold mine”), *teplé místočko* (“a cushy job”)

**2 ADJECTIVES:** *neslaný nemastný* (“wishy-washy”), *staří mladí* (“young and old alike”)

**3 VERBS:** *číst někomu levíty* (“read someone the riot-act”), *umět se narodit* (“have the luck of the devil”)

**4 ADVERBS:** *široko daleko* (“widely, in many places”), *kolem dokola* (“round and about”)

**5 PRONOUNS:** *nějaký ten* (“some, quite a few”), *moje maličkost* (“a little me”)

**6 NUMERALS:** *jeden dva* (“one or two, some”), *jeden za druhým* (“one after another”)

**7 PREPOSITIONS:** *co do* (“as for”), *v ohledem na* (“with regard to”), *až na* (“with the exception of”)

**8 CONJUNCTIONS:** *ba i* (“even, to be sure”), *kór když* (“let alone”)

**9 PARTICLES:** *co když* (“what about/if”), *kdyby tak* (“I wish”)

**10 INTERJECTIONS:** *jen klid!* (“(take it) easy”), *to zrovna!* (“no idea!”); this class includes an important part of sentence idioms, too.

There are some tendencies to be observed here which have their implications for the job of the dictionary compilation, too. Let us mention some of them.

- a – As a rule, the number of grammar idioms (5–9 above, with the obvious exception of numerals) is far greater than the number of single-form grammar words in each of the respective classes. This makes them a serious and important extension of the single-item grammar word stock which otherwise does not really grow any longer.
- b – A peculiar feature is a pronounced scarcity of functionally adjectival idioms.
- c – The Verb-Noun relation is very much different statistically from the relation of these two word classes in the vocabulary: there is several times more of the verbal idioms of many types than noun idioms. This relation, however, has its important counterweight in **terms**, which belong to another area of multi-word units.

Clearly, with the majority of multi-word terms in the area of functional nouns, terms can be viewed as a kind of opposite end on the scale whose one end is taken by idioms having its major concentration in verbal types.

Linguistically, it would be very important to compare statistical data here from more languages to see if there are some general tendencies behind this single-language picture.

Traditionally, a vast part of what is usually covered by the term of **collocation** (whatever that might be) consists of **verb-noun combinations**, mostly made up of abstract nouns. If one tries to look up relevant combinations offered in corpus (Word Bank 1995, Collins, 5 million words) of such a typical **abstract noun** as *ATTENTION*, what one gets is an amazing and perhaps confusing multitude of combinations based on such verbs, as these:

*arrest, attract, bring to, call a. to, catch, capture, centre, claim, come to, command, concentrate, crave for, demand, detract, develop, devote, direct, distract, divert, draw, escape, focus, gain, get, grab, have, hold, keep, lavish, love, need, pay, pull, receive, require, return a. to, share, shift, slip, take a. away, turn, vary, vie for, want, win (ATTENTION).*

Singling out as special idioms *stand at attention* and, perhaps, *call somebody to attention*, what one can use for classification are only frequency data, which are not really much help. Using them in descending order, they read here roughly like this: *pay, draw, focus, get, attract, turn*; oddly enough, they rather correspond to the choice made by many dictionaries, too, but that is hardly a way how to deal with the rest.

In the Czech theory of **idioms**, it has been shown that their substance should not be sought in deviant, non-compositional semantics, primarily, but in the **anomaly of the function** of their components (see, for example, Čermák 1994a). In this approach, based on a commutation test, the above-mentioned combinatory continuum of words can, indeed, be further split and many cases of what has been traditionally passed for collocations, may rather safely be classed as semi-idioms. It is purely on the basis of a limited inventory of stable combinations of such typical abstract nouns as *attention* that these combinations

begin to appear as a limited and small group of **semi-idioms** most of which a part of one or more larger structures.

Compare the simple **IN**choative, **DUR**ative and **TER**minative **Phases** (which may not always have a verb representative for every noun) and its parallel **Causative** series of expressions in three phases, too, where identity of the meaning of the noun and agents involved has to be retained throughout:

**ATTENTION<sup>1</sup>**: (two human agents)

<b>IN</b> pay, focus, turn, devote	<b>DUR</b> pay, devote	<b>TERM</b> ? [withdraw]
<b>C-IN</b> draw, attract, catch, bring to, call to, capture	<b>C-DUR</b> command, hold, have	<b>C-TERM</b> distract, divert, take away

**ATTENTION<sup>2</sup>**: (non-human agent)

<b>IN</b> come to	<b>DUR</b> command, have, require?	<b>TERM</b> ? [lose]
-------------------	---------------------------------------	----------------------

However provisional the inclusion of only some verbs may be, the goal here being to include the more stable and frequent ones, the picture shows a patterning rather clearly. Of course, a further corroboration (such as for the inclusion of *withdraw*) is necessary on a much larger corpus.

A successful attempt has been made to describe in this way Czech verbal semi-idioms of this type in a rather exhaustive way; it has entered the last volume of the Czech Idiom Dictionary (SČFI 3, 1994). At the same time, this approach pointed to the existence of both larger structures in the human mind and in human lexicon.

One might consider also other stable combinations which might be passed for multi-word units, or rather for semi-idioms, even though this is usually not the case with many dictionaries. Briefly, there are **valency** prepositional combinations, such as *insist on/that/0*. On a closer inspection, it is apparent that *insist* does not enter, as a rule, into any other combination. Why, then, is this information usually submerged inside the dictionary entry, if these two combinations could be viewed as obvious multi-word units, namely *insist on*, and *insist that*?

To conclude this part, let me stress that in the logic of this approach **sentence-type** multi-word entities should be given the same treatment in dictionaries as single words, whether they are idioms or belong elsewhere.

Most of the latter cases are subject to different treatments and represent very much open problems. It is, however, necessary to view them as separate entities, no matter what labels they might have taken, and treat them both systematically and in due detail. They are usually much more complicated than single words (lexemes) which make them up.

## 4. A Survey and Summary

To sum up, it seems that word combinations (C), given the present state of our knowledge, fall into several classes.

C-	<b>SYSTEM</b> (stable) /	<b>1 regular:</b>	terms ( <i>nitric acid</i> )
		<b>2 irregular:</b>	idioms ( <i>give a hand to</i> ) semi-idioms ( <i>turn attention to</i> )
\	<b>TEXT</b>	<b>3 regular:</b>	a – grammatico-semantic expressions ( <i>read a book</i> ) including some of traditional collocations b – grammatical ( <i>would read, has been</i> ) c – terms ( <i>nitric acid</i> )
		<b>4 irregular:</b>	a – idioms ( <i>give a hand to</i> ) semi-idioms ( <i>turn attention to</i> ) b – author metaphors c – accidental combinations representing no entities ( <i>[the general consensus in Russian] elite was that</i> <i>[the major threats]</i> ) d – gibberish, nonsense

Not all of these combinations are, however, language entities (as in 4b–d), some of which do not make much sense (4c–d). Some (as in 3b) are text entities only and belong, in the system, to a single-word lexeme. Since, however, there is no way yet how to incorporate in the dictionaries the combinatorial **potentiality** of lexemes, this picture may seem as consisting of more clear-cut classes than the language reality suggests.

### Shrnutí

#### Jazykové jednotky a textové entity: teorie a praxe

Lexikologie jasně uznává dnes za výchozí jednotku lexém, ale co je výchozí jednotkou při tvorbě slovníků, není ani zdaleka jasné. Jde tu o dvojí typ jednotek a pokud ano, jaký je jejich vzájemný vztah? Delimitace jednotek z korpusu není dnes o mnoho jednodušší.

F. de Saussure viděl mezi jednotkami langue a jazykovými entitami rovnítko, jakkoliv pro jednotku požadoval, aby byla uživateli vycitována. Na druhé straně nerozlišoval jasně mezi jednotkami systému a textu. Situace je dnes složitější o stále ne zcela jasný pojem kolokace.

Vyřešíme tento problém, když kolokace prohlásíme za víceslovné lexémy, které pak jsou i mathe-siovskými víceslovnými pojmenováními? A nakolik mezi ně patří frazémy jako *black coffee* nor *white coffee* (které nejsou sémanticky aditivními kombinacemi komponentů) a víceslovné termíny?

Vedle lexémů je však třeba uvážit i jejich tvary, např. angl. *I wonder*.



Na druhé straně jsou i entity nižší, než je slovo či lexém. Nápadné tu jsou slovanské prefixy, jinde sufixy (jako např. v indonéštině či finštině); v čínštině třeba najdeme prvky, které se vně svých lexémů užívají jen málo.

Nominativně jsou přitom všechny typy rovnocenné, i když může jít o prostý lexém, resp. derivát, kompozitum či kolokaci, srov. čes. *železnice*, fin. *rautatie* a fr. *chemin de fer*; přitom nemusí být zcela jasné, zda jde o kompozitum či kolokaci jako v čín. *tiěllù* (= železo + cesta). I když necháme problém gramatických slov stranou, je jasné, že ostrá hranice mezi lexikonem a gramatikou tu obecně není a že jde o několik úrovní relevantních jednotek.

V oblasti víceslovných jednotek je třeba počítat na jejich nevětné úrovni s funkční ekvivalencí lexému, pro češtinu tu je tedy 10 tříd víceslovných ekvivalentů, především frazémových, srov.

- 1 Substantiva:** *zlatý důl, teplé místočko*
- 2 Adjektiva:** *neslaný nemastný, staří mladí*
- 3 Verba:** *číst někomu levity, umět se narodit*
- 4 Adverbia:** *široko daleko, kolem dokola*
- 5 Pronomina:** *nějaký ten, moje maličkost*
- 6 Numerale:** *jeden dva, jeden za druhým*
- 7 Prepozice:** *co do, v ohledem na, až na*
- 8 Konjunkce:** *ba i, kór když*
- 9 Partikule:** *co když, kdyby tak*
- 10 Interjekce:** *jen klid!, to zrovna!*

Specifická je ale ještě struktura kolokací verbonominálních, založených na abstraktním substantivu. Korpusová analýza ukazuje, že např. substantivum *attention* se kombinuje s takovýmto souborem verb:

*arrest, attract, bring to, call a. to, catch, capture, centre, claim, come to, command, concentrate, crave for, demand, detract, develop, devote, direct, distract, divert, draw, escape, focus, gain, get, grab, have, hold, keep, lavish, love, need, pay, pull, receive, require, return a. to, share, shift, slip, take a. away, turn, vary, vie for, want, win (ATTENTION).*

Z nich se však jasně rýsují obecnější struktury označující fáze (začátek, trvání, konec), popř. kauzativitu, a jen tedy některé z nich vytvářejí takto struktury vícečlenné, srov.

#### **ATTENTION<sup>1</sup>:** (dva agenti)

<b>IN</b> pay, focus, turn, devote	<b>DUR</b> pay, devote	<b>TERM</b> ? [withdraw]
<b>C-IN</b> draw, attract, catch, bring to, call to, capture	<b>C-DUR</b> command, hold, have	<b>C-TERM</b> distract, divert, take away

#### **ATTENTION<sup>2</sup>:** (neživý agent)

<b>IN</b> come to	<b>DUR</b> command, have, require?	<b>TERM</b> ? [lose]
-------------------	---------------------------------------	----------------------

Typy kolokací (C) lze, jak se zdá, mj. členit na následující:

<b>SYSTÉM</b> (stable) / C- \ <b>TEXT</b>	<b>1</b> pravidelné:	termíny ( <i>nitric acid</i> )
	<b>2</b> nepravidelné:	frazémy ( <i>give a hand to</i> ) semi-frazémy ( <i>turn attention to</i> )
	<b>3</b> pravidelné:	a – gramaticko-sémantické výrazy ( <i>read a book</i> ), kam patří většinou tradiční kolokace b – gramatické ( <i>would read, has been</i> ) c – termíny ( <i>nitric acid</i> )
	<b>4</b> nepravidelné:	a – frazémy ( <i>give a hand to</i> ) semi-frazémy ( <i>turn attention to</i> ) b – autorské metaforý c – náhodné kombinace nepředstavující žádné entity ( <i>[the general consensus in Russian] elite was that</i> <i>[the major threats]</i> ) d – blábol a nesmysly

# Idiomatics

(In: *The Prague School of Structural and Functional Linguistics*, 1994, p. 185–195)

## 1. Idioms in Language and Linguistics

The phenomenon of idiom has always been partly recognized, though under different labels, in various languages. Though one may, in fact, today assume that this puzzling feature exists in all languages as a necessary and natural counterpart to the bulk of the rule-governed structures on more than one level, it is its substance and exact nature that has defied definition and general recognition. As a consequence, there has been a widespread disagreement about the identification, typology and distribution of idioms. Only very few criteria, such as stability and systemic character, have won general acceptance, while hitherto reliable identificatory criteria such as non-additive meaning (not regularly and directly composed from the component meanings) or various transformational deficiencies have failed. This problem becomes readily apparent once the academic approach concentrating on a few well-chosen items, so typical of many syntactic conceptions, is abandoned in favour of the systematic lexicographical description of the idiomatic aspects of a language (Čermák 1988).

In the following pages, a functional and structural theory of idioms is sketched that is based on the description of some 17,000 Czech idioms in dictionary form (*Slovník české idiomatiky a frazeologie* [A Dictionary of Czech Phraseology and Idiomatics], see Čermák – Hronek – Machač 1984, 1988, 1994). In its general form, the theory may apply to other languages as well. Most of the insights of this approach derive from the established theory of language nomination associated with Vilém Mathesius (Mathesius 1975), and, in a broader sense, with Ferdinand de Saussure's notions (1916) of syntagmatic and paradigmatic, and the later work by many scholars, especially Louis Hjelmslev (1972) and M. A. K. Halliday (1966). The only notion derived from the generative grammar, is that of anomaly (Chafe 1968, Fraser 1970), consisting in both generative and transformational deficiency of many idioms and in their generative anomaly, but this idea has, in fact, been around in linguistics for a very long time.

In much of the following, the rather neglected term of **analogy** (seen as the counterpart of anomaly), heavily employed in Saussure's time and before, will play a major role, though it may not be explicitly mentioned. There is no reason to reserve the term for historical and developmental linguistics only; in fact, all of the language functioning, usually referred to as

rule-governed, is based on analogy and the speaker's morpheme, word and clause combinations either follow old patterns and models or, by the force of anomaly, form new ones.

**The idiom** can be viewed as a *unique and fixed combination of at least two elements some of which do not function in the same way in any other combination (of the kind) or occur only in a highly restricted number*. This **operational definition**, easily convertible into a test for identifying idioms in general, is clearly based on the combinatorial, syntagmatic property of language items, an omnipresent and basic feature of any language unit. In this definition, generality and regularity are completely or almost completely lost and replaced by combinatorial **anomaly**. Thus, for example, idioms such as

*hold good, play truant, red tape, kith and kin, and as to,*

which are fixed combinations with a familiar nominative function and meaning, having, for the most part, no analogous counterparts in which their components recur with the same function; this can be tested by their commutation into the following tentative forms:

*\*hold bad, \*hold nice/fine; \*act/behave truant; \*green/blue tape; \*kith and family/relatives/clan; \*as from/through and others.*

Thus, a highly restricted **collocational paradigm** for all of the tested idioms, often consisting of a single item (*hold true, as for*), may be envisaged. It is, then, hard to maintain that the components of these unique combinations have independent meaning or existence outside the idioms in which they occur. This is the reason why the notion that some words have a specific phraseological meaning is rejected. Yet there is no denying the fact that one needs a criterion enabling one to decide when and how a new systemic meaning (sense) of an item arises, thereby permitting its new and relatively free combination and use. Until this new meaning is established beyond any doubt (current criteria badly fail us here), the field is open to idioms (for a further elaboration of this line, see Filipec – Čermák 1985, 170ff.). Parenthetically, cases such as *hold good*, and *as to* cannot be understood in terms of the current notion that idioms are not the sum of the meanings of their component parts.

In the examples above, the randomly chosen and tested components (*hold, truant, tape, kith, as*) should also be viewed on their paradigmatic axis. In the language system, every item must be seen as belonging to a class based on the same joint function (often represented as a set of synonyms of, for example, colours, vehicles, and professions). Such a class might be called **virtual paradigm**. Yet some of our examples seem to resist such classification, because no other members can be found for their hypothesized paradigms (*truant...?, kith...?* as against *red, green, blue,...*). One is forced to conclude that they do indeed form a one-member paradigm, which is a profound anomaly. In fact, this situation is true of all these components, once their polysemy is accounted for, but it is not true of all idioms; compare, for instance, the words *handy* in *come in handy*, or *party* in *stag party*, which function here in their literal senses.

It is the intricate interplay of the anomalous syntagmatic and paradigmatic positions of an item in the language system that makes a combination an idiom. Accordingly, a systemic

definition of the idiom might be given: *The idiom is a fixed syntagm at least one item of which is, if seen from the viewpoint of at least one other item of the syntagm, a member of a severely restricted and closed paradigm (always in syntagmatic function or paradigmatic membership, though often in meaning, too).*

As a consequence of the variety of component anomalies (that is, of the input), it is possible to claim that the combinatorial idiomatic output is, apart from functional and transformational deficiencies, often arbitrary in its semantic and semiotic aspects. Idioms are an inescapable resolution of the combinatorial blind alleys of any language, once the language deviates in its combinatorial structures from its regular possibilities and once its numerous anomalous combinations (of many kinds and types) become frozen and established as stable parts of the system of its nominations. Idiomatics is a residual anomalous counterpart to the regular, that is, rule-governed combinations of the language.

## 2. The Scope of Idiomatics

In accordance with the preceding discussion, certain conclusions are inevitable.

It is not possible to study the idioms of a language in isolation. Full knowledge and understanding of the syntagmatic rules of a language in their formal (including grammatical), semantic and functional aspects, and its paradigmatic properties is necessary. While the formal and semantic aspects of rules will help elucidate the anomalous, irregular processes involved in the constitution and structure of idioms, it is the **functional aspect** of rules and their contextual deficiencies that distinguishes particular idioms in their behaviour and function.

Language signs (items, units) on every level of meaning (to the exclusion of phonology, of course) figure in idiomatics, provided their (anomalous) combinations belong to some type of language nomination and are part of the language system. This provision includes, moving from the collocation-type idioms upwards and downwards, sentence-level idioms, perhaps also correlated communication pairs of sentences (such as some greetings and their answers), and anomalous morpheme combinations. It is clear that idiomatics, a linguistic discipline whose units cross most of the language levels and strata, should be viewed within a broad framework of language onomatology. Other aspects entering its scope – and it is hard to see which aspect might escape its attention entirely, as they all are projected into the field – exhibit relationships to other disciplines. Note that since the **semiotic** and semantic aspects of idioms are interconnected and the arbitrary semiotic character of idioms as signs is often absent from idioms having a concrete, material denotate, it is safe to say that the bulk of the idioms of a language resides in many types of abstract denotations.

The scope of **idiomatics**, viewed as both the set of idioms of a language and an autonomous linguistic discipline, overlaps with some non-linguistic disciplines as well, such as ethnology, psychology, philosophy, literary science, and history. Some of these relationships, primarily those related to etymology and logical structure, seem to be quite unique to this field (cf. Zich 1956, Mukařovský 1971, Büchmann 1946, Kuusi 1957, Permjakov 1979).

The inevitable fact that standard language rules and criteria of any kind become idiomatically twisted, limited, deformed or even cancelled requires both a search for new rules and a re-examination of the old ones against the combinatorial background of the language. In order to establish the point where the irregular ends and the idiomatic begins, it is necessary to know the exact nature of the rules of the language system and the extent of their application. Thus, advances in idiomatics are related to advances in the study of the regular language. Obtaining such knowledge will result in a new methodology of linguistics in general. It will provide us with better tools for analyzing the scalar nature of the rule-exception relationship.

### 3. The Idiom

As discrete units on most levels, idioms have integral meanings that usually do not permit any decomposition. Because of the marked discrepancy between the idiom's components and the resulting syntagm, that is, between the input and output, two correlated levels of analysis of three basic types, formal, semantic and functional, are necessary.

The **formal analysis** must include a description and classification of the type of components involved, of their paradigmatic membership (and constituent parts of a lower level, if any) and syntagmatic capability (collocability, if any, where nonce-words or monoclocable words of the type *kith*, *truant* may occur), and the structure, order, possible variants and reversibility of the form of the idiom itself. Furthermore, it is necessary to establish the formal links among the idiom's constituent parts (notably those due to grammatical words or inflection), what kind of component substitutes or modifiers are used (pronouns and numerals, primarily), and what type of formal valency the idiom has.

The **semantic analysis** must account for the type of meaning and denotate of both the idiom and its components, the type of semantic participation of the components (i.e. the extent and kind of their presence) in the idiom as a whole and for the **pragmatic component** of the meaning, which is a pronounced and general feature of almost any idiom. A special technique is needed to resolve the notable **semantic indeterminacy** of idioms, leading to their contextual referential vagueness, when they are put to use, and to their pronounced susceptibility to distributional shifts and non-standard usage.

The **functional analysis** will cover the syntagmatic capability or **collocability**, if any, of constituents, paying special attention to nonce-words of the type *kith*, *truant*, i.e. those empty forms or verboids (monoclocable words) that are to be encountered in idioms only. Next, the idiom's **structural function** or class collocability should be established, in order to determine the idiom's stable function in a higher-order language unit. The function of collocational idioms (those composed of word forms and not having, in the system, a propositional function) includes a variety of types, such as subject or predicate (but not object) (Čermák 1982, Filipec – Čermák 1985).

## 4. The Idiom System

The fact that idioms might form a system of some kind is not clearly recognized; in fact, our definition of the idiom seems to preclude the existence of such a system. However, the underlying combinatorial principle and its anomalies do provide a useful basis for classification. Thus, on each level of idioms a calculus of combinations or rather permutations of potential input constituent types can be given, providing a general inventory of possible structural types including idioms (i.e. after the nonexistent combinations have been subtracted). From this inventory, three types of combinations (in fact, permutations) will be selected: (1) regular language combinations and structures only (if any), (2) those occurring only in idioms (for example, *trip the light fantastic*), and (3) those in which both fields overlap (the majority of instances). It is a rarely stressed fact that the stock of possible combination types of a language is far greater than that which is actually used, and idioms do not automatically follow exactly the same path as regular language combinations.

Hence, the elements employed in these permutations (of the type 2 and 3) must be chosen on the basis of their relevancy to each level. A natural choice for collocational idioms would be functional word classes; for propositional idioms the option would be word and collocational classes; for multipropositional idioms (such as correlated greetings or larger proverbs) clause types can be used and, finally, for lexical idioms the type of morpheme and its combinations must be considered (belonging to the sphere of word derivation or inflection). Other levels of analysis are not obvious.

It is evident that the elements considered in these permutations are not of the same order; in their function and nominative power they form several subclasses. Thus, it is useful to split the traditionally favoured level of **collocations** (though under different names), wrongly posited as the only one where idioms obtain, into three subclasses: that of autosemantic elements (nouns, adjectives, verbs, adverbs), that of synsemantic elements (grammatical words, basically), and a third group made up of both types. A special status has to be given to nonce-forms which do not easily lend themselves to classification. For example: (a) *make believe, red tape, come true* (here, unlike the regular combinations, Czech idioms employ all of the sixteen possible binary permutations; Filipec – Čermák 1985); (b) *as if, so far, in all*; (c) *by and large, let up, put off* etc. Idioms in (c), in which one of the constituents is a synsemantic word, should not be confused with a subtype of (a), such as *come to the point*, or *catch one's breath*, where the same class of words serves as a relator connecting other components and where it is, then, used in an auxiliary role only. Other types of subclasses can be employed as well, such as that of substitutes.

## 5. Transformations

A number of quasi-transformation types can be given, which allows us both to describe systemic relations between various idioms and to distinguish transformations (being independent idioms usually) from variants (of a single idiom). The latter seems to be a problem plaguing numerous approaches to this field, though the basic confusion can be resolved by distinguishing transformations from mere variants.

**Systemic transformations** may mutually relate idioms

- (1) of various levels (a rarely perceived case, compare *stolen fruits/sweet are the sweetest, the sweet taste of stolen fruit*);
- (2) of various structural types of the same level (*blow one's own trumpet, trumpet-blowing, take one's ease – at one's ease*); and
- (3) of the same structural type (this type being identical with variants in most cases, however, for example, *be none/not any of somebody's doing, not darken somebody's door, not darken somebody's door again*).

Sometimes, within the same structural type, one can also distinguish

- (4) a special type of transformation that could be called **statutory transformation**,

as it covers different stages or states of an action or situation expressed by an idiom, such as inchoative, durative, terminative and, in another sense, causative, as illustrated by such examples as

*take it into somebody's head, put it into somebody's head, put it out of somebody's head;*

these transformations appear to be much more frequent in some languages than in others (such as Czech, perhaps due to its inflectional character). Any considerations concerning psycholinguistic reality of these transformations must be derived from the degree of their stability in memory and the degree of reusability of the transformational pattern.

**Textual transformations** including all permissible grammatical transformations (syntactical and morphological), such as passivization and many other may serve as a basis for one of the most reliable classifications of idioms. The inventory of these textual transformations (to be distinguished from those giving rise to new idioms) displaying many well-known anomalies is an expression of the idiom's limited structural function as well as of its specific meaning.

## 6. Functions of Idioms

Under the broad label of the **communicative function**, at least two distinctive spheres must be examined; the **structural** and **nominative** function of idioms. The former concerns the customary place of the idiom in the structure of a higher unit (for example, the idiom *dead end* occurring in both its predicative and attributive function, though usually not in the role of the subject), whereas the latter concerns the type of the basically semiotic role of the idiom in discourse, interrelated with but not identical to its meaning. Thus, most idioms seem to be primarily designators, others operators (such as quantifiers, commentors, deictic operators, etc.). Analogous functions must be specified for the propositional idioms, too, where a solid knowledge of text types, propositional attitudes, pragmatics and other areas must be at the disposal of a student of this field.

Due to the fact that most collocational idioms have a verbal character, it appears that the primary function of this type of idioms of a language is to provide its users with a stock of



**ready-made predicates.** This major structural function of verbal idioms is correlated with the prevailing predicative function of other idiom types as well. A major nominative function of idioms is to primarily serve as expressions with a high amount of subjective, **evaluative** and emotional load; in this function they seem to have no parallel in other parts of the language. Moreover, idioms are remarkably economic and conveniently vague means of linguistic expression in any language.

## Shrnutí

### Idiomatika a frazeologie

Frazém či idiom lze nazírat jako takovou *jedinečnou kombinaci aspoň dvou prvků, z nichž některé nefungují tímž způsobem v žádné jiné kombinaci téhož typu, anebo se vyskytují pouze v jejich krajně omezeném množství*. Tato definice se dá převést na definici operační a test, vycházející z deklarované syntagmatické anomálie, srov.

*hold good, play truant, red tape, kith and kin*, a také *as to*

které nelze nikterak ve smyslu této definice modifikovat, srov. pak

*\*hold bad, \*hold nice/fine; \*act/behave truant; \*green/blue tape; \*kith and family/relatives/clan; \*as from/through* atd.

Takto lze dospět k pojmu kombinatoricky a uzuálně vymezeného kolokačního paradigmatu, které je však v daném případě vždy radikálně omezené na jediný člen. Podobně ale ani virtuální paradigma, tj. to, do něhož každé slovo patří už v systému (srov. třídu barev, a např. *zelená, modrá, červená*), není u sledovaných prvků jasné (*truant...?, kith...?*). Jemná souhra anomálních aspektů syntagmatických a paradigmatických pak vede k zobecnění a systémové definici: *frazém/idiom je ustálené syntagma prvků, z něhož aspoň jeden je z hlediska jiného členem krajně omezeného a uzavřeného paradigmatu (vždy funkčně či v paradigmatickém členství, i když často v obojím)*.

Přijetím takového vymezení se dostáváme k určitým závěrům. Do oblasti frazeologie spadají takto všechny jednotky, které jsou nadány významem a pevnou funkcí. Obecně však frazeologie některými aspekty přesahuje do dalších disciplín. O většině frazémů lze také prohlásit, že sémioticky patří mezi takové jednotky, kde na úrovni celků vznikají jednotky s denotátem abstraktním.

I když je frazém jednotka zpravidla diskrétní, nedovoluje v důsledku svého celistvého významu žádnou dekompozici. Formální aspekty jejich popisu si všímají složení komponentů, variant, odlišení spojovacích prvků apod., jakkoliv ale jen na úrovni těchto komponentů. Sémantické aspekty postihují povahu komponentů a jejich (jen někdy) souvislost s povahou významu celku, ale také sémantickou vágnost a důležitou pragmatickou složku významu. Funkční aspekty si všímají především strukturní funkce frazému ve vyšší jednotce, např. verbálních frazémů ve větě v roli predikátů apod.

Postulovat pro frazémy v jazyce zvláštní systém kvůli povaze definic výše nelze. Kombinatorický kalkul základních typů komponentů, např. slovních druhů, ukáže na strukturní možnosti obecně, ať už realizované či ne, a frazémové struktury v nich. Takto nacházíme jednak pravidelné kombinace, resp. permutace, kombinace vyskytující se pouze ve frazeologii a nakonec kombinace sdílené jak jazykem pravidelným, tak frazeologií.

Struktury takto vydělené lze členit dvojím způsobem. Obecně na frazémy kolokační, kde jsou jejich struktury vytvářené kombinacemi slovních druhů, ale nemají větnou povahu, a struktury, které

povahu větnou mají. Obojí se uvnitř člení podle specifické funkce; ta má na úrovni kolokačních frazémů povahu analogickou slovním druhům.

Pro frazémy je příznačná i anomální existence několika druhů transformací, které nelze zaměňovat za varianty. Nejvýznamnější jsou systémové transformace různých typů spojující frazémy transformačně spjaté, ale odlišené funkcí. Proti nim stojí textové transformace, resp. transformace gramatické, které mají vždy individuální povahu, specifickou u každého frazému.

Funkčně lze frazémy třídit na základě jejich strukturní funkce (viz i výše) a funkce nominativní. Výrazná je např. predikátová funkce frazémů verbálních. K nominativní funkci lze počítat i výraznou evaluativní funkci.

# Substance of Idioms: Perennial Problems, Lack of Data or Theory?

(In: *International Journal of Lexicography* 14, 2001, p. 1–20)

## Abstract

In the following remarks, some of the major recurrent problems of the phenomenon of the idiom, or, alternatively, phraseme and, partly, of its discipline, i.e. idiomatics or phraseology, will be taken up. In so doing, the roots of these problems will be sought. A brief analysis of some cases and practices of how dictionaries deal with idioms is followed by a restatement of major aspects of the discipline and a list of desiderata.

## 1. An Introduction, a Brief Survey of the Situation and some Problems

It is evident that some of the basic problems are far from being clear and that two basic approaches to the **idiom** or **phraseme** can be spotted in various contributions. On the one hand, there are linguists who seem to be content with a rather limited intuitive view, while, on the other hand, other linguists keep asking uncomfortable global questions, trying to find out what idioms are really about, what is their **essential nature**, what kind of phenomena belong to this field and how to delimit them. It must be said that the following remarks will be related to this technical meaning of the term, leaving aside the broader and, perhaps, more popular meaning relating to a particular or unique manner of expression (not only in language). Neither will other, alternative, terms be discussed here, since their sheer number, if various languages were brought together, would go up, perhaps into the hundreds.

### 1.1 Some Phraseological Activities

The quest for a deeper understanding of the phenomenon of idiom is made no easier by the basic isolation of certain groups of researchers. There are, to take only the most notable examples, the Slavists with a long and perhaps richest tradition in Europe including regular conferences (with one of its latest platforms published and accessible in at least six volumes of the international series *Z problemów frazeologii polskiej i słowiańskiej* (*On the Problems of the Polish and Slavonic Phraseology*, cf. Basaj – Rytel 1994). There are also people involved in the activities of *Europhras*, who, having established the *European Society of Phraseology* in 1999 ([www.europhras.unizh.ch](http://www.europhras.unizh.ch)), are concentrating an increasing

number of people, languages and countries (with eight volumes printed in various countries and one more in press, cf. Sandig 1994), rather isolated events like those of the organizers of the Dutch Tilburg conference in 1992 (Everaert et al.), the so-far only three English-based *Symposiums* (the third of its kind held in Stuttgart in 1998) and, last but not least, contributors to the *Euralex* conferences. However, the mutual flow of information between these people has been limited, which is not only, presumably, because of the language barrier.

Obviously, such a situation of somewhat fragmented groups and individuals working in partial isolation has not helped the object of research, which is, of course, general for all languages. I do not propose to go here into the particular merits of particular schools, although it is astonishing, for example, to see how little scholars oriented towards English know about other schools of research, in particular about early Russian phraseological research, followed up by Polish or Czech studies, about German and Swiss activities in the field, etc., let alone about individual scholars. The phenomenon of **idiom** has also become the subject of study of some linguistic schools, including structural, generative or stratificational approaches, complemented recently by psycholinguistic and cognitive ones, yet very little communication across their borders has ever happened. One of the recent instances of this unnatural and vexing lack of information and contact was to be seen at the phraseological round-table (held in Paris, too) of the 1997 International Congress of Linguists.

This state of affairs is also reflected in the inability of linguists to even agree on a common term for the central unit of the field. It is **idiom** and **phraseme** which seem now to be becoming more and more acceptable but they are still far from replacing hundreds of alternative and mostly traditional native terms used, for the same thing, in various European languages. Moreover, this situation is made no easier by some new terms contributed by researchers. One must be also careful about what certain terms in use seem to suggest if contrasted with what they really denote. Thus, for instance, **free collocation** (or **combination**) should not be understood as free in any sense, as it is always based on formal and semantic rules and constraints (such as compatibility). Similarly, the term **restricted collocation** seems to point to an area of such types of collocations as *take care of*, *face the truth*, based on limited choice or selection restrictions. But in its literal sense, restrictions on choice can also be found elsewhere, in other types of word-combinations, such as idioms proper or multi-word terms where other kinds of restrictions operate. It is the specific sense restriction of the word 'restricted' which leads to some confusion, and it does cast a shadow on the discipline of phraseology itself, especially in the eyes of other linguists. In a sense, all word combinations are restricted in one way or another, since all words have a limited potential to combine with other words and **there is no such thing as a word without any collocational restriction** (indeed, even such words as the article *the* cannot be used freely with any noun, let alone with words of other word classes).

This seems to be some of the background, which clearly suggests, in my view, a serious lack of mutual information and which also explains why, from time to time, the same notion is discussed again and again, though under different labels.

## 1.2 An Approach and Problems Involved

The second general point I wish to make concerns **methodology** used in idiom research and the status of our (often vague) **argumentation** found in its interpretation, as well as their acceptability and applicability to the whole field. Let us have a look at a randomly chosen example which, being quite common and typical, could easily be repeated with reference to other authors. Thus, in his contribution to the Tilburg conference, Erbach makes the following statement:

‘Lexicalized<sup>4</sup> metaphorical<sup>3</sup> expressions (or analyzable<sup>1</sup> idioms) like *pull strings* ... differ from idioms proper in that the meaning of the expression can be modified<sup>2</sup> by modifying its constituent parts, e.g. *pull certain strings* ...’ (Everaert et al. 1992, p. 12).

It is amazing how most of the major problems of phraseology can be condensed into a single sentence. There are at least four serious problems (1–4) touched upon in such a wording.

### 1 Analyzability: How do we know that the idiom

#### (1) *pull strings*

is really **analyzable** and how does the author propose to analyze it? What does, in fact, analyzability amount to? What is even of greater importance is the implication, that is hard to accept, that such a general statement, not supported by any real argumentation, should be made a crucial criterion distinguishing two major types of idioms, namely those which are ‘analyzable’ and those which are not. A major problem, which especially cognitive linguists do not seem to be aware of, or pay much attention to, is that of methodology. First, one has to distinguish, in trying to analyze an idiom, between the form and the meaning. It is obvious that, because of the idiom’s discrete form on the level of collocations, one can dissect any idiom *ad libitum*. But does this also involve the possibility of an unproblematic and principled analysis into corresponding parts of meaning? For some reason, this kind of would-be analysis seems to be content with sticking to the form only. The second point to be made here is that the analysis should clearly be seen as being of two types. One is **based on analogy**, can be repeated (or is found in a number of analogical cases in a large corpus, for instance) and is thus a constant feature of the idiom, and hence part of **la langue**. The other is an individual and **idiosyncratic approach, based on no kind of analogy**, shared by others. The third point raises a different problem: it should be made quite clear **what the criteria for the would-be analysis are**, which is not the case here. It is certainly not true that such a dissection here is quite free and standard in the sense of any regular syntactic and semantic analysis. This raises the next point, modifiability.

**2 Modifiability:** Leaving aside the possibility which suggests itself, namely that **analyzable** is often made identical with **modifiable**, does the presence of *certain* really justify another overgeneralized statement that the idiom is really modifiable? In what sense is it modifiable and is it, then, still the same idiom? Are there any semantic or other restrictions? How many

other individual words and, in particular, types of words can one insert here and retain the original idiom? Can one freely improvise, using, for example:

- (2) \**pull long* (A) / *short* (A) / *strong* (A) / *weak* (A) *STRINGS*, or  
 (3) *pull interesting* (B) / *hidden* (A) / *political* (A) / *influential* (A) *STRINGS*

or, to take up a somewhat mixed choice of possibilities:

- (4) *pull iron* (C) / *nylon* (A) / *fish* (C) / *3 metre-long* (B) *STRINGS* etc?.

A native American informant suggests that one can, indeed, do this, although in a very limited manner and different sense. However, none of these possibilities is viewed as a stable, integral and organic part of the idiom (where the 'modifications' suggested were marked by the informant as A = possible, though not usual, B = remotely possible, C = impossible). It seems that all of these acceptable 'modifications' are pure textual phenomena only that do not alter the idiom's substance or produce a different lexeme combination. Having recognized the fact that such modifications are no part of the idiom, can we still maintain that whatever process has taken place here is really a modification of the original thing and not, perhaps, something else? What do these results amount to? Does the fact that a restricted range of quantifying adjectives (as in A cases) is possible as a 'modifier' here have any relevance for the idiom's status? Should one, perhaps, view (A) cases as mere variants? It is doubtful there are clear answers to such questions.

However, the obvious semantic restriction of the idea of modifiability is to be seen in the case of B only, which should be explored in more detail. Briefly, it seems that once an additional external semantic (or pragmatic) feature is forced onto or into the idiom (in most of the examples, being clearly **intrinsic** and therefore **potentially semantic features** of *STRINGS* in its dictionary meaning), the possibility to modify it is no longer so easy and becomes very remote, indeed.

Erbach's own example and similar views seem to point indirectly to yet another, little explored, possibility of insertion of **metalinguistic markers**, such as *certain*, *kind of*, *proverbial*, *true*, *veritable*, which he does not consider. These and their possibility could be viewed as a converse case, in fact, pointing to **typicality** and, accordingly, a typical non-analyzability. Should the term modification be applied also here, it would be a very different modification, indeed.

What one is dealing with here seems to be the creative potential of the language itself, operating anywhere in language, and that can hardly be taken for a specific criterion to be used in the idiom's analysis. It is its identity that is to be explored, not possibilities of its disintegration. Obviously, **any analysis must be based on well-defined criteria** and must be repeatable elsewhere.

**3 Metaphor:** Although there may not be much disagreement about our recognizing a metaphor here, a general question must be raised. How can one be sure to be able to safely recognize metaphor in *all* relevant cases and make it, therefore, a criterion?

Cognitivists, relying so comfortably on obvious, typical and central cases, do not seem to be interested in cases that are less typical and hard to define. Certainly, views on what metaphor is differ in various cases and an individual occurrence, about which there might be some agreement (or, perhaps, in a group of occurrences), does not guarantee one to be able to jointly distinguish the metaphor in every case that is being analyzed. **The notion of metaphor seems to have a rather blurred periphery and, definitely, a different capacity to view it with different users.** It seems to be one of the fairly established illusions that metaphor can be made a reliable and general criterion for the idiom. The tradition of trying, over and over again, to define what metaphor is, is a venerable one, dating from Aristotle and being, certainly, one of the topics linguists have written most often about. Yet, no definition of what the author thinks that metaphor should be is given here. However, some of the cognitivist approaches (see, for example, Dobrovolskij 1997), which are equally hard to accept for their vagueness, interpret even such idioms as *spill the beans* as metaphoric. In a sense, because of its being largely subjective and related to some parts of the field only, reference to metaphor seems to be rather counter-productive.

**4 Lexicalization:** Why is this particular type of idiom characterized by the attribute **lexicalized** while the second major type, which the author calls **unanalyzable idiom**, does not have this attribute? Is this to suggest that, in his view, perhaps some idioms are not lexicalized? Are, then, those idioms that are 'not lexicalized' free collocations? How is the label lexicalized defined here?

Many linguists do not now put much trust in the particular brand of linguistics associated with Noam Chomsky, because of the favourite one-example type of approach which was often made up to support a particular a priori theory or belated idea. If somebody was able to come up with a different counter-example, the theory was changed, and this has been going on for quite a long time. This no-data approach, so strongly opposed nowadays by corpus linguistics, seems to re-emerge, yet again, in some very limited cognitivist and psycholinguistic experiments (see, for example, a summary in Dobrovolskij 1997). Methodologically, it seems difficult to accept another, equally partial, solution, one that is evidently not general enough. Also the above-given type of argumentation, which I have tried to scrutinize, is of Chomskyan brand. One of the great problems of the theory of idioms and phrasemes still seems, after so many attempts and theories, to be a lack of criteria consistent with the bulk of data. To jump to conclusions on the basis of a few suitable examples leads nowhere.

The discipline dealing with idioms and phrasemes is a very intricate one where most of the 'safe' distinctions we know from elsewhere have become blurred and indistinct. Yet, admittedly, linguists do generally wish to cover the field decently, in an exact and unambiguous manner, leaving no large areas unattended, which means that this undertaking must be based on ample data and sound and consistent argumentation. It also means that, in view of the unique information offered by corpora, our criteria, very often mere labels and not properly data-founded, should be re-checked against this data and modified, if necessary. Certainly, one should nowadays be more critical towards such sweeping statements as those above.

## 2. Problems of Form, Meaning and Function

Some of the problems to be raised here have been suggested above. On a more general level, all of the problems in the identification and analysis of idioms should and can, hopefully, be handled in three inter-connected approaches, linking together their form (2.1), meaning (2.2) and function (2.3). Of the number of aspects involved here, only those that seem to be of a basically controversial nature will be briefly mentioned.

### 2.1 Form

As far as form, or rather formal aspects, of **the idiom** are concerned – let us use this single term here to cover a multitude of phenomena related to its form – at least five of these should be raised, including the idiom's identification, stability (fixedness), variability, multi-word character and derivations. All of these make up much of what is considered to be the basics and essence of the phenomenon of idiom.

As yet, there has been no way suggested how to handle the idiom's identification (and identity) in texts, such that would not leave many borderline cases open and even types of idioms unaccounted for and that would be able, at the same time, to draw a distinct line between various types of combinations suggested so far. Such cases of a type of restricted collocation as *black coffee* have never really been successfully and systematically accounted for. An attempt of this kind used in the Czech approach will be mentioned later (see Čermák 1982, 1988, 1994a, 1994b, 1994d and Filipec – Čermák 1985).

A prerequisite for this is to have and be able to use safe and criteria for establishing what **stability** (fixedness), fixed or institutionalized nature of some text combinations (fixedness) is and where it starts in the statistical continuum of various combinations. Stability seems to be one of very few basic attributes of the idiom that linguists may have some sort of consensus about. The obvious criterion to be used to determine this seems to be frequency, of course, but that is a scalar notion offering, in fact, no sharp boundaries in between the obvious extremes. Consider the fact that the 5 million Word Bank Corpus of English does not have a single record of *white coffee* (whatever that might represent) and only two occurrences of *black coffee*. Viewed in this light, are these two collocations found here to be declared really stable or do we point to the scarcity of data and avoid the answer? While it is not the inclusion of these particular cases in the category of idioms that should be argued here, the point to be made, next to their stability (which exists, at least in the case of *black coffee*), is a semantic **anomaly** involved in both (i.e. neither *black coffee* nor *white coffee* is really black or white, respectively), while there is no anomaly involved in other collocations (such as *brown coffee*, *green coffee*, etc.).

A complementary feature to that of stability or fixity is the idiom's **variability** (or variance) which is, sometimes, taken to be a decisive criterion for idiom classification. This is also the case of Chitra Fernando's recent (1996) book where, however, no really convincing argument for its relevance, or its scope and essence, is given and explained. Many other researchers have tried to tackle the phenomenon of variability, some of them (e.g. Moon 1996) offering a list of types (but see Filipec – Čermák 1985). Generally, it



seems that real usage offers an ever-changing and constant display of new variants, some of which may become stable in due time. In the meantime, one has to face a **scalar notion of variability** again, ranging from inconspicuous grades of it to those that are strikingly obvious. While **it seems hard to accept variability as a constitutive criterion for the idiom**, the real problem it seems to address may lie elsewhere.

Let us consider an abstract model idiom made up of three constituents. Current approaches do not hesitate to call such a case, where only *one* of the three constituents has a variant, a specimen of variability. But what about the case, in which *two* of the three constituents are variable, i.e. over 50 % of the constituents? Or, perhaps, two constituents are variable in a four-component idiom? The obvious question behind these putative cases, that are quite common in reality, is: how do we know when we are still dealing with the *same* single idiom and when it is already the case of two or more distinct idioms? The problem involved in this rather simplified example, where the degree of variability was taken to be the same in each case, that of the **idiom's identity**, has been given all sorts of ad hoc solutions so far, but probably never a principled one. In fact, most of the features I have mentioned so far, and some of those yet to follow, are all related to this very basic problem of identity.

It seems that most of the approaches to the idiom have stressed so far yet another feature, that of its **multi-word character**. In practice, this boils down to what might be called **collocational idioms**, i.e. non-sentential combinations of two or more word forms (Čermák 1994b). However, such a solution suggests at least two basic questions which are in need of reliable answers.

The first might be phrased quite simply: how many words does *multi-* really include, or, to put it differently, does it also include sentence idioms, or, perhaps, those made up of more than one sentence? The obvious answer, often given today, is that **proverbs**, for example, might be included in the idiomatic stock (although the history of rejection is quite long). Yet the problem behind the question seems to be much more broad and complex. To give an example of what else might be involved, I will mention the approximate figure of sentential idioms, of many other types, that our team has identified so far (in our corpus, following the test mentioned in section 4) and has been working with (to become the fourth volume of the *Czech Idiom Dictionary*, Čermák – Hronek – Machač, 1983, 1988, 1994). It now amounts to over ten thousand idioms i.e. to a number roughly equivalent to the number of verb idioms of the collocational (i.e. non-sentential) type, i.e. those that have traditionally been considered to form the bulk of the field. The field is rather vast, including such widely used idioms with a prominent pragmatic value as *You must be joking!*, *The penny has dropped.*, *Heads or tails?*, to give only some comparable English examples. Yet, except for some occasional remarks and isolated examples, not much thought has been given to this still rather new subfield of idioms, let alone to others.

In a sense, the second question is more important in that very few linguists ever stop to ask it, let alone try to answer it: **are there only these two levels** (i.e. **collocational and propositional ones**) where idioms are to be found? To answer this question, one must consider other, below-the-word, combinations of idioms (i.e. idiomatic morpheme combinations), such

as the English *backlash* (*lash* + *back*), *bottleneck* (*bottle* + *neck*), *breakfast* (*break* + *fast*), where the sum of the constituent meanings does not equal the integral one, or the Dutch *katjespel* meaning ‘quarrel’, which is a conjunction of *katje* ‘little cat’ and *spel* ‘game’. Or consider the German *Katzejammer*, one of the words for ‘hangover’, formed by a combination of *Katze* (cat) and *jammer* (misery), the Finnish *mustasukkainen* ‘jealous’, a combination made up of *musta* ‘black’ and *sukka* ‘stocking’ plus the adjectival ending (‘having black stockings’). Or, to use a slightly different type of example, one can point to the Czech *zákeřný*, an adjective corresponding to ‘treacherous’, which is made up of the prefix *zá-* in the sense of ‘behind’ and of a noun *keř* ‘bush’ (plus the adjectival suffix *(-ný)*, amounting to something like ‘behind-the-bush-lurking’). All of these, and many other examples, seem to be perfect **idiomatic combinations of morpheme forms** and our answer to the question above should be to include these as well.

It is really no use to try to discard these as marginal or non-existent cases, as one would discard, at the same time, much of what is so much alive and central in, for example, agglutinative languages, such as Finnish or Turkish, or in some isolating Asian languages like Vietnamese or Chinese (see also Čermák 1998a).

This requirement should be taken even more seriously, once we recognize that Chinese, for example, is notorious for its not being able to draw a sharp distinction between the morpheme, word (and syllable), and that the distinction between compound and collocation is difficult to establish there, too. Thus, to take a very simple example, it is not clear whether the Chinese *tielù* (‘iron + way’), standing for ‘railway’, is a compound or a collocation. A look at the Chinese terms (compounds, probably) for ‘past’, ‘present’ and ‘future’ and at their structural irregularity shows their idiomatic structure even better: *guòqù* (‘the past’ = ‘pass’ *guò* + ‘leave’ *qù*), *mùqián* (‘the present’ = ‘eye’ *mù* + ‘front part’ *qián*) and *jiānglái* (‘the future’ = ‘act’ *jiāng* + ‘come’ *lái*). Such familiar facts and intrinsic problems have to be accounted for in any non-idiosyncratic theory of idioms. It is, basically, the same combinatory problem, whether it is being recorded in writing as one word or a combination of several. Ultimately, all this boils down to the obvious fact that the general status of items that enter idiomatic combinations has to be reconsidered and extricated from what has been tailor-made to suit some European languages, or, worse, some theories only, so that it could include typologically different languages, too.

Finally, no clear criteria seem to be available to help one to distinguish what are sometimes called **derivations** (and what I think should be preferably called **transformations**, see also above), such as *corner-cutting* ( $\Leftarrow$  *cut corners*) or *arm-twisting* ( $\Leftarrow$  *twist somebody’s arm*), from the bases they are derived from. Are these particular cases variants, new idioms or something else? Presumably, also these display a prominent anomaly in their not being formed automatically. I will come to this point under Function (in 2.3).

## 2.2 Meaning

Semantic aspects, such as **analyzability**, **compositeness** or **compositional meaning**, **motivation** and **transparency** of the idiom, terms which are so often used (or misused),

seem to be equally a matter of controversy as are the relevant aspects related to the idiom's form, discussed above. As they are all interrelated, it is difficult to single them out one by one for discussion.

First of all, let us notice that there seems to be no end to attempts to throw some kind of doubt on the non-compositional or non-Fregean nature of the idiom. Judging the discrete appearance of many idioms, this might seem natural. Non-compositeness of the idiom's integral meaning seems to be a *conditio sine qua non* for the idiom's semantic substance. Yet, those who do not hold this view keep trying to show that the idiom can be semantically analyzed.

What does a 'semantic analysis' along these lines often look like? Approaches of this type to an idiom such as

(5) *rain cats and dogs*

often glossed as 'to rain very heavily or steadily' (*Random House Webster College Dictionary* 1991), would obviously, in these 'compositional' or 'analyzable' approaches, dissect it into *rain* standing, obviously, for 'rain', not really bothering to explain or mention the rest (*\*cats and dogs*). The claim made would amount to an observation like 'it looks as though the "rain" component is there, hence the idiom is analyzable, compositional and motivated'. Sometimes, this gives rise to a, further unanalyzed, notion of partially compositional idioms, as if this mere re-labelling of the uncomfortable fact offered a way out (see recently, for example, Dormeyer – Fischer 1998). It is very easy to show that this is not so in pointing to the rest of the expression – which is being left unattended and unexplained – and to the, perhaps, equally obvious fact that one may face a serious problem in saying that the rest, namely *cats and dogs*, amounts to, roughly, 'very heavily or steadily'. The seemingly incontestable logic behind such an approach is a typically pars-pro-toto one here. On a further reflection, it is apparent that no such metonymy obtains here and the procedure cannot be repeated to apply for the body of idioms. Hence, the procedure becomes faulty.

Methodologically, and in a very basic sense, **it is unacceptable to see any type of compositionality** between the meaning of that part of the integral form which seems loosely to correspond to 'some' type of meaning (familiar from the dictionary) and the rest of the form which does not correspond to anything in the realm of meaning. To put it very crudely, any potential comparison with simple arithmetical 'a number + zero' has to be refused, since zero is a figure, too. In idioms, one should rather consider the arithmetics (sum) of the type 'something + absence of something', should this type of reasoning be used. Of course, this does not prevent the idiom from having its regular **holistic effect of compositionality**, which is of a different nature.

Should we accept the notion of 'partial compositionality', it would be a different type of compositionality from the traditional one (some of the misunderstanding might stem from its not being elaborated enough). It would still leave us with the necessity of coping with the unexplained rest, and force us to explain how it fits into the internal structure, incontestable

for some, and what its role inside the idiom is. To be able to do this seems very doubtful – despite the fact that the search for relations and functions is the linguist’s business in any analysis, and one should question the wisdom of re-labelling the problem by a misleading name. Moreover, it is rather risky to rely on the apparent clear-cut correspondence between the seemingly motivated formal part and its corresponding part of the integral meaning. In how many cases can one safely do this, and is there any analogy behind it? Obviously, this is a matter that still has to be explored. In so doing, one would have to imply the possibility of a partial semantic analysis, which is difficult to accept for **lack of any reliable tools** – however, impressions and ad hoc ‘solutions’ are no arguments. By reversing the approach, however, and viewing the problem from the point of view of the idiom’s components, one may safely state that, at least in some cases, the component **makes some semantic contribution** – as yet not reliably studied and perfectly understood – and leave it there, not trying to fit it structurally into the overall structure (as any analysis would, in the reverse order, have to do anyway). The state of argumentation seems to be rather unbalanced here, ranging from mere declarations of faith to one-example analysis, presumed to hold elsewhere, without any really convincing large-scale analysis ever being presented. It is certainly difficult to accept seriously the kind of argumentation presented, yet again, in Nunberg et al. (1984) (quoted here from Dormeyer – Fischer 1988), for whom there exist parts of a class of idiom which ‘have identifiable meanings which combine to produce the meanings of the whole’. This peculiar and rather opaque pronouncement seems to suggest that some parts only (with identifiable meanings) combine with others (without any identifiable meaning) to form a whole.

This alleged subsense or meaning of a part, related to the notion of ‘partial compositionality’, which some Russian linguists used to be so fond of calling it a **specific phraseological meaning**, some time ago (Vinogradov 1947), appears to defy any rational or objective analysis, if executed exhaustively. To make the integral meaning really compositional would mean that both semantic elements must be objectively recognized as being in existence and supported by a sound **analogy**. The obvious conclusion is that this approach is not viable and one might spare oneself a lot of trouble if one does not pursue this line of thinking any further. There have been many variants to this approach, starting, perhaps, with V. V. Vinogradov’s rather preliminary and now hardly acceptable semantic classification of idioms and collocations as early as 1947, which was based on this idea and which has influenced a great many linguists in the field. Discarding it, the obvious conclusion is that **the idiom’s meaning is not compositional**, irrespective of all sorts of seemingly analyzable parts of it.

Semantically, the idiom has to be viewed as a **holistic, Gestalt phenomenon**, a feature which is often acknowledged, and which by definition precludes any possibility of an objective and exhaustive semantic analysis. Hence, it is also difficult to see any sense in using, in the strict terminological usage, other related labels, such as **transparent** and the like. Consequently, there seem to be no available criteria enabling us objectively and fully to distinguish here any basically clear-cut idiom types of this sort in an exhaustive manner. Obviously, this conclusion does not preclude cases, not always easily discernible, where some indeterminate semantic link or **participation** can be established between the meaning of a constituent and the integral meaning of the whole idiomatic combination. While it is not

possible to determine its exact status – the approach viewing it as a motivation arrived at by an analysis should be discarded, too, as there is no sufficient analogy for this – indeed, it may be much more useful to question the very wisdom of doing this, if its vague results can never serve as a reliable and systematic criterion for analysis and classification. Seeing the persistence and recurrence of such attempts (nurtured recently by cognitivists), one may wonder whether this is a kind of **linguistic analysis or a loose impressionistic literary interpretation** of a kind that one is after. Linguists who are still prone to pursue this line of approach will, sooner or later, inevitably become aware of limitations here and will have to be satisfied with rather vague and meagre results drawn from carefully selected areas and few cases only. Surely, such an eclectic approach is far from being part of the exact linguistic methodology which is so much needed in the field.

On a more specific level, one might also argue against this kind of reasoning present, in a sense, in the **cognitive metaphor** approach, which refuses, in fact, objective reference and testing as relevant criteria. Nonetheless, it is doubtful if George Lakoff (1988, p. 125–126) would be prepared to admit *cats and dogs* as corresponding to some sort of obvious metaphor and, hence, referring to a specific concept that it is based on, as it is **conceptualization** that seems to be at the core of his approach. It is difficult to match the form of *cats and dogs* and any concept here in any reasonable way. But again, the primary methodological problem, central to anything in the language, is the **relationship of analogy and anomaly** and degrees of both; in this case, it is the (sufficient) degree of analogy – an objective criterion – that one should be able to pin down, not individual and subjective aspects of the particular case in question.

What one is also faced by here is **vagueness**, an important aspect which is very much present, as a feature, in the idiom's meaning and function and analysis and which has, so far, never been given due attention in the field of idioms; it is related to a number of other features. This, however, should not be projected into vagueness of description.

## 2.3 Function

There seems to be an obvious difference between cases such as

- (6) *see or think fit to do something*
- (7) *corner-cutting or cut corners*

Clearly, this difference points to their communicative (textual) **function**, and such cases should thus be handled functionally, too. Accordingly, the former (6) may be taken as a case of two variants of the same single idiom with its identity kept intact, where only a marginal difference of meaning is to be considered. The latter case (7) represents two idioms which are transformationally related, but which have a different syntactic function, the first, broadly speaking, of noun phrase, the other of predicate. This difference in syntactic function implies other types of co-related idioms, too, including those linked by causative and non-causative aspect, such as *bring to a climax* and *come to/reach a climax* etc. Thus, the notion of the idiom's identity is related to some of its possible transformational tests proving a **functional difference**, and vice versa. Hence, this kind of **transformation**, which

is to be distinguished by a stable difference in textual function, may seem to be a most notable instance of what one may mean by stressing relevance of the idiom's function. There are many such types to be distinguished (which have been distinguished and treated in Čermák 1985, 1994a, 1994b and in Čermák – Hronek – Machač 1983, 1988 and 1994), some of them also structurally linking levels and types of idioms, which are different functionally, others linking idioms with traditional non-idioms (such as *backlash* ↔ *?lash back*, but see above).

There are at least two more types of function one should consider, a textual (communicative) function and a pragmatic one (Čermák 1985, 1988, 1994a, 1994b, 1994b). The **textual function**, referring to the textual behaviour of the idiom and its communicative situation, building heavily on the anomaly of the idiom, may be viewed as a set of possible standard formal textual functions an idiom might have in text (often expressed by specific grammar forms), such as a particular choice of person, number, mood or negation, etc. (for the verb type of idiom). This is what the generative grammar incidentally chose to call transformations, too, and the term might be borrowed from there, loosely. Yet, in fact, it seems that it is the converse that is much more convenient and revealing to record, namely to **list** and study the **anomalies of behaviour**, namely those functions that a particular idiom cannot have. Thus, it seems that the English collocational verb idiom *make oneself scarce* does not usually allow for a passive, and so on.

A key to a better and objective understanding and subsequent **classification** of idioms may (as argued in Čermák 1988, Filipec – Čermák 1985) be sought just here, in the variety of sets of anomalies of idioms, since functionally similar idioms tend to behave in a similar way in text, this set of these possibilities being different for each functional class. In the Czech approach, thousands of idioms have been analysed, classified and labelled in this way in the *Czech Dictionary of Idioms* (see above); an average number of functional indexes used is up to ten for the verb type idiom (whereas a similar approach by Cowie et al. 1975, 1983, usually has one or two only, though this number could go up to five). Of course, this depends on the functional type of idioms, theory adopted and type of the language being described in this way.

Thus, to take a Czech example, the verbal idiom *neříct v něčem ještě poslední slovo* ('have yet to say the last word', literally 'not to say in something still the last word') is characterized in the dictionary by this set of 'negative' transformational features, indicating in which of the given categories it cannot be prototypically used: in this particular case it is question, positive form, conditional mood, positive imperative, negative imperative, present tense, future tense, first person singular and first person plural. Those categories that are not mentioned, such as passive voice, are or may be used. By saying 'prototypically', the obvious fact is indicated of all sorts of individual types of usage and deviations from the norm. If, for some reason, this verbal type of idiom may also allow for nominal transformations, such as pluralization of the noun, the fact is also mentioned, though it is rare, as a rule; noun case forms cannot undergo any transformation at all here.

This seems a promising option to be pursued. It should be made clear, however, that there is, in fact, a prototypical approach applied behind such a labelling, allowing one to

concentrate on those cases of use that are amply attested and leaving aside minor and individual ones. Behaviour of the idiom and its anomalies is no absolute notion, of course. Difference must be sought both in the distribution of various transformational anomalies, which is due to differences in the **collocability** of idioms, and in the strength of their validity and presence (see also Fraser 1970). The latter point suggests that more insight is needed into the relationship of **typicality** and **potentiality**, which we do not know much about and which we take for granted, somehow, attributing its solution to frequency. Awareness of obvious gradations existing in and between both might throw some light on where to seek criteria for stability as well, enabling us thus to draw some kind of line between an idiomatic candidate and an idiom proper. There seems little doubt now that techniques of **fuzzy logic** will have to be employed.

The second functional aspect is to be sought in such **pragmatic aspects** as **evaluative function** of the idiom. Such idioms as *have one's heart in the right place*, contrasted with *have the heart to do something*, show a major difference in evaluation which is signalled when they are used, the former indicating, broadly, a positive attitude, praise etc., the latter indicating sensitivity and hence reluctance to act. The fact that idioms are primary means for expression of positive and negative **attitudes** and personal **evaluation** of whatever is made subject of one's talk, has been raised quite often. However, what one would rather like to have, instead of occasional individual remarks, is a set of instruments, criteria, telling us how to describe this function. And this is where we have drawn a blank, so far.

### 3. Lexicographic Treatment of Idioms

Lexicographic coverage of idioms and phrasemes, of varying quality and explicitness but focusing on one level mostly, is to be found both in specialized dictionaries and general monolingual dictionaries. To illustrate some of the points raised above, let us first look at three recent monolingual dictionaries of English, namely *The New Oxford Dictionary of English* (NODE), *The Random House Dictionary of the English Language* (second edition) (RH), and the *Collins Cobuild English Dictionary* (COBUILD), reflecting different approaches of their authors. (For a recent survey of phraseological dictionaries, see also Cowie 1998a and Moon 2000.) For purposes of brief illustration only, to use ***pull strings***, discussed above at (1), seems as good an example as almost any other. The information to be found in those works looks like this:

NODE: *make use of one's influence and contacts to gain an advantage unofficially or unfairly.*

RH: **a)** *to use one's influence or authority, usually in secret, in order to bring about a desired result.*

**b)** *to gain or attempt to gain one's objectives by means of influential friends, associates, etc.*

*He had his uncle pull strings to get him a promotion.*

COBUILD: *If you **pull strings**, you use your influence with other people in order to get something done, often unfairly.*

*Tony is sure he can pull a few strings to get you in.*

Obviously, all three approaches recognize implicitly that the idiom is lexicalized, hence its inclusion in the dictionaries. As it is not their business to do so, neither analyzability nor metaphor (points raised in 1.2) are touched upon. The problem of form (2.1) and, indirectly, of analyzability is represented here, though marginally, in two ways. RH mentions a variant here, namely *wires*, which is not given in the two British dictionaries. On the other hand, NODE uses the form *pull the strings*, one which most corpora attest, too. For instance, the 40 million word corpus of *The Times* (1995) attests 20 occurrences of the variant with the article *the* (in various morphological forms of the verb) and only 5 without the article. This information is not given in RH or COBUILD. It seems that the information about the article should find its way into the dictionary article in the form of a variant reading (*the*). To pin down the form more properly, more reliable corpus data should be used, which was the case with NODE, but also, surprisingly, with COBUILD. Note that the COBUILD illustration of use has a variant (*a few*), similar to the one discussed in 1.2. Yet corpora do not reveal any other frequent or standard type of variant (referred to above as modifiability) which would substantiate the generalized claim made by Erbach above.

In contrast to many others, the three definitions of the idiom's meaning (see also 2.2) seem to capture its essence quite well. Yet, the social role and relationship of the two actants, that of the instigator and subject and that of the real doer, should be made more clear. Both must have some influence but who has more, is there any debt to be paid, etc.? This is the social dimension between actants, often implicitly present in the idiomatic situation, which should be described. Also the fact that the goal of the action described by the idiom is to obtain something profitable for the instigator should be made more explicit: this is missing from the COBUILD definition entirely. Finally, no further subdivision of meaning is considered, except in RH (**a** and **b**). However, examples such as *There is little doubt who pulls the strings now* might suggest yet another meaning, in which the aspect of absolute power over a situation and general influence could be considered.

The aspect of function (cf. 2.3, also Čermák 2000) is not covered at all; this is the usual practice. Let us raise here at least what is evident in the case discussed. As idioms copy, structurally, word class functions and there are as many idiom function classes in this sense as there are word classes, on the level of collocations only, an explicit indication of the fact added to the description of each idiom would certainly help. This fact is traditionally recorded for words, so why not for their combinations as well? Functionally, *to pull (the) strings* is an obvious verb, or rather, verbal expression. It is hardly necessary to add that much of the confusion about the use of some idioms, especially grammatical idioms, could be avoided, were this type of information provided.

But there are other aspects of function as well. Those which are signalled through the attested morphological forms do, if covered lexicographically, signal an aspect of its usage. In view of the general acceptance now of the fact of the anomaly or deficiency of idioms (of many sorts), it is difficult to understand why this is not done; at least those anomalies seen in their form, or rather its absence, should be described. Thus, it seems that mention of the use of the idiom in question in its *-ing* form, which is not infrequent, would be helpful (6 occurrences in *The Times* Corpus). A similar case is that of passivization (which seems to be used) and use of the imperative (which has to be explored).



To complement and expand possibilities of the idiom treatment discussed above by referring to a Czech item in the large 4-volume dictionary (Čermák – Hronek – Machač 1994, the volume containing verb-type idioms only), let us briefly look at an entry which happens to have a near equivalent in English, namely *to pull someone's leg*. The general verbal type of function is uniform for all entries here and need not be indicated.

**tahat<sup>o</sup> někoho za nohu**

(kol; nepřízň) **0** ot, neg, pas, kond, imp, imp neg, fut, 1. sg a pl

(Čl. svým lákavým slibem, senzačním vyprávěním, jednáním, někdy škádlivě a dobromyslně, vůči důvěřivému druhému ap.:) klamat někoho n. ho falešně vzrušovat; balamutit někoho.

In English, this could read as:

**to pull<sup>o</sup> someone's leg**

(colloquial; unfavourable) **0** question, negative, passive, conditional, imperative, negative imperative, future, 1<sup>st</sup> person singular and plural

(*Man towards a gullible second man through his tempting promise, sensational story, behaviour, sometimes teasingly and good-naturedly, etc.:*) bluff, deceive someone or excite him unduly; bamboozle someone.

The first line is reserved for the entry form and its variants, if any, where also valency is given (through pronouns here). The superscript small circle <sup>o</sup> after the initial word indicates that this word may be paradigmatically changed in text and is also a reference to line 2. The rest of the entry form, which is unmarked, remains stable under all circumstances. The second line contains double information, registered in an abbreviated form. First, style and register are given in parentheses; in this case the idiom is considered to be of colloquial character and usage, whereas the label 'unfavourable' is a pragmatic marker pointing to the overall social acceptance of its use. Following the **0**, which is a marker for negative grammatical (transformational) information, there are all those categories listed, referring back to the <sup>o</sup> sign above, that may not be prototypically used (see also 2.3). These **negatively listed categories** amount to a set of individual anomalies of the idiom in question. It must be stressed that this set of negative transformations relates to Czech only, not to the English equivalent, given above. In this particular case, the idiom may not be passivized or used in the Conditional Mood, for instance. The third line gives in parentheses information about the context, situation of use, setting, type of the actants, their relationship, type of action, etc. Note that this is strictly separated from the description of the meaning proper; in a sense, it is a tool for an active use of the idiom. Only after this broad contextual information is given as a background, can the meaning be treated. It is evident that this is a much simpler matter to describe than the context.

However, the whole of the entry contains much more information than what has been briefly indicated so far. Also included is an illustration of typical use in context, followed by a large section listing selected synonyms, antonyms and more loosely related idioms, with a cross-reference to a separate semantic dictionary (or thesaurus) containing all of the dictionary's idioms grouped according to shared meaning. Some entries may have a specialized note on etymology, special usage, rare variants, special pragmatic features, etc.

Last but not least, equivalents of the idiom in English, German, French and Russian are given, opening up a broader field for comparative study (see also Čermák 1994d).

## 4. Methodology and Corpus Data

Basic and well-defined terms that one has to make use of in any systematic attempt to map the substance of the idiom and its field, should include: aspects of **combination** (collocation, syntagma, etc.), specifying the type of elements to be combined; the **paradigmatic and syntagmatic** framework on every level of analysis and aspect of approach (both formal and semantic); and **function** in the sense described above, including the interplay of **analogy and anomaly** within all of these. Most of these have been briefly mentioned above. In fact, it seems that **anomaly** is a ubiquitous feature of idioms in general. Accordingly, these notions have been made the basis of the Czech approach to theory and description of idioms. In this, the idiom is viewed as

*a unique and fixed combination of at least two elements for which it holds that at least some of these do not function, in the same way, in any other combination or combinations of the kind, or occur in a highly restricted number of them, or in a single one only*

A more detailed discussion of this approach and its consequences have been published elsewhere (see, e.g. Čermák 1982, 1988, 1994b). However, two points have to be briefly raised here: as this is an operational definition, it also allows for a set of tests of what idiom could be, a set of tests which have been successfully applied in the large multi-volume *Dictionary of Czech Idioms* (Čermák – Hronek – Machač 1983, 1988, 1994; Čermák 1994d). It should be pointed out, however, that the use of the term *elements* (forms) in this approach definitely implies more than one kind of form on more than one level. However, it does not imply, on the other hand, that real full lexemes (on the level of collocations) take part in a particular idiom; it is in fact only **forms** (certain **allosexes**) of lexemes, which cannot very often be linked to them, semantically or functionally, that enter the idiom in a given case.

To be able to do this, i.e. to identify idioms in a reasonably exhaustive and unambiguous way without overlapping or leaving large gaps unattended, one has to have sufficient data to examine and choose from. Today, there is no doubt about where to look for these data, as there is really no alternative to **corpora**. In fact very large corpora, such as the British National Corpus or the Czech National Corpus, are required, yet these might not be large enough, as both amount to 100 million words only. To illustrate how badly one is in need of data for some rare phenomena, note that the frequency of proverbs alone, as calculated for the Czech National Corpus (on an early sample of some 23 million words), is one proverb per some 80,000 words. It is evident that to arrive at a reasonable frequency, particularly for rare proverbs, one needs more data.

On the basis of such a corpus, among whose many positive aspects one should include the attributes of being authentic and vast, a really new type of study seems to be both possible and desirable. Thus, **major desiderata** for a **formal approach to the study of idioms** should include at least:

1. **Redefinition of terms used:** these should be made clear, unambiguous and testable; the sheer amount of data will require that. This seems to be a prerequisite for any further step.
2. **Large corpora** (perhaps of billions of words) **with known idioms tagged:** an enormous task to be accomplished only gradually. As yet, no reliable way has been devised of doing this automatically and exhaustively.
3. Automatic, both statistical and rule-based, **discovery procedures:** these should help to distinguish **regular collocations** (including multi-word terms) from **idioms**, primarily, which means that the problem of the **stability (fixity)** of linguistic entities must be solved first.
4. **Identity and identification of idioms made operational:** this must include a revival and redefinition of the status of anomaly and analogy, variability and function.
5. **Research into the function of idioms,** specifically of their textual behaviour, and incorporation of the results into their description.

If this could, eventually, be made the basis for such a revision of the theory and classification of the idioms, it might, at the same time, be viewed as a basis and prerequisite for any future large-scale semantic research. That could and should follow only after formal and functional aspects have been exhaustively covered and made clear.

## Shrnutí

### Podstata frazémů a idiomů: věčné problémy, nedostatek dat nebo teorie?

K roztržtosti názorů na podstatu frazeologie a idiomatiky přispívá mnoho faktorů, různost lokálních tradic aj., jakkoliv se založením Evropské frazeologické asociace *Europhras* v r. 1999 ([www.europhras.unizh.ch](http://www.europhras.unizh.ch)) se situace významně začíná stabilizovat. Základní problém je už ve vymezení samotné základní jednotky, frazému a idiomu.

Druhým hlavním zdrojem této situace je problematika metodologie a stav argumentace. Podívejme se na jeden z nich, ve kterém se tvrdí, že „lexikalizované metaforické výrazy (či analyzovatelné idiomy) jako *pull strings*... se od vlastních idiomů liší tím, že jejich význam lze modifikovat“ (Eve-raert). Je to ale pravda?

1 – Analyzovatelnost: jak víme, že *pull strings* se dá skutečně rozkládat? Na základě existující analogie by to možné bylo, ta tu však chybí.

2 – Modifikovatelnost: co to vlastně znamená? Lze přijmout modifikace typu  
*\*pull long (A) / short (A) / strong (A) / weak (A) STRINGS*, či  
*pull interesting (B) / hidden (A) / political (A) / influential (A) STRINGS* anebo  
*pull iron (C) / nylon (A) / fish (C) / 3 metre-long (B) STRINGS* etc?

Ukazuje se, že to možné není a případy *certain, kind of, proverbial, true, veritable* jsou krajně omezené a jen s jasnou metajazykovou funkcí. Opět tu chybí základová analogie.

3 – Metafora: na jejím pojetí se lingvisté shodnou jen těžko.

4 – Lexikalizace: proč by tento frazém měl být lexikalizován a další, který podle autora není analyzovatelný, lexikalizovaný není?

Frazém je třeba chápat komplexně, ze všech relevantních stránek. U formy k nim patří způsob stavení ustálenosti, určení vhodných variant v rámci všudypřítomné variability aj. Jinou otázkou je víceslovnost frazému, dělíci je na kolokační a propoziční, kterou je ale třeba modifikovat a rozšířit tak, aby zahrнула i kombinace morfémů uvnitř lexému, srov. české *zákeřný*.

Po stránce sémantické je třeba odpovědně posoudit pojmy jako analyzovatelnost, kompozičnost a motivaci. Ukazuje se, že je nelze přijmout. Pak bychom totiž museli např. u frazému *rain cats and dogs* jeho analýzu zastavit uprostřed, protože, v kontrastu ke komponentu a tedy i sémantice *rain*, si nebudeme vědět rady se zbytkem. Naopak je třeba posílit holistické, celostní chápání frazému i po stránce jeho sémantiky a bránit ji proti nezodpovědným pokusům ji nějak částečně krájet a poukazovat pak na analýzu a kompozičnost či motivaci.

Významná je konečně i funkce umožňující na základě propracovaného pojetí transformací propojovat případy jako *corner-cutting* or *cut corners* a další. Liší je právě jejich výrazně odlišná funkce, jakkoliv jsou k sobě vázány formálně i sémanticky. Zásadní je také poznání a popis anomálií frazému vycházejících z jeho textových restrikcí, tedy z chování v textu.

Mnohé z těchto problémů a aspektů se neodrážejí v lexikografickém popisu. Zmíněný *pull strings* nachází ve slovnících různé popisy, srov.

NODE: *make use of one's influence and contacts to gain an advantage unofficially or unfairly.*

RH: **a)** *to use one's influence or authority, usually in secret, in order to bring about a desired result.*

**b)** *to gain or attempt to gain one's objectives by means of influential friends, associates, etc.*  
*He had his uncle pull strings to get him a promotion.*

COBUILD: *If you **pull strings**, you use your influence with other people in order to get something done, often unfairly.*

*Tony is sure he can pull a few strings to get you in.*

které jsou všechny, každý však jinak, nedokonalé.

Jako protiváhu lze nabídnout ilustraci ze zpracování v českém slovníku

**tahat° někoho za nohu**

(kol; nepřízn) **0** ot, neg, pas, kond, imp, imp neg, fut, 1. sg a pl

(Čl. svým lákavým slibem, senzačním vyprávěním, jednáním, někdy škádlivě a dobromyslně, vůči důvěřivému druhému ap.:) *klamat někoho n. ho falešně vzrušovat; balamutit někoho.,*

kterému by mohlo odpovídat v angličtině zhruba toto:

**to pull° someone's leg**

(colloquial; unfavourable) **0** question, negative, passive, conditional, imperative, negative imperative, future, 1<sup>st</sup> person singular and plural

(*Man towards a gullible second man through his tempting promise, sensational story, behaviour, sometimes teasingly and good-naturedly, etc.:*) *bluff, deceive someone or excite him unduly; bamboozle someone.*

Pro uspokojivé řešení problémů je napřed třeba je redefinovat, opřít selekci dat o velké autentické korpusy a najít pro jejich další výzkum mj. i nová operační kritéria.

# Aktuální problémy frazeologie a idiomatiky a jejich odraz v slovníkové praxi (Nad bulharskou sbírkou přirovnání)

(In: *Slovo a slovesnost*, 1987, s. 320–327)

1. Je zřejmé, že stupeň dosaženého poznání v oblasti teoretické se odráží v aplikacích. Zvláště dobře to můžeme sledovat na vztahu lexikologie a lexikografie, ale také na poměru frazeologie (idiomatiky) a **frazeografie**, jak se této slovníkářské oblasti začíná říkat. A je také zřejmé, že tu teorie zůstává praxi leccos dlužna. Některých stránek této situace si povšimneme na pozadí bulharské sbírky přirovnání V. Kjuvlievové-Mišajkovové *Ustojčivite sravnenija v bălgarskija ezik* (Izd. na Bălgarskata akademija na naukite, Sofija 1986, 284 s.).

Autorka vychází z kartotéky Ústavu pro bulharský jazyk (přes 5 milionů dokladů různého druhu, včetně frazeologie), z níž čerpal také dvoudílný bulharský frazeologický slovník (Ničevová – Spasovová – Mihajlovičová – Čolakovová 1974–1975; rec. Machač 1977; srov. též Čermák, 1978). Kniha je rozdělena do přibližně dvou stejně rozsáhlých částí, teoretického výkladu (s. 5–100) a vlastního slovníku přirovnání (s. 100–227). Doplněna je indexem (s. 239–274), který eviduje přirovnání užitá ve slovníku, a soupisy literatury. Jde tedy o kombinaci slovníku a teoretické studie.

**1.1** Zásadní odklon od stanovisek, která zaujímá bulharská frazeologie na úrovni teoretické (srov. cit. frazeol. slovník a též Ničevová 1983), představuje v uvedené sbírce **výběr přirovnání**, který tu lze označit jako rozšířený. Dřívější úzká hlediska totiž z rámce frazeologie, a tudíž i z jeho slovníkového popisu frazémy typu *krokodýlí slzy*, *(být) bílý jako sníh* **vylučovala s poukazem na to, že jsou motivované** a že jeden z komponentů tu funguje ve svém doslovném významu (taková je praxe uvedeného bulharského frazeologického slovníku, Ničevová et al., 1974–1975). Ve své snaze, kterou autorky pracně, složitě a nepřesvědčivě dokumentují v dlouhém výčtu vyloučených a zavržených typů, se jejich stanoviska snažila opírat o kritérium **sémantické transformace** komponentů.

Původní podnět k takovému stanovisku, které však V. Kjuvlievová-Mišajkovová opouští, je dán v **dělení frazémů podle sémantického „sliť“ významu jednotlivých komponentů**, které je spjato se jménem V. V. Vinogradova. Tento přístup, v zásadě jako program zajímavý, nedokázal totiž nalézt, a to ani v pracích Vinogradových následovníků, dostatečně spolehlivé postupy, **jak objektivně odstupňovat a kategorizovat sémantiku** v oblasti **mnohonásobného splývání a redukce významů** u jednotlivých komponentů. Předpokládalo by to např. schopnost objektivně určovat **míru metaforizace** na straně jedné i **míru**

**desémantizace** na straně druhé. Protože každý takový postup je třeba v lingvistice opřít o spolehlivou analogii a protože právě ta ve frazeologii zásadně chybí a každý frazém je jinak anomální, byl uvedený postup prozatím lingvisticky neúspěšný. Sémantické klasifikace založené na předpokládaných stupních v uvedeném smyslu jsou jen ilustrativní a snad vhodné pro první, orientační vzhled do problematiky frazémů; rychle však narážejí na základní potíže dané splýváním komponentů a jejich neurčitostí, chceme-li přesně v širokém měřítku klasifikovat (srov. též Čermák 1982, zvl. s. 77–106).

V pozadí takových postupů, které porůznu přežívají, lze však vytušit ještě další, obecnější důvod epistemologické povahy. Hledání frazeologů, kteří spatřují podstatu frazému jednou v tom a podruhé v onom (srov. Čermák 1982, s. 18–19), pochopitelné v počátcích této disciplíny, vyústilo později – namísto očekávané syntézy a pozitivního konstatování – v **nihilistické prohlášení, že každý pokus o obecnou definici frazému je odsouzen k nezdaru** (Telija 1966, s. 187). **Mnohoaspektový přístup** se tudíž stal v té době nutností. Jeho zvláštností bylo, že **žádné z uváděných kritérií vymezujících frazém nebylo nijak specifické pro frazeologii jako celek**; bylo buď příliš široké a vágní, anebo jen dílčí a platné tudíž pouze pro část frazeologie. Akcentuje-li V. Kjuvlievová-Mišajkovová komplexní přístup, je to zároveň její polemická odpověď na jednostranné a dosud neúspěšné opírání se o kritérium jediné, prosazované u jejích předchůdců, zvláště o kritérium **obraznosti** nebo **expresivity**. (O možnosti obecné definice frazému založené na anomálii viz Čermák 1982, s. 115–116 a Filipce – Čermák 1985, s. 177.) Blíže však autorka v tomto smyslu mnoho neříká a formálně se přesto nakonec hlásí ke kritériu obraznosti a expresivity. Přirovnání, které považuje za typ frazému (idiomu), je pro ni jedním z logických prostředků poznání světa, který nás obklopuje. Vymezuje ho jako formální, sémantický a funkční celek, který je obrazně expresivní a jehož komponenty jsou sémanticky zčásti nebo zcela přehodnoceny. V tomto vymezení se tudíž vlastně připouští, že i typ (*být*) *bílý jako sníh* je třeba považovat za sémanticky přehodnocený (o tom viz už Čermák 1983, s. 480n.).

**1.2** Více implicitně a pod tlakem potřeb frazeografických vychází autorka bulharské práce o přirovnáních při zpracování z formy výrazu, a to – opět proti neuspokojivé předchozí praxi – z **formy relativně úplné**, včetně **tertia comparationis** na levé straně přirovnání (srov. však 2.6). Třebaže uznává přítomnost přirovnávací složky i u frazémů dalších typů, opírá se při výběru materiálu – podobně jako český slovník (*Slovník české frazeologie a idiomatiky. Přirovnání*, 1983) – jen o ustálené frazémy obsahující explicitní znak (komparátor) *kato* (*jako*), popř. jeho varianty. Staví tu opět do opozice proti dřívějším názorům (zvl. Ničevová 1983), v nichž se u přirovnání neústrojně vypouští právě *tertium comparationis*, a pak se chápé v nepřijatelně zjednodušeném pohledu jen jako adverbíale.

Zapojením pojmu **transformace**, který přebírá z české lingvistiky (Čermák 1982), dokáže se V. Kjuvlievová-Mišajkovová tomuto úskalí vyhnout a rozlišit celou řadu přirovnávacích typů po stránce formální a funkční – nikoli však důsledně. Uznává-li na jedné straně existenci transformačního vztahu mezi *bleden kato stena* a *poblenjavam kato stena*, není jasné, proč ho analogicky a důsledně nehledá i jinde, tedy i mezi případem typu *bledý jako stěna* a *být bledý jako stěna*. Sloveso *být*, třebaže bývá jakožto spona chápáno syntakticky odlišným způsobem, je frazeologicky (lexikologicky) řádným a nutným komponentem jednoho z obou frazémů. O tom svědčí i zásadně odlišná funkce: v prvním případě frazém fun-

guje jako atribut, v druhém jako predikát. V tomto je autorka stále ještě poplatná stanoviskům, proti nimž se jinak obecně staví.

Ve skutečnosti se tu opět skrývá další obecný frazeologický problém, který je dán **kla- mem diskrétní formy** a častou **neschopností lingvistů frazém přesně identifikovat**, především z hlediska funkčního. Postavíme-li se důsledně na stanovisko důsledně funkční, pak musíme jednotlivým transformačně spjatým frazémům přiznat vždy **odlišnou funkci**, a můžeme je tedy podle funkce identifikovat (na rozdíl od variant). Zároveň však musíme i bezpodmínečně **respektovat integritu**, celistvost frazému (idiomu) jakožto jednotky a **nepokoušet se** z ní některé **komponenty neústrojně vydělovat**, nechceme-li pracovat s nejasnými a těžko určitelnými torzy (srov. výše např. postoj Ničevové, srov. však i dále, zvl. 2.2). Tato obecně nedocenená otázka má ve frazeologii mnoho negativních důsledků, zvláště pro pojetí významu a motivace (viz 2.2 a 2.4), a je proto třeba předem pečlivě kontrolovat, zda např. zapojením běžné **syntaktické analýzy**, která je ve frazeologii principiálně **nemožná** (o tom viz 2.5), neztrácíme více, než získáváme. Charakteristickým rysem frazeologie a idiomatiky jakožto disciplíny je koneckonců to, že tu tradiční a osvědčené procedury, kritéria a metodologická měřítka selhávají, a to proto, že byly vypracovány pro **pravidelný jazyk** a jeho jevy (tj. v nichž platí pravidla), avšak pro frazeologii zásadně platí jen to, že je vždy nějak **anomální, nepravidelná** (zpravidla ve více ohledech syntagmatických a paradigmatických zároveň).

**1.3** Na základě parametrů z 20. a 19. stol. (a také mluveného jazyka) sestavila autorka bulharské sbírky přirovnání korpus asi o 1500 přirovnáních (český slovník jich má téměř 2500); tento materiál má reprezentovat současný stav bulharských přirovnání. Je zřejmé, že rozsah souboru (a šířeji i předmětu celé disciplíny) je dán i pojetím jednotek, které ho vytvářejí.

Vedle nedůsledného zapojení transformačního pohledu ukázala se jako další problém oblast tzv. **klišé**, která **jsou** z výběru plánovitě **vypuštěna**. Necháme-li stranou vlastní frekvenční podstatu klišé, která nebývá dostatečně docenována a která se odhadem určuje jen velmi nespolehlivě (o tom autorka mlčí), je tu rozpor i v samém pojetí věcném. Jako **klišé** autorka chápe (s. 49) taková přirovnání, která **nejsou** podle jejího pojetí **jednotkami jazyka** (jako ostatní přirovnání), ale mluvy, a **přesto jsou reprodukovatelná** (v určité sféře společenské komunikace). **Reprodukovatelnost** je však základní vlastností jednotek systémových, nikoli textových. Není jasné, jak tuto rozpornou představu, směřující jak langue a parole, tak synchronii a diachronii, dokázala autorka bez potíží aplikovat při výběru klišé. Ani další charakteristika věc neobjasňuje, spíše naopak. Klišé mají vznikat jako individuální tvůrčí akt (prostředek) k individuální charakterizaci jevů, příznaků apod. Má-li tu autorka na mysli diachronii, pak je třeba takový individuální akt, zprvu neustálený, předpokládat u všech frazémů (a nejen u nich), a není tedy pro klišé nijak specifický. A co se míní „individuální charakterizací“, není jasné, uvážíme-li, že běžnou komunikaci charakterizuje právě potřeba v různém smyslu individualizovat. Tak výrazy označené jako klišé, např. *privliča kato magnit*, *raboti kato automat*, se – přes uvedenou snahu odlišit je – nijak od přirovnání jako celku neliší.

Jde tu zřetelně o problém **vágních kritérií** obecně, specificky však o nedůslednosti v pří-  
stupu **synchronním**. De Saussure své pojetí jazykového stavu, na kterém synchronie spočí-

vá, sice nedovedl do tak exaktní formulace, jak by si leckdo dnes přál, ani později však nebylo nijak významně zpřesněno (zdá se to nemožné); a tak se užívá dosti volně a s přihlížením k synchronnímu kolísání (dynamismu), o kterém mluvil už V. Mathesius, když uvažoval o potenciálnosti jazykových jevů. Neznamená to však, že lze relativní pojetí hranic synchronie relativizovat ad libitum. Proto jsou výše zmíněná klíše buď reprodukovatelná, a tedy součástí synchronie a jejího systému, anebo jsou individuální, příležitostná, a tedy celonárodně neznámá, neustálená, a proto i nereprodukovatelná. Zda jsou součástí diachronie směrem do minulosti (archaismy apod.) nebo do budoucnosti (neologismy), je třeba studovat dále. Ústupek diachronii však představuje ve slovníku i další typ přirovnání, o kterém se mluví v klasifikaci podle historické perspektivy (s. 75–76), *chitār kato inglizin* (doslova „chytrý jako Angličan“), označovaný jako zastaralý; autorka ho ještě řadí do pasivního slovníku. I pro oblast idiomatiky a frazeologie tudíž platí rigorózní požadavek, aby se dodržoval synchronní přístup. To má své důsledky – nejen pro výběr materiálu – i pro další otázky, zvláště pro analýzu významu (srov. 2.4).

**2.1 Přirovnání má ve slovníkové části sledované bulharské práce tyto součásti:** a) lemma, b) valenci, c) zkratku podle příslušnosti k stylové rovině a emocionálně expresivní charakteristiky nebo nářečního původu, d) stručný výklad významu, e) exemplifikaci.

Zastoupení jednotlivých složek se může lišit, vždy přítomná je pouze složka první a čtvrtá, ostatní mohou chybět. U hojných výrazů nářečních se vždy udává označení nářečí, výklad a exemplifikace se však vypouštějí. V lemmatu, které začíná zpravidla slovem fungujícím jako *tertium comparationis*, se označují i **varianty**. Zde (stejně jako v řadě dalších prací) vadí **nedůslednost, s jakou se liší varianty syntagmatické od paradigmatických**, tj. varianty projevující se v zásadě buď ve změně rozsahu frazému, nebo v záměně jeho komponentů. Důsledky takových variant jsou jak pro chování, tak význam frazému odlišné; obojí se však označuje závorkou, srov. *stoja (sedja) kato dārvo, žlātvam (se)/žlātna (se) kato limon*, pro část paradigmatických variant se však užívá i lomítka, jak je vidět z posledního příkladu. Označení **valence**, a to pouze objektové, je velmi řídké, spíše výjimečné; někdy je naznačena jen ve výkladu (např. s. 176). Valence subjektová se přitom explicitně neoznačuje. Vzhledem k neexistenci infinitivu bylo třeba v bulharštině volit jinou **reprezentativní formu slovesa**; ve slovníku je to především 1. os. sg. prez. a v případě věcných subjektů též 3. os. sg. nebo pl. Avšak takto naznačená valence (subjektová) má daleko do spolehlivého popisu všech možných případů úzu. Je-li totiž možné jeden frazém užít v 1. os. sg., znamená to automaticky, že ho lze užít i v dalších osobách, zvl. ve 3. os., což se tu zvoleným záznamem nechtěně implikuje.

**Sémantická glosa** je spíše orientační a o vlastním významu pak specificky nevypovídá obvykle nic. Místy se zdá, jako by jejím cílem bylo jen různá přirovnání svést vždy do jedné synonymní skupiny, která má společnou sémantickou charakteristiku, srov. zřetelně odlišná hesla *mil mi e kato na decata prāčkata (tojağata)*, *mil mi e kato na kučeto kalta*, *mil mi e kato na ovcete vālka* aj.; všechna mají jedinou poznámku „není mi nijak milý“. „Jednoslovné“ lexikografii je přece dávno zřejmé, že u synonym je třeba vedle poukazu na rysy společné **vytknout i rysy rozdílné** a zabránit tak ztotožnění (které je pouze málokdy možné); není proto jasné, proč není tato zásada uplatňována i ve frazeologii. Nejvíce místa je v hesle věnováno **exemplifikaci** celonárodních přirovnání (až 90 % rozsahu hesla); obsa-



hují zpravidla více vět získaných z literatury (řídce dotvářela příklady i sama autorka, neuvádí však, ve kterých případech).

**2.2** Je zřejmé, že o **funkci** přirovnání, ať syntaktické či komunikativní, se uživatel z hesla mnoho nedoví jinak než nepřímo z exemplifikace. Uvědomíme-li si však, jak selektivní a individuálně omezená je funkce každého frazému, pak uváděné příklady mohou, jsou-li šťastně vybrány, ilustrovat nanejvýše jen a zčásti **úzus typický** (často však i netypický) – nikoliv tedy úplný. Cizinci pak už vůbec není jasné nejen to, které z příkladů má považovat za typické, ale i co už dané přirovnání funkčně nepřipouští, např. užití v některém modu, negaci, pasivu apod. **Otázka funkce je velmi důležitá** (a nejen pro cizince) a je jasné, že **bez přihlídnutí k ní nelze odpovědně popsat ani vlastní sémantiku**.

Silný důraz kladený na exemplifikaci na úkor ostatních částí přirovnání je ovšem sporný ještě jinak. Při selektivní pestrosti úzu frazémů nelze předpokládat, že se tento úzus úplně zachytí excerpací, byť seberozsáhlejší. Jde-li o **úzus typický**, pak je třeba excerpované příklady lexikograficky **standardizovat**, tj. zbavit autorských variant, sémantických posunů, zvláštností apod. Dále lze pochybovat o tom, zda literatura, tedy žánr převážně psaný (a v češtině navíc stále ještě převážně spisovný), dokáže zachytit to, co je v první řadě doménou úzu mluveného, jak ostatně připouští i autorka. Představa tzv. **dobrého autora**, u slovníkářů dříve běžná, zde je tedy silně narušena, především vzhledem k reprezentativnosti toho, co dokáže autor z jazyka zachytit (srov. zde i praxi českého frazeologického slovníku, Slovník..., 1983, s. 15–17; Čermák 1978, s. 50–51). Vedle prosté syntaktické funkce je však třeba uvažovat u každého frazému i o jeho **funkci komunikativní**, např. na pozadí mluvnických aktů a z hlediska zastoupení toho či onoho druhu příznačné **pragmatické složky** (srov. Čermák 1985). Tato otázka však není dosud uspokojivě vyřešena ani u běžných slovníků výkladových.

**2.3** Zřetelnou slabinou stavby přirovnávacího hesla recenzovaného bulharského slovníku je jeho složka **definiční**, pokud vůbec ještě lze – při vědomé rezignaci na bližší rozlišování významu – o definici mluvit. Výklad významu je tudíž, jak vyplývá z příkladů uvedených výše (2.1), shodný pro více výrazů a k **individualizaci hesel** nepřispívá. Je to přístup kupodivu rozšířený i jinde, **jako by cílem slovníku nebyla (relativní) individualizace jeho hesel, ale naopak jakási kolektivizace či synonymizace**. Už jen z hlediska rozsahu je výklad obvykle kratší než vlastní lemma. Při složitosti idiomatického významu to nutí k zamyšlení nad smyslem a cílem zpracovaného souboru: Jde jen o glosovaný soupis (heslář), nebo o slovník?

Hlavní problém však tkví podle našeho názoru jinde. Frazém je specifický útvar se specifickou funkcí, je však především specifickým **pojmenováním**, jehož úlohu v komunikaci je třeba vystihnout, nikoli obcházet. K jeho výrazné specifčnosti v tomto směru patří to, že jeho úzus a funkce mají nápadnou **pragmatickou složku**, kterou lze zpravidla dobře vystihnout jako **komunikativní reakci na podnět zcela určitého druhu** a za určitých podmínek. V praxi to znamená, že pro uspokojivé zachycení sémantiky je třeba vedle odstíněné **definice** (s vytčením typického jádra i možných hlavních potencialit, ve frazeologii velmi důležitých) nabídnout i popis, záznam typického **kontextu**, příslušného podnětu užití, okolností a dalších vztahů, pokud jsou zřejmé (srov. praxi českou, Slovník..., 1983, s. 18–19).

**2.4** Podstatnou část kapitoly o významu (v teoretickém oddílu) věnuje autorka staronové problematice **motivovanosti** významu útvaru prostřednictvím jeho komponentů, o níž již uvažovala v souvislosti s rozšířením přístupu k výběru materiálu (srov. 1.1). Vedle – v podstatě diachronní (a zjednodušené) – části o způsobech sémantického přehodnocování, a tedy idiomatice volného spojení (rozeznává jen podobnost, intenzifikaci a ironii), je to jediná sémantická klasifikace přirovnání jako útvaru, která se tu čtenáři nabízí (tematickou klasifikaci probírá autorka zvlášť). Opět se opakují staré a **nepřesvědčivé argumenty o třech stupních motivace** (plná průhlednost, částečná motivovanost a nemotivovanost), poplatné V. V. Vinogradovovi a dalším. Nejvíce pozornosti věnuje autorka tzv. plně průhledným přirovnáním, která se snaží oprávněně – proti svým předchůdkyním – rehabilitovat a zdůvodnit tak jejich zařazení do výběru. Zdá se však, jako by autorka sama cítila problematičnost takového dělení; nejspornější a nejobtížnější typ částečné motivace vcelku přechází a nesnaží se ho ani naplnit příklady.

V čem je podstata tohoto problému ve frazeologii? Jakožto obecný problém ho lze ukázat i na češtině, a to u kteréhokoliv strukturního typu frazému. Např. se běžně argumentuje tím, že je to zřetelný rozdíl mezi frazémy typu

a) *mluvit bez obalu, přijít s prosíkem, mít chlubné plíce* a typu

b) *vařit z vody, být v háji, mít namále*,

a že je tudíž můžeme takto klasifikovat. Frazémy typu a) obsahují údajně komponent s přímou, vlastní a individuální denotací i významem (*mluvit, přijít, chlubný*), kdežto frazémy typu b) nikoliv. Typ a) je tudíž určován jako částečně motivovaný, a proto sémanticky a de facto i formálně **dekomponovatelný**. Takovéto uvažování však není podle našeho názoru správné ani důsledné. Vydělíme-li uvedené komponenty, zůstanou nám zdánlivě více či méně samostatné útvary typu *bez obalu, s prosíkem a chlubné plíce*, což jako by uvedené dělení podpořovalo. Avšak už např. výraz *bez obalu* není v současné češtině v dané funkci a smyslu registrován v jiném ustáleném spojení (frazému) než v citovaném a ve výrazu *říct něco bez obalu* (jinak jde o homonymum s doslovným spojením *bez obalu*); samostatnou existenci v jazyce nemá. O statutu **monokolokabilního** slova *prosík*, které nelze považovat *senso strictu* ani za slovo, už bylo psáno dříve (Čermák 1982, s. 107–210; Filipec – Čermák 1985, s. 173n. a 182n.), jeho existence je výlučně podmíněna pevným sousedstvím předložky *s*.

Výraz *chlubné plíce* lze někdy chápat jako relativně samostatnou nominalizaci k verbálnímu frazému; je tedy třeba napřed analyzovat tento výraz, nikoli verbální frazém. Ani zde však komponent *chlubný* nelze bez problémů vydělit. Přiřadíme-li idiomu aspoň zhruba význam „schopnost nebo záliba se chlubit nebo být chlubný“ apod. a korelujeme-li ho se sémantikou adjektivního komponentu („chlubný, chlubit“), pak by mělo vycházet, že pro zbytek bude platit *plíce* = „schopnost nebo záliba“. K takovému chybnému závěru dojdeme, dovedeme-li snahu o jednostranné hledání motivace do konce. Na jedné straně (sice) zdánlivě dospíváme k motivaci, na druhé straně však zároveň vytváříme neústrojný **zbytek (torzo)** frazému, **který nelze z žádného hlediska určit ani obhájit. Dělení v neurčité a neurčitelné části není dělením**. Zároveň jsme však tím překročili definiční kánon samé podstaty idiomu, postulující vždy celistvou, integrlní **sémantiku**, kterou **nelze synchronně a pravidelně rozkládat**.

Obě uvedená adverbia jsou závažná. Statutu takového rozkladu z hlediska **synchronie** se ve frazeologii dosud pozornost nevěnovala; z nabízených analýz a úvah je často zřejmé, že

se jejich autor nepřipustně (nechtěně) uchyluje k diachronii, a to obvykle s poukazem na to, jaké sémantické poměry „musely“ ve frazému platit ve stadiu jeho idiomatizace, tedy dříve, v diachronii. A pokud by se mohlo zdát, že např. uvedené spojení *chlubné plíce* je řádným, pravidelným spojením, jednoduchý komutační test nás snadno přesvědčí, že to je kombinace synchronně anomální, a tedy i libovolně neopakovatelná (tím se ovšem nevylučuje případný vznik dalšího, nového frazému někdy v budoucnosti), srov. *\*chlubná hlava*, *\*chlubná ústa*, *\*chlubný mozek* apod. (druhý případ jako neustálený je možný).

Statut motivovanosti idiomu a frazému zůstává tudíž z metodologických důvodů prozatím nejasný. Jisté však je, že dosavadní pokusy, snažící se polovičatě a jednostranně určovat motivaci za použití tradičních přístupů, selhávají, a to opět na zásadním **nedostatku analogie** a v jiném smyslu i důslednosti. **Idiom nelze zároveň rozkládat a přitom trvat na tom, že má zůstat celý**. Pokud bychom přesto chtěli zde na pojmu **motivace** trvat, pak lze jistě souhlasit s **kvazimotivací** (stejně jako kvazitransformací apod.), ale tím se opět dostáváme zpátky k **anomálii**. Motivaci v běžném smyslu (tj. pravidelnou, opřenou o analogii a pravidla) tu však zatím nelze prokázat. A **techniky k přímému uchopení anomálie** nám dosud **chybějí**; lze dokonce pochybovat, zda jejich hledání má vůbec smysl.

Okrajovou pozornost věnuje autorka v sémantické kapitole i další často diskutované otázce – **modelovosti frazému**. Její přitakání modelovosti při tvoření frazému z volného spojení (vychází z pojetí Mokijenkova 1980, s. 40) je však poněkud rezervované (podotýká, že není automatická). Opět je zřejmé, že zdůrazníme-li v pojetí frazému jedinečnou konjunkci poměrů syntagmatických a paradigmatických, což je nutné, pak na úrovni individuálního frazému (idiomu) **je absurdní mluvit o nějaké modelovosti**. Avšak ani poukaz na zdánlivou modelovost (strukturních) typů, tříd frazémů není bez problémů. Hlavní námitkou proti přejetí modelovosti tu je nesamozřejmost, neprediktabilita, a tedy nedostatek běžné analogie v tvoření. Pokud bychom tedy modelovost chtěli hledat např. na úrovni struktury verbum – substantivum (v evropských jazycích nejvíce zatížené idiomy), měli bychom analogicky mluvit o ní i u tvoření slov (kde se raději mluví o *typu*, srov. Dokulil 1962, s. 69n.), u univerbizace, růstu slovní polysémie atd.; a to se běžně a ze stejných důvodů nedělá.

**2.5** Za klad bulharské sbírky lze považovat i to, že se – ať už z toho či onoho důvodu – nesnaží přirovnání analyzovat syntakticky. S ojedinělými zmínkami o syntaxi lze souhlasit, pokud se týkají syntaktické funkce přirovnání jako celku.

Otázka **syntaktických vztahů a syntaktického popisu frazémů** je natolik ožehavá, že se jí většina frazeologů prostě vyhýbá. Ani syntaktikové nejeví velkou chuť se do této problematiky pouštět; standardní **syntaktické popisy frazeologické příklady** zřetelně **obcházejí**. A to je jistě v éře, v níž syntax v lingvistice dominuje, resp. donedávna dominovala, zvláštní a svědčí to zřejmě o tom, že tu snad je něco až příliš problematického, nač teorie zatím nestačí. Je to možné chápat i jako uznání platnosti zásadního argumentu, že **neplatí běžná, a tedy i syntaktická pravidla pravidelného jazyka v jazyce nepravidelném**, ve frazeologii a idiomatice.

Samozřejmý, až banální předpoklad – často nebývá vůbec vyslovován – je **pevná** (znaková) **vazba mezi formou a významem** libovolného komponentu při jakémkoliv syntaktic-

ké operaci. Při jejím porušení máme co dělat buď s „čistou“ formální syntaxí bez významu, anebo se sémanticko-sémiotickou kombinatorikou bez formy, pokud je vůbec možná. **Přijmeme-li platnost teze, že sémantiku idiomu nelze dekomponovat**, protože patří na úroveň vstupních dat hloubkové struktury (srov. úvahy o motivaci, 2.4), **pak musíme od dekomponování upustit i při zamýšlené syntaktické analýze** poměrů uvnitř idiomu, čímž upouštíme od analýzy samé (povšimněme si přitom i toho, že komponent *chlubný* a význam „chlubit se“ si v příkladu uvedeném výše slovnědruhově neodpovídají). **Každý syntaktický vztah, ať už v jakémkoliv referenčním rámci, předpokládá vždy mnohostrannou a rekurentní vazbu** (často operacionalisticky pojatou) **ke vztahům dalším** (srov. subjekt – objekt při pasivizaci, formální vyjádření valence na objektu, schopnost objektu apod. přijímat atribut aj.). Nic z toho ve frazeologii neplatí pravidelně, a tedy není pravidelné ani prediktabilní (Čermák 1982, s. 42n. a 181n.).

Občas se ovšem objeví výjimky, které však lze obvykle chápat jako **nechtěný lapsus ze strany syntaktiků**. N. Chomsky v posledních letech idiomům pozornost sice nevěnuje, ale jeho žáci se jimi zabývají v pracích vznikajících za jeho přímého vedení. Jedním z nich je A. P. Marantz, který ve své práci o gramatických vztazích (1984) na Chomského (1982) přímo navazuje. V úvahách o tom, že idiomy vznikají generováním z logicko-syntaktické roviny do syntaktické povrchové struktury (s. 7, 23, 27), mluví o tom, že např. „objektové idiomy jsou kombinacemi verba plus objektu“ a mezi jeho komponenty vidí normální gramatický vztah (funkci). Když pak má vyložit, jaký je vztah mezi idiomy *throw a party* („udělat mejdan“), *throw a fit* („dostat záchvat“) či *kill a bottle* („ztrestat flašku“) a pravidelným spojením *throw a baseball* („hodit baseballový míč“) a *kill a cockroach* („zabít švába“), nedokáže nabídnout nic než mlhavou poznámku o „**určitém vlákně významu**“, které „váže“ tato užití *kill* v sobě“ (s. 25–26). Pokusíme-li se zpřesnit a zjistit, co se za oním „vlákem významu“ (**thread of meaning**) skrývá, zjišťujeme, že je to jen **falešná analogie a snaha hledat motivovanost** tam, kde synchronní motivace chybí. Jeho přístup je pouze variantou přístupu rozebíraného výše (2.4); **jednak idiom rozkládá, jednak se pokouší uchopit neuchopitelné v domnění, že tu nějaký význam je**.

Aby takový význam mohl existovat, musel by totiž být, jako všechno pravidelné v jazyce, **rekurentní**, což není, srov. *\*kill a mug*, *\*kill a glass*, *\*kill a decanter*, *\*kill a cask* atd. (tj. *\*zabít hrnek/sklenici/karafu/sud*), kde už se analogicky zopakovat nemůže, třebaže jde o analogické kontexty (na rozdíl od pravidelného významu, který by toho musel být schopen). **Domnělý význam (komponentu) tudíž neexistuje** (a proto od něj také nevede „významové vlákno“ k pravidelnému významu slova *kill*), existuje pouze (jediný) idiom jako celek, který ho obsahuje. Zdá se, že právě v takovém nedomyšleném kvazisyntaktickém přístupu mohl vzniknout i onen nepřijatelný pojem **významu** (tj. významu komponentu) „**frazologicky vázaného**“, běžný v některých příručkách. Marantz zde operuje jen s formou, ale **určovat v prázdné, významu zbavené formě objekt, predikát apod. je absurdní**. Konečně i jeho chápání procesu generování je v daném kontextu problematické a lze ho přijmout jen jako možné **diachronní** vysvětlení (o což lingvistům generativistům jistě nejde), v **synchronii** takovéto **generování žádnou rekurentnost nemá, a není tedy generováním**.

**2.6** K charakterizaci bulharské knihy o přirovnáních zbývá ještě se stručně zmínit o některých dalších věcech. Za **diskutabilní** lze považovat to, že autorka připouští, že **některá při-**

**rovnání existují bez své levé strany**, např. *kato ništo, kato sveta* aj. Zvláštnost těchto výrazů není totiž v tom, že levou stranu postrádají úplně, ale v tom, že je významově značně široká a zahrnuje celou (synonymní) třídu slov, popř. i více tříd – nikoliv však všechny (pak by šlo o pouhou formální valenci široké kategoriální povahy). Tuto situaci lze řešit tedy např. tak, že se vytkne – s příslušným upozorněním – **hyperonymum** celé třídy, srov. českou praxi (Slovník..., 1983, s. 226–227 apod.): *řít něco jako nic* (a implicate další verba dicendi), *udělat něco jako nic* (a např. verba faciendi *sníst, napsat, vzít* aj.), vedle pouze individuálního *tvářit se jako nic* (kde *tvářil se* nezastupuje žádnou třídu dalších verb) atd. Zabrání se tak opět, vedle dalších věcí, nemožnému či neustálenému, a tedy nefrazeologickému úzu (u českého příkladu např. *\*zezelenat jako nic* aj.).

Zvláštní problém v bulharském slovníku představuje mechanické **řazení hesel podle prvního komponentu**, které je pro nestabilitu a variabilitu levé strany přirovnání i slovosledné komplikace nespolehlivé. Autorka tento nedostatek zčásti vyvážila zvláštním rejstříkem, kde jsou hesla řazena podle pevného komparata; to by bylo v hlavním slovníku řazení jistě vhodnější.

Z posledních kapitol teoretické části knihy zaujme především informace, že na **devadesát procent bulharských přirovnání se vztahuje k člověku**, což dobře koresponduje se zjištěním českým. Připojené úvahy o klasifikaci podle významu komponentů či podle tématu jsou jen ilustrativní a materiál nevyčerpávají.

**3.** Posuzovaná práce o bulharských přirovnáních je především cennou a zřejmě reprezentativní sbírkou materiálovou (až na vypuštěná klišé), třeba je úrovní nabízené informace poněkud chudší a zaměřená tradičně. Uvedené nedostatky nebrání tomu, aby bulharské veřejnosti dobře posloužila. Sympatická je na ní snaha vyrovnat se odpovědně s jazykovými fakty, tj. frazémy (i mluvenými), a to v zásadě synchronním způsobem, navzdory tradici domácí. Teoretická východiska práce, rozebíraná zde poněkud širě, než je zvykem, měla posloužit zároveň jako základ k ilustraci některých tendencí a problémů v současné frazeologii. Zdá se, že nedořešenost či spornost některých otázek obecně se promítala především do teoretické části knihy, kterou jsme podrobili analýze.

## Summary

### Current Problems of Phraseology and Idiomatics and Their Reflection in Dictionaries

Against the background of a recent Bulgarian dictionary of similes *Ustojčivite sravnenija v bălgarskija ezik* (Izd. na Bălgarskata akademija na naukite, Sofija) by V. Kjuvlieva-Mišajkova, some basic problems of phraseology are discussed. Though the author basically abandons the perennial dogma of Russian linguists, starting with V. V. Vinogradov, that idioms can be classified on the basis of their meaning into several types, she still sticks to the semantic criterion of metaphoricity both in her selection of similes and their treatment. Thus, upholding this criterion, it is not clear what has been left out of consideration while selecting data for her dictionary. Another drawback to be found in the data selection is an a priori refusal to include clichés in the body of her dictionary, without any good reason. Moreover, as there is always a lot of subjective views to be found in clichés' delimitation, one would expect that this has

become the author's working limitation. It is a kind of paradox that she refuses to consider them for inclusion because they seem to her to be phenomena of text only (as in *privliča kato magnit*); because of their currency the very opposite seems to be true.

In the dictionary's lemma, some of its items call for comment. On the one hand, the idiom's transformations are usefully distinguished from its variants. On the other hand, syntagmatic variants need to be separated from the paradigmatic ones, too, and this is not the case here. The simile's meaning, deserving a specific treatment on a par with any good dictionary at least, is reduced, however, to a mere brief gloss, which is quite general and may be true of a whole class of other expressions. Similar, rather superficial, attention is paid to the simile's function which one has to decipher from the examples only. Moreover, should the dictionary represent standard use, the examples should reflect this, being standardised, too, i.e. used without interfering textual and authors' variation. There is no information given on the simile's pragmatics, either.

It is not clear to what extent the author has applied, in her dictionary, tripartite semantic division of similes according to degrees of motivation that she mentions in a separate study; this problematic, originally Russian, approach mentioned above, does not seem to be so easily abandoned. In defending this, one has to be consequent: while he/she seems to be able to decompose an idiom in one part only (as in Czech *mluvit bez obalu* where *mluvit* seems to be separable), he/she has to account for the rest of the idiom, too. That is, however, never done in such an approach. Hence, this seemingly partial analysis (or, to put differently, partial motivation) is no analysis at all as it always stops halfway. The same problem keeps coming back in similar (and futile) attempts, such as *chlubné plíce* and a vague speculation of a possible relation of *chlubný* and *chlubit se* does not yield a full semantic analysis. In a final analysis, synchrony and diachrony are mixed up here. One cannot insist on the idiom's integrity trying, at asme the time, to cut it up.

Another general objection may be raised against the notion of model formation of idioms, popular with some linguists. Not only does this notion denies the idiom's specific character, suggesting that it may be multiplied ad libitum, but there is no proof, yet again, given that it is a viable one.

If one accepts that the idiom may not be semantically decomposed and analysed, it follows that, likewise, no syntactic analysis of it is possible; that, too, is not based on any analogy. Following Chomsky, A. P. Marantz (1984, pp. 27, 7, 23) says that "object idioms are combinations of verb plus noun", seeing normal syntactic relations here. If confronted, however, with a necessity to account for the idiomatic *throw a party* and *throw a fit* on the one hand and regular expression such as *throw a baseball*, he has nothing more to offer by way of explanation than there is "a certain thread of meaning". This kind of a vague and undefinable remark (what does *certain* and *thread* exactly mean?), where contrary to his intention to stick to syntax he switches over to semantics instead, does not prove anything; there is only a false analogy to be found here. For such explanation to be able to hold, a solid analogy holding for a number of analogous cases must exist, but there is none to be found here. As there is no analogy to be detected in a number of recurrent identical processes here, there is no justification in seeing a generative process here either.

# On the Identification of Idioms

(Vyšlo jen španělsky jako Identificación de las expresiones idiomáticas v *Léxico y fraseología*, eds. Juan de Dios L. Durán, Antonio Pamies Bertrán, 1998, p. 1–18.)

## 1. Idioms and Idiomatics (Phraseology)

Many years devoted to research in many places have revealed a large number of often widely different views as to what **phraseology and idiomatics** are about. Both terms, sometimes interchangeable, sometimes different, as well as the fields they cover, have been defined in many ways; the same may also be said of their units, that is, the **idiom** and **phraseme**, and of the criteria that are assumed to constitute them. Some approaches concentrate, in part at least, on the form of idioms, others are more semantic, while functional aspects are generally overlooked, unless one elects to view some of the conclusions arrived at by transformationalists as functional (cf. Fraser 1970, Chafe 1968). Yet none of these approaches is devoid of serious shortcomings.

### 1.1 Criteria for Idioms

Sometimes it is possible to discern a specific meaning in the components of phrasal units of all kinds. Here it is fairly safe to adhere to one of the most favoured semantic approaches to idioms, based on the fact that their meaning is not amenable to analysis, decomposition; the meaning of their component parts is not additive (that is, the sum of their meanings is not amenable to analysis because these meanings do not add up to the semantic total of the idioms as a whole).<sup>1</sup> However, any attempt to explain in this way such components as are semantically depleted and devoid of any distinct meaning runs into trouble, as in the case of

- |                                  |                 |
|----------------------------------|-----------------|
| (1) <i>beyond my k e n</i>       | : <i>ken</i>    |
| <i>at somebody's b e h e s t</i> | : <i>behest</i> |

or that of such prepositional idioms as

- (2) *as to* (sb/sth)

The semantic criterion is no longer sufficient here. The two cases are rather different: one points to a complete lack of meaning (as in *ken*) or, on the other hand, to a full presence of a meaning (as in *as to*). Both cases limit, or nullify, the validity of the semantic approach, based on semantic composition. It is clearly inconceivable to posit any kind of semantic

additiveness, compositionality here, whether regular or irregular, of the component meanings when one or other component is devoid of any meaning, as in (1). Attempts at discarding cases containing such unique constituents as not idiomatic (among others Weinreich 1969, p. 44) do not solve the problem of the identity of the idiom, nor do they contribute to the desirable generality of a definition of 'idiom' that ought to cover all types. Indeed, this kind of inconsistency is usually why the search for an overall, integrated view of the field and discipline and the general substance of the idiom is abandoned. On the other hand, examples such as (2) represent the opposite case: there may be little doubt that their meaning is additive and also that no semantic transformation is involved.

Perhaps the second most frequent view of idioms, under whatever guise, is that which is related to their transformational-syntactic deficiency, anomaly or restriction<sup>2</sup>. Although this approach undoubtedly seems to go to the heart of the matter and leaves hardly any type of idiom unaccounted for, it is not satisfactory in a primary sense, required of any criterion: it is not specific enough in that it embraces not only idioms but also many other perfectly "normal" words which display a cluster of transformational deficiencies; consider, for example, the impossibility of passive formation for such verbs as modals (*can*, *must*). Whether this criterion can ever be tightened up so as to apply exclusively to idioms is a matter for a further research.

It is hardly possible here to go into other criteria proposed in various schools or by various scholars. On closer inspection they all produce unsatisfactory results, similar to those above, being either too specific (and partial), or not specific enough. The situation is somewhat paradoxical: linguists may recognize idiomatics and phraseology as a complex and specialized field of study, as "a basic theoretical stumbling block" (Weinreich 1969, p. 23), but they have only ever able to produce a handful of limited and often loose criteria applicable only to selected cases or parts of the field. In order to arrive at the essence of idioms and delimit their field, **idiomatics**, as a specialized discipline and branch of linguistics dealing with a kind of phenomena not to be found elsewhere in language, an area-specific and comprehensive theory is called for. Some of the protracted argument as to the proper place of idiomatics within the scope of linguistics, as well as some insecurity as to the adequacy and reliability of its current methods and insights, are direct consequences of the present state of affairs.

The kind of problems likely to be encountered here requires a complex and problem-sensitive treatment, starting from the point where all the relevant relations and criteria, derived from the phenomena described, interact. To arrive at this, clear answers to at least three principal questions are of primary methodological importance:

- 1: Is it possible to define the discipline of idiomatics by a configuration of familiar criteria, valid elsewhere in language? Or is it necessary and/or possible to look for some other specific criteria ? (*autonomy of criteria*).
- 2: Is the set of criteria to be applied specific to this field only? In other words, might it not cover other areas of language as well? (*specificity and non-substitutability of criteria*).
- 3: Do the proposed criteria cover the field as a whole and not some segments of it only, however large those segments might be? (*adequacy of criteria*).



## 1.2 Syntagmatic and Paradigmatic Relations

In the following remarks, I propose to examine the problem of **general criteria** for the idiom in the light and framework of what has been around for a long time. Specifically, I hope to show that, despite a profusion of later concepts and theories, it is, somewhat paradoxically, Ferdinand de Saussure (1916) who is still able, after eighty years or so, to provide us with crucial insights. I shall also draw on a specific line in the development of his ideas by certain other linguists, namely de Groot (1949), Firth (1957, p. 194–214), Halliday (1976, p. 73–83), Gak (1972, p. 373–395) and others in the field of syntagmatic relations, and, above all, Hjelmslev (1972 and 1970, p. 110), in the field of paradigmatic relations.

For the delimitation and description of idioms and their substance, a number of criteria have been suggested. Of these I will list only those that have gained some currency, in some schools at least, and claim to cover all types of idioms. Other criteria, though familiar, are not considered here for a number of reasons: either they do not cover the field as a whole, or they are too vague, or have turned out to be erroneous (such as the dictum that idioms are peculiar to a people, they are not translatable, have a metaphorical or an expressive meaning only, or they are simple equivalents of words). The nine features which follow may not be entirely representative and, in their present form, they are certainly somewhat artificial in that they are interconnected and partially overlapping. Thus an idiom is claimed:

- A** to be a combination of at least two words, that is, it has a discrete, analytical form (a generally accepted view; its generality, though, is limited by those advocating the existence of idioms on other levels, e.g. within word boundaries, compare, for example, Coates 1964),
- B** to be a unique concatenation of its components (Chernysheva 1970),
- C** to display a syntactic frozenness (Arkhangelskii 1964, p. 77, Fraser 1970 etc),
- D** to be formed by such components the meanings of which do not add up to its integral meaning (a generally held view; though, by implication, not accepted by, for example, Shmelev (1973, p. 259–273), who does not concede the existence of unmotivated components),
- E** to require the presence of a bound or “re-evaluated (transformed)” meaning in a component (Kunin 1972, s. 8 etc),
- F** to be stable (in form and meaning), which, in turn, allows it to be reproduced at liberty (a generally accepted view, see, for example, Rojzenzon 1973, p. 101–119),
- G** to have a nominative integrity which cannot be decomposed (Axmanova 1957, p. 169 and many others),
- H** to have an integral meaning (e.g. Machač 1967, p. 137–149, Mokijenko 1980, 4 etc.),
- I** to contain a transformational deficiency, an anomaly (Chafe 1968, Fraser 1970 etc.).

On closer inspection, it is not difficult to translate most of these and similar formulations, some of which have already been mentioned above (in 1.2), into a broader framework of (a) **syntagmatic**, or combinatorial features (views listed under A–E), and (b) **paradigmatic**, or class features (views under F–I, at least in part). It has to be stressed, though, that it is only their special **interaction** (of many types) which enables us to determine what is an idiom. At the same time, a characteristic expression of this state of affairs is to be found in some **anomaly** in the behaviour and structure of idioms. In both its behaviour and structure, this anomaly (irregularity) is manifested along many different lines.

By arguing that an idiom can neither be fully generated, nor is it amenable to certain transformations, some generative linguists have said that, by the same token, from the given bases (= components) and by the given rules it is not possible to derive an idiom's surface form and meaning. Both these aspects, that is, the bases and the rules, are important. Either these componential bases themselves are not regular (this appears to be the less frequent case), or it is the generative and transformational rules that do not hold here. However, most of analyses stop at this point and linguists seem to be content with the idiom viewed as being somehow exceptional, anomalous, or irregular. In consequence, a special idiom vocabulary is posited, viewed as an inventory of ready-made, not decomposable expressions (e.g. Weinreich 1969, p. 73) which cannot be generated.<sup>3</sup>

The discipline of idiomatics is quite young and as such it has had to go through a difficult and often contradictory set of stages, marked by a plurality of often synonymous labels used for the same concepts, each attributable to particular competing theories and the reference frameworks used within them. In an attempt to translate some of the terms used into broader ones, it may be useful to point out that not even the generative conclusions mentioned above were, in fact, new in their substance, and that they can be correlated with the much earlier insights of de Saussure. He, too, had spoken about *locutions toutes faites* as belonging to the language system (de Saussure 1816, p. 148), and thus illustrated the idea that they are not formed, generated by any combination for and in *parole*.

This is, without a doubt, important. Yet, against the intellectual background of those of his contemporaries and predecessors who also paid attention to idioms, and some long before him (e.g. Paul 1880), it does not seem to be particularly new. Perhaps also because of this, any attention paid to de Saussure in this respect unfortunately stopped at that isolated remark and was subsequently transferred to his former student and younger colleague, Charles Bally (Bally 1905). Thus this has become another instance of forgotten priority and of interpreting things outside their original framework, a peculiar misfortune, also in many other respects, of de Saussure's theories ever since his time.

The cornerstone of de Saussure's view of language is the concept of **relation** (*rapport*) and **opposition**, which he felt existed everywhere, and it is out of their types and configurations that everything else grows. For de Saussure, the language **system** (*langue*) is a totality of facts, interconnected by a network of relations, syntagmatic and paradigmatic (the latter were originally called *associative* in his work), including what today would be called generative rules. Hence any configuration of signs, especially a textual one (i.e. a structure, form etc.), was, for him, based in principle on such relations. This is also why he once pointed out that in **speech** (*langage*) there is no single thing that would exist in one term only, in isolation (Engler 1967, p. 264),<sup>4</sup> and called this a "final law", though this formulation never entered traditional editions of his *Cours*.

Because de Saussure had to see idioms as complex formations where such relations were neutralized and frozen, he inevitably had, in consequence, to view idioms as integral language facts and classify them as part of the language system.

Since the appearance of the critical edition (Engler 1967 *et seq.*), containing the actual manuscript notes on which the canonical form of his *Cours* is based, it has been noticed that most of the examples of idioms in the book were supplied by the editors, Charles Bally and Alfred Sechehaye, and it has been suggested that de Saussure never paid any attention to idioms at all. But this conclusion, which seems to follow from the Engler edition, is far from correct. It is true that the editors did add most of these examples to the canonical edition of 1916 themselves, but what they in fact did – no doubt with justification – was a mere extrapolation of the de Saussure's general remarks, well within the spirit of his theories and intentions. That he did have the issue on his mind and had given it some thought is revealed by the following recorded remark (accorded to E. Constantin, see Engler, *ibid.*, p. 284): "In any case, there are also syntagms to be found among the facts belonging to language (= *langue*)." However, his sole illustration of the point was an example from morphology (derivation) example. Yet de Saussure did say, at the same time, something of much greater importance, which his editors left out completely, altering the passage concerned: "Among other things there are numerous sentences which are ready-made for the language" (Engler *ibid.*)<sup>5</sup>. It makes one wonder whether this statement, had it been sufficiently familiar to linguists before, might not have brought to an end the long discussion for and against the inclusion of propositional, that is, sentential, idioms into idiomatics.

This short excursion into history is meant as a background to much that follows below and as a counterpoint to the current view of the discipline as being, in its principles, associated with Charles Bally only. If de Saussure had not elaborated his thoughts further, while Bally did, this does not mean that Bally was not influenced by de Saussure. But the main point of the argument lies elsewhere, in the adequacy and suitability of the *general* theory of language, as proposed by de Saussure, to the *particular* field of idioms.

I implied above that the existence of **idioms** is to be viewed as **a special, anomalous case of interdependence and a function of both syntagmatic and paradigmatic relations**. It is necessary, when studying idioms, to take into account both their internal and external syntagmatic and paradigmatic relations as manifested, above all, in the structure of idioms, their function and class membership respectively. However, I shall, in the following, concentrate almost entirely on their internal relationships, that is, on the relations between the components and the whole. Moreover, I shall limit the scope of these remarks to the traditional sphere of collocational, that is, non-sentential idioms only (for other types, see also Čermák 1982, Filipec – Čermák 1986).

## 2. Idioms: A Structural Approach

### 2.1 Paradigmatic Aspects of Idioms

Let us first turn to the paradigmatic aspects of the components *kith*, *spick* and *umbrage* in the following English idioms:

- |  |                  |
|--|------------------|
| (3) <i>k i t h</i> and <i>kin</i> "friends and relations"        | : <i>kith</i>    |
| <i>s p i c k</i> and <i>span</i> "clean and tidy"                | : <i>spick</i>   |
| (4) <i>give u m b r a g e</i> , <i>take u m b r a g e</i> at sth | : <i>umbrage</i> |

In these examples, anyone sufficiently familiar with the English language becomes readily aware of a striking distributional restriction of the above-mentioned components. In the first case (3), the form *kith* and *spick* are attested only for their particular idiom and are thus effectively their sole mode of existence in the language. In the second (4), the form *umbrage* occurs in two instances only, in *give umbrage* “offend someone by suspecting or disregarding him” and in its transformational correlative *take umbrage at* “feel offended and disregarded by”. There is no other way in which these forms can be used, or vice versa, how they can be understood, beyond the particular expressions in which they are found, which may in turn be viewed, in a sense, as their specific contexts.

## 2.2 Collocational Paradigms: the Syntagmatic Aspect

Even in a very superficial approach, the examples above (2.1) evince some of their syntagmatic aspects. Let us note some others.

In the light of de Saussure’s dictum that there is nothing that consists in a single term in speech, the nature of the lexical classes (paradigms) underlying these forms stands out as rather peculiar and unusual. There seem to be at least two types of classes involved here. The first has already been hinted at above: it is the class, or rather classes, with which the English forms *kith*, *spick* and *umbrage* specifically and usually collocate. This is why I propose to call such a class a **collocational paradigm**. However, several points in our examples deserve further comment.

- (A) For any real use of a word, its **collocational paradigm**, related to the word though not including it, is essential. Moreover, I shall assume that in fact no word can exist without its collocational paradigm. Such a paradigm is specific to every word in that it represents its specific and usual, standard textual distribution, that is, the ways in which the given word co-occurs with other words. This means that for each word in a language a different set or sets of words forming its specific collocational paradigm (paradigms) has to be postulated.
- (B) The scope of the paradigms illustrated above is clearly small and the forms defined by them are thus severely restricted in usage; in the extreme case (here, in (3)) it is reduced to a one-member class only. Accordingly, such a **paradigm** is a **closed** one, to use the term employed by Hjelmslev (1970, p. 110).
- (C) There is no doubting that its collocational paradigm determines a word’s meaning. Let us recall here one of the axioms of modern semantics, corroborated by logicians, that the **meaning of a word is to be sought in its uses and contexts**, its logical extension (Wittgenstein 1953, p. 139, Lyons 1977, p. 158 etc).

In (A) above, I digressed from the earlier examples in considering the state of affairs of any word of the vocabulary. Actually to prove that statement, namely that the collocational paradigm or paradigms of any word, for example *dog* or *cat*, is specifically different, is practically impossible. Instead, I shall refer to similar views held by other linguists (for example, Halliday 1976), though some indirect corroboration can also be found in the distributional criteria applied in the study of synonyms; these are based on the assumption that there are, other things apart, no two words with an identical distribution (Lyons 1968, p. 448).

The examples given above seem to be special and unique through, above all, the feature described in (B). In this case, the collocational paradigms are extremely small and usually so limited that they might seem to contradict de Saussure's dictum quoted above. Does this mean that de Saussure was wrong? I do not believe so. While these cases may border on conflicting with his view, they do remain exceptional. **For the bulk of the vocabulary**, on the contrary, **one has to reckon with paradigms that are very large**, not readily enumerable and without any distinct boundaries, and not providing us with any notion as to the actual number of their members. Thus the collocational paradigm, relative to the word *dog*, can only be indicated, though not exhaustively listed, by the following examples:

- (a) {*bark, growl, whine,...*}, but not by, e.g., *mutter, whistle, mew...*
- (b) {*lie, run, stand, jump, dig,...*}, but not by, e.g., *wave, rock...*
- (c) {*sniff, track, nose about,...*} etc.

It goes without saying that to obtain a full array of the combinatorial capacity of the lexeme *dog*, these verb-based collocational paradigms should be further complemented by relevant adjectival lexemes as well.

These and similar examples, relative to the word *dog*, are characteristic of it in two different ways, individually and generally. **Individually** it holds that some of these verbs collocate with *dog* only, and not with, for example, *cat*. But it is the **general** aspect that is perhaps more important here. By virtue of the specific and unique composition and structure of these paradigms, all words (including *dog*) appear to be fully individualized and sufficiently characterized for the purposes of any subsequent classification within the lexical system.

It is evident that the independent existence of such forms as Eng. *kith, umbrage*, etc. is highly doubtful and that they are formally fitted into the same mould as any other "normal" word merely by analogy. The reason for this analogical view, even though there is no specific analogy to be used here, is in the fact that they do not exist distributionally by themselves. It is not commonly or readily admitted that a form viewed superficially as independent, may not exist as such.

Let me suggest here that it exists only by virtue of its single and highly specialised, restricted collocational paradigm, **which happens to be identical with the rest of the idiom**.

## 2.3 Virtual Paradigms and the System

In contrast to collocational paradigms, outlined above, there is another, but fundamentally different, type of paradigm in language that one knows quite well intuitively. It suits our habit and need for a taxonomic classification. I shall call it here **virtual paradigm**. It is a substitution class of which a word, by virtue of its function and other properties (including meaning and form), is a member; such membership is based both on the typical, often intuitive and also logical relationships, that belong to a group. Thus the noun *dog* belongs to the class of domestic animals (mammals, along with *cow, cat, pig, horse...*), the adjective

*wooden* belongs to those denoting materials ((of) *stone, metallic, plastic...*), the verb *run* among verbs of motion etc. The magnitude and properties of this kind of paradigm depend, clearly, on the hierarchy of the vocabulary and on its relative place in it. Yet the forms given above – *kith, umbrage*, etc. – seem to be again quite special here. If one sought to determine their virtual paradigms, one would find none that would readily invite them for membership. Thus, again, this is a case of isolated, one-member classes; more generally, this is another aspect of the anomalous character of idioms.

## 2.4 Idiomatic Syntagms

Reconsidering the idioms above (3–4), let it be recalled that combination of such forms as *kith, umbrage*, etc. with a member of their respective collocational paradigms is the *sine qua non* of their existence, notwithstanding the highly restrictive scope for such combination. This feature is shared by all the types of syntagms mentioned so far. It will also be recalled that the *kith* and *spick* of their respective syntagms belong to a highly restricted collocational paradigm, consisting of a single word only, namely *kin* and *span*, respectively. From the point of view of their counterparts, *kith* and *spick*, these syntagms (idioms in this case) are unique, and in fact the sole possible forms of existence for these components.

Having established this point, let us consider the state of affairs of the bulk of idioms, viewed in this light.

Of the following English examples:

- (5) *hit the nail on the head*
- (6) *set the Thames on fire*
- (7) *be born with a silver spoon in one's mouth*

it can safely be claimed, without going into detail, that *some* of their components (but *all* the components in certain other idioms) seem not to preserve their semantic identity as familiar from the dictionary. Thus here *nail, Thames* etc. do not mean what one usually takes them for in their normal usage. Neither of these forms has preserved any of its familiar meanings.

There seems to be no positive conclusion to be drawn here, for it is not, in principal, possible, in this view, to arrive at any positive delimitation of the nature of such components, if viewed as language entities, facts. This point is rather crucial, given that it is impossible to take both approaches – that of the components and that of the whole – simultaneously. It follows, the idiom being (semantically) an indivisible whole, that one cannot divide or segment it and then continue asking, as frequently occurs, questions of the type: what is the significance of this or that constituent? This is a very simple and consequential view and in order to appreciate and apply it to the fullest extent its consequences have to be considered. One major consequence is the rejection, as quite unsubstantiated, of the entire swathe of lines of thought that have evolved here, notably the one associated with the search for phraseologically bound lexical meaning (see criterion E, 1.3), so popular with some Soviet and other linguists (Mel'chuk 1960, Weinreich 1969, p. 44) in fairly recent times.

What one can and must do is to resort to a negative statement that is so basic and important that the very conception of idioms can be based upon it. Before going on, however, we must deal with certain outstanding matters regarding the earlier examples.

Synchronically, constituents such as *kith*, to take but one example, have no discernible meaning (sense) at all; they are semantically void. Since their independence is only contingent, and more alluring in writing than in speech, and since they have no existence beyond the idiom of which they are a part, it is impossible to tear them out the idiom. Semiotically, if they represent any sort of sign at all (which is questionable), they are a special kind of zero sign. This semantic nature of theirs can be readily verified by native speakers. It is this fact that, among other things, leads one to reject or, at least, restrict the allegedly general validity of the traditional definition of idiom as something not derivable semantically from the meanings of its constituent parts (components). If *kith*, etc., in isolation (a condition to which they cannot aspire anyway) have no meaning, one cannot base one's conception of idiom on a semantics that views it through the prism of its constituents. This means that one cannot seek to find or generate any viable semantic relation between the semantic output and input, so to speak, even though the form may suggest that this may be possible.

To appreciate this point, one must be ready to recognize that any new meaning of a word can be discerned and determined only by virtue of its being invariably present in a sufficiently large number of collocations entered into by the word – certainly not in one or two only, or, for that matter, even a couple more than that. In other words, there must be a **sufficient analogy** upon which to base any semantic generalization. No meaning or sense is given to a word *a priori*; it can be discerned by language users only by virtue of analogy, based on the fact that a given meaning is encountered under like functional conditions with sufficient frequency. Any new meaning of a word obviously owes its existence to just such collocations, which evolve (as a rule) quite spontaneously through metaphorical, metonymical or other types of transfer of existing meanings.

Though different in other aspects, cases like *take umbrage* and *set the Thames on fire* all share this common feature: there is no sufficiently large background of analogous collocations for them to support the claim that *umbrage*, *Thames*, *fire* as occurring in them have acquired some specific meaning which they might have in isolation or in a dictionary, that is, a meaning which would fit their specific idioms only. In fact, if one is ready to recognize criteria H and G above, one must, in a different manner, arrive at the same conclusion: that there is only an **integral idiomatic meaning** of the idiom that is not to be reached via the standard meanings of its components, and the very notion of a special meaning or sub-sense present in its components (Weinreich *ibid.*) has to be abandoned. If this were true and if one pursued the argument to its logical end, it would have to be possible to make the integral meaning of the idiom decomposable into its distinct parts. One cannot have both, and it is clearly the integral meaning of the idiom that is incontestable. To conclude, it is **not before** the level of the constituents combination is reached, not before a syntagm is formed, that an idiomatic meaning arises, thus its exhaustive decomposition into distinct parts is impossible<sup>6</sup>. Any attempt so to do seems to border on metaphysical endeavour.

## 2.5 Restricted Collocability of the Unique Constituents

When considering the collocability of such by now familiar idiom constituents as the forms *kith*, *umbrage*, etc., one inevitably arrives quite early on at their radical combinatorial restriction to a single or very few other elements (its combinatorial partners, so to speak) as well as the impossibility of entering into any other analogous combinations. Because of its special status, this kind of collocability may, therefore, be called **restricted collocability** (though the collocability of any word is restricted in a broader sense). Within the framework of this basically combinatorial approach one cannot but arrive at conclusions that may seem rather surprising: for the forms *kith*, *umbrage*, etc., it is impossible to define clearly any sort of specific meaning or denotate, nor any clear category membership (here, within the system of word classes). It seems advisable, therefore, to consider these forms as mere word analogues or **verboids**, not as full-fledged words. The combinatorial property of the language can, then, become a basis for a hypothesis under which, historically, all language categories arose in such a framework; and it is this framework that has, in a roundabout way, been evoked and reconsidered here. That also explains why the elements *kith*, *umbrage* and the like have been referred to as to forms, and not words. In fact, this usage seems to be gaining currency in modern theories of phraseology (Molotkov 1977). Hence, idioms and phrasemes are not combinations of words etc., but of their forms, and these have very little to do with the real words (or textual word-forms) they might resemble; they are merely homonymous with them.

## 2.6 A Summary and a Definition

To summarise briefly: There is the **virtual paradigm** which may be understood as a class of elements (items, words or other units) with a broad equivalence related to their common function and meaning above all, a class that is given in the language system *in abstracto* and is definable by its place in the hierarchy of the system. The broad functional equivalence of its members is manifested in their ability, at a given place of the same context, to replace one another. Then there is the **collocational paradigm** which is understood as the class of elements with which the given external element usually and meaningfully collocates; however large, this class represents the use and distribution of the element.

In both kinds one can further distinguish their quality (nature) as given by their position in the system, their structure, size and function. Moreover, it is also useful, especially in the case of the collocational paradigm, to distinguish whether it is a closed or open paradigm, that is, whether it is capable of receiving and assimilating other, new members, or not. This last quality, if projected onto its constituent members, could be called their high or low (limited) **paradigmaticity**, respectively. One has to assume, then, that every word of the lexicon is a member of at least one virtual paradigm and is, in turn, defined by at least one collocational paradigm. And, in consequence, it may also be assumed that the evolution of a language, specifically of its lexicon is, in a sense, immediately manifested by changes in the extent and content primarily of its collocational paradigms, as these instantly reflect changes in usage.

This, then, leads us to a revision of the traditional view and definition of idioms, based on the **function** of their component parts related to both types of paradigms, rather than on their



meaning. A useful concept to be recalled here is de Saussure's *value*, of which the present approach might be reminiscent. With this modification, one can, then, view **the idiom** as

*a unique and fixed combination of at least two elements of which it holds that at least one or more of these do not function, in the same way, in any other combination or combinations of the kind, or occur in a highly restricted number of them, or in a single one only.*

In this definition, which can be made still more general, the emphasis is clearly on combinatorial anomaly. Any syntagm based on such a combination must always, therefore, be unique and anomalous.

At this point it seems necessary to introduce some other distinctions. Unlike the field of idioms, characterized by the ever-present and central concept of **anomaly**, the kind of language (and its use) which is based on normal, familiar rules that have to be observed may be called **regular**. **Idiomatics**, however, **suspends or cancels some of these rules** (these include primarily, as indicated, combinatorial or syntagmatic rules). In regular language, there are also many limitations though of a different order: it is impossible, as we know, for an average word to combine with just any other. Viewed in this light, the term **free combination** or collocation becomes inevitably misleading.

### 3. Idioms: Identification and Test

How, then, does one identify an idiom in text, for the purposes, of, say, analysis or while excerpting texts for a dictionary? Recalling the above definition, the procedure following from it can be split into three steps, of which only the first two are really necessary. Thus,

- (1) Any textual combination encountered in a text and found to be **stable or fixed** in the sense of either having been heard or read before, usually more than once, is a candidate for inclusion among the language's idioms (though this is no substitute for introspection, and a large corpus also helps). Because all idioms are stable or fixed combinations they must occur repeatedly in text and speech. This is, of course, not sufficient in itself and so the second step has to be answered in the positive.
- (2) For all of the candidate combinations chosen (in step 1) it has to be true that **one may not replace ad libitatem** all of its constituents in their slot function and still get acceptable, that is, meaningful (although perhaps not so stable) combinations. Consider, for example, such combinations as *talk through the/one's door/window/wall...*, or *talk through one's cap/coat/sweater/shirt... etc*; however bizarre they may seem, they are all acceptable as variations of a single nominal constituent. There is no doubt, that these randomly chosen variations could be further extended, depending on one's invention and need. But precisely because of this combinatorial, seemingly unlimited liberty, these examples do not represent idioms. By contrast, there is another and, in fact, the sole combination, or variation along the lines suggested by the examples, namely *talk through one's hat*, which is an idiom. Clearly, no other idiom is yielded by substituting any member(s) of the same virtual paradigm for any component of this idiom.

This substitution or commutation is functional in that it operates on a specific pattern and within the same syntagmatic position or distribution, that is, within the domain of the same general function. It also automatically covers the cases adduced in example 3 above (grammatical idioms such as the English prepositional expression *as to*).

- (3) The third step in the identification procedure involves checking the **presence of metaphor**, in a broad sense of the word, in the combinations selected, as this is a very frequent but **not obligatory feature** of idioms. Thus *talk through one's hat* may be viewed as metaphorical, but the prepositional idiom *as to* does not fit the case.

It is evident that this kind of three-step procedure can usually be undertaken only by a native speaker with his/her unique introspection. Such a person effectively taps into his/her previous experience of the language in question. This also explains why idioms can be so difficult for foreign learners of a language to master.

## 4. Conclusion

The definition of 'idiom' given above could be made somewhat more abstract in order to highlight the points raised about the nature of paradigms. Thus the idiom might be viewed as

*a fixed and stable syntagm of elements where at least one element is, as seen from the viewpoint of (at least) one other element, a member of a paradigm that is highly restricted, closed (in form always, though often also semantically).* (Čermák 1982, p. 116, Filipec – Čermák 1985, p. 177, Čermák 1988)

In conclusion, let us return briefly to the problems raised earlier. The answers to the questions in 1.1 are *yes* for No. 1 and 3, but *no* for No. 2. Although there are very many types of anomaly to be dealt with, two general and broad ones are of primary importance: those which are based on transformational behaviour and those which are mirrored in the degrees to which they are ossified, that is, in the constraints on their generative aspects.

However, these and all other types seem to have their solution and explanation in the nature of complex paradigmatic and syntagmatic aspects and a kind of inherent anomaly.

## NOTES

- <sup>1</sup> This is the essence of the definitions given by, for example, A. P. Cowie and R. Mackin (1979, Vol. 1, IX, "which often have implications quite unconnected with the normal meanings of the words themselves"), Ch. F. Hockett (1968, p. 222, "... a grammatical form – single morpheme or composite form – the meaning of which is not deducible from its structure"), B. Fraser (1970, p. 22, "a constituent or series of constituents for which the semantic interpretation is not a compositional function of the formatives of which it is composed), A. Makkai (1972, p. 128, "a polylexemic construction whose aggregate literal meaning derived from its constituent lexemes functions additionally as the realization of an unpredictable sememic network"), D. Crystal (1980, p. 179, "the meanings of the individual words cannot be summed to produce the meaning of the idiomatic expression as a whole"), Dictionnaire de linguistique Larousse (1973, p. 249–250, "on appelle expression idiomatique toute forme grammaticale dont le sens ne peut être déduit de sa structure en morphèmes et qui n'entre pas, dans la constitution d'une forme plus large..."), C. Heupel (1975, p. 95, "habituelle Kollokationen, deren Gesamtbedeutung

nicht aus der Bedeutung der Einzelwörter erschlossen werden kann. Durch die Sinnverlagerung auf die Summe der Elemente sind die Einzelelemente demotiviert“), *Wörter und Wendungen* (1979, p. 30–31, “Ein Glied der Wendung [...] ist abgeblasst oder teilweise umgedeutet [...], aber die Bedeutung der gesamten Einheit ist dennoch nicht direkt erschliessbar [...]. Die gemeinsame Bedeutung ist [...] unbegründet und nicht aus den einzelnen Gliedern abzuleiten“), Molotov (1977, p. 26 “vyrazhenie, imeyushchee leksicheskoe znachenie, kak postoyannyy svoi priznak, i sostoyayushchee ne iz slov, a iz komponentov, kotorye utratili priznaki slova (leksicheskoe znachenie, formy izmeneniya slova)”), Mokijenko (1980, p. 4, “... otnositelno ustoyichivoe, vosproizvodimoe, ekspressivnoe sochetanie leksem, obladuyushchee (kak pravilo) celostnym znacheniem”) etc.

- <sup>2</sup> See Chafe (1968, p. 112, “Peculiarities of idioms – their anomalous meanings, their transformational deficiencies, the ill-formedness of some of them, and the greater text frequency of ill-formed idioms relative to their literal counterparts...”), Fraser (1970, p. 23, “the recalcitrance of idioms to undergo particular syntactic transformations”), ibid. (32), where he quotes Weinreich (1969, “there is no idiom which does not reflect some transformational defect”), Crystal (1980, p. 179, “... a sequence of words which is semantically and often syntactically restricted, so that they function as a single unit”) etc. Note that of Chafe’s four features only the first two are posited as universal.
- <sup>3</sup> For a recent view here and some of its consequences, see note 8.
- <sup>4</sup> “La loi tout [...] fait finale du langage est [...] qu’il n’y a jamais rien qui puisse résider dans un terme...”
- <sup>5</sup> “Il y a entre autres toute une série de phrases qui sont toutes faites pour la langue.”
- <sup>6</sup> This, then, might help to resolve the much contested problem of whether idioms are motivated or not. It evidently boils down to how motivation is defined. Without going into details, let us note that in idioms such as *to rain cats and dogs* the constituent *to rain* has retained its original lexical meaning and can thus be said to be motivated with respect to this meaning. Clearly, this is to be distinguished from the meaning of the idiom whose integrity must be preserved; an opposite approach would, then, have to explain what the meaning of the *cats and dogs* part is, once detached from the allegedly transparent *to rain*. The point recalled here in 1.3 (criterion D) in reference to Shmelev (1973), is certainly an intricate one, but it has to be kept apart from meaning proper. In sum, only some idioms may be partially motivated, but their rather indistinct motivation does not amount to their meaning. Indeed, this very phenomenon, readily observed by many in the past, lay behind some of the early attempts at idiom classification, hence the tendency among some scholars, such as Weinreich (1969), who follow certain Soviet linguists in this respect, to draw a terminological distinction between *idiomatics* and *phraseology*. Because the distinction is both crude and vague and, in fact, a merely a matter of degree, it has become nugatory.

## Shrnutí

### Identifikace frazémů a idiomů

Populární kritérium říkájící, že celkový význam frazému není sumou významu komponentů, neplatí všude a není tedy obecné, srov.

- (1) *beyond my k e n* : *ken*  
*at somebody's b e h e s t* : *behest*  
 (2) *as to* (sb/sth)

V případě (1) nelze pro vyznačené komponenty nalézt význam žádný, v (2) jako by byl význam komponentů zcela zachován, a přesto se tím existence frazému nevyloží (viz ale dál). Alternativní a častý pohled na frazémy vychází z jejich nesporné transformačně-syntaktické anomálie či restrikce. Objevují se však i kritéria mnohá další. Obecně ale je paradoxní, že sice hrubá shoda co do pojmu frazeologie a idiomatika je, ale ta přesto chybí co do kritérií jejího vymezování. Musíme se tudíž ptát

- 1: je možné frazeologii definovat konfigurací kritérií známých odjinud? Anebo je třeba hledat kritéria specifická? (autonomie kritérií)
- 2: je soubor kritérií uplatňovaných zde specifický jen pro tuto oblast? A nevztahuje se také na jiné oblasti? (specifičnost a nezaměnitelnost kritérií)

3: pokrývají zvolená kritéria celou oblast, a ne pouze vybrané úseky, jakkoliv velké? (adekvátnost kritérií)

Analýzou a zobecněním nejčastějších přístupů lze dospět k jejich hlavním rysům, aspektu syntagmatickému, paradigmatickému, a to v podobě jejich různě význačné, ale vždy přítomné anomálie. Už v detailech lze základy tohoto pohledu spatřovat u F. de Saussura a frazém tudíž vnímat jako speciální a anomální případ vzájemné závislosti a funkce vztahů syntagmatických a paradigmatických současně.

Paradigmatické aspekty lze ilustrovat např. na

- (3) *k i t h* and *kin* „friends and relations“ : *kith*  
*s p i c k* and *span* „clean and tidy“ : *spick*  
 (4) *give u m b r a g e*, *take u m b r a g e* at *sth* : *umbrage*

kde vyznačené komponenty lze nalézt pouze v jedné či dvou kombinacích. Můžeme pak tudíž postulat pojem kolokačního paradigmatu sestávajícího z prvků (slov) uzusálně se kombinujících s daným (externím) prvkem (slovem); v uvedených případech je kolokační paradigma drasticky omezeno na 1–2 členy; obecně však, a to je pro celou frazeologii základní, je velmi malé a vždy uzavřené. To ostře kontrastuje se zbytkem pravidelného jazyka, založeného na pravidlech, kde je obvykle takové paradigma otevřené a značně velké, intuitivně i nepřehledné. Situaci si lze ukázat na příkladu *dog (pes)*, který obvykle

- (a) {*bark, growl, whine,...*}, ne však např., *mutter, whistle, mew...*  
 (b) {*lie, run, stand, jump, dig,...*}, ne však např. *wave, rock...*  
 (c) {*sniff, track, nose about,...*} etc.

Víme, že úhrn členů kolokačních paradigmat (je jich obvykle víc) udává spojitelnost slova a obecněji vymezuje jak jeho chování, tak i význam. V našem případě však tento úhrn plus daný externí prvek je totožný s celkem frazému.

Proti kolokačnímu paradigmatu lze spatřovat v jazyce jemu interní virtuální paradigma, které je vymezeno substitucí svých členů ve stejné funkční pozici, sémanticky a funkčně funkcí a významem. Takto je zmíněný *pes* členem paradigmatu domácích zvířat (spolu s kočkou, krávou, prasetem ap.) a širě savců aj. Ale ani z tohoto zorného úhlu nejsou zmíněné formy *kith*, *umbrage* aj. zařaditelné; nejsou členy ničeho. V obou smyslech tu narážíme na zcela principiální anomálii.

V příkladech jako

- (5) *hit the nail on the head*  
 (6) *set the Thames on fire*  
 (7) *be born with a silver spoon in one's mouth*

některé z komponentů jako *nail*, *Thames* nemají žádný z běžných slovníkových významů. Obecněji se tedy chovají jako komponenty v předchozích případech, tj. nepatří do žádného kolokačního ani virtuálního paradigmatu. Je zřejmé, že pokusy přiřadit jim nějaký význam (o to se marně pokoušeli někteří Sověti) selhávají, není tu dost analogie a je třeba se tudíž spokojit s tím, že frazém má integrální, nedělitelný význam. Obecně lze tudíž frazém/idiom pak chápat jako

*jedinečnou a ustálenou kombinaci aspoň dvou prvků, pro kterou platí, že z nich jeden či více nefunguje stejným způsobem v žádné další kombinaci či kombinacích tohoto druhu, anebo se vyskytuje pouze v krajně omezeném jejich množství, popř. jen v jediném.*

Uvedené zakládá možnost identifikaci frazému a idiomu rozložit na test o třech stupních. V prvním je třeba zjistit, je-li daná textová kombinace ustálená. V druhém musí platit, že aspoň jeden z jejích členů nelze podle libosti nahradit jiným a zachovat přitom význam a funkci spojení. Srov.

např. *talk through the/one's door/window/wall...*, či *talk through one's cap/coat/sweater/shirt...* etc. nedospívající k frazému v pozadí *talk through one's hat* a nevytvářející ani žádné jiné či podobné frazémy. Ve třetím kroku, který je však jen opční a platí jen někdy, je třeba hledat přítomnost metafory. Tu např. nelze zjistit u příkladu výše *as to*.

Je zřejmé, že tento druh testu uplatňuje spíše intuitivně pouze rodilý mluvčí. Uvedené lze zobecnit ještě více a postulovat, že *frazém/idiom je ustálené syntagma prvků, z nichž aspoň jeden je z hlediska jiného členem paradigmatu, které je krajně omezené a uzavřené (vždy formálně, i když často i sémanticky).*

# Statistical Methods for Searching Idioms in Text Corpora

(In: *Phraseology in Motion I Methoden und Kritik*, 2006, p. 33–42)

## 1. Idioms and Their Recognition

A major problem, which still seems to remain unsolved, is how to design a procedure for automatic idiom recognition in text. In the age of corpora, this happens to be a major challenge. On the basis of his/her knowledge, a native speaker may recognize a new idiom or phraseme intuitively, but modern corpora are too large to allow any manual approach that could be efficient. A trained linguist may know better, however. Two features of an idiom, its (1) anomaly or restriction (Čermák 1988, 1994, 2001 et al.) and (2) recurrence (viewed generally as the basis for its stability and fixity) do stand out and it seems that it is precisely these two aspects that any automatic procedure has to involve. As suggested earlier, a commutation test (Čermák above) for “manual” identification of idioms is, indeed, a working solution behind intuition, which may be applied by any native student for text data.

A large corpus, such as *Czech National Corpus*, is just large enough to record reliably the recurrence of some idioms at least. On the other hand, the **anomaly problem**, being a complex one and involving many aspects, primarily of syntagmatic and paradigmatic nature, may require a stepwise approach, based on more than one procedure. In practice it should imitate the commutation test above (1988, p. 433–434). To know how to solve these, one has to recognize what sort of anomalies are involved and what they are based on and this is where theory comes in. Hence, it is the author’s view that no pure and blind automatic method of discovery is possible without a priori theory. In what follows, a first and modest step through the multitude of possible cases of idioms is suggested, aiming at their identification, concentrating, admittedly, on rather simple types of these. The very identification is up to the researcher, then, and his/her theory applied.

## 2. A Recognition and Discovery Approach

Czech National Corpus (CNC, see <http://ucnk.ff.cuni.cz>), or, rather, its 100-million specimen SYN2000, is made up of written contemporary texts which are POS-tagged and lemmatised. Hence, phrasemes and idioms which are typical for the spoken language, may be underrepresented here.

Despite a variety of corpus tools which have been suggested so far, i.e. statistical formulas based on frequency in various ways, two of the oldest ones, being highly employed in corpus linguistics, namely MI-score and t-score (see, for example, M. P. Oakes 1998), have been found useful in discovering word collocations. T-score is very useful in determining and, for practical purposes, finding highly frequent collocations, which are quite regular and often grammatical in their nature (including difficult and poorly-researched multi-word particles, for example, cf. Czech *a vůbec* “and anyway” having MI-score 1.464 and t-score 32.748). Thus, it is the other score which offers greater yields. MI-score (or mutual information) is a well known statistical procedure for discovering genuine relationships between words. It compares the probability of finding a word “x” and “y” together (their joint probability) with the probabilities of finding words “x” and “y” independently of each other (chance). Among other things, MI-score is well suited to discover idioms in certain sections of the range of collocations obtained, if one needs to study their syntagmatic anomaly or restrictions or any other aspect. To illustrate this, the Czech idiom *krokodýlí slzy* (Engl. *crocodile tears*, Germ. *Krokodilstränen*, Fr. *larmes de crocodile*), having a rather low corpus frequency of 12, is discovered thanks to its MI-score which is rather high, namely 12.513 (while the absolute frequency of its constituents is different, 70 and 3545, respectively). To appreciate this, one has to see this against the background of the full range of MI-scores, however, which start with the top value of 25 going down the scale.

Useful preliminary to application of such an approach based on very large annotated and lemmatised corpus data is setting the limits or threshold for the area to be explored, should one not want to get MI-scores for the whole corpus and face a laborious analysis of all corpus data. Usually, it is useful to work with bigrams and a list of frequencies for a list of bigram combinations delimited in advance. It turns out that the evaluation of the full scale of results of MI-score for the whole structure chosen, if it is applied to results obtained from a large corpus as a whole, is largely empirical. Accordingly, one has to find out which part of the full range offers more plausible or, rather, appropriate idiom results, in other words to determine the subrange relevant for the given purpose offering real gains. Thus idioms having a monocollable constituent (see Čermák above), i.e. one which is to be found in a single combination only or in very few belonging to a clearly closed class, score very high on the list. Here, such cases belong as e.g. *šerý dávnověk* (Engl. *the ancient past*, where the noun *dávnověk* is attested only here), whose MI-score 16.33 is higher than the absolute frequency (14), or *zastírací manévr* (Engl. *camouflage manoeuvre*) with MI-score of 16.24 (and absolute frequency 33). Now, should one compare idioms found by MI-score method in CNC with those recorded in a large Czech Dictionary of Idioms (*Slovník..., Výrazy neslovesné 1988*), it appears that over a quarter of idioms having the Adjective-Noun structure in the idiom dictionary has a corpus frequency higher than 10. Having a corpus which is fully tagged (also for POS) and lemmatized, such as CNC, it is easy to make a query on a higher level, such as combination of parts of speech.

Along these lines and following the MI-score approach, a research has been undertaken for six simple binary classes and structures of idioms. By way of illustration, only examples from the top of each structure (where the first figure is MI-score and the second its overall corpus frequency) are given, including combinations of

- (1) **Adjective-Noun**, *Augiášův chlív* (Engl. *Augean stable*) 18.16 and 22,
- (2) **Noun-Noun**, *techtle mechtle* (Engl. *an affair, hanky-panky*, Ger. *Techtelmechtel*) 22.88 and 10,
- (3) **Adjective-Adjective**, *neslaný nemastný* (Engl. *wishy-washy*) 10.44 and 22,
- (4) **Adverb-Adverb**, *lážo plážo* (Engl. *leisurely time*) 23.65 and 7,
- (5) **Adverb-Adjective**, *narychlo spíchnutý* (Engl. *makeshift*) 15.84 and 10,
- (6) **Preposition-Noun**, *pod taktovkou někoho* (Engl. *under the baton of*) 10.19 and 438.

In sum, the relevant section of the MI-score range where idioms obtain may be different for different structures. Thus, for the first structure, idioms usually obtain an MI-score of around 18 or below, but for the fourth structure these are concentrated at the very top of MI-score values, mostly. On the other hand, the first structure seems to have a much wider and balanced distribution over the full scale of the values found.

Some idioms have a very low absolute frequency in the corpus, but their high MI-score safely points to them as candidates for inclusion in the idiom list (or dictionary), cf. *pavlačová drbna* (*backyard gossip*) with 17.42, *lukulské hody* (Engl. *Lucullan banquet/repast*) 15.72, etc. On the other hand, the lowermost MI-score limit for idioms seems to be around 7 where idioms are blended and overlap with other, standard collocations, cf. *nezvaný host* (Engl. *uninvited guest*, MI 7.62, frequency 117) and *logický důsledek* (Engl. *a logical consequence*, MI 7.89, frequency 111). More generally, expectations about low frequency of idioms can thus be confirmed, since most of those belonging to the type investigated in CNC seem to have a frequency well under 100.

Although this has not been tried yet, the MI-score working on binary combinations so far can easily be adapted and used for a similar identification of of collocations and idioms made-up of three constituents as well.

To get a more detailed view, in what follows, an investigation has been narrowed down to a single structure, namely Adjective-Noun.

### 3. An Experimental Probe into a Corpus

The Adjective-Noun combination and idiom structure chosen is very simple, supported by grammatical congruence, and it is also very common. The query of 100 million words of CNC gave some 45 420 collocations of this type. These results have then been easily narrowed down by their relevant MI-score, verified earlier on a sample. The relevant range where idioms obtain most appears to be 22.7 or lower (down to 7), supported by a minimal absolute frequency of 10. On the MI-score level 7, a marked transition is to be observed from idioms to standard, normal collocations. In this way, some 1700 collocations as candidates for idioms have been delimited. Subsequently, these have been manually checked-up, too. Obviously, since results obtained in this way are much broader, they can easily be made subject of a special study of collocations, etc.

Results show that idioms are to be found along a range of MI-score values, the highest being around 18, such as *Augiášův chlív* (18.16, Engl. *Augean stable*, Ger. *Augiasstall*), i.e.



those made of a proper name (*Augean* in English, while German is different here) or *Pandořina skříňka* (17.65, Engl. *Pandora's box*), etc. This way, some old and familiar idioms, though never recorded in any dictionary before, could have been found, such as *pěnivý mok* (17.04, Engl. appr. *frothy drink*, i.e. beer, *bermudský trojúhelník*, Engl. *Bermuda triangle* 16.02, or *chytrá horákyně*, Engl. “*sitter-on-the-fence*” 15.42). Of course, this procedure does not automatically guarantee that only idioms are discovered, since similar MI-scores might be found for other interesting and important collocations, too, and these have to be filtered manually, using introspection and native knowledge. Thus, virtually identical values are to be found also for *diakritické znaménko* (Engl. *diacritical mark*, 17.41) or *kropicí konev* (Engl. *sprinkling can*, 17.10), showing that this approach may be made much more general and help to find out other sorts of collocations, as well, a topic widely discussed in corpus linguistics.

Going down the scale and having a look at a lower MI-score of, for instance, 7, one finds collocations and idioms with adjectives of a very high frequency, such as *malá dušička* (Engl. *heebies-jeebies*, *jitters*, lit. *small soul*, 7.202) or *silný kafe* (Engl. *staggering act/feat*, 7.17, lit. *strong coffee*).

Going into even finer details, on the MI-score levels that we have selected to follow, namely 7 (A), 8 (B), 10 (C), 12 (D) and 15 (E), but looking at the overall frequency of the collocation higher than 10, one gets a glimpse at the rate of idioms chosen from each particular value range. Compare, then, the results for

MI-score	Collocations Found	Identified as Idioms	Example
<b>A 7.0–7.9</b>	11 896	495	<i>ohnivá voda</i> (fire-water) <i>dobrá vůle</i> (good will) Note: only 35 idioms with frequency 10, while 438 idioms have frequency under 100
<b>B 8.0–8.9</b>	9 628	205	<i>širé pole</i> (wide field) <i>čisté svědomí</i> (clear conscience) Note: only 14 idioms with frequency 10, while 184 idioms with frequency under 100
<b>C 10.0–10.9</b>	5 745	250	<i>nekalé úmysly</i> (dishonest intentions) <i>drtivá většina</i> (overwhelming majority) Note: only 11 idioms with frequency 10, while 214 idioms with frequency under 100
<b>D 12.0–12.9</b>	2 984	140	<i>hořká příchut'</i> (bitter taste) <i>železná opona</i> (iron curtain) Note: only 13 idioms with frequency 10, while 117 idioms with frequency under 100
<b>E 15.0–15.9</b>	803	26	<i>dojná kráva</i> (milk-cow) <i>odrazový můstek</i> (jumping-off point) Note: only 1 idiom with frequency 10, while 24 idioms with frequency under 100

It is evident that increased MI-score values yield diminishing returns and that highest yield for

(A) **idioms** is to be found, in this case, around value 7. Going still further up, i.e. above the given MI-score range, i.e. above the value of 15, one finds an increase in other types

of collocations. Thus, while only 2.5% of idioms could have been found above 15, such as *kategorický imperativ* (Engl. *categorical imperative*, MI 16.13 with frequency 52), some 56% of collocations could be labeled as

- (B) **multi-word terms**, such as *lexikalizované přirovnání* (Engl. *lexicalized comparison*, MI 16.89) and over 20% as
- (C) **multi-word proper names**, such as *Gibaltarský průliv* (Engl. *Straits of Gibraltar*, MI 17.56), the rest being
- (D) **neutral (normal) collocations**.

If one goes in the opposite direction along the MI-score scale, ad hoc normal collocations seem to prevail, both textual and systemic (44.3% and 37.5%, such as *veřejná omluva*, Engl. *public apology* with MI-score 7.17 and frequency 48). Only 2% of the collocations found here belong to idioms, such as *osmý div* (Engl. *eighth wonder* MI 7.05 and frequency 13).

#### 4. A Comparison of Dictionary and Corpus Data

Should we compare idioms found by MI-score in CNC with those recorded in a large Czech Dictionary of idioms (*Slovník..., Výrazy neslovesné* 1988), it appears that about half of 600 idioms, having the Adjective-Noun structure, chosen for inspection from this dictionary, have a corpus frequency higher than 10. This sample of 600 idioms is made up of different types, the largest part (7–11%) being formed by those having MI-I score 10 to 4. However, those idioms that have their MI-score lower than 7, **overlap statistically** with standard collocations, cf. *celkový dojem* (MI-Score of 6.84, Engl. *general impression*) as against *pocitivá hra* (Engl. *fair-play*, 5.58) and *šťastný konec* (Engl. *happy-end*, 4.60). Hence, searching in this part of the MI-score scale becomes difficult due to the sheer body of data. Thus it has been confirmed that the **heuristic choice of the MI-score range between 18 and 7** for the given structure is suitable.

#### 5. A Comparison with English

To verify these findings and results for the Czech language, English has been investigated, too, in a limited probe of 20 comparable (identical) idioms in British National Corpus, being of the same size as Czech National Corpus. In general and apart from specific differences, it is evident that results are similar and this approach can be applied to English, too, cf. the Czech *stříbrný táč* (MI-score value 10.618 and frequency 16) and the corresponding English idiom *silver platter* (MI-score 11.263, frequency 14).

Some, rather obvious, differences are due to the typological difference of both languages, such as *anoun* in English used where Czech has a full-fledged Adjective (*lion* vs. *lví*, in *lion's share* and *lví podíl*), whereby frequencies of constituent parts are naturally influenced.

Almost identical results are observed for such classic idioms as *Gordian knot* and *Trojan horse*, however.

These twenty Czech idioms and their English equivalents, as found in CNC and BNC, give the following table:

ADJ	NOUN	MI	Idiom freq.	Adj. freq.	Noun freq.	ADJ	NOUN	MI	Idiom freq.	Adj. freq.	Noun freq.
křišťálový	koule	12.253	69	631	2 708	crystal	ball	8.709	66	1 997	8 788
fackovací	panák	17.41	44	55	555	whipping	boy	10.432	14	55	20 494
krokodýlí	slzy	12.513	12	70	3 545	crocodile	tears	10.072	26	376	7 148
dominový	efekt	13.942	39	58	5 167	domino	effect	8.207	22	239	34 649
mrtvý	bod	6.134	198	10 347	32 939	dead	end	4.268	123	11 794	60 204
velká	ryba	3.821	163	181 776	7 674	big	fish	5.342	103	24 853	11 367
železná	pěst	7.763	23	6251	2 048	iron	fist	8.421	24	5 190	1 501
darovaný	kůň	9.256	16	270	11 720	gift	horse	5.307	20	4 581	12 267
první	pohled	6.977	3 504	138 317	24 309	first	sight	6.255	597	116 940	7 434
pokusný	králík	12.549	55	595	1 865	guinea	pig	13.219	130	617	2 458
stará	vesta	5.627	21	58 579	877	old	hat	4.577	42	52 755	3 710
chladná	hlava	5.467	49	3 333	40 189	cool	head	3.421	20	4 569	45 455
zlomené	srdce	10.932	319	1 340	14 732	broken	heart	7.580	70	2756	14 766
trojský	kůň	10.894	95	515	11 720	Trojan	horse	11.217	37	141	12 267
slonovinová	věž	11.438	7	56	5 448	ivory	tower	11.308	82	826	4 355
gordický	uzel	16.060	66	76	1 537	Gordian	knot	15.121	9	20	1 405
lví	podíl	10.962	149	458	19 724	lion's	share	3.773	7	2 084	27 337
kostelní	myš	11.714	32	592	1 946	church	mouse	4.629	14	23 474	2 681
stříbrný	tác	10.618	16	4 494	274	silver	platter	11.263	14	3 987	159
horké	křeslo	6.077	19	5 301	6 421	hot	seat	5.858	53	8 827	11 515

Note that idioms have been sorted by the Czech MI-scores.

## 6. Conclusion

It seems that most idioms have the absolute frequency below 100 in CNC, but those with a very low frequency (1–3) seem to be obsolete and falling out of use. It is important to note that most idioms seem to be lumped together somewhere on the scale of the MI-score values.

Although there is still some necessity of a manual evaluation and filtering of the results which have been obtained automatically, this seems to be a rewarding approach. Beside a much-needed numerical and frequency features that enable one to compare idioms on a more solid basis (for lack of space, this can only be noted here), it might seem to be a step to an **idiom discovery procedure**, enabling one to find out new idioms in a language corpus, not recorded before and help to automatically identify these. Of course, due to the cline or scalar quality of results obtained, there is an ample possibility and a vast field for a study of transitory phenomena, too, ranging from individual metaphors to idioms, from idioms to terms and proper names, and from idioms to neutral collocations. Obviously, only on the basis of such an approach, dealing with massive and objective data, criteria for discrimination between idiom and non-idiom can be sought.

One of the limitations of this MI-score approach is that it does not yield idioms made up of frequent constituents only, such as the Czech *tvrdá hlava* (Engl. lit. *hard head*, “a stubborn person”, *pig-headed person*, Ger. *Dickschädel*) where both *tvrdý* (hard) and *hlava* (head) are highly frequent words. The MI-score of this idiom is very low, only 2.07, in fact.

In general, it holds that idioms with frequent constituents, such as *hlava* head (having absolute frequency over 40 thousand in CNC) differ statistically due to a varying frequency of the adjective accompanying it. Thus MI-score of *korunovaná hlava* (Engl. *crowned head*) is 8.80 since *korunovaná* (*crowned*) is less frequent (208), while *čistá hlava* (Engl. *clear head*) has MI-score of 2.61, since absolute frequency of the Adjective *čistý/á* (*clear, clean*) is almost 14 000. For more, see the Appendix on *head*-idioms (*hlava*) and their MI-Scores.

Finally, due to a different and often much higher use of some idioms in spoken language, it has to be stressed, yet again, that this approach has been limited to written language only, so far.

On a more general level, it is apparent that searching more complex structures, a more elaborate approach would be necessary. The part played by humans, in fact natives preferably, in evaluation of this kind of results obtained automatically is considerable and there is no avoiding it. However, a very large corpus is a safeguard that an objective picture of most of existing idioms could be made available to choose from. It seems to be a first approach ever to offer us an exhaustive (or almost exhaustive) coverage of the field giving us, for the first time, assurance that there is almost nothing left out of our data and field examined. This is exactly the feeling lexicographers and linguists are looking for, wanting to be sure they do have all their data. Of course, even a very large corpus may not have everything today, but corpora do grow by the day.

## Appendix

Adjective-Noun Idioms containing the same constituent **hlava** (head) whose MI-score is usually lower than 7:

Idiom	MI-score	Frequency	English Equivalent
<i>korunovaná hlava</i>	8.809	31	crowned head
<i>kočičí hlavy</i>	8.046	47	cobble-stones
<i>skopová hlava</i>	7.211	10	muttonhead
<i>bystrá hlava</i>	6.508	22	quick/ready wit
<i>chytrá hlava</i>	5.982	52	clever head
<i>rozpálená hlava</i>	5.465	9	hot head
<i>telecí hlava</i>	5.458	9	litt. calf head, i.e. idiot
<i>dutá hlava</i>	5.184	8	lackbrain
<i>děravá hlava</i>	4.889	4	a memory like a sieve
<i>horká hlava</i>	4.737	47	hothead
<i>moudrá hlava</i>	4.629	20	wiseacre
<i>hloupá hlava</i>	3.438	9	scatterbrain
<i>prázdná hlava</i>	2.908	16	empty vessel
<i>čistá hlava</i>	2.617	28	a clear head
<i>jasná hlava</i>	2.585	40	a clear head
<i>tvrdá hlava</i>	2.074	16	pig-headed fellow
<i>otevřená hlava</i>	0.302	5	good head

Note. Much of data research is due to Renata Blatná who has thus been a co-author of this contribution.

---

### Shrnutí

Tento příspěvek se do značné míry překrývá s příspěvkem *Statistické metody hledání frazémů a idiomů v korpusech*.

# Statistické metody hledání frazémů a idiomů v korpusech

(In: *Studie z korpusové lingvistiky 2, Kolokace*, Praha 2006)

## 1. Frazémy a jejich identifikace

Velkým nevyřešeným problémem při automatickém zpracování textu je, jak navrhnout takový účinný postup, který by v něm rozpoznal mj. i frazémy a idiomy. Toto je, jak se zdá, nyní v korpusovém věku zásadní a obecná výzva vybízející k hledání cest, jak se spolehlivě a počítačově dobrat obecně zadaného cíle; ten však na první pohled nenese přímočaře nějaké závazné formální znaky, o které se dá opřít. Na základě své znalosti rodilý mluvčí většinou idiom či frazém dokáže v konkrétním textu jistě zhruba rozpoznat, ale moderní korpusy jsou natolik velké, že jakýkoliv manuální přístup, který by byl efektivní, jasně vylučují.

Avšak lingvistu školeného v korpusové problematice určitá řešení napadat můžou. Nápadné už na první pohled tu jsou aspoň dva rysy frazémů, které lze při hledání řešení uvážit a pokusit se použít je pro automatický přístup: (1) **anomálnost** či restringovanost frazému (Čermák 1988, 1994, 2001 et al.) a (2) jeho opakovaný výskyt (**rekurence**) v textu, obecně chápáný jako základ jeho ustálenosti. Pro manuální a „intuitivní“ identifikaci frazémů uplatnitelnou každým rodilým mluvčím je možné, jak už bylo popsáno dříve (viz Čermák výše), použít komutační test; ten se zdá, že lze tudíž použít nejen pro ověření správnosti nalezených výsledků, ale je svým způsobem v pozadí i možného techničtějšího řešení. V základu tohoto řešení je takové pojetí anomálnosti, které operuje s intuitivní znalostí jazyka rodilého mluvčího a tedy i jeho schopností rozpoznat, který typ kombinací (kolokací) je běžný (přičemž daná kombinace se skládá z „běžných“ komponentů) a který není. Toto rozpoznávání uživatel intuitivně dělá na pozadí znalosti svého idiolektu, tedy celého svého jazyka a kombinatorických poměrů v něm. Pokud je však některá kombinace (srov. např. *nechat na holičkách, vzít do zaječích*) „neběžná“, neobvyklá a tedy anomální například v tom smyslu, že některý komponent se jinde vyskytuje omezeně, resp. jen s jiným typem komponentů a je tedy jedinečný, uživatel to intuitivně většinou vytuší. Je nasnadě, že úhrnnou znalost slovníku (v tomto případě je nepřesně ztotožněný se slovníkem daného velkého korpusu) a jeho kombinací zjistí snadno potřebné korpusové nástroje; ty také rozumně můžou ukázat na „neběžnost“ hledaných kombinací. Bude tedy užitečné tento předpoklad, převádějící intuici do formalizovaného přístupu, na korpusu skutečně ověřit.

Velký korpus, jakým je i Český národní korpus (zvláště SYN2000), je už natolik velký, že zachycuje spolehlivě rekurentnost aspoň běžných frazémů a jejich chování, takže se souvšsky

jejich komponentů tak vyjeví poměrně snadno. **Anomálnost** je naproti tomu komplexní otázka syntagmaticko-paradigmatické povahy složená z řady konkrétních stránek a může si vyžádat přístup postupný, založený na vícero krocích (výše byla naznačena jen její jedna stránka). V praxi by však zmíněný komutační test měly počítačové přístupy aspoň obecně sledovat a napodobovat (1988, s. 433–434). Pro formulaci tohoto přístupu je ovšem zapotřebí si uvědomovat, o jaké anomálie se jedná, na čem jsou založené a v jaké fázi je třeba se uchýlit k řešení nabízenému teorií. V tomto smyslu je autor přesvědčen o tom, že žádná čistá a slepá automatická metoda frazémy objevující neexistuje a bez ověřené předchozí teorie ani není možná. V následujícím se pokusíme navrhnout první skromný krok ukazující cestu skrze množství různých možných případů frazémů, který bude usilovat o jejich rozpoznání; je třeba otevřeně přiznat, že se přitom soustředíme jen na poměrně jednoduché typy. Vlastní identifikace však pak bude záležet na lingvistovi a také na tom, jakou teorii k tomu použije.

## 2. Rozpoznávání a objevování

Český národní korpus, resp. jeho stomilionový vzorek SYN2000, je vytvářený současnými psanými texty, je morfologicky označován na úroveň slovních druhů a jejich kategorií a je lemmatizovaný. Je mj. už předem jasné, že frazémy a idiomy tu nemusejí být zastoupeny nejlépe; hodně jich je mluvené povahy a v mluvených textech časté nejsou, pokud tam jsou vůbec.

Přestože nástrojů, tj. statistických formulí založených různým způsobem na frekvenci, které se teoreticky navrhuje k hledání (a nejsou tedy v praxi ověřené), je celá řada, zdá se, že dva z těch nejstarších, MI-score a t-score (viz např. Oakes 1998), se k hledání slovních kolokací hodí docela dobře. T-score je velmi užitečný pro určování a prakticky i nacházení vysoce frekventovaných kolokací, které jsou jinak (sémanticky) pravidelné a svou povahou gramatické, takže zahrne i obtížné a málo zkoumané částice, srov. např. české *a vůbec* mající MI-score 1,464, ale t-score 32,748. Jinými slovy pro jádro frazémů vydatnější výsledky dává zřejmě MI-score, čili index vzájemné informace. Tato známá statistická míra, která se obvykle užívá ke zjištění skutečné souvztažnosti mezi dvěma slovy, se zakládá na tom, že se srovnává pravděpodobnost nalezení slova (komponentu) „x“ a „y“ v korpusu ve vzájemné společnosti (jejich společná pravděpodobnost) s pravděpodobnostmi nalezení slova „x“ a slova „y“ nezávisle na sobě (náhodný výskyt) a z obojího se vyvozují signifikantní závěry o povaze nalezených kombinací. MI-score se tedy mimo jiné dobře hodí k nacházení frazémů v určité části získané škály všech kolokací, na níž chceme studovat jejich syntagmatickou anomálnost a restriktce či jiný jejich aspekt. K ilustrování si zvolme český frazém *krokodýlí slzy* (angl. *crocodile tears*, něm. *Krokodilstränen*, franc. *larmes de crocodile*), který lze takto objevit právě díky MI-score, jehož korpusová hodnota 12,513 je zde dosti vysoká (zatímco absolutní frekvence jeho komponentů je odlišná, tj. 70 a 3545). Aby se takové výsledky nicméně daly plně pochopit a docenit, je třeba je spatřovat a posuzovat proti pozadí celé škály MI-scorů dalších výrazů a kolokací v celém SYN2000, které začínají vrcholovou hodnotou okolo 25 a dál klesají.

Nutným krokem předcházejícím tomuto přístupu, který se zakládá na velmi velkém anotovaném a lemmatizovaném korpusu, je potřeba stanovit si hranice či prahové hodnoty pro

zkoumanou oblast, resp. škály, pokud nechceme pracně vytvářet, resp. vypočítávat MI-score pro celý korpus a ještě pracněji pak analyzovat veškerá korpusová data. Takové hranice lze v zásadě stanovit jen empiricky a často je pak vlivem analýzy dat i zpřesňovat.

V takovém (i tomto konkrétním) případě je pak obvykle obecně užitečné použít předem získané bigramy celého korpusu (doplňené informací o slovních druzích). Spolu s dalšími získanými údaji, tj. s frekvencí bigramů prvního i druhého jejich slova a také jejich MI-scorů, tak bylo možné vytríděním „nepotřebného“ z celého výsledku sestavit seznam relevantních frekvencí pro seznam bigramových kombinací daného typu, který byl rozhodnutý předem; v daném případě volba padla na kombinaci adjektivum-substantivum. Nutnou omezující podmínkou už na začátku postupu bylo samozřejmě vymezení minimální frekvence jednotlivých slov, kterou vyžadují výpočty MI-scoru (takto bylo rozhodnuto, že jak první tak druhé slovo má mít každé frekvenci minimálně 5) a obdobně musela být (pragmaticky) vymezena i minimální frekvence celého bigramu (v daném případě 3). Dosažený výstup se pak už jen přetřídil podle hodnoty MI-scoru. Obecně je zřejmé, že stanovení hranic, odkud nás budou výsledky zajímat, a tedy „odříznutí“ (cut-off) nezajímavých, je věcí osobního rozhodnutí a závisí to mj. na sledovaném cíli.

Ukazuje se, že vyhodnocení plné škály výsledků MI-score získaných z celého korpusu pro celou zvolenou strukturu je do značné míry empirické. Takto je třeba pak zjistit, která část z dané škály výsledků skýtá přijatelnější, resp. užitečnější a případněji frazeologické případy. Jinými slovy je pak třeba určit takový podrozsah, úsek dané škály, který je relevantní pro daný cíl a nabízí plodné výsledky toho typu, o který usilujeme. Při tomto pohledu se ukazuje, že frazémy s **monokolokabilním** prvkem (viz Čermák výše), tj. takovým, který nacházíme v kombinaci jediné nebo jen několika málo náležejícím ke zřetelně uzavřené třídě, se na dané škále umísťují velmi vysoko. K takovým případům zde patří *šerý dávnověk*, kde slovo *dávnověk* je doloženo jen pro tuto kombinaci, jejíž MI-score je 16,33 a absolutní frekvence 14. Specifičtěji substantivum *dávnověk* se pojí 21x s předložkou *v*, tj. *v dávnověku*, 14x s adjektivem *šerý*, tj. *šerý dávnověk*, dále 7x s předložkou *z*, tj. *z dávnověku*, 5x s předložkou *od*, tj. *od dávnověku*, a 3x s předložkou *do*, tj. *do dávnověku*. Všechny další kombinace tohoto substantiva s adjektivy jsou aktualizacemi, srov. následující tabulku:

Adjektivum	Frekvence v kombinaci se substantivem <i>dávnověk</i>	MI-score kolokace
<i>šerý</i>	14	16,33
<i>český</i>	3	4,41
<i>antický</i>	2	11,39
<i>pohanský</i>	2	12,75
<i>automobilový</i>	1	8,46
<i>barbarský</i>	1	12,08



Adjektivum	Frekvence v kombinaci se substantivem <i>dávnověk</i>	MI-score kolokace
<i>cikánský</i>	1	11,05
<i>iberský</i>	1	14,65
<i>ideální</i>	1	8,17
<i>keltský</i>	1	10,86
<i>lidový</i>	1	6,82
<i>mytologický</i>	1	13,06
<i>počítačový</i>	1	7,83
<i>praslovanský</i>	1	15,55
<i>prehistorický</i>	1	13,23
<i>rockový</i>	1	8,68
<i>šachový</i>	1	10,12
<i>tajuplný</i>	1	11,31
<i>temný</i>	1	8,46
<i>železniční</i>	1	7,65

I jiná adjektiva než jen *šerý* mají vysoké hodnoty MI-score, ale kvůli nízkým frekvencím v kombinaci s daným substantivem nemohou být vnímána při výběru frazémů jako relevantní. Situaci substantiva *dávnověk*, resp. kolokaci *šerý dávnověk*, je podobný *zastírací manévr*, který má MI-score 16,24 (s absolutní frekvencí 33). Pokud bychom chtěli srovnávat frazémy nalezené takto pomocí MI-score s frazémy uvedenými v hesláři *Slovníku české frazeologie a idiomatiky*. Výrazy *neslovesné* (1988), ukáže se, že přes čtvrtinu frazémů se strukturou adjektivum-substantivum z tohoto slovníku má v korpusu frekvenci vyšší než 10. Pokud máme k dispozici plně označovaný a lematizovaný korpus jako ČNK (SYN2000), je pak snadné pokládat dotazy i na vyšší úrovni, tj. úrovni kombinace slovních druhů, a to nejen těchto dvou.

Tímto způsobem a za využití MI-score jsme takto prozkoumali, resp. sondažovali šest prostých binárních tříd a zároveň frazeologických struktur a zjišťovali, zda mezi výsledky pro každou takto obecně vymezenou kombinatorickou strukturu jsou někde ukryty i frazémy, resp. idiomy. Nalezené české případy budou v dalším pro srovnání ilustrovány a srovnávány většinou s anglickými, které vycházejí ze srovnatelného korpusu anglického BNC. Pro ilustraci si uvedme jen příklady z vrcholu každé struktury (kde první číslo udává MI-score a druhé celkovou frekvenci bigramu v korpusu), reprezentované kombinacemi

- (1) **Adjektivum – Substantivum**, *Augiášův chlív* (angl. *Augean stable*) 18,16 a 22,
- (2) **Substantivum – Substantivum**, *techtle mechtle* (angl. *an affair, hanky-panky*, něm. *Techtelmechtel*) 22,88 a 10,
- (3) **Adjektivum – Adjektivum**, *neslaný nemastný* (angl. *wishy washy*) 10,44 a 22,
- (4) **Adverbium – Adverbium**, *lážo plážo* (angl. *leisurely time*) 23,65 a 7,
- (5) **Adverbium – Adjektivum**, *narychlo spíchnutý* (angl. *makeshift*) 15,84 a 10,
- (6) **Prepozice – Substantivum**, *pod taktovkou někoho* (angl. *under the baton of*) 10,19 a 438.

Shrme-li, lze říct, že relevantní úsek celé dané škály podle MI-score, kde se frazémy vyskytují, se může pro ten který typ značně lišit. Tak například zjišťujeme, tj. po studiu výsledků, že se frazémy u první struktury obvykle vyskytují kolem MI-score 18 a níže, avšak u čtvrté struktury se většinou soustřeďují na samém vrcholu hodnot MI-score. První struktura, jak se zdá, má naproti tomu distribuci frazémů mnohem širší a vyrovnanější, která se rozprostírá přes celou škálu nalezených hodnot.

Některé frazémy mají v korpusu absolutní frekvenci velmi nízkou, i když jejich MI-score na ně bezpečně ukazuje jako na kandidáty k začlenění do seznamu frazémů či slovníku, srov. *pavlačová drbna* s MI-score 17,42 a *lukulské hody* s 15,72. (Důvodem může být například jejich mluvená povaha, zmíněná výše, ale třeba i stále malý rozsah korpusu apod.) Naproti tomu se ale zdá, že **nejnižší limit MI-score pro frazémy je kolem hodnoty 7**, kde se už počítají mísit a splývat s jinými, běžnými kolokacemi, srov. *nezvaný host* s MI-score 7,62 (absolutní frekvence 117) a *logický důsledek* s MI-score 7,89 (absolutní frekvence je 111). Obecně se ukazuje, že naše očekávání, že frekvence frazémů bývá nízká, se takto potvrzují; většina frazémů patřících do zkoumaného typu má totiž v korpusu frekvenci nižší než 10.

Jakkoliv se o to dosud nikdo nepokusil, zdá se, že MI-score fungující na binárních kombinacích lze zřejmě snadno upravit a rozšířit tak, aby se ho dalo užít pro obdobnou identifikaci kolokací a frazémů vytvářených i ze tří komponentů, ale o to se zde nepokoušíme.

K získání podrobnějšího vhledu se v dalším naše pozornost zúžila a omezila pouze na jedinou strukturu, totiž na první kombinaci, adjektivum-substantivum.

### 3. Experimentální sonda do korpusu

K běžné nabídce korpusového manažeru Bonito nepatří tvorba seznamu bigramů, který by také udával frekvenci dané kombinace (kolokace), frekvence dílčích lemmat i hodnotu MI-score; proto se bigramy získaly z korpusu SYN2000 pomocí zvláštního programu, jehož výsledkem byl seznam bigramů složených z adjektiva a substantiva s minimální frekvencí dané kombinace (kolokace) 3. Ukázalo se, že v takto málo frekventovaných bigramech jsou často kombinace autorské nebo aktualizovaná užití, popř. šlo o frazémy užívané v mluveném jazyce, které se v rámci korpusu psaných textů proto vyskytovaly v omezené míře. Seznam byl z toho důvodu redukován na kombinace s minimální frekvencí 10 výskytů daných bigramů. Tento materiál byl pak dále analyzován z hlediska zastoupení frazémů po manuální analýze.

Kombinace a zároveň frazémová struktura adjektivum-substantivum je velmi prostá, podporovaná gramatickou shodou, a je také velmi běžná. Zkoumání stomilionového korpusu dalo na 45 420 kombinací tohoto typu. Takto široké výsledky se daly pak snadno omezit podle relevantních MI-score hodnot, které se zjišťovaly ručně na vzorcích. Relevantní **úsek, ve kterém se frazémy vyskytují nejčastěji, se zdá začínat u indexu 22,7 či nižšího, až k 7**; to podporuje vedle MI-score i minimální absolutní frekvence 10. Na úrovni MI-score 7 lze pozorovat nápadný přechod od frazémů k běžným kolokacím. Takto se tedy dalo vymezit na 1700 kolokací kandidujících na frazémy. Ty bylo pochopitelně pak nutné projít manuálně. Protože tímto způsobem získané výsledky jsou mnohem širší, je zřejmé, že se dají využít i pro speciální studium kolokací vůbec a podobně.

Výsledky ukazují, že frazémy lze najít po celé škále hodnot MI-scoru, kdy nejvyšší je kolem 18, tj. frazémy budované na vlastních jménech, např. *Augiášův chlév* (18, 16, angl. *Augean stable*, něm. *Augiasstall*) (*Augean* v angličtině a *Augiášův* v češtině, zatímco v němčině je z něj část kompozita) nebo *Pandořina skříňka* (17,65, angl. *Pandora's box*) atd. Tímto způsobem se daly najít i některé ze starých a známých frazémů, které do té doby žádný slovník nezaregistroval, jako např. *pěnivý mok* (17,04), *bermudský trojúhelník* (16,02) či *chytrá horákyň* (15,42). Tento postup ovšem nezaručuje automaticky, že se jeho pomocí najdou pouze frazémy, ježto podobné MI-scores by se mohly najít i pro jiné zajímavé a důležité kolokace, které však je pak třeba odfiltrovat ručně, a to na základě introspekce a znalosti rodilého mluvčího. Protože se takto dají nalézt stejné hodnoty pro takové kolokace jako *diakritické znaménko* (17,41) a *kropící konev* (17,10), naznačuje se tu tím, že takovýto přístup lze zobecnit a napomocť pak nacházení i jiných druhů kolokací. A to je téma v korpusové lingvistice hojně diskutované.

Sestoupíme-li po dané škále níže a podíváme se například na MI-score 7, nacházíme kolokace a idiomy s adjektivy o značně vysoké frekvenci, jako např. *malá dušička* (7,202) či *silný kafe* (7,17).

Pokud se pustíme do ještě jemnějších rozdílů na úrovních MI-score, které jsme se rozhodli sledovat, tj. 7 (A), 8 (B), 10 (C), 12 (D) a 15 (E), a budeme-li přitom sledovat celkovou frekvenci kolokace vyšší než 10, spatříme aspoň část frazémů a sílu jejich zastoupení na každé ze škál těchto jednotlivých úrovní. Srovnáme výsledky pro

MI-score	Nalezené kolokace	Identifikované jako frazémy	Příklady
<b>A 7,0–7,9</b>	11 896	495	<i>ohnivá voda, dobrá vůle</i>
		Pozn.: jen 35 frazémů s frekvencí 10, ale 438 frazémů má frekvenci nižší než 100	
<b>B 8,0–8,9</b>	9 628	205	<i>širé pole, čisté svědomí</i>
		Pozn.: jen 14 frazémů s frekvencí 10, ale 184 frazémů s frekvencí nižší než 100	
<b>C 10,0–10,9</b>	5 745	250	<i>nekalé úmysly, drtivá většina</i>
		Pozn.: jen 11 frazémů s frekvencí 10, ale 214 frazémů s frekvencí nižší než 100	
<b>D 12,0–12,9</b>	2 984	140	<i>hořká příchuť, železná opona</i>
		Pozn.: jen 13 frazémů s frekvencí 10, ale 117 frazémů s frekvencí nižší než 100	
<b>E 15,0–15,9</b>	803	26	<i>dojná kráva, odrazový můstek</i>
		Pozn.: jen 1 frazém s frekvencí 10, ale 24 frazémů s frekvencí nižší než 100	

Je zjevné, že narůstající hodnoty MI-scoru dávají kvantitativně klesající výsledky a že nejvyšší výsledky pro

(A) **frazémy a idiomy** se nacházejí v daném případě kolem hodnoty 7. Stoupáme-li výše, a to jak po vlastní škále MI-scoru, tak i nad hodnotu 15, nacházíme nárůst jiných typů kolokací. Pod hranicí 7 se sice vyskytuje značná část frazémů taky, ta ale není podchytitelná statisticky, protože jde o kombinace frekventovaných adjektiv s frekventovanými substantivy, srov. např. frazémy se substantivem *hlava* v tabulce uvedené v závěru článku. Tímto způsobem splývají běžné kolokace s frazémy – srov. např. *horká hlava* s přímým významem „horečkou rozpálená hlava“ a „člověk jednající ukvapeně“ apod.

Frazém	Frekvence frazému	Frekvence adj.	Frekvence subst.	MI-score frazému
<i>horká hlava</i>	47	5 301	40 189	4,737
<i>moudrá hlava</i>	20	2 431	40 189	4,629
<i>hloupá hlava</i>	9	2 498	40 189	3,438
<i>prázdná hlava</i>	16	6 413	40 189	2,908
<i>čistá hlava</i>	28	13 727	40 189	2,617
<i>jasná hlava</i>	40	20 060	40 189	2,585
<i>tvrdá hlava</i>	16	11 429	40 189	2,074
<i>otevřená hlava</i>	5	12 203	40 189	0,302

Zatímco se tu (A) takto dá najít jen asi 2,5 % frazémů (jako je *kategorický imperativ*, MI 16,13, frekvence 52); další naproti tomu lze chápat jako

(B) **víceslovné termíny** jako je *lexikalizované přirovnání* (MI 16,89), na 56 %

(C) **víceslovná propria** jako je *Gibaltarský průliv* (MI 17,56), přes 20 %

(D) **běžné kolokace**, zbytek.

Postupujeme-li po škále MI-scoru obráceným směrem, zdá se, že tam převládají už běžné kolokace (D), a to jak systémové, tak textové (44,3 % a 37,5 %), srov. například *veřejná omluva* (MI 7,17 a frekvence 48). Jen kolem 2 % kolokací tu patří frazémům, jako například *osmý div* (MI 7,05 a frekvence 13). Statistické rozdělení výsledků do (A)–(E) tedy není a ani nemůže být zcela přímočaré a čisté, ale lze takto pro každý typ (a tedy i frazémy a idiomy) získat rozhodně více než jen pouhé jádro.

#### 4. Srovnání slovníku a korpusových dat

Když srovnáváme frazémy nalezené pomocí MI-scoru v ČNK (SYN2000) s frazémy uváděnými ve frazeologickém slovníku (SČFI2 1988), zdá se, že kolem poloviny z 600 fra-

zémů, které byly se strukturou adjektivum-substantivum pro zkoumání z tohoto slovníku zvoleny, má korpusovou frekvenci vyšší než 10. Tento slovníkový vzorek o 600 frazémách se však skládá z různých typů, přičemž největší (7–11 %) tvoří frazémy s MI-score mezi 10 a 4. Avšak ty frazémy, jejichž MI-score je nižší než 7, se statisticky už překrývají s běžnými kolokacemi, srov. *celkový dojem* (MI-score 6,84) a naproti tomu *pocitivá hra* (5,58) nebo *šťastný konec* (4,60). Zkoumání této části MI-scoru se tudíž stává obtížnějším. Takto se potvrzuje, že heuristická volba pásma MI-scoru mezi 18 a 7 je pro danou strukturu vhodná.

Takovýmito korpusovými rešeršemi se dospělo k cca 1200 kolokacím v rozsahu MI-scoru 8–17, které byly vytipovány jako frazémy. Z nich SČFI2 zachycuje celkem 318, tj. asi jedna čtvrtina. Statisticky jsou takto dobře zachytitelné frazémy se

- a) **sémantickou anomálií**, např. *křiklavá barva* (MI-score 9.19, frekvence frazému 32, frekvence adj. *křiklavý* 418, frekvence substantiva *barva* 15 811) nebo *bílá vrána* (MI-score 9.56, frekvence frazému 69, frekvence adjektiva *bílý* 25 046, frekvence substantiva *vrána* 441), zpravidla mají MI-scoru pod hranicí 10;
- b) **formální anomálií**, např. *pěkné nadělení* (MI-score 11.63, frekvence frazému 10, frekvence adjektiva *pěkný* 6147, frekvence substantiva *nadělení* 62) nebo *ranní kuropění* (MI-score 13.30, frekvence frazému 13, frekvence adjektiva *ranní* 3382, frekvence substantiva *kuropění* 46), zpravidla mají MI-scoru vyšší než 10. Formální anomálii se tu pro jednoduchost míní anomálie, nepravidelnost ve vzájemném souvyskytu.

Problém představují frazémy utvořené z kolokací, které mohou mít jak přímý, tak frazeologický (metaforický) význam (typ *mýdlová bublina*). Uvádíme 10 frazémů tohoto typu (abecední řazení podle substantiva):

frazém	MI-score	frekvence frazému	frekvence adj.	frekvence subst.
<i>mrtvý brouk</i>	10,33	99	10 288	899
<i>tlustá čára</i>	10,34	158	3 341	4 396
<i>horká kaše</i>	10,44	20	68	1 012
<i>bílý límeček</i>	10,51	88	25 046	291
<i>vedlejší kolej</i>	10,51	81	3 824	3 906
<i>zmenšené měřítko</i>	10,12	11	341	3 495
<i>široká paleta</i>	10,34	165	15 181	957
<i>zdvížený prst</i>	10,55	36	343	8 412
<i>ledová sprcha</i>	10,92	40	2 127	1 169
<i>stříbrný tác</i>	10,61	16	200	1 750

U tohoto typu frazémů nezdídky dochází k tomu, že jejich komponenty mají srovnatelné frekvence, srov. *tlustá čára*, *vedlejší kolej* nebo *ledová sprcha*. Statisticky proto zůstávají na úrovni běžných kolokací typu *bílé auto* (celková frekvence 11, frekvence komponentů ale je *bílý* – 25 046, *auto* – 21 154).

## 5. Srovnání s angličtinou

K ověření těchto zjištění a výsledků získaných pro češtinu se zkoumala i angličtina, jakoliv jen v omezené sondě (20) srovnatelných, resp. identických frazémů v British National Corpus, který je rozsahem stejný jako ČNK, resp. SYN2000. Obecně a s odhlédnutím od specifických rozdílů je zřejmé, že výsledky jsou podobné a že daný přístup lze uplatnit i na angličtinu, srov. český *stříbrný táč* (MI-score 10,618 a frekvence 16) a odpovídající anglický frazém *silver platter* (MI-score 11,263 a frekvence 14).

Některé z rozdílů jsou zřejmé a jsou dané typologickými rozdíly mezi oběma jazyky, jako je například substantivum v angličtině užívané tam, kde čeština má plnohodnotné adjektivum (*lion* vs. *lví*, srov. *lion's share* a *lví podíl*). Jimi je pochopitelně ovlivněna i frekvence komponentů dané struktury. Téměř totožné výsledky lze však pozorovat u takových klasických frazémů jako *gordický uzel* a *Gordian knot* a *trojský kůň* a *Trojan horse*, kde se i v angličtině užívá adjektivum.

Těchto dvacet českých frazémů a jejich anglické ekvivalenty, ověřované v ČNK a BNC, dává následující tabulku:

ADJ	NOUN	MI	Idiom freq.	Adj. freq.	Noun freq.	ADJ	NOUN	MI	Idiom freq.	Adj. freq.	Noun freq.
křišťálový	koule	12,253	69	631	2 708	crystal	ball	8,709	66	1 997	8 788
fackovací	panák	17,41	44	55	555	whipping	boy	10,432	14	55	20 494
krokodýlí	slzy	12,513	12	70	3 545	crocodile	tears	10,072	26	376	7 148
dominový	efekt	13,942	39	58	5 167	domino	effect	8,207	22	239	34 649
mrtvý	bod	6,134	198	10 347	32 939	dead	end	4,268	123	11 794	60 204
velká	ryba	3,821	163	181 776	7 674	big	fish	5,342	103	24 853	11 367
železná	pěst	7,763	23	6251	2 048	iron	fist	8,421	24	5 190	1 501
darovaný	kůň	9,256	16	270	11 720	gift	horse	5,307	20	4 581	12 267
první	pohled	6,977	3 504	138 317	24 309	first	sight	6,255	597	116 940	7 434
pokusný	králík	12,549	55	595	1 865	guinea	pig	13,219	130	617	2 458

ADJ	NOUN	MI	Idiom freq.	Adj. freq.	Noun freq.	ADJ	NOUN	MI	Idiom freq.	Adj. freq.	Noun freq.
stará	vesta	5,627	21	58 579	877	old	hat	4,577	42	52 755	3 710
chladná	hlava	5,467	49	3 333	40 189	cool	head	3,421	20	4 569	45 455
zlomené	srdce	10,932	319	1 340	14 732	broken	heart	7,580	70	2756	14 766
trojský	kůň	10,894	95	515	11 720	Trojan	horse	11,217	37	141	12 267
slonovinová	věž	11,438	7	56	5 448	ivory	tower	11,308	82	826	4 355
gordický	uzel	16,060	66	76	1 537	Gordian	knot	15,121	9	20	1 405
lví	podíl	10,962	149	458	19 724	lion's	share	3,773	7	2 084	27 337
kostelní	myš	11,714	32	592	1 946	church	mouse	4,629	14	23 474	2 681
stříbrný	tác	10,618	16	4 494	274	silver	platter	11,263	14	3 987	159
horké	křeslo	6,077	19	5 301	6 421	hot	seat	5,858	53	8 827	11 515

Poznámka: V tabulce se frazémy třídí podle MI-scorů získaných pro češtinu.

## 6. Závěry

Zdá se, že **většina frazémů má v ČNK absolutní frekvenci nižší než 100** a že ty, jejichž frekvence je velmi nízká (1–3), zastarávají a přestávají se užívat. Důležité však je si uvědomit, že **většina frazémů se** zřejmě někde na škále hodnot MI-scoru **shlukuje**, jakkoliv stále nevíme, jaká část jich zůstává mimo.

Ačkoliv je stále do jisté míry nutné výsledky získané automaticky ověřit manuálně, přesto se naznačený přístup zdá plodný. Kromě toho, že takto získáváme tolik potřebné numerické a frekvenční rysy umožňující nám frazémy srovnávat na solidnějším základu (toto zde lze pro nedostatek místa jen naznačit), může to být i krok k **automatickému objevování frazémů** (discovery procedure) umožňujícímu nacházet v jazykovém korpusu nové frazémy, dosud nezaznamenané. S ohledem na škálovitou povahu získaných výsledků naskytá se tu i bohatá možnost zkoumat i široké pole přechodných jevů, které sahají od individuálních metafor k frazémům, od frazémů k termínům a propriím a od frazémů k běžným kolokacím. A zřejmě jen na základě takového přístupu, který pojme obrovská a objektivní data, lze **hledat i kritéria pro odlišení frazému od nefrazému**.

Jedno z omezení tohoto přístupu založeného na MI-scoru je to, že zřejmě nepochycuje frazémy složené jen z **frekventovaných komponentů**, jako je *tvrdá hlava*, kde jak *tvrdý* tak *hlava* patří k vysoce frekventovaným slovům. MI-score tohoto frazému je velmi nízký, pouze 2,07. Obecně platí, že frazémy s tak častými komponenty, jako je *hlava* (ta má v ČNK

absolutní frekvenci víc než 40 000), se statisticky liší podle měnící se frekvence doprovodného adjektiva. Takto MI-score frazému *korunovaná hlava* (8,80) odpovídá nižší frekvenci (208), zatímco *čistá hlava* je téměř 14 000. Podrobnosti viz v příloze u frazému s *hlavou* a jejich MI-scorů.

Konečně je vzhledem k odlišnému a často mnohem vyššímu úzu určitých frazémů v mluveném jazyce třeba znovu zdůraznit, že tento přístup se omezuje zatím jen na psaný jazyk.

Na obecnější úrovni je zjevné, že kvůli prohledávání složitějších struktur (nejen frazeologických) bude třeba složitějšího přístupu. I tady bude při ověřování a vyhodnocování výsledků získaných automaticky role člověka, resp. rodilého mluvčího značná a nedá se zatím obejít. Avšak značně velký korpus je sám o sobě už jistou zárukou, že z něj získáme objektivní obraz existujících frazémů. Daný přístup je zřejmě první, který nabízí (téměř) **vyčerpávající pokrytí dané oblasti** a který nám dává i jistou záruku, že z našich dat i zkoumaného pole se nevynechá téměř nic. Jde tu o pocit, o který lexikografové a lingvisté usilují dávno, totiž být si jistý, že všechna relevantní data jsou pohromadě. Samozřejmě ani tak velký korpus nemusí obsahovat všechno, ale korpusy se zvětšují téměř denně.

### Poznámka

Tento příspěvek je upravená a rozvedená verze anglicky psaného příspěvku Čermák *Statistical Methods for Searching Idioms in Text Corpora*, na kterém zvláště v přípravě dat spolupracovala Renata Blatná. Za pomoc s přípravou podkladů autor děkuje i M. Křenovi.

## Příloha

Adjektivně-substantivní frazémy s komponentem *hlava* (head), jejichž MI-score je obvykle nižší než 7:

Idiom	MI-score	Frequency	English Equivalent
<i>korunovaná hlava</i>	8,809	31	crowned head
<i>kočičí hlavy</i>	8,046	47	cobble-stones
<i>skopová hlava</i>	7,211	10	muttonhead
<i>bystrá hlava</i>	6,508	22	quick/ready wit
<i>chytrá hlava</i>	5,982	52	clever head
<i>rozpálená hlava</i>	5,465	9	hot head
<i>telecí hlava</i>	5,458	9	litt. calf head, i.e. idiot
<i>dutá hlava</i>	5,184	8	lackbrain



Idiom	MI-score	Frequency	English Equivalent
<i>děravá hlava</i>	4,889	4	a memory like a sieve
<i>horká hlava</i>	4,737	47	hothead
<i>moudrá hlava</i>	4,629	20	wiseacre
<i>hloupá hlava</i>	3,438	9	sccaterbrain
<i>prázdná hlava</i>	2,908	16	empty vessel
<i>čistá hlava</i>	2,617	28	a clear head
<i>jasná hlava</i>	2,585	40	a clear head
<i>tvrdá hlava</i>	2,074	16	pig-headed fellow
<i>otevřená hlava</i>	0,302	5	good head

---

## Summary

The paper that largely overlaps with this is *Statistical Methods for Searching Idioms in Corpora*, elsewhere in this volume.

# Funkce frazémů a idiomů

(In: *Česká slavistika 1993*. České přednášky  
pro XI. mezinárodní sjezd slavistů Bratislava 1993, s. 331–336)

## 1. Funkce obecně a ve frazeologii

V Evropě se pojetí funkce vyvíjelo především v rámci dvou strukturalistických škol, pražské a kodaňské (nověji ovšem i jinde, zvl. v systemické lingvistice, spojené se jménem M. A. K. Hallidaye). Kodaňské pojetí funkce jako vztahu závislosti mezi dvěma funkcivami, entitami, jednotkami (Hjelmslev 1972) má blízko k pojetí čistě matematickému; obecně se tedy soustřeďuje na vztah mezi různými prvky, entitami či jednotkami. Naproti tomu pražské pojetí, které se soustřeďuje spíše na konkrétní prvek či jednotku a na jejich vztah k jiným, je takto formalistické jen v první ze svých dvou linií, v níž se uvažuje mj. o distinktivní funkci rozlišující jednotky (téhož paradigmatu), popř. o delimitativní funkci vymezující hranici mezi jednotkami (téhož syntagmatu).

Ve své druhé linii, ve které funkční trichotomii K. Bühlera (1934) rozvíjí R. Jakobson (1960) do systému svých šesti funkcí, tj. **expresivní, konativní, fatické, referenční, poetické a metajazykové**, se funkce pojímá jako funkce sdělení v komunikativním rámci, a přechází se tak od funkce jednotky až k funkci celého sdělení (message). O propojení obou linií se dosud nikdo nepokusil; přesto se nabízí otázka, zda i pro „jednotku“ druhé linie, tj. sdělení, nelze uvažovat, podobně jako pro foném a slovo, určitou analogii distinktivní funkce, protože šest Jakobsonových funkcí tvoří přece také určité paradigma. Jednou z prvních překážek, na které při takovém pokusu narazíme, je vágnost jak hranic jednotlivých funkcí, mezi nimiž je mnoho přechodů, tak podmínek jejich zapojení, realizace.

V **pražském** pojetí se však funkce chápe ještě z jiného aspektu (Skalička 1948, nověji i Daneš 1987), tj. zdůrazňuje se zvláště její instrumentálnost, resp. pojetí jako **role či úlohy jednotky ve struktuře**, která je zpravidla typově vyššího řádu, v níž pak uvažovaná jednotka má uplatnění jako komponent. Toto pojetí funkce jako role, které je nutně volnější než její pojetí jako závislosti, je ovšem vlastně jen zobecněním pojetí funkce distinktivní, resp. delimitativní a Jakobsonových šesti funkcí:

u distinktivní funkce lze roli chápat místo v mikroparadigmatu,

u delimitativní místo v syntagmatickém řetězu a

u Jakobsonových funkcí jejich místo v celku dané konkrétní komunikace, tedy vždy místo a roli jednotky, entity v širším rámci. Mezi funkcí fonému a celého sdělení je ovšem široký

diapazon jednotek dalších úrovní, a ty mají své funkce v daném smyslu také. Nověji se jim obvykle říká **strukturní funkce**.

Protože jazyková jednotka má jakožto znak vztah i vně, ke svému denotátu, je třeba uvažovat i její funkci **sémiotickou** (Hjelmslev o ní mluví jen jako o závislosti mezi obsahem a výrazem, tj. uvnitř jazyka). A protože je tato funkce základem jazykového pojmenování (nominace), lze ji nazývat funkcí **nominativní**.

V odlišném, nejazykovém smyslu se konečně někdy mluví ještě o **fregeovské či kompoziční funkci** jednotky, když se charakterizuje pravidelné skládání, slučování jednotek v komplexní významový útvar.

Je zřejmé, že u frazémů a idiomů lze uvažovat o stejných třech typech funkce, byť v dosti odlišném smyslu. Protože pozornost věnovaná této oblasti z hlediska funkce je dosud velmi malá (srov. Čermák 1985 a Filipec – Čermák 1985), budou mít následující poznámky a pokusy o funkční klasifikaci především povahu diskusní. Třebaže v rámci tohoto příspěvku nelze všechny aspekty, s nimiž funkce souvisí, popř. na které naráží, ani integrovat ani řešit, lze na ně aspoň poukázat, a to především ve smyslu potřeby určitého řešení. První věcí je potřeba precizace a explicitní kvalifikace **povahy funkčního vztahu** (vztah, závislost, role ap.), která je nejasná, resp. různá už v obecné teorii. Druhou věcí je potřeba precizace a explicitního stanovení **podmínek, kdy k realizaci funkčního vztahu dochází**; na úrovni Jakobsonových úvah o komunikativní funkci se jen velmi vágně naznačuje textově-literární, popř. psychologická motivace, nikoliv tedy podmíněnost jazyková. Třetí věcí je potřeba zjistit, **zda je systémová funkce jednotek**, v našem případě frazémů a idiomů, nějak **signalizována formálně**; jako zjevný prostředek takové signalizace, nikoliv však jednoznačný, se nabízí valence. Konečně čtvrtou věcí, která může souviset s kompoziční funkcí, ale také nemusí, je **potřeba zapojení pragmatické složky funkce** do celkového rámce; je jednou z mála konstant, které se ve frazeologii nabízejí.

## 2. Funkce kompoziční

Při tvorbě, tj. skládání, kombinaci významových jednotek ve větší a vyšší celky, platí všude v jazyce, že se významové komponenty v takovou vyšší strukturu skládají v zásadě pravidelně (v rámci fregeovské funkce) a že význam takové vyšší struktury je sumou významů svých komponentů, i když naproti tomu podle principu Gestalt není útvar, struktura pouhým mechanickým souhrnem komponentů. **Žádný frazém či idiom na druhé straně pravidelnou kompoziční funkci mezi svými komponenty nemá ani ze sémantického ani strukturního hlediska**, a je tudíž z tohoto hlediska (a nejen něho) výrazně **anomální**, srov. takto např. frazém *hodit něco za hlavu*, který není ani sémanticky ani syntakticky prostým úhrnem komponentů *hodit + za + hlava* (a podobně *hlava* tu není objekt). Pokud komponenty nějakou funkci z hlediska sémantického v celku frazému mají, je nezřetelná a nástroje k jejímu popisu dosud chybí; různé úvahy o částečné motivaci frazému skrze některé komponenty jsou jen vágní a nepřesvědčivé (srov. Čermák 1987). Mezi komponenty idiomu, resp. frazému kompoziční, **fregeovská funkce tudíž není**, lze nanejvýš v některých případech mluvit o blíže nepoznané sémantické participaci komponentu ve významu celku. Ve

smyslu teorie informace dochází tedy při přechodu od komponnetu k celku k maximální entropii.

### 3. Funkce strukturní

Strukturní funkci, kterou lze chápat jako roli frazému či idiomu ve vyšší jednotce a která je obvykle stabilní, má každý typ frazému pochopitelně jinou, pro frazémy a idiomy kolokační, tj. nevětné povahy, je to funkce v zásadě v rámci věty, resp. propozice, pro frazémy a idiomy propoziční je to funkce obvykle v rámci promluvy. Dodejme, že vedle tohoto aspektu role ve struktuře je však vhodné uvažovat funkci ještě z užšího druhého aspektu konkrétního syntagmatického vztahu k jednotce, se kterou se frazém či idiom kombinuje (a kterou obvykle kvalifikuje ap.). Takto se zde pak kombinuje pojetí funkce pražské a kodaňské školy, ba jeví se jako komplementární.

Pro frazémy kolokační (I) lze podle druhého aspektu, který je deskriptivnější, lišit strukturní funkci ve vztahu ke (A) kolokacím (nevětným kombinacím lexémů), ale i (B) propozicím (větám), naopak pro frazémy propoziční (II) lze uvažovat strukturní funkci jen podle prvního aspektu, a to z hlediska tradičních modů věty. Možnosti klasifikace propozičních frazémů podle druhého, tj. kombinatorického aspektu, se zdají málo plodné.

Lze tak, tj. podle druhu funkčního vztahu, dospět (srov. Čermák 1985, Filipec – Čermák 1985) **pro frazémy kolokační** k většinovému funkčnímu typu adkolokačnímu (IA) a menšinovému typu adpropozičnímu (IB). Oba typy se blíže člení na subtypy podle své syntaktické funkce, a to zvl. na

- (IA) **adkolokační** *verbální* (*chytat lelky, hodit něco za hlavu*),  
*nominální* (*zlatý důl*),  
*modifikační* (*neslaný nemastný, široko daleko*),  
*relační* (*na úkor něčeho*) a  
*substituční* (*co... za?*, *čertví kdo*) a na
- (IB) **adpropoziční** *modifikační* (*pro jednou*) a  
*relační* (*že by, co – to*) aj.

K ještě užší strukturně-funkční klasifikaci lze pak dospět uvnitř každého subtypu, sledujeme-li jednotlivé kombinatorické možnosti komponentů a tedy jejich kombinatorické modely.

**Pro frazémy propoziční** lze naproti tomu spatřovat strukturní funkci v nutnosti mít jen určitý modus, podle něhož lze lišit frazémy

- (1) *indikativní* (*To je stará písnička!*),  
(2) *imperativně-vokativní* (*Ruku na srdce!*) a  
(3) *interogativní* (*A já jsem vosk?*),  
a to vždy jak v pozitivní, tak negativní podobě.

Obecně platí, že mezi různými typy frazémů jsou rozdíly i co do počtu funkcí, které mohou mít; zatímco např. frazémy verbální jsou podle prvního aspektu ve větě obvykle jen

predikátem (rozvitým), ale podle druhého, kombinatorického aspektu se mohou spojovat jak s nominálními, tak adverbialními částmi, členy věty, je tomu u frazémů nominálního typu téměř naopak: ve větě, tj. podle prvního aspektu, mohou mít více „nominálních“ rolí (ale ovšem také nemusejí), avšak podle druhého aspektu se zpravidla kombinují jen s predikátovou částí věty. Tyto kombinace s jinými jednotkami, stejně jako povaha širší, resp. vyšší jednotky, jsou však striktně nefrazeologické; strukturní funkce frazémů a idiomů se tudíž zásadně vztahuje k nefrazeologickému okolí, a je tedy prakticky (až na řídké zvláštní případy) vyloučeno, aby se běžně kombinovaly dva frazémy navzájem.

Obligatornost, pevnost funkce je ovšem daná územ, a je tudíž proměnitelná, jakkoliv ne nijak snadno. Pokud úzus určitou proměnu nesankcionizuje, vyvolává vždy v posluchači dojem improvizovanosti, srov. neustálenou transformaci *hození něčeho za hlavu* (k frazému *hodit něco za hlavu*) a tedy i změnu funkce z verbální na nominální. **Povaha funkce je tudíž výrazně stálá** a případ jako vícerost nominální funkce, zmíněný výše, neznamená její oslabení, ale pluralitu.

#### 4. Funkce nominativní

Chápeme-li pod funkcí strukturní roli v určité jazykové struktuře, resp. relaci ke specifické složce této struktury vyššího řádu, je pod funkcí nominativní třeba chápat navenek spíše jednoznačnou relaci, popř. stabilní roli než roli proměnlivou, tj. k vnějšímu **denotátu**. Protože však lze výrazy jako *kočičí hlavy* ve smyslu „velké a nerovné dlažební kameny“, tj. výrazy s fyzickým denotátem konkrétní povahy, chápat jen jako periferní, resp. kvazifrazémy, je třeba pojetí **atributu** „vnější“ ve vztahu k denotátu chápat především jako „**abstraktní, nemateriální či myšlený**“; nelze ovšem ani takto omezený denotát, který je univerzálně charakteristický pro všechny frazémy, jen prostě zjednodušeně ztotožnit s významem frazému, je tu zásadní rozdíl.

Analogicky funkci strukturní lze pro obě roviny frazémů pak uvažovat nominativní funkci různého typu, která se rozpadá obecně do dvou složek, **deznativní**, založené na nocio-nálním, pojmovém vztahu k denotátu, a **pragmatické**, založené na (subjektivním) vztahu uživatele k frazému i cíli, partnerovi promluvy ap. Protože deznativní složka je charakteristicky u frazému vágní a co do svých kontur a hranic jen obtížně vymežitelná, zdá se, že univerzální přítomnost složky pragmatické, a to obvykle v její **evaluativní**, hodnotící podobě (pozitivní nebo častěji negativní), je zřejmě potřebná a nikterak tedy náhodná; jsou navzájem komplementární. Ne náhodou jsou z tohoto pragmatického, resp. evaluativního hlediska právě frazémy nejvýraznější a nejbohatší částí slovníku na úrovni složitějších autosémantických lexémů a propozic (tj. vedle jednodušších autosémantických podtříd adjektiv či adverbíí a partikulí, pokud jde o evaluativa synsémantická). Dodejme, že však relativně **snadná aktualizovatelnost** a také denotativní vágnost typického frazému, zesilovaná jeho pragmatickou složkou, odlišnost pojetí funkce jakožto pevného vztahu vlastně poněkud oslabuje a jakoby přibližuje povaze funkce strukturní pojeté jako role.

Pro úroveň **kolokačních frazémů** (I) lze uvažovat, v zásadě analogicky k povaze slovních druhů, dvě široké funkční třídy frazémů. I když mezi nimi ostré hranice nejsou, lze tu spatřovat aspoň přibližná analoga autosémantik a synsémantik, tj.

- (A) **deznátory**, které denotát primárně označují, a  
 (B) **operátory**, které deznátory (byť ne frazeologické, ale z běžného, pravidelného jazyka) spojují, nahrazují, komentují či na ně ukazují ap. (Weinreich 1966).

Pro úroveň **frazémů propozičních** (II) lze pak uvažovat funkční typ

- (A) **vokativní**, vyjadřující vztah mluvčího a posluchače,  
 (B) **tematický**, vyjadřující vztah mluvčího k tématu,  
 (C) **mediální**, vyjadřující vztah mluvčího k médiu, resp. prostředku komunikace, a aspoň ještě typ  
 (D) **situační**, vyjadřující vztah k situaci ap.

V takto daném rámci můžeme pak blíže jednotlivé funkční třídy konkretizovat pro kolokační frazémy na základě kvazilogického, a to u **deznátorů** zvláště na podtřídě

- (IA) **argumenty** (*Ze starého bicyklu se stalo jablko sváru*),  
**predikátory** (*Po chvíli zapírání musel kápnout božskou*),  
**kvalifikátory** (*Andělsky trpělivá babička to zkoušela znovu a znovu*) či  
**cirkumstativa** (*Široko daleko nebyla jediná hospoda*). Pro  
 (IB) **operátory** lze postulovat v praxi aspoň podtřídě různého typu  
**komentorů** (co do úspěšnosti, zaujetí, pravdivosti ap., např. *Přednáška byla k zbláznění nudná*),  
**deiktik** (*Nějakou tu korunu ti půjčit můžu*),  
**kvantifikátorů** (*Přispět teď musí jeden každý z vás*) a  
**relátorů** (*S ohledem na nedostatek času se to musí zkrátit*).

Pro frazémy propoziční, pro něž je základním kritériem, jímž se funkce specifikuje, **druh vztahu mluvčího k různým aspektům frazému v rámci celkové komunikace**, lze analogicky určovat konkrétní podtřídě také, i když bližší celkové zmapování oblasti propoziční dosud chybí. Takto vyjadřují různé podtřídě

- (IIA) vokativního typu např. frazémy *To mi může být ukradený!*, *Koukej mazat!*, *Tak se měj!* a  
 (B) tematického typu pak frazémy jako *Vzduch je čistý.*, *To zrovna!*, *Jak se věci mají...* ap. Řídký  
 (C) funkční typ mediální lze ilustrovat frazémem *Div jsem si na tom nezlomil jazyk.*, *Jak bych to řekl...* a  
 (D) typ situační frazémem *Den se dlouží.*, *Teď anebo nikdy!* ap.

Je zřejmé, že v důsledku své nevětné povahy není možné frazémy kolokační, tj. deznátory a operátory s Jakobsonovými funkcemi korelovat přímo; vyskytují se nespecificky ve funkcích všech. Takováto **korelace s komunikativními funkcemi**, i když jen volná, je **naproti tomu možná u frazémů propozičních**. V různém poměru takto plní frazémy **vokativního typu** především **funkci emotivní a konativní**, resp. **expresivní a apelovou**, zatímco frazémy **tematického, popř. situačního typu** mají nejčastěji **funkci referenční** a konečně **frazémy typu mediálního** vyjadřují monopolně **funkci metajazykovou**. Nevyhraněný charakter má co do své distribuce funkce **fatická a poetická**: tu první lze najít u frazémů v první řadě vokativních a tematických, tu druhou nevyhraněně pak u všech typů frazémů.

Obecně je třeba ovšem zdůraznit, že **funkční povaha velké většiny frazémů je smíšená** a pokud se tu o ní mluví jako o takové či takové, znamená to jen tolik, že se z funkčně smíšené povahy jednostranně zdůrazňuje pouze její dominantní složka.

Protože i teorie mluvních aktů je svou podstatou funkční a bude užitečné ji integrovat s funkčním rámcem Jakobsona, pokusme se zde aspoň ve stručné poznámce poukázat na některé souvislosti s frazeologií propozičního typu; pro frazémy kolokační je tato souvislost málo jasná.

Zdá se, že jako **lokuční akty** funguje velká většina frazémů různých typů, která prostě jen něco konstatuje, resp. pojmenovává, a situace je zde tedy stejná jako u běžného, pravidelného jazyka. Čistě **ilokučních aktů** je naproti tomu zřejmě ve frazeologii velmi málo, i když **aktů s ilokuční silou** lze nalézt velké množství. Výrazná je tu skupina frazémů charakteristická svou formou imperativní (a tedy zvláště imperativního typu) a svou formou žádosti, rady, hrozby ap., která má zřetelnou ilokuční sílu, protože obvykle vede k zamýšlenému účinku, srov. *Nenechte se prosit!, Dej si bacha!*. Jako **perlokuční akty** můžou fungovat typy méně formálně vyhraněné, ale jen příležitostně, nikoliv obvykle, tj. nikoliv do té míry, aby se u nich mohlo mluvit o pevné funkční souvislosti, srov. frazém *To je, jako když blecha štípně* jako prosté informativní konstatování (což je obvyklé a není perlokuční povahy) a jako konstatování s následnou reakcí např. ztráty strachu a odhodlání se přijmout injekci (což je pak perlokuční akt, srov. např. Allan 1986). Zřejmé pole pro další výzkum tu tedy skýtá možnost studia častých typů ilokuční síly u různých typů frazémů a jejich korelace s jednotlivými funkčními typy.

## 5. Závěrečné poznámky

Výše zmíněná anomálie je ve frazeologii přítomná v mnoha různých manifestacích paradigmatického či syntagmatického druhu (u kompoziční funkce jsme uvažovali zatím jen o jednom syntagmatickém typu anomálie). Obecně ji je možné zjišťovat ve všech **třech obecných směrech**, které jsme uvažovali pro funkci, tj.

„**uvnitř**“ **frazému** v různých vztazích komponentů navzájem a vůči celku (ilustrována a pojmenována tu byla jen kompoziční funkce),

„**navenek**“ **frazému**, kde ji lze spatřovat především v přítomnosti pragmatické složky v rámci nominativní funkce, a konečně i

v jeho „**úzu**“.

V tomto posledním směru, pokrývaném zde co do funkce poznámkami o komunikační funkci, se anomálie výrazně projevuje především v mnoha různých **transformačních defektech** frazémů, ilustrovaných už mnohokrát dříve.

Nabízí se tudíž otázka, zda tento výrazný **paralelismus výskytu různých typů anomálie a funkce** je možné a užitečné dále propracovávat. Domněnka, že povaha funkce je takovouto anomálií ovlivňována a specificky modifikována, je ovšem nasnadě.

Dodejme závěrem, že právě tato myšlenka se stala podkladem návrhu (Čermák 1985, s. 69–70 a Filipec – Čermák 1985, s. 235) vzít pro **základ** komplexní, funkční a vůbec

nejdetailnější **klasifikace frazémů**, která by byla založena objektivně na jejich skutečném chování, právě různé **typy projevu transformačních anomálií** v širokém slova smyslu.

## Summary

### Function of Idioms

Most major and basic functions of idioms are discussed here. First, some attention is paid to the non-compositional or non-Fregean function denying presence of the compositional function between its constituents. Thus, it is not possible to find it in an idiom such as *hodit něco za hlavu*, which is not semantically and functionally equal to a mere sum of its constituent parts, namely *hodit + za + hlava*; hence the idiom exhibits a pronounced anomaly, one of many, typical for idioms in general.

The structural function, based on the “syntactic” role of the idiom in a larger structure, allows for a classification into functional types of idioms based on it, where collocational idioms might be viewed as

(IA) **adcollocational** *verbal* (*chytat leky, hodit něco za hlavu*),  
*nominal* (*zlatý důl*),  
*modifying* (*neslaný nemastný, široko daleko*),  
*relational* (*na úkor něčeho*)  
*substitutive* (*co... za?, čertví kdo*)

(IB) **adpropositional** *modifying* (*pro jednou*)  
*relational* (*že by, co – to*) etc.

On the other hand, propositional idioms allow for a different view of idiom types, a simple one being at least

- (1) *indicative* (*To je stará písnička!*),
- (2) *imperative-vocative* (*Ruku na srdce!*),
- (3) *interrogative* (*A já jsem vosk?*),

all of them existing in both a positive and (much rarer) negative form.

It has to be stressed that this structural function is basically stable.

Nominative function refers to the type of denotate, which is mostly of an abstract nature in idioms. On the level of collocational idioms, two major functional types of idioms may be seen, designators and operators. While the former designate the idiom's denotate directly, the latter link, replace, point to or comment the former.

Further, both may be subdivided into classes of

(IA) **designators:**

*arguments* (*Ze starého bicyklu se stalo jablko sváru*),  
*predicators* (*Po chvíli zapírání musel kápnout božskou*),  
*qualifiers* (*Andělsky trpělivá babička to zkoušela znovu a znovu*)  
*circumstatives* (*Široko daleko nebyla jediná hospoda*).

(IB) **operators:**

*commentors* (as to do success, involve, truth etc. *Přednáška byla k zbláznění nudná*),  
*deictics* (*Nějakou tu korunu ti půjčit můžu*),  
*quantifiers* (*Přispět teď musí jeden každý z vás*)  
*relators* (*S ohledem na nedostatek času se to musí zkrátit*).



Propositional idioms (II) suggest the following functional classes

- (A) **vocative**, expressing the speaker-listener relationship,
- (B) **thematic**, expressing the speaker-topic relationship,
- (C) **medial**, expressing the speaker-medium (means of communication) relationship,
- (D) **situational**, expressing the speaker-situation relationship etc.

Illustrations of II these may be seen in

- (A) *To mi může být ukradený!, Koukej mazat!, Tak se měj!*
- (B) *Vzduch je čistý., To zrovna! Jak se věci mají...*
- (C) *Div jsem si na tom nezlomil jazyk., Jak bych to řek.,*
- (D) *Den se dlouží., Ted' anebo nikdy!*

It is evident that these classes rely heavily on the Prague school tradition and R. Jakobson's system of communicative functions. It can easily be shown, then, that there is a link between (A) and Jakobson's emotive and connative or Appell function, etc. Nature of most idioms is, however, mixed from this point of view, idioms being thus blends of various functions in most cases.

# Termín a frazém. Případ překrývání a periferie dvou nominativních oblastí

(In: *Termina 2000*. Sborník příspěvků z II. konference 1996  
a III. konference 2000, s. 31–36)

## 1. Úvod

Žádná jazyková oblast není zcela přesně co do svých hranic vymezena, i když přesahy a překrývání s oblastmi jinými bývají různě rozsáhlé a zřetelné: u gramatických oblastí a disciplín je obecně menší, u nominativních, resp. lexikálních oblastí, na kterou se teď přenášší naše pozornost, však obecně větší. Ke známým případům, u nás notoricky neřešeným, patří především v důsledku nedobrého pojetí spisovnosti uměle kladená hranice mezi termíny vlastní a to, co se v češtině poměrně unikátně nazývá **slang**, což **je** však bytostně **terminologie také**. V méně zřejmých směrech lze však periferii terminologie a tedy i překrývání s jinými jevy hledat u leckterého pro termín postulovaného rysu dalšího. Zajímavé by bylo v tomto smyslu prošetření např. vztahu exaktnosti a nociónálnosti a povahy některých názvů filozofických (i když lze namítnout, že sami filozofové svou disciplínu za vědu nepovažují), terminologie a onomastiky (překrývají se např. v astronomii u názvů pohyblivých umělých nebeských objektů) či, a to je otázka velmi aktuální, přesné souvztažnosti stránky terminologické a neterminologické u valné části těch nejobyčejnějších, všedních substantiv, jako je *chléb* či *voda* (ty jsou přece pro pekaře či chemika termíny také a stojí za nimi definice).

Právě absence bližšího poznání v této poslední oblasti výrazně brání lexikografovi v dosažení nesubjektivní možnosti položení **hranice mezi termín a netermín** ve slovnících. Ne všechny stránky periferie terminologie představují ovšem přechod do disciplíny či oblasti jiné, některé tvoří naopak, především z důvodů vývojových a frekvenčních, jakési předstupně i postfáze do neterminologie, popř. do nebytí. Řešení otázek **periferie terminologie** a zapojení mj. principu škály je bezesporu jedním z velkých mank této disciplíny.

Jednou z takových závažných oblastí, definitoricky z termínu vylučovaných, je přítomnost metaforičnosti, resp. přenesenosti významu apod. Setkáme se s ní porůznu leckde a prakticky ve všech oborech je tato periferní část terminologie shodná s přechodem terminologie do frazeologie a idiomatiky. Tam se takové výrazy jako *kuří oko* (v lékařství), *rajské jablko* (v botanice, resp. zelinářství), *zlatý déšť* (v botanice, resp. zahradnictví), *vličí zuby* (v technice, u pily), *bludný kámen* (v geologii), *sladké dřevo* (v hudbě, kytara), *bílá kniha* (v politice aj.) či *valná hromada* (v bankovníctví, podnikání aj.) mohou nazývat **kvazifrazémy**, resp. **kvaziidiomy** (viz mj. Čermák 1982). Nemusí přitom jít jen o takovéto poměrně

jednoduché případy, do hry překrývání a periferního oslabování platnosti můžou přistupovat i faktory další, např. slang, popř. argot a jeho expresivita či tabuovost, kombinovaná dokonce s pseudoonomastičností, srov. příklady ze Suka (1993) a vězeňského slangu: *Mařenka Hrstková* (onanie), *medvědí ucho* (ženský pohlavní orgán), *zelená láska* (marihuana) aj.

Je-li tu obecně situace zkoumání na jedné straně sevřenější a jednodušší tím, že jak terminolog, tak frazeolog se může prakticky omezit jen na substantiva a výrazy nominální (výjimky jsou řídké), objevuje se tu na druhé straně ovšem i podstatná překážka, většinou nereflektovaná, a tím méně poznaná a popsaná. Je jí zásadně odlišná povaha kolokability abstrakt a konkrét. Ta první oblast, abstrakta (tj. substantivní abstrakta), se v zásadě v textu realizuje většinou pomocí víceméně pevných struktur, kde příslušná verba jsou selektovaná jejich sémantikou (srov. např. *věnovat pozornost*, *uzavřít mír*, jejich relativně vyčerpávající popis z hlediska ustálenosti podává 3. díl SČFI), ta druhá, konkréta, však jeví zásadně více kolokační, kombinatorické volnosti, bez výraznějších a zjevných kombinatorických struktur. (Je otázkou, nakolik se tu podaří dospět ke specifitějším poznání na vznikajícím počítačovém *Českém národním korpusu*.) Toto má i své důsledky další. Na jedné straně se v terminologii jeví jisté názvy, např. notorický lingvistický výraz *doplňek* či *intence* se chápou velmi různě a bez valného konsensu (platí tu téměř doslova, že co autor, to pojetí), ba nulo- vě (jiní autoři je přímo negují) právě v důsledku své abstraktní povahy, kterou nelze definovat jinak než opět velmi abstraktními opozicemi, nikoliv denotativně. Přitom **termíny s konkrétním denotátem tuto míru a možnost zpochybnění a interpretativnosti nemívají** a jasný denotát mají; jinými slovy *bicyklem* sice může být leccos, ale ne do tak velké míry pochybnosti jako mnohé u abstrakt.

Na druhé straně v dnešní frazeologii jsou svým způsobem problémy obdobné řešeny nejspokojivěji tak, že frazémy a idiomy s denotátem konkrétním (tj. frazému jako celku, útvaru) se z této oblasti vyrazují na periferii, tj. mezi kvazifrazémy a kvaziidiomy. Sémioticky tu jde tedy o situaci do značné míry k terminologii opačnou.

V tomto zúženém rámci, tj. souvztažnosti frazeologie a terminologie, kterému bych zde rád věnoval pozornost, je jednou z poměrně dobře vymezených oblastí takto se překrývajících dvojí periferie **oblast lidových názvů rostlin**, kde jde tedy o poměrně homogenní výrazivo botanické, odborníky pro některé své nežádoucí rysy většinou už opouštěné, které má zároveň některé rysy frazeologické. Praktickým důsledkem existence takových výrazů jako *vraní oko* či *kozí brada*, založených tradičně řečeno na výrazné **metafoře**, frazeologicky pak na různém typu **anomálie**, pak bývá, že nejsou v centru zájmu ani terminologů ani frazeologů a soustavnému popisu ke škodě věci většinou unikají, ve frazeologii úplně, v botanice zčásti. Podívejme se tudíž stručně, jak tu vypadá situace z obou hlavních východisek, terminologického i frazeologického. Za základ zkoumání, relativně vyčerpávajícího, byla vybrána známá synoptická dvousvazková kniha Deyla a Híska *Naše květiny I-II* (1973).

## 2. Souvztažnost terminologie a frazeologie

Dosavadní literatura u nás obě oblasti co do jejich souvztažností příliš nezkoumala, byť si jich je zčásti vědoma (srov. Čermák 1982, Filipec – Čermák 1985, popř. Čermák – Blat-

ná 1995; Horecký 1965); pozornost zahraniční literatury tu není lepší, pokud vůbec nějaká je či není zavádějící. Tak lze za překonané považovat už starší sovětské názory Šanského, který byl do frazeologie ochotný počítat téměř cokoliv (výrazy typu **socialistická soutěž**), zatímco např. Molotkov (1977) s termíny ve frazeologii pracuje spíše implicitně a jen opatrně naznačuje, že je to oblast nezkoumaná, i když o meritů zařazení některých jeho řídkých příkladů do průniku obou oblastí lze pochybovat (ve výrazu *krutit' baránku*, tj. „šoférovat, dosl. otáčet volantem“ má slovo *baránka* jako jeden význam význam „volantu“ už ve slovníku, a zřejmě tu nejde tedy o frazém, jakkoliv o jakýsi profesní výraz by jít mohlo). Podobně se řešení této otázky vyhybají i další: např. Fleischer (1982), který tu pro němčinu také konstatuje jen jisté styčné body, zvl. metaforičnost, nic však víc neříká.

Ve snaze zjistit, jaký tu je vlastně **průnik** a při pokusu schematicky shrnout zastoupení hlavních aspektů v obou oblastech, dostaneme asi takovýto obrázek:

	TERMÍN	FRAZÉM/IDIOM
Ustálenost	+	+
Nominativnost	+	+
Monosémie	(+)	–
Sémantická průhlednost	+	–
Pravidelnost tvoření/neanomálnost	+	–
Denotativní přesnost	+	–
Kontextová jednoznačnost	+	–
Oborové zakotvení	(+)	(–)
Pragmatické aspekty	–	+
Původní metaforičnost	(–)*	+
Nesynonymičnost	(–)	–

Jakkoliv jde jen o jedno z možných zachycení situace (některé další aspekty jsou přitom implikovány přítomností jiných, zvl. ve frazeologii mnohoaspektová anomálnost), je i tak zřejmé, že průnik je velmi malý, zvláště odečteme-li první dva aspekty, které jsou platné obecně a nejsou tudíž nikterak specifické jen pro tyto dvě oblasti. Mj. i proto se další úvahy soustředí na oblast metaforičnosti (\*), jakkoliv by bylo možné, zřejmě však zcela okrajově a povrchně, ještě uvážít možné souvztažnosti v oblasti synonymie a oborové zakotvenosti, popř. i vztahu monosémie-polysémie.

### 3. Analýza

Autoři Deyl a Hísek ve zkoumané práci uvádějí celkem 901 názvů květin a k nim se pro populárnější zaměření knihy snaží uvádět, vedle standardních názvů, i názvy lidové. Důležitým předpokladem, zajímavým co do svých proporcí i obecněji, je to, že přes 90 % ze všech evidovaných lidových názvů jsou termíny víceslovné, v nichž především lze studovat některé aspekty frazeologické. Jak už o tom padla zmínka výše, výrazným rysem těchto výrazů je synchronně pojímaná, tj. zjevná metafora, resp. **sémantická anomálie** v širokém slova smyslu. Ze zkoumání bylo přitom pro synchronní nemotivovanost, popř.

kořenovost (nederivovanost) některých komponentů víceslovných lidových termínů nutno některé vyloučit, např. *jmelí*, *chrpa*, *rýt*, *sléz*, *vachta* aj.; diachronní analýza tu cílem nebyla.

Svou formou jsou kvazifrazémy (a zároveň tedy i periferní termíny), zde pro stručnost nazývané **názvy lidové**, bezvýjimečně substantiva, a to obvykle v kombinaci s adjektivem. Po bližším pohledu se však ukazuje, že je lze dále dělit do dvou formálně výrazně odlišných oblastí, do (A) kolokací vlastních a do (B) kompozit, u kterých může navíc, ale nemusí figurvat i atributivní adjektivum.

**(A) Kolokační kvazifrazémy**, kterých je evidováno 19 a jsou užity ve 21 názvech (tj. některé se zčásti opakují), zahrnují tyto případy (v závorce je pro jednoznačnou identifikaci uváděn jejich latinský ekvivalent, i když v řadě případů existuje i ekvivalent český):

*čarovník pařížský* (Circaea lutetiana)  
*čekanka obecná* (Cichorium intybus)  
*dáblík bahenní* (Calla palustris)  
*hadí mord nízký* (Scorzonera humilis)  
*chudina rolní* (Stenophragma thalium)  
*ježatka kuří noha* (Echinochloa crus-galli)  
*kandík psí zub* (Erythronium dens canis)  
*kozí brada pochybná* (Tragopogon dubius SCOP)  
*kozí brada větší* (Tragopogon dubius SCOP)  
*kozí brada východní* (Tragopogon orientalis)  
*kozí pysk tetucha* (Aethusa cynapium)  
*kuří noha ježatka* (Echinochloa crus-galli) = (*hloh kuří noha*)  
*len matky boží* (Linaria vulgaris)  
*netýkavka nedůtklivá* (Impatiens noli tangere)  
*oměj vlčí mor* (Aconitum lycoctonum)  
*plamének plotní* (Clematis vitalba)  
*rdesno hadí kořen* (Bistorta major S. F. Gray)  
*tchynin jazyk* (Sansevieria trifasciata)  
*tygřice obecná* (Tigridia pavonia)  
*vousy sv. Ivana* (Stipa joanis)  
*vraní oko čtyřlisté* (Paris quadrifolia)

První typ je kupodivu, tj. při daných terminologických zvyklostech a početním omezení, poměrně dosti pestrý, konstituuje se celkem v 6 strukturách. Podle početnosti zastoupení jde o struktury s těmito slovnědruhovými komponenty:

S-A : *plamének plotní*, *tygřice obecná* aj. (celkem 7)  
 S-A-S : *rdesno hadí kořen*, *kandík psí zub* aj. (celkem 5)  
 A-S-A : *kozí brada pochybná* aj. (celkem 5)  
 A-S-S : *kuří noha ježatka* (celkem 2)  
 S-S-A : *len matky boží* (1)  
 A-S : *tchynin jazyk* (1)

Obecně tu platí, že adjektivní komponent vždy stojí v postpozici, tedy jako např. v dnešní chemické terminologii. Je zjevné, že jen první tři typy struktury jsou výraznější i to, že poslední je hypotetický, resp. výjimečný a jeho slovosled je daný jen díky pevnější roli posesiva v atributivní pozici. Většina trojmístných, resp. trojkomponentních kvazifrazémů se dá přitom běžně užít, mj. v důsledku snahy o jazykovou ekonomii, i bez třetího, obv. substantivního komponentu, který nebývá ani součástí vlastní (kvazi)frazelogické kombinace, tj. např. *rdesno hadí kořen* → *hadí kořen*.

(B) **Kompozita** (většinou v kolokacích), kterých je 12 a jsou užita (opět v důsledku částečného opakování) v 17 názvech, zahrnují tyto případy:

*bělozářka liliovitá* (Anthenicum liliago)  
*bělozářka větvitá* (Anthericum ramosum)  
*černohlávek obecný* (Brunella vulgaris)  
*čertkus luční* (Succisa pratensis)  
*devětsil bílý* (Petasites albus): devět sil, magická moc  
*devětsil lékařský* (Petasites hybridus)  
*jitrocel kopinatý* (Plantago lanceolata):  
*jitrocel prostřední* (Plantago media)  
*jitrocel větší* (Plantago major)  
*lomikámen vždyživý* (Saxifraga aizoon JACQ)  
*lomikámen zrnatý* (Saxifraga granulata)  
*mateřídouška časná* (Thymus praecox)  
*potměchut'* (Solanum dulcamara)  
*sedmikráska chudobka* (Bellis perennis)  
*všedobr* (Chenopodium bonus-henricus)  
*zemědým lékařský* (Fumaria officinalis)  
*zeměžluč obecná* (Centaurium minus MOENCH)

U kompozit je třeba si všimnout především vlastního idiomatického komponentu jednoslovného, který je daný **anomální kombinací morfémů**. Pokusíme-li se tedy o analogický obrázek na této morfematické rovině a pokusíme-li se stanovit i původní slovnědruhovú východiska, lze dospět k takovému strukturnímu, leckdy však už poněkud zastřenému obrázku:

A-S : *černohlávek* (4)  
 S-A : *jitrocel* (3/1)  
 S-S : *zemědým* (3)  
 Num-S : *sedmikráska* (3)  
 V-S : *lomikámen* (2)  
 ADV-S : *potměchut'* (1)  
 Pron-S : *všedobr* (1)

Ve své plné podobě, tj. jako kombinace dohromady se svými nefrazelogickými, nicméně však terminologickými komponenty, však dávají podstatně jiný obrázek. Je výrazně sevřenější a co do pestrosti (proti typu A) prakticky homogenní:

- S-A : *zemědým lékařský, jitrocel kopinatý* (celkem 14)  
 S : *potměchut', všedobr* (2)

Sémanticky je nesporně hlavní zápornou vlastností lidových názvů v terminologii právě jejich **metaforičnost** v širokém slova smyslu, která tudíž **vylučuje deskriptivnost** a naznačuje možnosti expresivního chápání; širše tato vlastnost odporuje, zvláště v nové době, i snaze po dodržení ústrojnosti a výrazné stereotypnosti a modelovosti tvoření, která, jak vysvítá z pestrosti formální stránky těchto lidových výrazů (viz výše), představuje tudíž nežádoucí vybočení.

Jistě by bylo možné detailně stopovat původ a etymologii řady metafor zde obsažených a leckterou nabízí už i etymologický slovník; např. vedle Erbenovy básnicky ztvárněné interpretace *materídoušky* lze poukázat na starobylost původu a tudíž nečeskost, resp. kalkovost tohoto výrazu a na starou pověst za ním, jak ji už signalizuje alternativní latinský ekvivalent *matris animula* (pověst byla známa už Římanům) či na magickost čísla i reálnou léčivost *devět-silu*, který se opět vyskytuje v řadě jazyků; zřejmě plně český je naproti tomu původ na první pohled neprůhledné *potměchuti*, kde ve skutečnosti o tmu ani adverbale původně nejde, a výraz vznikl radikální kontrakcí z delšího (podle Machka) „po chuti nechut“ ve smyslu „po prvním nepříznivém chutovém dojmu se dostaví později příznivější, zřejmě v důsledku zvyknutí si“. Proti individuální etymologizaci však bude zajímavější a závažnější si položit otázku, zda se tu neprojevuje, resp. v minulosti neprojevovala nějaká obecnější nominativní tendence.

Ze skrovného materiálu kompozit se žádná jasnější doložitelná tendence vytknout zřejmě nedá (jsou tu nicméně zjevně zapojeny vizuálně barevné motivace), jinak se však chová oblast kolokací. Stranou je přitom napřed třeba nechat některé případy neprůhledné či zvláštní (u názvu *netýkavky nedůtklivé* jde přitom stylisticky o krajně vzácný pleonasmus, lingvisticky pak o nominativní redundanci založenou na opakování synonym). Ze zbytku je patrná určitá tendence k metaforizaci čerpající jako z východiska poměrně vydatně z názvů pro živočichy a části (lidského) těla, tedy frazeologicky řečeno o komponenty animální a somatické (v širším smyslu); druhý, popř. třetí komponent nevykazuje žádnou homogenost (nejde-li o kombinaci těchto zmíněných dvou). Lze tedy aspoň v podmnožině lidových názvů v oblasti kolokací mluvit o tendenci typu

#### ANIM +/-nebo SOMAT → PLANTA

srov. *hadí mord, psí zub, kozí brada, kuří noha, vlčí mor, hadí kořen, vraní oko, tchynin jazyk, vousy sv. Ivana*.

Přítomnost komponentů jiného typu je z hlediska původní metafor, jakkoliv široce chápané, sporná a ukazuje spíše na dodatečné spojení, obvykle jen deskriptivní povahy (často lokální): *dáblík bahenní, chudina rolní, plamének plotní*.

Konečně svou funkci se obě oblasti lidových výrazů nikterak od termínů obecně neliší, zastávají všechny běžné strukturní funkce substantiv. Odlišnosti, bez kontextového zkoumání ovšem jen tušené, lze hledat v jisté míře expresivnosti užití, popř. i v pragmatickém aspektu evaluativnosti, je-li některý z jejich komponentů tak aktualizován.

## 4. Závěr

Shrňme-li, lze konstatovat, že zkoumané výrazy vykazují formálně jistou tendenci k heterogenosti zapojených pojmenovávacích struktur, sémanticky pak se jako jediná poněkud zřetelnější tendence jeví metaforický přenos od animálních a/nebo somatických komponentů k metaforicky vzniklému obecnému významu květiny.

Obojí typ kvazifrazémů z dané oblasti zahrnuje celkem 4,2 % z celku 901 zkoumaných názvů.

### Summary

#### Terms and Idioms.

#### A Case of Overlapping and Periphery of two Areas of Nomination

There is an overlapping between terms, used in various disciplines, and idioms; many of these are to be found in slang as that, too, is in Czech a kind of (unofficial) terminology. Mostly, these cases might be viewed as special cases of idioms called quasiidioms. A major obstacle preventing these border cases to acquire the status of a full idiom is their exactly and explicitly defined meaning and often a concrete denotate. On the other hand, some of these terms, usually older and of folk origin display obvious metaphors and point to a kind of anomaly also here. For an examination of at least part of the field a well-known botanical compendium *Naše květiny* (1973, by Deyl and Hísek) has been chosen.

For a survey of features of terms and idioms, namely their presence in these, see below

	TERM	IDIOM
Stability	+	+
Nominative power	+	+
Monosemy	(+)	–
Semantic transparency	+	–
Regular formation	+	–
Denotatively precise	+	–
Contextual unambiguity	+	–
Membership in a discipline	(+)	(–)
Pragmatic aspects	–	+
Original metaforicity	(–)	+
Absence of synonyms	(–)	–

Out of 901 names of plants, over 90% have a multi-word character, some of which (4,2 %) displaying features of quasiidioms, based mostly on a semantic anomaly. Thus, there are 19 quasiidioms of a collocational nature, such as

*čekanka obecná* (*Cichorium intybus*)

*čarovník pařížský* (*Circaea lutetiana*)

*dáblík bahenní* (*Calla palustris*)

*hadí mord nízký* (*Scorzonera humilis*)



*chudina rolní* (Stenophragma thalium)  
*ježatka kuří noha* (Echinochloa crus-galli)  
*kandík psí zub* (Erythronium dens canis)  
*kozí brada pochybná* (Tragopogon dubius SCOP)  
*kozí brada větší* (Tragopogon dubius SCOP)  
*kozí brada východní* (Tragopogon orientalis)  
*kozí pysk tetucha* (Aethusa cynapium)  
*kuří noha ježatka* (Echinochloa crus-galli) = (*hloh kuří noha*)  
*len matky boží* (Linaria vulgaris)  
*netýkavka nedůtklivá* (Impatiens noli tangere)  
*oměj vlčí mor* (Aconitum lycoctonum)  
*plamének plotní* (Clematis vitalba)  
*rdesno hadí kořen* (Bistorta major S. F. Gray)  
*tchynin jazyk* (Sansevieria trifasciata)  
*tygřice obecná* (Tigridia pavonia)  
*vousy sv. Ivana* (Stipa joanis)  
*vraní oko čtyřlísté* (Paris quadrifolia)

But there are some collocations containing interesting compounds, such as

*bělozářka liliovitá* (Anthenicum liliago)  
*bělozářka větvitá* (Anthericum ramosum)  
*černohlávek obecný* (Brunella vulgaris)  
*čertkus luční* (Succisa pratensis)  
*devětsil bílý* (Petasites albus): devět sil, magická moc  
*devětsil lékařský* (Petasites hybridus)  
*jitrocel kopinatý* (Plantago lanceolata):  
*jitrocel prostřední* (Plantago media)  
*jitrocel větší* (Plantago major)  
*lomikámen vždyživý* (Saxifraga aizoon JACQ)  
*lomikámen zrnatý* (Saxifraga granulata)  
*mateřídouška časná* (Thymus praecox)  
*potměchut'* (Solanum dulcamara)  
*sedmikráska chudobka* (Bellis perennis)  
*všedobr* (Chenopodium bonus-henricus)  
*zemědým lékařský* (Fumaria officinalis)  
*zeměžluč obecná* (Centaureum minus MOENCH)

In many cases, an interesting etymology may be found.

# Somatic Idioms Revisited

(In: *EUROPHRAS 95 Europäische Phraseologie im Vergleich: Gemeinsames Erbe und kulturelle Vielfalt*, 1998, p. 109–119)

## 1. Language Nominations, Universality and Human Culture

The topic of somatic idioms addressed yet again in this paper is a rather old one and well worn-off, for that matter. I do not propose to sum up all that has been written up to now here as the very number of contributions made here is vast. In fact, it seems that there is no escaping the field of somatic idioms once any attempt is made at a more general description of idioms in any language. Also, even a cursory glance at the field readily suggests that it is very difficult to find a language which does not have this sort of idioms: it is not really relevant whether one looks at Chinese, Arabic, Indonesian, Quechua, Turkish, Finnish, Czech, German or English. Obviously, a tentative **universal** seems to offer itself here for formulation and a subsequent corroboration but that is a goal to be pursued elsewhere or by others (Čermák 1993a, 1994a).

Yet, on a closer look, there are obvious differences here, too. To be able to say something about these, one can either resort to repeating, yet again, the few commonplaces usually mentioned, or to draw not very far-reaching conclusions based on few randomly chosen examples, or, finally, to undertake a systematic and exhaustive contrastive analysis of at least two languages, which is, however, not available so far. In turn, I will try to say a few words on each of the possible approaches.

It has often been stressed that (1) it is necessary to study language in the *context of culture* (Dell Hymes 1964 and many others), which implies that it is not possible to have any beforehand universal conception, a view very much rejected by such linguists as N. Chomsky. I do not propose to go into this matter now, let me only suggest that it very much boils down to the very essence of what language is made of; obviously, it is a question going beyond the scope of our current concern. Yet, if the idea is reformulated somewhat more narrowly as the relation of **language nominations** (names), of which idioms are a prominent part, to the **culture** in question it does become meaningful. In fact, all one is faced with here is still another version of the Saussurean idea of arbitrary segmentation of reality by the language, this being a reformulation of a much older approach which leads us back to ancient Greeks.

It is, however, the **arbitrary aspect** of this which has to be stressed and inspected together with somatic idioms: human body, being generally the same for everyone, offers,

so to speak, a ready-made segmentation in its parts which leaves very little option for the language variation here (rarely, there may be, however, such unique and exceptional names as Chinese *zang* being a name for the group of five organs, namely heart, liver, spleen, lungs and kidneys). As constituent parts of somatic idioms names of body parts are used together with most of their traditional functions and symbolism: *nose* as instrument for smelling but also spotting eventual danger, *heart* is viewed as centre of love, courage and other feelings, etc. Indeed, in its firm anchorage in our body reality, this part of language nominations reminds us, surprisingly, of interjections, although semiotically the phenomenon is not quite the same; one cannot observe here any motivation by reality.

However, once the body parts function is separated from its symbolism, differences in the latter field appear, both minor (also among closely related languages) and greater (as in Chinese where *heart*, i.e. *xin*, is traditionally a centre of intellectual, thinking processes, while *gall(-bladder)*, i.e., *dan*, is a centre of courage or devotion etc.). It is here where one finds beginnings of all subsequent types of variation and differences. On the whole, it does seem, however, that, apart from frequency of use of body-parts names as idiom constituents, peculiarities in their use may be explained simply by the chance factor: it is hard to see from our point of view why there are so few idioms anywhere using as constituents such words as *brain*, *ankles*, *kidney* or *appendix* where idiomatics clearly goes against functionality; thus in Czech, *brain* (*mozek*) is attested in only 4 non-sentential idioms while, for example, *nose* (*nos*) is found in 65 idioms. This chance factor related to the external and more prominent body parts to the exclusion of internal ones can be reformulated as historical one, basically. The obvious conclusion in this respect, then, is that **no direct outstanding relation between body-parts names as idiom constituents and culture can be sustained.**

Yet there is another, very general dependence one can observe here with a high degree of certainty, namely a more than chance occurrence of somatic idioms of verbal type combining with human type of subjects; in this, verbal somatic idioms seem to follow the bulk of verbal idioms, in fact. For the Czech language, the statistics show that the overwhelming 95.93% of all verbal idioms combine with human subjects (subject arguments), the meagre rest being distributed among subjects representing animals, concrete matters, abstract matters and place names (0.09%, 0.57%, 3.22% and 0.19% respectively), with some overlapping. Now, this kind of figure has not been calculated for the bulk of somatic idioms, but a single illustration will do. The Czech somatic verbal idioms based on the component *ruka* (hand), representing the total of 222 separate idioms, follow, basically, the general line already just outlined: there are 92.3% taking “human” subjects. To give some comparison and hint at a possible generality of this feature, let us look at other languages: in French (Dubois) the figure, based on some 132 verbal idioms containing *la main*, is very similar: 93.19% of these idioms do require a human subject; only such rare cases as *changer de main*, *fondre entre les mains*, or *passer de main en main* do not take this kind of subject. One could go on and arrive at similar results elsewhere. Thus, Barron’s dictionary suggests, due to a single idiom with non-human subject (*hang heavy on one’s hands*), 97.14% of this kind of subject for English, Russian (Molotkov) 94.1% (*schodit’s ruk, gorit’v rukach, popadat’v ruki, prochodit’čerez ruki*), while Taylor-Gottschalk registers a somewhat lower percentage of 87.5% for German, perhaps due to a special character of the dictionary (*weder Hand noch Fuss haben, auf der Hand liegen, durch viele Hände gehen* etc.) etc.

There seems to be at least one more rather general relationship between culture and somatic idioms which is worth examining, namely that between some somatic idioms and paralanguage of gestures, but I shall come to that later (in 3.). To conclude this part, an even more prominent relation between somatic idioms and **anthropocentric principle** of the language **anthropomorphism**) has to be mentioned and duly stressed. Thus somatic idioms may not be such a surprising and unique phenomenon if they are viewed on a par with and within the framework of all other and well pronounced anthropomorphic features one finds in language in general, such as major and general tendencies of polysemy growth of human vocabulary, which is based on metaphor and metonymy, cf. *time goes*, *the door leads somewhere*, *the foot of a hill*, or *the shoulder of a river*, to give only a few English examples, limited, moreover, to the lexicon only.

However, a note of caution is due here. As I have not deliberately stated what exactly I may mean by **somatic idioms** I must do it now and for a good reason which must seem to be quite evident by now: **somatic idioms are simply idioms (phrasemes), or idiomatic (phraseological) combinations of various functions containing at least one obvious body-part name**. By stressing the quality of *obvious* I want to exclude such historically less-known human organs as *appendix*, *uvula* or *hypothalamus* as well as minor parts of other, better-known organs. Now, to relate the field, which has been thus limited, to the anthropocentric principle, one has to realize that “human presence” in idioms, so to speak, is not limited to somatic idioms only. On the contrary, it is extremely hard to limit the universe of idiomatic components only to those which would *not*, one way or another, be related to *man* and to what might be called *human*. An extensive analysis including, next to proper somatic parts of idioms, such noun types as types of persons (such as *idiot*, *gentleman* etc.), human products and utensils (*bed*, *beer*, *ruler* etc.), abstract priori notions related to human culture and civilization (such as *honesty*, *truth*, *Friday*) and abstract states (such as *pension*, *need*, *comfort* etc.), has been undertaken in this sense for the Czech language, and it gives some unexpected results. It may come, perhaps, as no surprise, that somatic idioms proper, i.e. verbal idioms in this case only, amount to over 1700 idiomatic combinations in Czech, i. e. slightly less than 20% of all Czech verbal idioms. Yet what may be really surprising is that over 75% of all Czech idioms do belong to one or another of the anthropomorphic types outlined above (Čermák 1994d). Thus idioms containing other types of noun constituents, such as animals, plants, insects, nature formations and the like, are in clear minority amounting to less than a quarter of the whole.

## 2. Meaning-Form Relation

To ask whether there is any relation between the language idiomatic form and its meaning may be both misleading and rewarding. A comparison of some carefully chosen items across languages may give encouraging results and tempt one to suspect a deeper relationship, if not a downright motivation of some sort. Thus, the idiomatic meaning “*to help somebody in his/her difficult situation*” does yield highly **comparable items based on HAND constituent** in most of the Germanic, Romance and Slavonic languages, such as

Engl. *give/lend someone a hand*,  
 Germ. *Hand an etwas anlegen*,  
 Dutch *de behulpzame hand iemand bieden*,  
 Swed. *räcka någon en hjälpande hand*,  
 French *donner/prêter la main à qn*, or *tendre une main secourable*,  
 Ital. *dare mano forte a qc*,  
 Span. *dar la mano a un persona*, *soltar la mano*,  
 Russ. *podát'/protjanút' ruku pómošči*,  
 Pol. *podac komu pomocną dłoń* (*palm being part of hand*, in this case) or  
 Czech *podat pomocnou ruku někomu*,

but Finnish, despite its coexisting with and having been influenced by Indo-European languages, does not have a comparable equivalent (the Finnish equivalent *auttaa miestä mäessä* "help man/s.o. in the hill" suggests quite a different solution). It has to be conceded, however, that this rather simple picture could be made even richer, should one include other, complementary terms such as English *arm* or Czech *paže* existing next to *hand* and *ruka* (= arm).

But this rather coherent picture begins to disappear if we go on with the comparison of the adjacent or opposite meaning of "*not to help somebody*" where the same languages differ either in the verb used or even drop the HAND constituent altogether, e.g.

Engl. *not lift/raise a hand* or *sit on one's hands*,  
 Germ. *keine Hand rühren* or *die Hand von jm abziehen*,  
 Dutch *zijn hand voor iets niet omdraaien* or *geen hand uitsteken voor iemand*,  
 Swed. *ta sin hand ifrån någon*,  
 French *ne pas en ficher une ramée*,  
 Ital. *non muovere un dito per qualcuno*,  
 Span. *quedarse con las manos cruzadas*,  
 Russ. *sídet' složá/podžáv ruki*,  
 Pol. *nawet nie ruszyć dla kogo palcem*, or  
 Czech *nehnout prstem pro někoho*.

The Finnish language has a *calque* here probably: *ei heilauta kättään* ("not wave a hand") but it prefers to use an equivalent very similar to Czech *ei sormeansa liikuta*.

The lesson to follow from this is obvious: while the former meaning examined seems to have come, by having been translated gradually, from the same source and to be probably of an older date, the latter, opposite meaning did not have any such common source, as the verbs seem to suggest ("*lift/raise, move/stir, turn/put out, take from, knock in, move, fold, move*") or choice nouns used here (*hand, foliage, finger*). Thus, one can hardly look for any motivation or well-established relation between the idiomatic form and meaning here, despite the seemingly apparent similarities due to historical reasons, in most cases. Apart from common history and its influence, there is, in fact, very little in common here (as the second set of examples suggests) and even the apparent cases where the actual movement or gesture may seem to be the only possible or imaginable one may turn out to be deceptive or are not mirrored by actual idioms at all.

Thus, it is hardly worth saying that even such prominent fields of idioms as those inspected above **abound in specific, original and idiosyncratic formations not to be found elsewhere**. Let us take a couple of more examples from the same field of HAND which seem to be language-specific:

Engl. *sit on one's hands*, *try one's hand*,

Germ. *auf der Hand liegen* (be as clear as daylight), *etw. kurzer Hand ablehnen* (reject sth out of hand),

Dutch *iets achter de hand hebben* (have sth in reserve), *met de handen in het haar zitten* (be at one's wits' ends),

Swed. *ha ingen hand med något* (be clumsy with sth),

French *ne pas avoir trente-six/quatre mains* (have only one pair of hands), *avoir des mains de beurre* (to be butter-fingered),

Ital. *avere le mani bucate* (spend unreasonably), *avere le mani di pasta frolla* (be clumsy dropping things),

Span. *dar una mano a una persona* (scold somebody, give somebody a good talking-to),

Russ. *vzjat'sebja v ruki* (pull oneself together), *nagret'sebja ruki na čem* (gain money in sth),

Pol. *mieć lekką rękę* (spend unreasonably), *urabiać sobie ręce (po łokcie)* (plod, sweat one's guts out),

Czech *mít volšový ruce* (be a butterfingers, be very clumsy),

Finn. *hakea jotakin käsiinsä* (fish out something, find) or, not surprisingly,

Chinese, as in *zuo shou jiao* (cheat/trick someone, lit. make hands and feet) etc.

There is no doubt that the field is made up, primarily, of idiomatic constructions of considerable diversity of finer, less conspicuous, or larger proportion, such as those above; such examples could be easily multiplied.

### 3. Gestures and Idioms

If minute analysis may not yield much, one may ask whether there are at least some statistical tendencies between form and meaning. Now, if Czech idioms with the HAND constituent which happen to be also names for gestures are examined, over 20 meanings, closely-knit together, may be discerned, such as

capitulation (*zvednout ruce nad hlavu*),

congratulation (*potřást někomu rukou*),

despair (*spínat ruce/lomit rukama*),

esteem (*políbit někomu ruku*),

expectancy (*zamnout si ruce*),

lack of interest (*mávnout nad něčím rukou*) and the like.

It remains to be seen, against the background of other languages, whether there is a more general tendency to be stated here or whether this state of affairs is a language-specific one. The examples above are not to be mixed, however, with others, that are not linked with a gesture (such as */po/zvednout ruku proti někomu*), where a variety of different meanings can be found.

Some tendencies, then, are quite apparent, even more so, should they be compared with those for other somatic constituents:

- (1) 100% of the *hand*-idioms of this type are, in various degree, **oriented towards a second agent**, usually human expressed as an object, i.e. none of them has an expressive function only, e.g. friendship, as in *podávat si ruce*.
- (2) Over 97% of *hand*-idioms express **intellectual attitudes or reactions**, e.g. recognition, as in *potřást někomu rukou* (the exception being piousness as in *sepnout ruce*), i.e. although some emotional colouring may be present with some of the idioms (e.g. indignation as in *dát ruce v bok*), none of them is of a primary emotional nature.
- (3) Over 97% of these idioms seem to be **reactions to abstract, human or cultural stimuli** (the only exception is that of feeling cold: *mnout si ruce*).
- (4) Although some attitudes expressed may be quite strong, none of the *hand*-idioms may be labelled as expressing aggressive meaning, implying attack or threat to someone else.
- (5) All idioms express **processes**, and none of them designates a mere state of affairs, static relation or physical/local relation or, even, a real action.

Of course, much can be added here on the nature of these somatic idioms which have the function of **gestures**. One can, basically, speak here of at least three important features:

- (A) These idioms have a **double denotate**, a concrete (the movement itself) and an abstract one (attitude or stance), and they are semiotically of symbolic nature (Čermák 1982, 1988). The existence of their double denotate is supported by a possibility of using these gestures without any verbal counterpart. In fact, there seems to exist a neat distributional difference: while pure gestural use tends to be limited to **direct speech, idiom use prefers indirect speech**; also, while direct gestures may not be, as a rule, used with the other than 1st person, idioms are mainly used with other than the 1st person of the verb (they seem to be non-performative, in fact).
- (B) These idioms belong to **paralanguage**, or kinesic part of it, accompanying the verbal communication proper; as a rule, there seems to be no other way how to link the verbal and non-verbal part of communication than by idioms of this type.
- (C) In general, these idioms are very much **culture-dependent**, both in degree of gestures used and their selection.

#### 4. Czech and Chinese and Some Open Questions

To round up these remarks, I propose to return back to a relatively exhaustive analysis of the field. A recent and particularly revealing comparison and analysis of the Czech and Chinese language of this sort has been undertaken (Song Fen-Yuen) based on the same

methodology (Čermák 1982, 1988) which makes it possible to offer more than fragmentary information, so typical of other approaches. On the basis of the overall number for somatic idioms of all types encountered in Czech and Chinese, i.e. 2101 and 2861 respectively, it has been found that some 72% of Czech and 52% of Chinese somatic idioms belong to the field of verbal collocations, i.e. to the verb-noun structure and its variants, primarily.

It is quite surprising, given the great dissimilarity of Czech and Chinese, that a basic general correspondence of more than 70%, both in general function and general meaning, could be established between the two languages as far as their somatic idioms are concerned.

This overall correspondence is to be seen, among other things, in the 70% general match of **the first 10 somatic noun constituents**, ordered by their frequency:

**Czech:** *hlava, oči/oko, ústa, nohy, nos, jazyk, ruce/ruka, krk, srdce, nervy*  
*(i.e. head, eye(s), mouth, feet, nose, tongue, hand(s), neck, heart, nerves)*  
**Chinese:** *heart, eyes, head, mouth, body, hand, face, feet, ears, skin*

The obvious differences are in the use of *nose, tongue, nerves* in Czech and of *body, ears* and *skin* in Chinese. However, if this general figure is compared, e.g., with that of the Czech verbal somatic idioms only, a somewhat different sequence obtains: *hand, eye, head, mouth (huba), foot, nose, heart, finger, ear, tongue*. At the same time, the *hand* constituent happens to be the most frequent one Czech verbal idioms use here (unlike, for example, in comparisons, where it is *dog (pes)* which is most heavily employed).

## 5. Conclusion

To conclude, a number of questions, rather than answers, offers itself here:

- (1) Are somatic (or anthropomorphic) idioms universals?
- (2) Are somatic idioms always linked to gestures?
- (3) Are there gestures (completely) outside somatic idioms, too?
- (4) Are there new somatic idioms still being formed or is it a closed, very old historic chapter?
- (5) Can one observe here a microsystem of some sort?

---

### **Shrnutí**

#### **Znovu o somatických frazémeh**

Somatické frazémy se považují za univerzální; jejich podoba je v různých jazycích však nutně různá. Je to vidět už např. na srovnání češtiny a čínštiny (*dan*, tj. název pro žlučník, se chápe za centrum odvahy a oddanosti, v češtině se však téměř neužívá); jiné se užívají jinak nebo se neužívají v somatických frazémeh vůbec. Nelze tedy stanovit žádný obecný vztah mezi částmi těla a jejich zapojením do kultury, zprostředkované zde především frazeologií.



V češtině bylo zjištěno, že přes 95 % verbálních frazémů se pojí s lidským subjektem, přitom z nich valná část obsahuje somatické komponenty. Takto frazémy obsahující komponent *ruka* (222 z verbálních frazémů) toto obecné zjištění potvrzují: 92 % jich má lidský subjekt také a podobnou situaci lze očekávat i u dalších. Podobná situace je i u jiných jazyků. V širším pohledu tu jde o uplatňování principu antropomorfismu, známého i odjinud z jazyka.

Některé případy ve srovnání jako by naznačovaly souvislost mezi formou a významem, srov.

angl. *give/lend someone a hand*,

něm. *Hand an etwas anlegen*,

hol. *de behulpzame hand iemand bieden*,

švéd. *räcka någon en hjälpande hand*,

fr. *donner/prêter la main à qn*, či *tendre une main secourable*,

ital. *dare mano forte a qc*,

špan. *dar la mano a un persona, soltar la mano*,

rušt. *podát/protjanút ruku pomošči*,

pol. *podać komu pomocną dłoń*,

češ. *podat pomocnou ruku někomu*,

zatímco finština analogický výraz nemá (*auttaa miestä mäessä* pomáhat někomu v kopci); jsou tu však i další problémy, jako je vztah *paže-ruka* aj.

Naproti tomu téměř obrácený pohled skýtají jazyky při vyjadřování významu „nepomocť někomu“, srov.

angl. *not lift/raise a hand* či *sit on one's hands*,

něm. *keine Hand rühren* či *die Hand von jm abziehen*,

hol. *zijn hand voor iets niet omdraaien* či *geen hand uitsteken voor iemand*,

švéd. *ta sin hand ifrån någon*,

fr. *ne pas en ficher une rameé*,

ital. *non muovere un dito per qualcuno*,

špan. *quedarse con las manos cruzadas*,

rus. *sidet' složá/podžáv ruki*,

pol. *nawet nie ruszyć dla kogo palcem*,

češ. *nehnout prstem pro někoho*.

Zdá se tedy, že nějaká obecná souvztažnost mezi názvem části těla a významem frazému spíš není. Ukazují to i výrazy idiosynkraticky omezené jen pro daný jazyk, srov.

angl. *sit on one's hands*, *try one's hand*,

něm. *auf der Hand liegen*, *etw. kurzer Hand ablehnen*,

hol. *iets achter de hand hebben*, *met de handen in het haar zitten*,

švéd. *ha ingen hand med något*,

fr. *ne pas avoir trente-six/quatre mains*, *avoir des mains de beurre*,

ital. *avere le mani bucate*, *avere le mani di pasta frolla*,

špan. *dar una mano a una persona*,

rus. *vzjat' sebja v ruki*, *nagret' sebja ruki na čem*,

pol. *mieć lekką rękę*, *urabiać sobie ręce (po łokcie)*,

češ. *mít volšový ruce*,

fin. *hakea jotakin käsiinsä*,

čín. *zuo shou jiao* (podfouknout někoho, dosl. dělat ruce a nohy) atd.

Somatické frazémy ale souvisejí také úzce s gesty a šíře patří k parajazyku. Srov. význam a frazém/gesto kapitulace *zvednout ruce nad hlavu*,  
blahopřání *poťrást někomu rukou*,

zoufalství *spínat ruce/lomit rukama*,  
 úcta *políbit někomu ruku*,  
 očekávání *zamnout si ruce*,  
 nezájem *mávnout nad něčím rukou* aj.

Prozatím, zvl. na základě studia *ruky*, lze nabídnout některé závěry, zvl.

- 1 – 100 % frazémů má vztah k druhému objektu, člověku,
- 2 – přes 97 % vyjadřuje intelektuální postoj či reakci, přičemž žádný nevyjadřuje primárně emoce,
- 3 – přes 97 % je tu reakcí na abstraktní, lidské či kulturní podněty, výjimka: *mnout si ruce (zimou)*,
- 4 – žádný z nich nevyjadřuje vysloveně agresivní postoj,
- 5 – všechny vyjadřují procesy, žádný stav, statický vztah ap.

Srovnání frekvenčního zapojení nejčastějších somatických komponentů v češtině a čínštině naznačuje i rozdíly, srov.

**čeština:** *hlava, oči/oko, ústa, nohy, nos, jazyk, ruce/ruka, krk, srdce, nervy*

**čínština:** *srdce, oči, hlava, ústa, tělo, ruka, tvář, nohy, uši, kůže*

Závěrem lze zformulovat několik otázek:

- 1 – jsou somatické, resp. antropomorfické frazémy univerzální?
- 2 – vážou se somatické frazémy vždy ke gestům?
- 3 – existují i gesta vně somatických frazémů?
- 4 – tvoří se i nové somatické frazémy, anebo jde o starou a uzavřenou kapitolu?
- 5 – lze v nich spatřovat nějaký mikrosystém?

# On the Nature of Universality in Idioms

(In: *Z problemův frazeologie polské i slovianské VI*, 1994, p. 15–19)

## 1. Situation and Basic Notions

Relating the concept of **idiom/phraseme** to universality is impossible without stating first what is meant by both terms, i.e. by idiom and by universality. It is readily admitted that in the first case, the idiom is viewed by linguists differently, but this difference is usually not recognized as being so fundamental as it really is. Both the unit of idiom and its respective field, idiomatics, are so vastly different in various approaches in their constitutive features and scope and have become blurred that it is no longer clear what really belongs here, under these two headings. If the extreme approaches are mapped onto each other their intersection is much less in size than what is different in them.

As a result, an almost disastrous misunderstanding may follow, since the label of **idiom** may be used for almost anything. The outstanding border regions, not sufficiently supported by the theory are often not mentioned at all or are labeled differently as collocations in English linguistics, at least partly. An enumeration and discussion of what is meant by the term would be a major topic in its own right. This situation of not clearly delimited boundaries of the field and the essence of its unit is often paradoxically further worsened by suggestions for internal divisions or classifications; as a rule, these suggestions are based on vague and useless criteria as in the case of the Russian Šanskij and others following the old hazy remarks of Vinogradov.

It is therefore necessary to state one's own position positively. It has been published on various occasions and used as the entrance test for the entries of the Czech idiom dictionary (*Slovník české frazeologie a idiomatiky I–III*), one of the greatest of its kind. Let me add that the term idiom will be used here for the sake of brevity as a general term for all kinds of phraseological and idiomatic units and. Thus as a general background of following remarks on universality of idioms, I propose to view the idiom as follows:

*The idiom is such an unmodelled and fixed syntagm of elements, of which at least one, from the perspective of the other, is a member of an extremely restricted and closed paradigm (formally and often semantically, too).*

Also the latter of both terms, **universality**, seems to be lacking in many respects, the general one being a lack of factual corroboration and evidence for the alleged universals,

proposed by various authors. My general view will be based on older conceptions of universals (implicational universals), many of which can, however, be translated into Chomsky's formal universals as well (Chomsky 1965). Specific objections against other universalist approaches to idioms may be summed up in four points:

- (1) They are **diachronic** in nature and do not, in fact, account for the present state of the field or unit (such as Thass-Thienemann/Fabricius principle of suppressed taboo meaning, suggested by Makkai 1978, which is projected into secondary senses of words).
- (2) They are **in no way specific to idioms only** (Dobrovolskij 1988, but also partly Makkai and Coulmas 1981), some being outright commonplaces (such as existence of polysemy or synonymy in idioms, Dobrovolskij). It is evident that such features cannot delimit the specific field of idioms in the vast territory of language phenomena.
- (3) They are statements about the **semiotic nature** of the idiom's constituent parts only and, accordingly, not about the idiom proper (Kopylenko – Popova 1972, partly also Coulmas). However, what readily suggests itself, i.e. their primary abstract semiotic nature, is, for some reason, never mentioned at all. This abstract denotate of most idioms stands in sharp contrast to their components having, mostly, concrete denotates.
- (4) Some of their premises might be wrong (such as Dobrovolskij's claim of the regularity of the idiomatic system being dependent on the degree in which language is analytical). In fact, the very idea of the **idiomatic system** being in existence is wrong; such an alleged system cannot be stated in positive terms since there exists only a residuum of combinatorial anomalies, as implied by the above definition, which is complementary to regular combinations obtaining in language. What these four statements do have in common is their hypothetical nature not corroborated by any research based on any substantial language material, a fact which holds for Chomsky's claims, too, contrasted by, e.g. Greenberg's substantive universals who did undertake some research to support his claims.

## 2. Ten Universals: A Proposal

It is therefore necessary, among other things, to avoid any diachronic approach and stick to the matter at issue, i.e. be specific enough as to really characterize idioms **synchronically**, and not only their constituents.

In what follows, ten basic (absolute) universals will be proposed, in some instances with related implicational universals based on them. However, no feature will be admitted as universal if it happens to be a general property of language nomination, such as synonymy etc., and which does not follow from the presence or constellation of other features. Should we start any **diachronic** investigation as to idiom universals, an interesting question arises: should such a consideration be purely historic (and clearly not belonging here) or should it

be pursued in a framework similar to that of word-formation? If we opt for the latter, an objection may be readily seen in the nature of misused term productivity (which is not purely synchronic) which does not obtain in idioms.

To come back to our synchronic idiom universals, let us briefly comment on each of them. I will deliberately restrict my remarks to a general form only and propose not to go into any details or illustrations as these would always be partial to some language only.

### (1) Idioms exist in every natural language.

Here, I would like to suggest that even the most far-fetched views of idioms do not go completely wrong and do cover at least part of idiom field in the language in question. Thus no language without idioms has been reported so far; on the contrary, even artificial languages may acquire idioms, if they go through some development (Esperanto). An additional, though quite banal, implicational universal, following indirectly from the above definition may be formulated as follows:

*If a language has metaphors, then it has idioms too.*

It may be useful to remind ourselves that idioms are by and large entities whose predecessors were once metaphors which are of universal nature, too; in due time they became stable and entered the vocabulary of the language as idioms.

### (2) Idioms exist on all combinatorial levels of language elements related to meaning; some of these combinatorial results may enter as idioms into the language system.

It is clear that combinatorial, or rather permutational possibilities of a language are far greater than those in actual use which are of necessity regular, governed by rules; idioms always draw on the unused possibilities and, in some cases, even on those which seem to be quite impossible because of structural reasons, i.e. they are not rule-governed.

It is not quite clear whether one is justified in looking for idioms in la parole, text, and, indeed, whether it is possible to distinguish them from newly-coined metaphor or metonymy. It may seem that such an alleged text idiom, i.e. one formed ad hoc, and idiom are identical and cannot be distinguished. Yet, the difference is the idiom's acceptance and familiarity. An extension of this is the tricky problem of **stability or fixedness** (there seems to be no good term for this in English), i.e. such a degree of fixed existence of a combination which allows for its inclusion into the vocabulary. Both this notion of stability and of **variation** is in need of further elaboration and precision. Let me just make one remark here: only by variation and variants such oscillation and/or alternatives is understood that preserves the idiom's identity and syntactic function. Here, some widely different views are to be found, too. Thus an implicational universal may be formulated:

*The higher the idiom's frequency the more the idiom is fixed (stable) and the less the chance that there might be variants of it.*

One may observe that this particular universal implies convergence with diachrony. This is true both of (1) fresh neologisms, new idioms, of (2) overused idioms or clichés, as well as of (3) idioms with small frequency of use getting slowly obsolete. A consequence of this is loosening of the idiom's form and/or meaning which gives rise to variants. In general, this is, then, a peculiar specimen of the centre-periphery relation.

### **(3) Idioms are results of a chance formation and are not based on any model.**

Past use of the term *model*, *model-like* etc. has been rather loose and might wrongly suggest that there is a possibility to coin new idioms at liberty (perhaps following some item coined previously) and thus manipulating the language as libitum. It is dangerous nonsense, of course, and, accordingly, such use of a model is to be avoided. At the same time, this view might also give some grounds to a search of regularity in idioms, so popular with some linguists. On closer inspection, based on the above definition, it is evident, however, that no aspect of the idiom may be seen as regular and generally recursive; in fact, it is only the language elements entering the idiom as its components that might be regular in the very beginning, before this entry. But upon this entry, they both stop being regular and, indeed, standard and analyzable language units. The problem of extremely limited analogy in the idiom formation may be resolved by the existence of **closed paradigms** postulated in our definition of idioms (see Čermák 1982, Filipec – Čermák 1985). More generally, the basically chance existence of idioms in language is due to its entropy, which should be sought, however, in its system as well as in its use, i.e. communication (Čermák 1991, 1994).

### **(4) Idioms are based on various kinds of either internal syntagmatic or paradigmatic anomaly of their constituents, or on both.**

This central **defining feature of anomaly** of many kinds may be judged and evaluated only against the background of all relevant paradigmatic and syntagmatic relations (formal, semantic and functional), on their regularity, potentiality and, eventually, their impossibility. Thus an exception or shift from regular possibilities in the behaviour or status of a component of the idiom, i.e. a deviation from the regular analogical situation of all the other analogical forms must be viewed as anomalous and idiom-formative. The actual number and typology of these irregularities is not known; it is definitely vast. Every single idiom exhibits a number of these anomalies.

It is evident that a reliable judgement of this, then, makes it a task for the native speaker only, unfortunately, and this is also why idioms are usually seen as a specifically difficult part of a learning syllabus by non-native speakers.

At this point, four implicational universals may be formulated, if the frequency and meaning of constituents and homonymy and transparency of the idioms are taken into account, too:

*A – The higher the frequency of a constituent, the more remote it is semantically to senses given in its dictionary definition.*

- B – The less frequent the monocollocable form (archaic, nonce-form) of some of the idioms's constituents (if any), or the less transparent the constituent is, the higher the frequency of some other constituents of the idiom.*
- C – If the constituents are semantically compatible, the idiom has an exact homonymous counterpart.*
- D – If the constituents are only monosemous or only polysemous, the idiom has a homonymous counterpart.*

The term **monocollocable** used in B above refers to the fact that some word-forms are not to be found outside the boundaries given by a single idiom or a small cluster of them (English kith and kin). The term is coined to conform to others, such as **collocability** and designates a form which forms a collocation, combination with a single other form or with a sharply specified and restricted set of such forms only.

**(5) The idioms are based on a non-Fregean (non-additive) function or meaning of its constituents (non-compositionality).**

This universal is based, among other things, on the assumption that two of the frequent notions, cherished by many linguists, be abandoned: the first is the view that idioms are **semantically motivated** through their components, the second is the belief that there is such a thing as a **phraseologically bound meaning** of some components (some of the Soviet schools). Proposals trying to substantiate the claim that idioms are semantically motivated do nothing else but concentrate on some components only leaving the rest of the idioms components unaccounted for: on a closer glance, what is left, after the seemingly clear and motivating part is cut out, is an indefinite, artificial torso which cannot be analyzed by any of the familiar techniques and which is, basically, not to be matched with anything. This consequence is due to the linguist's erroneous view that the idiom is a sum of its component parts in meaning also and that it can be regularly **decomposed** just like any other regular language unit. One might, perhaps, draw a parallel here with a mutation in biology where, however, all attempts of a direct and mechanical decomposition are unthinkable. Semantic motivation of the idiom seems to be nothing more than one of the myths some linguists like to cherish.

An equally problematic myth is the existence of phraseologically bound meaning or sense of some components. A seeming lack of methodological criteria and analytic tools for idioms does not mean that they are to be handled just like normal, regular expressions. A meaning, sense of a component, lexeme, word etc. can only be postulated against a **sufficiently great analogy** to be found in sufficiently repeated, recurrent identical uses of it in different distributional patterns; however, such a proposed meaning of an idiom component has never been witnessed and proved. In fact, one might refer here to what F. de Saussure had to say here about the isolated, non-analogous forms.

Plainly, to use such an approach amounts to a *contradictio in adiecto*, since words always lose their identity upon entering an idiom as its constituents and the alleged special

idiomatic meaning is nothing else but a half-baked, inconclusive attempt at a semantical analysis of the idiomatic whole, which is impossible. Generally speaking, such approaches show a remarkable disregard of the precious few analytical tools left to the linguist when analyzing idioms, namely the degree of (sufficient) analogy, which should always back it up and tell one whether there are grounds for stating an existence of a phenomenon, item, relation etc. for the system, or not.

**(6) Idioms are functionally analogous to corresponding regular combinations of regular elements of the same level.**

Idioms are not really quite analyzable and represent a typical Gestalt or holistic phenomenon also psycholinguistically (see, e.g., Schweicert and Moates 1988). They often acquire, generally speaking, the same type of syntactic **function** as analogous normal, regular combinations of expressions. Yet there are important differences. Thus a verbal type of idiom, for example, may function as predicate (or even as a verb) in sentence, but the idiom's noun (if any) is just not analyzable as an object of such a sentence. This is also the case of the adjective-noun type of idiom not being analyzable into head and an attribute and so on. Such an analysis would presuppose, again, both a semantic and syntactic analysis, commented on above in No. 5, which does not seem to be viable. In fact the very notion of the word classes becomes blurred here. Also such aspects might be, if necessary, reformulated into an implicational universal.

**(7) Idioms are functionally and semantically always defective to a degree, exhibiting restrictions in their behaviour.**

This universal being a consequence of the fifth is yet another manifestation of many types of anomaly in idioms, to which restrictions of their functions, limited possibilities to acquire a specific grammatical category or even irregular forms etc., are due. These include, to give an example, restrictions in the syntactical functions of nominal idioms, i.e. either in their capacity to be used in subject or predicate position etc. These restrictions stand in sharp contrast, then, to behaviour of **regular** combinations where they do not obtain. (Regular is here and elsewhere to be understood simply as "conforming to usual grammatical and semantical rules of the language in question".)

This point has been raised quite often; what has not been stressed, however, is that the typology and ultimate functional classification of idioms on the basis of the types and degree of these restrictions seems to be not only feasible but, in fact, the most exhaustive and reliable one, too. This, yet again, is a consequence of our not being able to decompose the idiom while its behaviour is easily observable and can even be measured by comparison with regular language units. One of the greatest tasks for a future research is to investigate relationships between this and the fourth universal.

**(8) Idioms are transformationally always defective and restricted to a degree.**

Also this feature has been repeatedly raised before (about transformational restrictions, i.e. possibilities of a conversion of the idiom into a different functional entity, mentions



are to be found as early as in 1970, Fraser). Basically, this suggests, for example, that not all verbal-type idioms allow for a nominalization to be formed, etc. However, it is important to view this feature, just like the preceding one, as scalar in its nature, which might help explain differences between various types of idioms. A popular but rather vague label used in past here is rigidity of the idiom. Restated as an implicational universal it might go like this:

*The fewer functional and transformational anomalies (defects) of the idiom there are, the closer it is to a regular expression (if there are none, it is not idiom).*

It might well be that a future research will lead us to other implicational universals here, especially in connection with types and number of these transformational anomalies. An interesting possibility, open to a further investigation, is, of course, a potential unification of this and the preceding universal.

**(9) In their use in communication, idioms are language nominations which are both economical and denotationally as well as referentially vague.**

Being a natural counterpart to terminology whose meaning is sharply defined, idioms serve the speakers for the very opposite goals to those of terms, namely, the need to be deliberately **imprecise**, non-specific, broad, noncommittal and flexible in one's formulations as well as to be **readily at hand** in spontaneous, unprepared and often fast communication. Much of this corresponds to the Jakobson's phatic function, in fact (Jakobson 1960), which is basically nonreferential. Observation supports also the view that, for communication's sake, there is a direct relation between the idiom's form and its nature. Rephrased and expanded as an implicational universal this universal says:

*The shorter and/or more archaic (monocollocable, "nonce-word-like") the constituents are, the greater the idiom's polysemy and/or its vagueness.*

**(10) Idioms are systemic means (nominations) which are most readily at hand for pragmatic communication of the speaker's evaluation above all.**

This universal – in contradistinction to commonly held view that certain functions arise in the text only – postulates existence of **ready-made pragmatic means** in the system; more broadly, it specifies the particular systemic part of the language (vocabulary) corresponding to Bühler's Apell function (or connative function of Jakobson), whose character is clearly nondesignational.

Let me add that this particular universal might refer to basically the same thing as other approaches under the label of *expressiveness* which itself is definitely not universal in idioms. But any discussion and further elaboration of this point would depend, again, on a careful investigation of the unique correlations of all the syntagmatic and paradigmatic aspects involved here and on discerning their particular anomalies for each idiom.

### 3. Conclusion

A necessity to corroborate these and other universals has been stressed above; this is no different from what is needed for any other alleged universals anywhere in the language. At present, only very little experience has been accumulated from more than one language which is partly due to difficulties in idiom's identification across language borders, an undertaking a foreigner is specifically handicapped in. Next to corroboration, another task we are likely to face in future in this respect, is to reconcile views of universals with those of typology, being more realistic and closer to actual languages at hand, and trying to fit the idioms into this unified framework.

---

#### **Shrnutí**

#### **O povaze univerzálního ve frazémech**

Srov. velmi podobnou českou verzi *Povaha univerzálního ve frazeologii a idiomatice*.

# Povaha univerzálního ve frazeologii a idiomatice

(In: *Frazeológia vo vzdelávaní, vede a kultúre. Phraseology in Education, Science and Culture*, 1993, s. 45–52)

## 1. Úvod: Frazém a univerzále

Už i jen zběžné zamyšlení nad tématem, resp. vztahem **frazém a univerzalita**, navozuje dvě otázky, na které je třeba dát odpověď před každou další úvahou, totiž otázku pojetí frazému, resp. idiomu, a pojetí univerzálního a univerzále v jazyce.

Je víc než zřejmé, že v prvním případě tu mezi lingvisty nejsou jen pouhé rozdíly v konvenčním slova smyslu, ale rozdíly tak zásadní, že se různými lingvisty zcela různě vymezovaná oblast frazeologie a idiomatiky neobyčejně liší a její hranice se jak proměňují, tak se ve většině teorií stávají nezřetelné a není jasné, co sem patří; v krajních případech je pak společný průnik oblastí v různých koncepcích mnohem menší než to, co je rozdílné. Má to pak za následek téměř katastrofální nedorozumění, protože pod stejnou nálepkou frazému či idiomu se skrývají věci diametrálně odlišné. Pro pásma sporného, pokud je vůbec jako sporné přiznáváno, se pak užívají taková označení jako *collocation* v britské lingvistice; většinou se však sporné, problematické zamlčuje či zamlžuje. Nemám přitom na mysli problematičnost, která vzniká v důsledku dobře pochopitelné, ale marné snahy špatně vymezenou oblast frazeologie a idiomatiky ještě dál uvnitř dělit podle nepoužitelných a zcela vágních kritérií typu významu, jako to dělá například Šanskij a po něm řada jiných.

V důsledku této situace se pro další úvahy přidržím vlastního vymezení frazému, publikovaného už víckrát a ověřeného na vyčerpávajícím frazeografickém popisu celé oblasti frazeologie, který se promítl do tří svazků Slovníku české frazeologie a idiomatiky (Čermák et al. 1983, 1988) a začíná se promítat do svazku čtvrtého. **Frazém a idiom** je pak *takové nemodelové a ustálené syntagma prvků, z nichž aspoň jeden je z hlediska druhého členem extrémně omezeného a zavřeného paradigmatu (formálně a většinou i sémanticky)*. Dodejme jen, že z tohoto vymezení mj. vyplývá jak předpoklad široké existence frazémů a idiomů na různých jazykových úrovních, tak primární důraz na funkci, a nikoliv sémantiku, jak se dosud běžně dělá.

Pokud jde o druhý výchozí pojem univerzálnosti, budou se mé úvahy točit spíše kolem staršího pojetí univerzálií; řadu z nich lze převést i na Chomského univerzále formální (Chomsky 1965).

## 2. Přístupy k posuzování univerzálnosti frazému

Dosavadní pokusy o vystižení univerzálních rysů u idiomu a frazému neuspokojují z různých důvodů (především Makkai 1978, Coulmas 1981, Kopylenko – Popova 1972, Dobrovolskij 1988), a to ještě jiných než neprokazatelnosti tvrzení jednotlivých autorů, která by měla být podporována reprezentativním materiálem – taková je řada univerzálií postulovaných i mimo frazeologii, jakkoliv i tam zůstává takové ověření fakty požadavkem zcela základním a v tomto smyslu pak jeví stejné nedostatky i Chomsky, který taky jen postuluje a nedokazuje (na rozdíl od staršího Greenberga 1966, který určitý výzkum provedl). Námitky proti těmto přístupům lze shrnout takto: Tyto přístupy

- (1) jsou **diachronní povahy**, pokud se mluví např. o fenoménu potlačeného tabuového významu, který se promítá do vedlejších významů lexému (tzv. Thass-Thienemann/Fabricsiův a Kovacs/Kronasserův fenomén, Makkai a zčásti i Coulmas) – přitom však mají být povahy synchronní;
- (2) jsou **frazeologicky nespecifické** a tudíž nikterak omezeně a konkrétně aplikovatelné jen na frazémy či idiomy samotné, ať už v izolaci či kombinaci (Dobrovolskij, zčásti Makkai i Coulmas). Některá postulovaná univerzále mají pak nebezpečně blízko k banalitám (jako když se např. u Dobrovolského konstatuje jako univerzální výskyt polysémie, homonymie či synonymie). Naopak mají frazém vymezovat na rozdíl od jiných jazykových útvarů dostatečně a jednoznačně;
- (3) jsou **orientovány na povahu komponentů**, a nikoliv vlastní frazém (Kopylenko – Popova a zčásti i Coulmas) – očekává se však naproti tomu, že budou vypovídat o idiomu jako celku, útvaru. Co se naopak v této souvislosti nabízí jako téměř samozřejmé univerzále po stránce **sémiotické**, se nekonstatuje: totiž to, že přestože velká většina komponentů frazému má povahu konkrétní, především substantivních, je obrovská většina na vlastních frazémů naopak povahy nekonkrétní, abstraktní a denotáty konkrétní tu vyvolávají vždy u jednotlivých teorií rozpaky či potíže co do jejich statusu.
- (4) jsou sporné už svou podstatou (jako např. Dobrovolskij, když postuluje závislost pravidelnosti frazeologického systému na míře analytičnosti jazyka). Vychází se tu z mylných předpokladů existence **frazeologického systému**, nedefinovaného pojmu pravidelnosti frazeologie, který se zdá být protimluvem už sám o sobě a zcela se ignoruje nutnost úměrného posouzení paradigmatických poměrů frazeologických komponentů na základě analogie s pravidelnými (o tom viz např. Čermák 1982 a 1987, srov. i opačné stanovisko vyjádřené níže v univerzále třetím). Obecně pak ukazují na samo nezřetelné východiskové pojetí frazeologie a idiomatiky.

## 3. Deset univerzálií: návrh

Je tudíž zřejmé, že povaha univerzálií, která hodláme ve frazeologii hledat, musí být jiná. Metodologicky je pak možné je vymezit jako **univerzále synchronní**, jevově specifické a postihující frazémy a idiomy samotné, a nikoliv jejich předpoklady, ať už jde o kompo-

nenty či výchozí podmínky. Pokud vůbec chceme o jejich vzniku uvažovat, pak – má-li to být úvaha synchronní – jen v rámci obdobném slovotvornému, i když i v slovotvorbě synchronně pojímané jsou nepřiznaná rezidua diachronní (zvláště v neřešeném, téměř magickém pojmu produktivnosti), kterým je ovšem třeba se pak vyhnout.

V dalším se pokusím velmi stručně navrhnout a prezentovat deset základních univerzálií patřících typově k absolutním. Podle potřeby se k nim pokusím, kde se to bude nabízet, zformulovat i některá univerzália implikační jako jejich určité konsekventy. Za univerzále nebudou však považovány takové jevy, které se vztahují obecně ke každé nominaci (kam patří i frazém a idiom), jako je synonymie ap., pokud specificky nevyplývají z dané přítomnosti či konstelace určitých rysů jiných. I když jsou v mém pojetí termíny **frazém** a **idiom** vzájemně zaměnitelné (zatímco termín *frazém* se váže na formu a funkci útvaru, termín *idiom* respektuje v poněkud užším pohledu jeho stránku anomálně sémantickou), v následujícím se bude pro stručnost mluvit vždy jen o jednom z nich.

## 1 – Frazém existuje v každém přirozeném jazyku.

O životaschopnosti frazému a jeho potřebnosti i o tom, že může dokonce vzniknout i v jazyce umělém, svědčí případ esperanta, kde je nyní už jeho existence prokázána. Co naopak nebylo zatím prokázáno, je neexistence frazému v nějakém konkrétním jazyku. Je užitečné si připomenout, že frazém je ve většině případů – což je poznámka diachronní – jen specifickým pokračováním existence metafor v jazyce, které se v něm ujaly a zůstaly v něm, protože i metafora má v jazyce univerzální výskyt. Bylo by jistě možné už zde postulovat implikační univerzále, které by znělo asi tak, že „*Má-li jazyk metaforu, má i frazémy*“, bylo by však velmi banální.

## 2 – Frazém existuje na všech úrovních kombinatoriky jazykových jednotek, které jsou nadány významem a jejichž rezultativní kombinace mohou jako ustálené vstupovat do systému.

Obecně platí, že kombinatorické, resp. permutační možnosti jednotek libovolného jazyka jsou teoreticky, potenciálně mnohem větší než možnosti realizované a využívané pro skutečné a pravidelné kombinace v tom kterém jazyce. Frazém pak čerpá vždy principiálně z těch možností, které jsou nerealizované a v některých případech i z těch, které nejsou ani realizovatelné (např. u kombinací dvou slovních druhů, které se obvykle navzájem setkat nemůžou).

Nejasná je v této souvislosti možnost uvažovat o frazému, idiomu parolovém, vytvořeném pro potřebu ad hoc a odlišit ho přitom od prosté metafory či metonymie. Zdá se, že pojmově takový parolový idiom a metafora jsou identické a není ani třeba ani možné je lišit. Obecně se tak potom nastoluje potřeba vymezení a precizování pojmu ustálenosti a v případě její neúplnosti i pojmu variantnosti. Dodejme, že **variantností**, kolem které je mnoho nedorozumění, je třeba rozumět takové oscilování formy frazému, které neporušuje jeho identitu, tj. především syntaktickou funkci, jinak je nutné už uvažovat o **transformaci** a tím i o přechodu do frazému jiného. Implikační univerzále by tu tudíž mohlo znít:

*Čím frekventovanější frazém je, tím je ustálenější a tím menší je i pravděpodobnost výskytu jeho variant.*

Je zřejmé, že právě v této stránce, a to jak (1) u nejčerstvějších neologismů či (2) frazémů nadužívaných, tak (3) frazémů velmi řídkých a zastarávajících, sem zasahuje diachronie a rozkolísává i mění jak úzus, tak formu. Obecně je to tudíž jev souvztažnosti centra a periferie jazyka.

### 3 – Frazém je výsledkem náhodného, nemodelového tvoření.

V minulosti se s pojmem **modelu** a modelovosti zacházelo dost nezodpovědně a vznikal při jejich užití dojem možného manipulování s jazykem podle své vůle a tvorby frazému, což je ovšem nesmysl. Nikdo nedokáže frazém utvořit à la thèse, kdy se mu zamane, třeba jen pouhou nápodobou jiného, a domnívat se, že ho ostatní lidé nutně přijmou a ustálí, čímž se dostane do slovníku. Toto je čirá utopie a je proto před územ slova *model* v tomto smyslu varovat. Od mylného pojetí modelovosti pak je už jen krok k absurdnímu tvrzení, že ve frazeologii lze hledat nějakou pravidelnost. Pravidelná jsou pouze frazeologii externí východiska, na kterých buduje, nikoliv však v ní. Řešení zdánlivého problému krajně omezených analogií tu přináší pojetí **paradigmatu zavřeného** (srov. Čermák 1982, Filipec – Čermák 1985). Něco jiného je ovšem autorské alternování či variování formy nebo sémantiky či funkce frazému ustáleného, které se blíží výše uvažovanému idiomu parolovému; do systému ovšem nepatří. V širším pohledu je pak frazém, který náhodně vznikl, přirozeným výrazem všudypřítomné entropie v jazyce; je totiž nesmyslné o entropii uvažovat pouze ve sféře komunikace a tedy parole a nevidět ji i v systému, resp. podmínkách vzniku jeho jednotek.

### 4 – Frazém je založen na nejružnější vnitřní syntagmatické, paradigmatické nebo obojí anomálii svých komponentů.

Posouzení tohoto definitorického rysu frazému, **anomálie**, je možné jen na složitém a vnímaném či analyzovaném pozadí všech vztahů a poměrů paradigmatických a syntagmatických, na pozadí členských i kolokačních paradigmát každého komponentu i jeho syntagmatické kapacity; ty mohou být navíc jak formální, tak sémantické povahy. Případné vybočení, výjimku, posun či porušení v chování či postavení frazémového komponentu z těchto **pravidelných** poměrů, které může mít mnoho podob a být obvykle i vícené, je třeba tudíž vnímat a posuzovat jako anomální a pro frazém konstitutivní. Ne nadarmo je oblast idiomatiky tak zdůrazňována jakožto obtížná právě při cizojazyčném vyučování a působí v učení těžkosti, protože složité předivo vztahů obojího řádu a identifikace anomálie na jeho pozadí je analyticky rozpoznatelné jen rodilým mluvčím (srov. Čermák 1987 aj.).

Aspoň pro ty nejzřejmější případy se pak nabízí možnost zformulovat následující implikační univerzália postihující vztah mezi frekvencí a typem významu komponentu na jedné straně a sémantickou „průhledností“ a homonymií na druhé straně:

*1 – Čím je komponent frazému frekventovanější, tím nepravděpodobnější je jeho sémantická souvislost s jeho slovníkovými významy v izolaci.*

- 2 – Čím je monokolokabilní komponent frazému (pokud ho má) řidší a/nebo méně sémanticky průhledný, tím je jeho další komponent frekventovanější.
- 3 – Je-li mezi komponenty frazému sémantická kompatibilita, je frazém homonymní se svým doslovným protějškem.
- 4 – Jsou-li komponenty frazému jen monosémní či jen polysémní, je frazém homonymní se svým doslovným protějškem.

**5 – Frazém je dán neaditivní, nefregeovskou funkcí (popř. významem) svých komponentů.**

Toto univerzále předpokládá zřeknutí se dvou snad nejoblíbenějších a řadou lingvistů přímo hýčkaných témat: **sémantické motivace** frazému skrze jeho komponenty a **frazeologický vázaného významu komponentů**. Všechny úvahy o prvním, tj. sémantické motivaci, totiž demonstrují motivaci vždy jen na některých komponentech a nenamáhají se stanovit povahu komponentů zbývajících, které komponentovou, sémovou či jinou analýzou jsou exaktně neurčitelné a neuchopitelné a tvoří zcela umělé a neurčité sémantické torzo. Tento stav je dán prostě tím, že lingvista mylně předpokládá prostou sumu významu tam, kde být nemůže (to by to nebyl idiom), a frazém mylně rozkládá jako každou jinou jednotku. Neznám lepší obdobný případ z jiné disciplíny, který by situaci idiomu přiblížil, než je biologická mutace: ani tam přece nikoho nenapadne mutace a mutanty rozebírat přímo a mechanicky. Sémantická motivace vycházející ze vztahu komponent-útvár je tudíž mýtus a zjevné kvazimotivace je třeba řešit jinak.

Stejný mýtus, odrážející stejně jako u motivace smutný stav naší neznalosti a absence přesných metodologických a analytických nástrojů, je i mylně postulovaný frazeologický vázaný význam komponentů (který by pak snad takový lingvista mohl ztotožnit s výše naznačeným sémantickým torzem, reziduem). Význam komponentu, tj. samostatného komponentu a tím i tento komponent sám, lze stanovit pouze na základě dostatečné analogie v opakovaném výskytu a i když na tomto základu stojí všechny jazykové jednotky včetně slov, nikdo dosud tuto oblast bližšímu zkoumání nepodrobil; i tak je však jasné, že izolovaný a tedy neanalogický výskyt neznamená v jazyce nic, o čemž mluvil už de Saussure (de Saussure 1989, s. 145).

**6 – Frazém je svou funkcí analogický odpovídající funkci pravidelné kombinace pravidelných prvků své úrovně.**

Frazém je specifický, pevný a nedekomponovatelný útvar, gestalt a jako takový je taky vnímán jako holistický celek, o čemž svědčí i psycholingvistické výzkumy (srov. např. Schweicert a Moates 1988). Jde tedy u něj především o funkci syntaktickou, která se analogicky zachovává, takže frazémy verbální se chovají ve větě jako verbum ve své primární funkci, tj. nositele predikátu, frazémy nominální jako nomen či frazémy interjekční jako propozice apod., a to i v tom případě, že komponentem frazému může být útvar z vyšší úrovně, zvl. pak u frazému kolokačního, nevětného, se může jako komponent vyskytnout i věta. Je zřejmé, že je tak zároveň i dána funkční klasifikace frazémů, jejichž typů je tolik, kolik rozeznáváme slovních druhů a vyšších jednotek. Z univerzále pátého (výše) vyplývá pak ovšem řada omezení, především nemožnost frazém syntakticky analyzovat na komponenty

a určovat např. u frazému verbálního objekt či u nominálního atribut apod. Nejasná zůstává otázka vedlejších funkcí slovních druhů a možnost jejich výskytu u frazémů. Tyto a další aspekty by pochopitelně bylo možné formulovat i jako implikační univerzále.

### 7 – Frazém je funkčně a gramaticky vždy do určitého stupně defektní a restringovaný.

Toto univerzále je specifikací univerzále pátého a vystihuje známý poznatek, že všudypřítomná anomálie se projevuje i tak, že frazém nemůže v rámci své funkce realizovat některé funkce, které bychom analogicky očekávali (např. roli nominálního frazému jak v subjektu, tak predikátu věty nebo tvorbu pasiva nebo plurálu, pokud je nebudeme považovat za transformace, o těch však dál). I zde by jistě bylo možné jednotlivé případy rozepsat do formy implikačních univerzálií. Dodejme, že právě v tomto rysu, který představuje tisíce konstelací, se nabízí nejsolidnější **základ pro klasifikaci** frazémů vůbec. Jedním z největších desiderat budoucího výzkumu frazeologie a idiomatiky je myslím prozkoumání souvislosti tohoto univerzále s frazeologicky konstitutivním univerzále čtvrtým.

### 8 – Frazém je transformačně vždy do určitého stupně defektní a restringovaný.

Jasně omezená schopnost frazému transformovat se do něčeho (funkčně) jiného už byla konstatována víckrát (zvl. Fraser 1970) a připomeňme si, že toto i předchozí univerzále se skrývá za často užívanou vágní nálepkou ustrnulosti frazému. Protože jde v obojím případě o stupeň, resp. **škálu**, lze dospět k následujícímu implikačnímu univerzále (jde tu vlastně o dvě v jednom):

*Čím má frazém méně funkčních a transformačních anomálií (defektů), tím blíže má k pravidelnému výrazu (nemá-li žádné, není to frazém).*

Dodejme jen, že bližší výzkum může přinést možnost formulace dalších, dílčích implikačních univerzálií i zde, především v souvislosti s typem a množstvím transformačních defektů a anomálií. Obecně je u tohoto univerzále ovšem úzká souvislost s předchozím a není zatím jasné, zda je nelze přeformulovat do jediného.

### 9 – Frazém je svým územ v komunikaci nominace ekonomická a denotativně i referenčně vágní.

Frazém je v systému ve smyslu, o kterém mluví toto univerzále, pravým opakem termínu. Právě proto, že je tak **výhodně vágní** a nabízí jak možnost si v komunikaci příliš nezadat, příliš se neangažovat vysokou přesností formulace, tak možnost pružné adaptace a posunu individuálního smyslu, je dominantní v mluvené a tudíž nepřipravené komunikaci. Zároveň frazém touto stránkou zřetelně vyhovuje potřebám fatické komunikace, jak o ní mluvil Malinowski (srov. zvl. Jakobson 1960), která je primárně nereferenční.

I toto univerzále je tedy zřejmé a třebaže by ho bylo možné potvrdit prostým statistickým porovnáním s analogickými nominacemi pravidelnými, není to cíl nejnaléhavější. Implikační univerzálie se tu zatím nabízejí tato (opět jsou dvě v jednom):



*Čím víc jsou komponenty frazému krátké a/nebo monokolokabilní, tím víc polysémie frazému roste a/nebo se oslabuje ve prospěch zesilující se vágnosti frazému.*

#### 10 – Frazém je nejdisponibilnější systémový prostředek (nominace) k pragmatickému vyjadřování, především evaluativního typu.

Toto univerzále vystihuje skutečnost, že jazyk už v systému má prostředky k pragmatickému vyjadřování postojů a hodnocení a tedy aspektů nedezignativního typu, obecněji pak je výrazem potřeby realizovat apelovou či konativní funkci jazyka (srov. Jakobson 1960). Třebaže je vzhledem k východiskům tohoto pojetí frazeologie třeba univerzální existenci expresivity ve frazeologii odmítnout, není jasné, nakolik by bylo možné ji jako výraznou tendenci, resp. relační univerzále, které je jiného řádu, propojit právě s tímto, popř. předchozím univerzále. Její statut by ovšem bylo nutné precizovat a objektivizovat a opět opřít o pojem anomálie, odchylky jak paradigmatického, tak syntagmatického rázu.

### 4. Otevřený závěr

O problému **ověření** navržených univerzále, která lze dále rozmnožovat, jak jsem občas také naznačil, jsme už hovořili, je stejný jako u všech jiných univerzálií; v současné době se opírá jen o dílčí zkušenost s různými jazyky. Závažnou, a to i jako program do budoucnosti, ovšem zůstává dvojí problematika: ověření, resp. precizace vstupní koncepce frazeologie, která se musí zbavit balastu a propojit se zbytkem lingvistiky v některém z převládajících paradigmat a zkoumání souvislostí univerzálií s teorií typologickou.

---

### Summary

See *On the Nature of Universality of Idioms*, elsewhere in this volume, covering much of its substance.

# Universals, Typology and Idioms

(In: *Proceedings of LP '90 Conference Linguistics and Phonetics: Prospects and Applications*, 1991, p. 261–269)

## 1. Introduction

Since idioms seem to be present in every language investigated so far, a rather recent idea of their universal existence has been no surprise. This empirical conclusion is supported by the fact, not so often admitted, that every single language theory has, at some point, run into insurmountable difficulties which are then brushed aside as either exceptions or idioms. First theoretical approaches to the universality of idioms (Makkai 1978, Coulmas 1981, Kopylenko – Popova 1972, Dobrovolskij 1988) rather differ from each other; however, all of them try to capture the reason or effects rather than the **substance** of what might be universal in idioms.

Thus Makkai's **multiple reinvestability principle**, based on diversification and neutralization is both too broad, general and vague a concept to be specific for idioms only; moreover, it seems to suggest an exchange of one type of meaning for another which is not the case here; his choice of the Thass-Thienemann/Fabricius Kovacs-Kronasser phenomenon of **repressed taboo meanings** being obsessively repeated in adjacent senses as the explanation of the likely underlying psycholinguistic mechanism seems too psychopathologic for such a common phenomenon as the idiom. While this basically psycholinguistic and diachronic approach is in need of a further research, the approach of Kopylenko – Popova, which is allegedly universal and thus suitable for any language, explains only the **semiotic status** of the idiom's constituent parts and raises, basically, no doubts in this respect. What this approach does say, however, is little more than a commonplace truth about apparent semiotic stratification in the denotation region (which could be, no doubt, further elaborated and refined).

Dobrovolskij's pretentious book on universals in language does very little to justify his general claims and postulates by argument or analysis. Indeed, some, being patent commonplaces, **might not be called idiom-specific universals** at all (polysemy, homonymy, synonymy, irregular distribution of idioms in the referential region etc.); other, as his claim that the more a language is analytical in its nature the more regular its idiomatic status is, are badly in need of proof. In fact, such claims like this raise a number of basic theoretical questions as to the status of idiom and of the idiom system, if any, as to **what is regular and irregular** and so on. Some of these points are, finally, raised by Coulmas. His answers are,

again, quite **broad and indirect**; idiomaticity is due to the language sign being arbitrary, while its universality is to be sought in diachrony, in the **unpredictable development of language**. Coulmas's contribution is important, however, in that it frankly admits the state of our ignorance in many of these matters and raises many open questions stimulating further research.

This brief review of major approaches is by no means exhaustive; in fact, passing comments on the issue can be found with other linguists as well, none of which is offering, however, much more than a claim, expectation or outright guess. It seems, thus, that the very notion of idiom is pretheoretical, being used in virtually every major school of thought or trend, sometimes under different headings, such as **phraseme** in most of the Soviet views. A major disadvantage of a pretheoretical term is its misleading familiarity which must be painstakingly dissipated by defining it over and over again, yet always differently.

## 2. The Idiom: Some of Its Basic Features

There is precious little, except for one feature, that could be viewed as a kind of consensus here as to what idioms actually are. Yet the very single exception, namely the claim that the idiom's essential quality is to be sought in its inability to allow for a semantic decomposition, or vice versa, in the impossibility to derive its meaning from its constituent parts, can be questioned or, to be more precise, basically reformulated. Elsewhere (Čermák 1988) I have shown that an alternative and more general functional approach is to be preferred, as it both explains cases left unattended and tackles, above all, the core of the idiom's problem itself. Since there are many idioms functioning as or being made of grammatical words (*as to*, *as for* = idiomatic prepositions) as well as idioms based on constituents that do not exist as words (*kith and kin*, *beyond my ken*) it is the **function of the constituents**, not their meaning, that is put to test in the examination of the interrelationship of the relevant syntagmatic and paradigmatic aspects of the idiom's constituents.

I have tried to demonstrate that the defining property of any idiom is always some kind of **anomaly** of at least one of its constituents in so far as its paradigmatic and/or syntagmatic aspects are concerned. It is postulated that the normal status of words and other items of the system usually allows for a rather high degree of combinatorial capacity (= syntagmaticity) and (functional) membership (= paradigmaticity) of these words (items) of the same order. Any pronounced deviation, limitation of these properties, seen as a system anomaly and as a very limited combinatorial capacity or collocability of such an item, makes it thus a part of an idiom. Indeed, a consequence of an observable anomaly is the item's suppressed, or rather, undeveloped word status, meaning etc. making one uncertain as to what its status, meaning and behaviour might be, once the "word" is viewed outside of its "idiomatic context". To sum up, the **idiom** is, in this approach, *such a unique and fixed combination of at least two elements for which it holds that at least some of these do not function, in the same way, in any other combination of the kind, or do function in a highly restricted number of them only*. This approach, relying heavily on combinatorial analogy and anomaly, is open to a very simple test of substitution, and leaves thus no areas of idioms unattended, specifically those which might escape more traditional approaches.

As these combinatorial or syntagmatic and class or paradigmatic aspects are primary and universal in any language system it seems that it is their anomalous interplay which is in the substance of the idiom and which can become, consequently, the universal criterion for the judgement on **idiomaticity**. This very basic property gives thus rise to all of the other idiomatic qualities, such as the lack of compositional function between constituents' meanings, their transformational deficiency, or formation based on no model which allows then, basically, for no exact repetition or replica in another idiom. It is in this sense only that idiom may be viewed as a pretheoretical concept, i.e. as independent of any theory, though specifiable by it; the syntagmatic and paradigmatic relations happen to be general in their nature despite the fact that some major theories seem not to have any use for them.

Let me mention some of the consequences of this approach. As the number of possible anomalies, both syntagmatic and paradigmatic, based either on the idiom's form, function or meaning, is very large and they occur in most peculiar combinations, the resulting scale and scope of a possible idiom typology is very large, indeed, ranging from semi-idioms (*swallow one's pride, white lie*) to fully frozen ones (*kick the bucket, red herring, travel the light fantastic*). On the other hand, as only stable idioms based on fixed combinations are currently viewed as idioms, they are, as specific naming units or denominations, part of the language system. Because idioms are language denominations just like words or multi-word technical terms etc. they **do not form an independent system** of their own, which seems to be one of many popular misconceptions held about idioms. This is supported by the very fundamental combinatorial principle consisting of both analogous and anomalous combinations, the former being reserved for the regular language use in la parole mostly, the latter being the basis for idioms in la langue. An alternative view of a system here would run into trouble at every single stage of its establishment, as there is no analogy to rely on anywhere: a multitude of exceptions is hardly a system. Here again, our attention is returned to the state of affairs lying actually outside the scope of idiomatics proper, to the anomalous status of the constituents which, in turn, can be viewed as anomalous only by virtue of the contrast with the regular and normal syntagmatic and paradigmatic situation of the language.

The idiom, in general, is an entity, a whole made up of its form, meaning and function. Of these three, it is the **meaning**, which has traditionally drawn most of the attention to itself, its nonadditive, non-motivated, but integral character which is not be derived from the constituent meanings, being a basis for a widely accepted definition. Let us recall, however, that this may not be a general feature and has to be replaced by a functional alternative, as explained above. The trouble with anomaly is that it gives the impression of being both unique and, perhaps, uniform, which is a false impression, of course. In fact, there are many kinds of anomaly, different in both quality and quantity, as well as concatenations of it, which also explains why some idioms may be semi- or fully transparent, but some, further down the scale, may not. Perhaps, the idiom may be compared, by approximation, to mutation in biology, which, like analogy, results in the creation of an irregular and often higher entity, too. Likewise, it is a product of an accident, which is linguistically based on a new, accidental and originally ad hoc (metaphorical) combination of free forms. It is, then, hard to see it, as Makkai (1978) does, as a psychopathological phenomenon of repressed taboo meanings (usually sexual) which are being obsessively repeated in adjacent senses, the claim which

may be, it seems, statistically disproved if the etymology of the most frequent constituents is gone into (which, by the way, is the case of Makkai's own example of the most idiom-prone English verb *to get*). Semantically, there seem to be three basic types of idioms:

- (1) idioms based on metaphor,
- (2) idioms based on metaphor and allowing for ambiguity and homonymy, and
- (3) idioms based on non-metaphors in the traditional sense, i.e. such as not allowing for any meaning to be superficially read into them.

These three types are not identical with another tripartition grouping idioms by their **denominative function** and power:

- (A) idioms basically synonymous with a single word (other type of item), both of which are currently used,
- (B) idioms synonymous with words that are rare, which makes the idiom the primary denomination (*take over: assume*),
- (C) idioms having no natural synonym (*to beat around the bush, to get even, to force one's hand*). To fully appreciate this means to acknowledge the fact that idioms, especially those belonging to the last type, are neither something secondary, nor an unnecessary complication in language, something one can do without. This, in turn, suggests how difficult the idiom interpretation might be and how poor and oversimplified its definition given in dictionaries usually is.

As their semantic status is neither always nor completely clear-cut because idioms allow for a lot of textual, actualized reading – which is not to be mistaken for some kind of shift of meaning, that is to say, as idioms are characteristically vague up to a high degree – their notional and denotational aspect is, in general, far from clear, too.

### 3. Idiom: Its Function

This last feature is closely connected with the idiom's **semiotic function**. By virtue of its partially vague meaning, making possible a number of potential and specific uses, and of the potential echoing of the constituents' meanings, which both require the speaker's awareness awakened with a fine tuning in the desired shade of meaning, idioms are ideal, **ready-made and handy pragmatic tools** of every language. Though their **evaluative function** is familiar and dominant (the prevailing negative evaluation, though conspicuous, is often exaggerated, it seems), it is not the only one. In fact, the whole domain of human customary and honorific contact and communication, corresponding to the phatic communion in Malinowski's term, is heavily burdened with idioms. It is hard not to see here a relation between the semi-automated nature of this kind of communication, being repeated over and over again, and the automated nature of these ready-made, non-generated idiomatic labels. If compared with the alternative, non-idiomatic ways of expressing the same meaning, i.e. in so far as these alternative ways do exist, idioms come out as very handy, relatively short and **economical** tools, which are, moreover, often elegant, witty or interesting, to add the least.

In their **structural function**, idioms, being lexicologically denominations and on a par with any other, come, basically, under yet another tripartition. Either they are

- (a) ready-made utterances of a propositional character (= **propositional idioms**) and are thus equivalent to sentences;
- (b) they are combinations of word forms (= **collocational idioms**), which is the traditional core of the idiomatic stock of any language functioning, primarily, as ready-made predicates (a fact easily confirmed by statistics) based on verbs in the first place. The third functional type
- (c) based on **intra-word** anomalous combinations of its constituents, consists of words (**lexical idioms**), most often compound words; this type, least explored and still discounted by many, seems to partially merge with the second in such languages as Chinese where there is no clear-cut boundary between collocation and compound. It seems that further exploration of this offers some interesting possibilities. Any higher type of idioms, i.e. above the propositional type (a) is, generally, not postulated.

Next to **variants** (*beat about/around the bush, fall/land on one's feet*) preserving the idiom's identity, there are also **transformations**, creating either a new idiom out of an old one, or changing one and the same idiom grammatically. The former type of transformation (as in *keep one's head – lose one's head, put it into one's head – take it into one's head*, see Čermák 1988, Filipec – Čermák 1985) leads to a pair or group of structurally and functionally correlated idioms while the latter takes place when the idiom is used in different structural ways, such as in plural, passive voice, past tense, imperative mood, or as a subject, adjunct, attribute etc. Here again, anomalies of many kinds are recorded (cf. Fraser 1970, for example) and this, indeed, seems to be another universal phenomenon, as there is always some kind of deficiency or anomaly involved. Yet, it is an anomaly of a higher order, one based on the idiom's function as opposed to the syntagmatico-paradigmatic anomaly of its constituents.

This, finally, takes us to the idiom's **form**. If the tripartition relating to the idiom's function is to be upheld, one of the major, often not verbalized views that idiom is a **discreet** combination of constituents is, it seems, no longer valid. Without going further into this, it is enough to acknowledge that this may not be a general feature, after all. Indeed, since all the typical formal features, such as idiom's relative shortness, i.e. both of it as a whole, and of its constituents, and its stability enabling one to reproduce it ad libitum, are in no way special for idioms only, there seems to be no single formal feature to hold universally. This finding supports, again, the priority accorded to the idiom's function, both internal and external.

Due to idioms being functionally complementary to other, regular denominations in language, they do not seem to form a formal system of any kind; this specific subsystem is neither a closed one, nor – due to its pronounced semantic potentiality (see above) – a stable one; also its internal structure of any kind, apart from infrequent structural transformations, is non-existent. What seems to be systematic about idioms, is, paradoxically, the very fact that they exist.

#### 4. Idiom: Its Formation and Capacities of the Language System

This should not be surprising. Although there are no reliable measurements of language redundancy and entropy available so far, it is clear that they are very high in many respects.

The actual number of combinations (permutations) which words of a language do contract, to take only one example, is much smaller than the potentialities never or seldom used. If, then, an accidental, entropic combination occurs and happens to become generally accepted and stabilized in use – all of this as a n-th successful trial out of many unsuccessful and discarded ones – the event does resemble a kind of natural selection, so to speak. In other words, out of many trials, failures and dead-ends resulting in nonsense or nonce-form, few are lucky enough to acquire a new meaning and win a general approval. Of course, idioms are results of language creativity, in a sense, but it is a creativity of a hit-or-miss kind rather than a straightforward and conscious one. Idioms are, thus, both **redundant** once a hitherto unused combination pattern is employed next to the usual ones, and **entropic** in the haphazard manner in which they are formed, a manner following no plan. This, then, explains the well-known fact that idioms cannot be made, coined and tailored to suit the needs of a moment.

What happens once an idiom is accidentally formed, i.e. once out of combinatorial impossibility a distinct and promising possibility is created and a nonsense makes suddenly sense, is an equally remarkable and universal process. Because a new idiom is, by consensus, given a meaning and a function, it has become **a new sign**. But it is a very peculiar sign where the services of both standard sign orders of the language at disposal are avoided, and an arbitrary sign of a new order is formed. Even though this process employs signs of the second order of the language symbolization, i.e. words as its input, in the output these signs cease to exist and give way to an amalgam, a sign belonging, in most cases, to the **tertiary symbolization of language** which is not to be found elsewhere. By the same virtue, one might consider this process to be a way how to establish a rudimentary treble articulation in language, to paraphrase the Martinet' term.

The bulk of linguistic descriptions and theories seems to be concerned with **regularity**, rule-governed processes and suchlike phenomena, i.e. with the Fregean principle of combination where its meaning is the function of the constituent meanings. While I do not subscribe to Coulmas (1981, s. 32) in putting the opposite, i.e. the non-Fregean functioning on the same footing and seeing this possibility as equally important, I do think that this area as a whole – being, incidentally, very much larger – requires a systematic and concerted approach. In the wide scale ranging from downright nonsense (as of today), purely theoretical but not feasible possibility, to practical possibility, or possibility already employed by some, though not accepted by other, idioms seem to occupy a rather modest, though a special place. It seems, therefore, that the idiom is a defining property of natural language with its familiar properties based on omnipresent anomaly. It is a result of an interplay of several factors:

- (I) of too large combinatorial capacity unused,
- (II) of language creativity drawing on analogy and similarity, above all, and
- (III) of the human capacity to view a functionally recurrent item as a sign.

In a sense, this phenomenon may, too, be regarded as a mark of the language **economy**: leaving so many possibilities unemployed would seem to be too great a luxury. Of course, this general kind of economy is not to be identified with the semantic and functional economy of a particular idiom.

## 5. Idiom: Universals and Typology

Above, it has been suggested that due to the idiom's sharing, basically, some of the same structural functions, such as that of predicate, and due to the idiom's undergoing some transformations to the exclusion of other, idioms can also be tested, which appears to be a universal possibility. Statistical analysis of a large-scale testing of idiom behaviour might reveal much more in this respect. So far, it seems that only a limited number of features may be suspected, with some reason, to be **(statistical) universals**, including the following:

- 1 – underlying (syntagmatic and paradigmatic) anomaly,
- 2 – deficient and restricted function and transformation capacity,
- 3 – entropic, accidental, non-modelled formation,
- 4 – functional analogy to the word-class and sentence function,
- 5 – economical, abstract, as well as vague denominative quality,
- 6 – ready-made pragmatic tool quality,
- 7 – non-additive (non-Fregean) function (and meaning) of constituents.

This list of suspected universal features is certainly open to much criticism, which might be partly due to a different view of the idiom itself and general theory of language applied, as well as to some, marginal idioms displaying this or that feature in, so far, no discernible way. The same kind of problem involving the choice of the theoretical framework appears in typological approach to the idiom (see, e.g., Skalička 1966, Čermák 1978). It is evident that **idioms are typologically conditional**, just like much of other phenomena, and employ, basically, all major structural features of each language type. Whether these features find a specific expression in the form of idioms is still very much open question. By this I do not mean the fact that if a language has cases, the cases come to be employed in the idiom's form, though in a deficient manner. What I mean is the kind of dependency in which the existence of cases in a language loosens, substantially, idiom's word order.

It is, in fact, far from clear whether the existence of cases brings about, accordingly, any substantial loosening of the idiom's constituent order, if any at all. While any typological relationship of the latter kind, bordering, in fact, on universals, has to be explored as yet, the former kind is, if need be, easy to exemplify. Thus there are idioms employing aspectual distinction in verbs, giving rise to either a variant of the idiom or employing only one aspect of the two, mostly, as in Slavonic languages; idioms employing complex participial constructions as in Finnish; idioms employing numeratives as in Chinese or articles as in English etc. As such a listing does not reveal much, attention should be turned rather to such border cases as that of Chinese, mentioned above (see 3), where the functional analogy and similarity of lexical and collocational idioms as between two distinct types seems to lead to merging of both types and making the boundary difficult to draw.

In general, then, due to the idiom being denomination and part of the language system, those areas of language are drawn upon, primarily, that serve the **construction of denominative units** in language. This explains why – next to the lexicon – it is the word formation, specifically in the sphere of compounds, rather than the inflexion which is mostly employed. The syntax, due to its primary combinatorial nature, seems to be employed quite exhaustively



here, i.e. all of the syntactic and word order patterns serving for the denominative formation, including, on a higher level, sentential patterns as well. In fact, it seems that idioms may set up new syntactic patterns sometimes, unknown in the regular language use. This general (syntactical) feature would seem, then, to correspond to the idiom's tendency to be formed out of hitherto new individual (lexical) permutations as well.

## Shrnutí

### Univerzálie, typologie a frazémy

Univerzálnost existence frazémů a idiomů ve všech jazycích se pokoušejí zdůvodnit různí lingvisté, ať už jde o málo průhledný princip mnohonásobné reinvestovatelnosti A. Makkaie, sémiotické deklarace Kopylenka – Popové či jen soubor tvrzení Dobrovolského, blíže nedokládaných. Nejserióznější se tu zdá přístup Coulmasův, který přiznává, že důvody známé vlastně nejsou, jakkoliv poukazuje na neprediktabilnost jazykového vývoje ap.

Na tomto pozadí se připomíná česká teorie frazémů a idiomů vycházející z mnohostranné anomálie a vedoucí k definici frazému nazírajícího ho jako takové *jedinečné a ustálené kombinace aspoň dvou prvků, o níž platí, že aspoň jeden z nich nefunguje tímž způsobem v žádné jiné kombinaci takového typu či funguje pouze v jejich krajně omezeném a malém počtu*. Z toho vyplývá řada důsledků v různých směrech. Pro význam platí neaditivnost komponentů, kterou tradiční pohled opírající se jen o metaforu nedokáže vyložit.

Funkčně je pro frazém specifická jeho sémiotická funkce a v ní zvláště její pragmatická složka, funkce evaluativní. Z jiného hlediska je funkcí frazému být hotovým a velmi ekonomickým nástrojem v oblasti pojmenování.

Strukturně nazíráno lze pak funkčně frazémy dělit na kolokační, propoziční a lexikální, podle povahy celku, která je ekvivalentní promluvě, kombinaci lexémů či lexému jedinému. Jen pro první dva typy pak ale platí diskretnost formy frazému.

Na pozadí úvah o dvojí artikulaci jazyka lze podmíněně u frazému uvažovat o jakési třetí artikulaci.

Obecně pro povahu frazému platí aspoň tyto vlastnosti:

- 1 – anomálie různého druhu, syntagmatická a/n paradigmatická
- 2 – defektní a restringovaná funkce a schopnost transformace
- 3 – entropické, náhodné a tedy nemodelové tvoření
- 4 – funkční analogie ke slovu a větě
- 5 – ekonomická, abstraktní a vágní povaha
- 6 – kvalita hotového pragmatického nástroje
- 7 – neaditivní (nefregeovská) funkce a význam složek.

