STUDIES IN FUNCTIONAL AND STRUCTURAL LINGUISTICS 54

Meaning Predictability in Word Formation

Pavol Štekauer

John Benjamins Publishing Company

Meaning Predictability in Word Formation

Studies in Functional and Structural Linguistics (SFSL)

Taking the broadest and most general definitions of the terms functional and structural, this series aims to present linguistic and interdisciplinary research that relates language structure — at any level of analysis from phonology to discourse — to broader functional considerations, whether cognitive, communicative, pragmatic or sociocultural. Preference will be given to studies that focus on data from actual discourse, whether speech, writing or other nonvocal medium.

The series was formerly known as *Linguistic & Literary Studies in Eastern Europe (LLSEE)*.

Volume 54

Meaning Predictability in Word Formation: Novel, context-free naming units by Pavol Štekauer

Meaning Predictability in Word Formation

Novel, context-free naming units

Pavol Štekauer
Prešov University

John Benjamins Publishing Company
Amsterdam/Philadelphia

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Pavol Štekauer
Meaning Predictability in Word Formation : Novel, context-free naming units / Pavol Štekauer.
p. cm. (Studies in Functional and Structural Linguistics, ISSN 0165–7712 ; v. 54)
Includes bibliographical references and index.
1. Grammar, Comparative and general--Word formation. 2. Semantics. 3. Onomasiology.

P245.S783 2005
401/.43--dc22 2005042125
ISBN 90 272 1563 4 (Eur.) / 1 58811 633 6 (US) (Hb; alk. paper)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Acknowledgements xi

List of abbreviations xiii

Introduction xv

Chapter 1
Literature survey 1
1.1 General 1
1.2 The morphological tradition 1
1.2.1 Lees 2
1.2.2 Levi 4
1.2.3 Van Lint 7
1.2.4 Zimmer 8
1.2.5 Downing 10
1.2.6 Allen 13
1.3 Basic psycholinguistic models 13
1.3.1 Slot-filling models 14
1.3.1.1 Concept Specialization Model 14
1.3.1.2 Hampton's Model of Attribute Inheritance 16
1.3.1.3 Selective Modification Model 17
1.3.2 Relation models 19
1.3.2.1 Coolen et al.'s semantic representation model 19
1.3.2.2 Gagné & Shoben's thematic relation model 21
1.3.3 Analogy-based models 23
1.3.3.1 Ryder's cognitive approach 23
1.3.3.2 Coolen, van Jaarsveld, and Schreuder 26
1.3.4 Combined and other models 27
1.3.4.1 Finin's artificial intelligence framework 27
1.3.4.2 Meyer's knowledge representation model 28
1.3.4.3 Wisniewski's two-process model 30
1.3.4.4 Libben's semantic transparency model 32

1.3.5 Non-compound interpretation models **34**
1.3.5.1 The Clarks' 'contextuals' **34**
1.3.5.2 Beard's transpositions **36**
1.3.5.3 Kiparsky's principle of canonical features **36**
1.3.5.4 Dokulil's meaning predictability **38**
1.4 Summary **39**

Chapter 2
General word formation framework **43**
2.1 An onomasiological model of word formation **43**
2.2 Onomasiological Types **51**

Chapter 3
A theory of predictability **55**
3.1 Why context-free meaning predictability? **55**
3.2 Predictability – lexical meaning – conceptualisation – extra-linguistic knowledge **58**
3.3 Predictability and the native/non-native speaker factor **61**
3.4 Predictability and seme level **62**
3.5 The meaning-prediction process **70**
3.5.1 Predictability and the Onomasiological Type **72**
3.5.1.1 Onomasiological Type 1 **73**
3.5.1.2 Onomasiological Type 2 **74**
3.5.1.3 Onomasiological Type 4 **75**
3.5.1.4 Onomasiological Type 3 **76**
3.5.1.5 Summary **79**
3.5.1.6 Idiosyncrasies of Onomasiological Type 5 **81**
3.5.1.7 Predictability of recategorised proper names **82**
3.6 Onomasiological Structure Rules **85**
3.7 Predictability and productivity **88**
3.8 Predictability and typicality **91**
3.9 Predictability Rate **94**
3.10 Objectified Predictability Rate **95**
3.11 Hypotheses **97**

Chapter 4
The Experiments **99**
4.1 Method **99**
4.2 Experiment 1 **102**

4.2.1 Sample naming units 103
4.2.2 Experimental data and their analysis 104
4.2.2.1 baby book 104
4.2.2.2 dog spade 109
4.2.2.3 flower hat 115
4.2.2.4 ball hammer 119
4.2.2.5 hill star 124
4.2.2.6 apple-juice seat 128
4.2.2.7 game wheel 133
4.2.2.8 shape cloth 136
4.2.2.9 garden whisky 139
4.2.2.10 age bag 143
4.2.3 Summary 1 147
4.2.3.1 Native speakers vs. non-native speakers 147
4.2.3.2 Predictability Rates 151
4.2.3.3 Meaning predictability and dominant readings 153
4.2.3.4 Objectified Predictability Rate 154
4.2.3.5 Seme level 156
4.2.3.6 Tendencies 157
4.3 Experiment 2 159
4.3.1 Sample naming units 159
4.3.2 Experimental data and their analysis 159
4.3.2.1 to boy 159
4.3.2.2 to lion 162
4.3.2.3 to tulip 165
4.3.2.4 to cableway 168
4.3.2.5 to planet 170
4.3.2.6 to river 172
4.3.2.7 to conference 174
4.3.2.8 to triangle 177
4.3.2.9 to courtyard 179
4.3.2.10 to morning 182
4.3.3 Summary 2 184
4.3.3.1 Native speakers vs. non-native speakers 185
4.3.3.2 Meaning predictability and dominant readings 186
4.3.3.3 Seme level 190
4.3.3.4 Main group vs. check group 190
4.3.3.5 Objectified Predictability Rate 191
4.3.3.6 Tendencies 193

4.4 Experiment 3 **195**
4.4.1 Sample naming units **195**
4.4.2 Experimental data and their analysis **196**
4.4.2.1 blondesjoker **196**
4.4.2.2 feather-dialler **198**
4.4.2.3 contactee **200**
4.4.2.4 refusnik **201**
4.4.2.5 anthraxist **203**
4.4.2.6 smile-man **204**
4.4.2.7 removage **206**
4.4.2.8 befoot **208**
4.4.2.9 obsess$_N$ **209**
4.4.2.10 leave-behind$_N$ **210**
4.4.3 Summary 3 **211**
4.5 Experiment 4 **216**
4.5.1 Sample naming units **216**
4.5.2 Experimental data and their analysis **217**
4.5.3 Discussion **221**
4.5.4 Summary 4 **229**
4.6 Meaning predictability and associative meaning: The experimental results in the light of free association of words **230**
4.6.1 Meaning predictability of conversions and the associative principle **233**
4.6.2 Summary 5 **235**
4.6.3 Meaning predictability of two-constituent naming units and the associative principle **237**
4.6.4 Summary 6 **240**

Chapter 5
Conclusions **241**
5.1 General **241**
5.2 Conclusions **245**
Conclusion 1 Meaning prediction capacity **245**
Conclusion 2 No 'correct' reading **246**
Conclusion 3 Prototypical semes **246**
Conclusion 4 Crucial role of extra-linguistic knowledge and experiences **249**
Conclusion 5 Preference for stable relationships **251**

Conclusion 6 Acceptable Word Formation Types 252
Conclusion 7 Word formation types and meaning predictability 254
Conclusion 8 Onomasiological type and Predictability Rate 254
Conclusion 9 Objectified Predictability Rate 256
Conclusion 10 One (two) dominant reading(s), or potential vs. predictable readings 257
Conclusion 11 Single-occurrence readings 257
Conclusion 12 Analogy 258
Conclusion 13 Meaning predictability and associative meaning 259
Conclusion 14 Interplay of factors 259
Conclusion 15 Meaning predictability boosting and reducing conditions 260

Notes 265

References 275

Author index 283

Subject index 285

Acknowledgements

There are many to whom I owe a debt.

First of all – as in my previous books – I express my gratitude to my intellectual 'fathers' – Professor Ján Horecký and Professor Miloš Dokulil – whose works have given me a lot of research inspiration.

I owe a special debt, however, to (in alphabetic order)

Harald Baayen
Laurie Bauer
Robert Beard
Jon Don
Joachim Grzega
Rochelle Lieber
Leonhard Lipka
Karl Zimmer
and two anonymous reviewers of my manuscript

who gave the various earlier versions of the draft manuscript detailed reading, and provided me with their invaluable comments and recommendations that helped me correct many errors and improve the draft manuscript. Any remaining infelicities, of course, reflect my own limitations.

I am grateful to Salvador Valera for having provided me with a lot of valuable data, and to my students and their friends in English-speaking countries, without whose help I would not have been able to collect the required number of completed questionnaires.

And last but not least my thanks go to Jack Lala, Joseph Feinberg, and an anonymous reviewer who gave the manuscript the flavour of genuine English.

List of abbreviations

CN	Complex Nominal
DR	Discourse Representation
EL	Extra-linguistic
FO	Frequency of Occurrence
MSAP	Morpheme-to-Seme-Assignment Principle
NA	Necessary Association
NS	Native Speaker
NNS	Non-native Speaker
OM	Onomasiological mark
OPR	Objectified Predictability Rate
OSR	Onomasiological Structure Rule
OT	Onomasiological Type
PR	Predictability Rate
PRG	Predictability Rate Gap
RDP	Recoverably Deletable Predicate
SLC	Seme Level Combination
WF Type	Word Formation Type

Introduction

New naming units come into existence almost every day.[1] They reflect the progress in human knowledge and understanding, and meet the naming demands of a speech community. Each new naming unit results from a particular naming act performed by one particular member of a speech community who coins a new naming unit with one specific meaning in mind. It means that each such coinage is or approaches an ideal linguistic sign, i.e., the unity of a unique form and a unique meaning (biuniqueness). Stated simply, on the coiner's side, a new form corresponds to a single meaning. The position of a language user who first encounters such an 'ideal sign' is far from being so ideal. No doubt, each of us knows the feeling of hesitation and uncertainty connected with the effort to figure out the meaning of a word (s)he has never come across before. To that end, the purpose of this book is an examination of the predictability of meaning(s) of novel naming units under the conditions of their context-free interpretation.[2] Therefore, the focus will be on the listener's/reader's (meaning the interpreter's) side of the communication channel.

A new naming unit may refer to existing as well as non-existing, tangible as well as non-tangible, conceivable as well as inconceivable (to the majority of mankind – see, for example, the theories of astrophysics) objects (in the widest sense of the word) of extra-linguistic reality. Whatever can become an object of human thought can become an object of the naming process. And whatever becomes an object of the naming process comes to be interpreted by language users. Certainly, this work will not discuss the probability of occurrence of extra-linguistic objects. Rather, the discussion will deal with the question of which of the (usually) multiple possible readings of a new naming unit, always coined (obviously) with one specific meaning in the coiner's mind, becomes the best candidate for the interpretation of that naming unit from the listener's/reader's point of view. Put differently, which of the number of the possible readings of such a naming unit comes most readily to the interpreter's mind as the most acceptable one. Meaning predictability is therefore defined in relation to all the possible meanings of a new naming unit, in particular, as the degree of probability that a particular meaning of a naming unit, encountered

Chapter 1

Literature survey

1.1 General

The meaning predictability of naming units is an integral part of a broader topic covering the whole interpreter's side of the communication channel, i.e., it is a part of the whole process of reception, processing, representation, and interpretation of naming units in varying context-free or co-textual and contextual conditions.

In addition, it covers a wide range of partial issues which directly or indirectly pertain to the main topic of this work. Interestingly, while the indirectly related issues have been discussed mostly by morphologists, the more central aspects of the topic, including the processing and the representation models of 'lexical units', have been developed mostly within the framework of psycholinguistic research. This chapter is not aimed at providing an exhaustive review; rather, its focus is on outlining the central topics and basic approaches in the relevant literature. It will be shown that the vast majority of relevant books and articles concentrate on compounds, in particular, Noun + Noun, and partly, Adjective + Noun compounds. On the other hand, one of the major contributions in this field (Clark & Clark 1979) covers the issues of converted naming units.

1.2 The morphological tradition

Let us start with several relation-based theories which establish the necessary foundation for a theory of meaning predictability, that is, those theories which – via the *classifications* of unequal degree of complexity and detail – attempt to *generalise possible meanings* of compound naming units. Their effort is motivated by the fact that "the overwhelming majority of the derivations... gravitate toward a limited class of core functions" (Beard 1981:345). Then, attention will be shifted to various factors conditioning the interpreta-

tion of novel naming units, mostly, but not exclusively, developed within the psycholinguistic framework.

1.2.1 Lees

The vexed problems related to primary compounds have been discussed frequently over the last forty years within the framework of both transformationalist and lexicalist hypotheses. R. B. Lees (1960) gave a strong impetus to the discussion of *generating* such compounds from *kernel sentences* by a sequence of *transformations*. Lees is aware of a number of problems connected with derivation of compound nominals from kernel sentences. One of them is the multiplicity of 'grammatical form' (in fact, the multiplicity of meanings) in this kind of compounds, for example:

(1)

puppydog	(= dog which is a puppy)
bulldog	(= dog which is like a bull)
shepherd dog	(= a shepherd's dog)
watchdog	(= dog which watches something)
police dog	(= dog used by the police)
sheep dog	(= dog which herds sheep)
prairie dog	(= dog which inhabits the prairie)
etc.	

It follows that "English nominal compounds incorporate grammatical forms of many different sentence types, and of many different internal grammatical relationships within sentences, such as subject-predicate, subject-verb, subject-object, verb-object, etc." (1960: 119). Lees maintains that many compounds can be derived in a number of different ways, for example, *snake poison* may be derived from 'the poison is from snake'; but it also may be interpreted as derived from 'the poison is for snakes'; *snake flesh* is derived from 'flesh of a snake', *snake meat* from 'meat from a snake', and *snake food* from 'food for a snake'. This suggests that the identification of the actual meaning of a novel compound is an extremely complicated matter which cannot be explained from the underlying kernel sentence or from the formal surface structure. Therefore, Lees' approach, however innovative, failed to answer the question of the particular compound meaning identification.

Consequently, it faced fierce criticism. Probably the most profound analysis of Lees (1960) was given by Marchand, primarily in (1965a), (1965b), and (1974) and by Scalise (1984). The criticism, relevant to the present topic, can be summarised as follows:

a. Grammatical criteria alone are insufficient to describe the semantic aspect of compounds; a semantic description is indispensable. Marchand (1974: 298) expresses a view (formulated by Dokulil as early as in (1964)) that "the aim of word formation is the production of new lexical units, not just the formation of new entities on grammatical patterns."

b. The underlying structures are arbitrary; they do not specify accurate criteria which make it possible to generate individual compounds from various underlying structures. This issue has been noted by many authors. Thus, Bauer (1983: 160) asks:

> There may be a number of verbs which could have been deleted from any given compound. For example, should *police-dog* be derived from an underlying 'the dog serves the police', 'the police use the dog', 'the dog works with the police', 'the police work with the dog' or from some other structure entirely?

And in (1978: 74), he stresses, in reference to compounds like *blackmail*, *easychair*, and *shortbread*, that while

> The mail which is black
> The chair which is easy
> The bread which is short

are acceptable, they are not paraphrases of the compounds, which calls into question Lees' method of compound generation.
Scalise (1984: 16) arrives at the same conclusion. While the compound, for example, *windmill* is paraphrased as 'the wind powers the mill' nothing excludes other possible paraphrases such as 'the wind activates the mill', 'the wind makes the mill function', etc., or even a 'passive' paraphrase such as 'the mill is activated by the wind'. For a discussion on this point also see Bauer (1978: §4.2), Bauer (1983: 159–163), and Motsch (1970).

c. Marchand (1965a) points out that Lees does not explain why identical sentence structures generate entirely different compounds, for example, 'we push buttons' – *pushbutton,* 'we shed blood' – *bloodshed,* 'we eat apples' – *eating apple*, etc. Moreover, he does not explain why semantically identical types feature formal differences (see different *determinata* in *deer shooting, bloodshed, alcohol consumption, alcohol intake*). And finally, Lees does not explain why the same sentence structure corresponds to different compound structures: in *wading-bird* the subject of the sentence becomes the *determinatum* ('The bird wades') whereas in *population growth* it is the verb which becomes the d*eterminatum* ('The population grows'). In addition,

one and the same type of sentence may yield different compounds: 'we eat apples' – *eating apple, apple-eater, apple-eating.*[1]

d. The non-recoverable deletion of meaningful elements by means of transformations is also faulted by Scalise (1984: 10–12), Gleitman & Gleitman (1970: 91–94), Zimmer (1971: C3), Bauer (1978: 81, 1983: 160), Allen (1978: 87), Hammond and Noonan (1988: 3), and others.

Nor does Lees' revised, more semantically oriented version of compound generation and classification (1970) bring widely accepted answers to these questions.

1.2.2 Levi

J. Levi (1978) proposes a small set of *Recoverably Deletable Predicates* for the process of formation of complex nominals. The set includes predicates CAUSE, HAVE, MAKE, BE, USE, FOR, IN, ABOUT, and FROM. It is only these predicates that may be deleted in the process of transforming an underlying relative clause construction into the typically ambiguous surface configuration of the complex nominal. These predicates are of a piece with the more traditional terms as follows (Levi 1978: 77)

(2)

CAUSE	causative
HAVE	possessive/dative
MAKE	productive; constitutive, compositional
USE	instrumental
BE	essive/appositional
IN	locative (spatial or temporal)
FOR	purposive/benefactive
FROM	source/ablative
ABOUT	topic

and may be illustrated as follows (Levi 1978: 76–77):

(3)

CAUSE	*tear gas*	*viral infection*
HAVE	*picture book*	*government land*
MAKE	*honeybee*	*snowball*
USE	*voice vote*	–
BE	*consonantal segment*	–
IN	field mouse	–

FOR	*horse doctor*	–
FROM	*olive oil*	–
ABOUT	*tax law*	–

Levi maintains that a complex nominal is potentially nine-ways *semantically ambiguous* because any of the nine Recoverably Deletable Predicates can be deleted. In addition, the first three predicates are two-ways syntactically ambiguous because the modifying constituent can be derived from either the subject or the object of the underlying predicate. In total, there are twelve potential readings for complex nominals. This multiple ambiguity, however, does not imply – in contrast to Chomsky's view (1970) – a fully idiosyncratic nature for complex nominals. As emphasised by Levi "novel CNs [complex nominals] are frequently coined by speakers and understood by hearers with great ease precisely because of the predictable aspects of CN grammar" (1983: 188). This is thanks to the so-called disambiguating strategies, including (1) the knowledge of the regularities of complex nominal formation; (2) the knowledge of naming patterns based on semantic class of head and modifier nouns, such as naming artefacts by their purpose (*study lamp*); living things by their habitat (*field mouse*); human activities by time (*morning lectures*), place (*urban riots*), instrument (*shock therapy*), subject (*royal orders*), and object (*child abuse*); people by sex/age (*boy genius*), habitat (*mountain tribes*), and occupation (*clerical enemy*) (1983: 240); (3) pragmatic disambiguation employing our extra-linguistic (encyclopaedic) knowledge to identify the most plausible reading of a certain complex nominal.

Van Lint (1982) notes that Levi's system of twelve Recoverably Deletable Predicates is not applicable universally, because there are a number of problematic fuzzy cases admitting more than one deletable predicate, which are pointed out by Levi herself (1983: Chapter 7).[2] Levi therefore suggests abandoning the idea of a limited set.[3]

Furthermore, as aptly noted by Finin (1980: 34), the Recoverable Deletable Predicates "are extremely vague", and therefore "should not be the stopping point of the semantic analysis." It is for this reason that Murphy (1988: 533) proposes – in addition to the step of concept specialization in interpreting complex concepts – an additional step, the *conceptual elaboration*, based on *extensional feedback* (Hampton 1987, 1988) reflecting our knowledge of the world. Both of these steps heavily rely on the language user's *world knowledge*. While in the first step world knowledge helps a language user identify a slot to be filled by a modifying feature, in the second step it is used to refine the interpretation. For example, for *apartment dog*, the first step helps a language user

to realise that it is more likely that a dog inhabits an apartment than it looks like an apartment or that it bites an apartment, etc. The second step proposed by Murphy provides a refined analysis based on the interpreter's knowledge of dogs, apartments, and their possible interaction, implying that an apartment dog is usually smaller, quieter, and better behaved than a farm dog, for instance.

In fact, Levi herself is aware of the fact that too much generalisation may miss the point. For her, however, this overgeneralization concerns relationships like 'x is related to y' rather than her Recoverably Deletable Predicates. On the other hand, she admits that the RDP-based analysis "does not (and cannot) specify *all* that we know about the meanings of individual CNs " (1978: 84).[4]

From the point of view of my research, a crucial observation is that meaning generalisation and meaning predictability are concepts that are based on two different degrees of meaning generalisation. The Leesian and Levian level of generalisation covers a number of naming units with the same semantics, thus necessarily disregarding the individual and idiosyncratic character of each naming unit. It may be shown that the notion of meaning predictability is far from being exhausted by any general predicates for the simple reason that one such RDP can subsume several specific meanings of unequal degree of predictability. For illustration, the interpretation of one of my sample compounds *baby book*, analysed in Chapter 4 below, cannot be exhausted by saying that it is the FOR predicate that has been deleted in the process of transforming the underlying (kernel) sentence into a complex nominal. From the meaning predictability point of view such a statement is of little relevance. The generalised Recoverably Deletable Predicate FOR cannot discriminate the subtleties crucial to the meaning-prediction process. Thus, the following meanings can be subsumed into the more embracing general reading 'a book FOR baby': 'a book bought for a baby', 'a book written for babies (to be read by parents)', 'a book with pictures for babies (babies cannot read a text but can perceive drawings and pictures)', a book which a baby likes to fall asleep with', ' a book with the essentials about bringing up babies', 'a book with records about one's baby growth', etc.

In a similar vein, both predictable readings of another sample naming unit *game wheel*, i.e., 'a wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games' and 'a wheel which is a part of a game equipment, a wheel with which a game is played' can be represented by the RDP FOR.

Furthermore, Levi's classification is not applicable to converted naming units. From this it follows that, first, the generalisation step must be completed with the individualising step, and, second, as will be shown below, the identification of possible meanings of a novel naming unit requires, *inter alia*, relevant

and fairly detailed extra-linguistic knowledge and experiences. The amount of the extra-linguistic knowledge required for the identification of a particular meaning cannot be determined by a single general rule, because it differs from case to case.

Moreover – apart from several interpretations potentially subsumed by a single RDP – some other predicates are acceptable in accounting for the generation of such compounds in Levi's fashion: in my above-mentioned example, they include HAVE, ABOUT, and, possibly, USE. The problem of ambiguity, for which Lees (1960) was so much criticised, reappears in Levi with no less urgency.

In general, both Lees' and Levi's works are about *generation* rather than about interpretation of naming units. It may therefore be concluded that while Lees and Levi aptly demonstrate the heterogeneous nature of meaning relations between compound constituents, their more specific theoretical conclusions cannot be used as the exclusive basis for the meaning-prediction process. At best, they can be understood as an initial step which must be completed with a subtler analysis, for example, in the vein of the above-mentioned Murphy's conceptual elaboration. The discussion of the individual predictable readings of experimental naming units in Chapter 4 provides numerous examples of the relevance of this assumption.

1.2.3 Van Lint

The indisputable contribution of T. van Lint (1982) – despite my negative attitude to the transformationalist account of compound generation by a series of transformational steps (Štekauer 1998, 2000) – is her elaboration of the role of *semantic* and *pragmatic* factors in accounting for the interpretation of non-lexicalised isolated compounds. These two factors as a key to a correct interpretation of compounds represent the starting point of her rule system. Van Lint distinguishes between 'Relator' compounds and 'other' compounds. The basic idea underlying her interpretation model is that the 'Relator' compounds are generated from structures the predicates (= Relators) of which are selected on the basis of a match between relevant features of the compound elements, and 'the other' compounds are generated from structures which result from the incorporation of the first compound element (Specifier) into the semantic structure of the second element (Head). It follows from this that van Lint combines the *transformationalist principle of predicate identification/deletion* with the *matching principle of the slot-filling model* introduced by the lexicalists.

For illustration, the compound *embankment traffic* is generated from the underlying concept (a linguistic representation of the mental concept) represented by a relative clause 'traffic which is on the embankment' by a series of transformations. In this process, the predicate (Relator) 'is on', standing for the feature 'location', is *incorporated* in *embankment* and is *excorporated* during the process of interpretation. As indicated above, the operation of rules is related to matching the semantic features of compound constituents: one of the central features of *embankment* is LOCATION, that of *traffic* is a PHENOMENON which in itself is LOCATED. Since the feature LOCATED is intrinsic to *traffic*, van Lint speaks of the *Necessary Association* (NA) feature. Consequently, the relation between the two compound constituents follows from the matching of LOCATION and LOCATED, permitting the incorporation of 'is on' into 'embankment'.

In addition to NA-features, *pragmatic features* (based on our knowledge and experiences) play an important role. It will be shown below that pragmatic and semantic features can be advantageously incorporated into a non-transformationalist theory, which – like van Lint's model – relies on the interaction between the processes of word formation and word-interpretation. However, contrary to van Lint's scepticism (1982:139) assuming that "[t]here is no way in which a linguistic theory can deal with the 'translation' of a mental concept (the meaning) into a linguistic construction (the compound)", I believe that the cognitive onomasiological model (Štekauer 1998, 2001a), outlined in Chapter 2 below, answers van Lint's 'translation' problem in a viable way.

1.2.4 Zimmer

Within the framework of the lexicalist hypothesis the semantics of primary compounds has been given much attention, too. Considerable effort has been exerted to define the *conditions for the interpretability* of primary compounds.

K. Zimmer (1971) prefers – as the most promising approach to the acceptability problem – the *negative characterisation* proposed by Heidolph (1961) and subsequently furthered by Motsch (1962). Their approach is based on defining those relations which cannot underlie compounds. By implication, Zimmer suggests that all compounds are acceptable that are not forbidden by the specified inadmissible relations, irrespective of whether or not they fit the established types of compound formation.

Zimmer uses some examples to illustrate this idea: *knife box* cannot refer to a box which typically has no knives. Therefore the spatial and temporal relations require coincidence, i.e. location of the referent of one compound constituent at or within the area or period referred to by the other compound

constituent; space adjacency will also do (*sea town, river road, Wednesday lecture*). On the other hand, transitory locations in time and space of unique objects are not classificatory features. Therefore, he argues, one does not refer to a cat that happens to be sitting in a tree at the time it is observed as a *tree cat*, because sitting in a tree at a particular time is not relevant to the category of cats. This is one of the basic (general) observations made about the semantic interpretation of English compounds: "The first stem of any compound, either root or synthetic, is nonreferential in interpretation" (Lieber 2004: 46).[5]

Certainly, Zimmer's assumption, and its generalisation, must be viewed with reservation. As will be explicated below, one *can* refer to such a cat as a *tree cat*, in the same way as one can call a book placed on a radio a *radio book* (time-bound), and a man in a team of researchers a *Harvard man* if he is the only one to have graduated from Harvard (situation-bound), etc. And to adduce some more recent examples, L. Lipka (personal communication) drew my attention to Truffaut's film *Fahrenheit 451* where there are *firemen* who 'set fire to books' and (in the German version) *Buchmenschen* who *are* books. What matters, therefore, is the fact that the Predictability Rate of this sort of reading is, for obvious reasons, very low in context-independent interpretation conditions.

Also forbidden, says Zimmer, is the relation of rejection if not explicitly expressed by one of the compound constituents: *war hater* is acceptable, but *war man* in the sense of 'man who dislikes, denounces, etc. war' is not.

Zimmer then goes on to propose the necessary conditions that must be fulfilled for compounding to occur appropriately. One such condition is that the relation between compound constituents must be 'appropriately classificatory' (for example, the look-like relation, the being-an-instrument-for relation, the being-a-part-of relation). This, however, does not exclude individual variations conditioned by cultural differences, different beliefs, and other factors. Zimmer emphasises that the *classificatory relevance* has to do with the distinction between naming and describing: "Anything at all can be described, but only relevant categories are given names..." (1971: C15).

These ideas are further developed in Zimmer (1972) where the notion of an *appropriately classificatory* (AC) relationship, applied to endocentric non-idiomatic N + N compounds, is defined as follows:

(4) A noun A has an AC relationship to a noun B if this relationship is regarded by a speaker as significant for his classification – rather than description – of B (1972: 4).

Furthermore, and this seems to be of crucial importance, the AC relationship exists between components of a situation referred to by linguistic forms rather than between the linguistic forms themselves. While this relation is central to his classification, Zimmer himself admits that it may be rather vague. It is, however, intended to emphasise that the AC relationship cannot be defined without reference to a speaker's view of a situation. For me, however, this implies that the speaker's view of the situation, being highly subjective, need not be accepted by other language users; consequently, given this criterion for compounds, various speakers may differ in their evaluation of the compound status of individual constructions. Zimmer is aware of this fact and maintains that even the same (or a very similar) relationship between two nouns may be, for the same speaker, an instance of an AC relationship in one situation but not in another.[6]

1.2.5 Downing

In my view, the first significant attempt to do justice to the *meaning predictability* of primary compounds itself is Downing (1977). Downing demonstrates the futility of former attempts (Lees 1960, 1970; Levi 1974, 1975; Li 1971) to reduce the possible meanings of primary compounds to several broadly defined semantic classes and aptly assumes that a number of interpretations of novel compounds "are at best *reducible* to underlying relationships as suggested by Li and others, but only with the loss of much of the semantic material considered by the subjects to be relevant or essential to the definitions" (1977: 826). Furthermore, her experiment demonstrates that readings rated by informants as possible or likely are very difficult to reduce to any of these general classes, and gives an example of a compound admitting readings of *various levels of specificity* which can be reduced to the same general underlying relationship. And – even more importantly – Downing's informants unanimously judged a more specific reading to be acceptable or even likely, while a more generalised reading was judged unlikely or even impossible.

Downing imposes an important *constraint upon the combinability* of compound constituents by assuming that *semantic redundancy* in the modifier of a compound prevents the existence of a compound (the modifier is semantically redundant if it does not bring any new information). Thus, *head hat* is hardly acceptable because all hats are designed to be worn on the head, *egg bird* is said to be unacceptable because all birds come from eggs; similar considerations apply to *lad-boy*, and *book-novel*, etc. (1977: 832). Certainly, this is not a hard-and-fast rule, as illustrated by Downing herself with *palm tree*, which is

a case of so-called pleonastic compounds, where the head is a hyperonym of the non-head.

The same idea was later reiterated, for example, by Meyer (1993: 102) as the first productivity restriction on NN-compounds: "... novel NN-compounds with a modifier denoting a superset of the head extension are ruled out since they do not denote a specialization of the head noun extension." Meyer's German examples include **Möbelstuhl* ('furniture chair'), **Kleidungshose* ('clothes trousers'), and **Metalleisen* ('metal iron'). Thus, for example, chairs are pieces of furniture, and therefore *furniture* in *furniture chair* does not add new information.[7] Another set of examples is given in Bauer (1978: 86), including **humanman*, **animal-horse*, **buildinghouse*. Bauer points out, however, that the opposite situation with the head implicit in the determinant is found in tautological compounds like *elm tree*, *tuna fish*, etc., thus rejecting the conclusions of Downing's informants who put instances of hyponym-hyperonym structure, like *truck-vehicle* and *horse-animal*, on the 'blacklist' along with the above-mentioned hyperonym-hyponym compounds *lad-boy* and *book-novel*.[8]

These facts are also confirmed within the discussion of a different issue – one of attribute inheritance and dominant concepts – by Hampton (1987), who demonstrates that in some conjunctions of concepts only one concept significantly contributes to the conjunction. Thus, for example, 'buildings', and 'games' do not significantly contribute to their respective conjunctions, 'dwellings that are also buildings' and 'games that are also sports', because the attributes of one concept are largely subsumed by those of the other:

> Since typical sports *are* games, typical dwellings *are* buildings... (whereas the converse holds much less well, if at all), the attributes of the latter concept in each conjunction listed above are largely redundant. This is the case even though the concepts in fact overlap and are not related by class inclusion... (1987: 64).[9]

A similar view is presented by Finin (1980: 16) who believes that "the relationship between the nouns should not be totally predictable." Therefore, for instance, we do not use *milk butter* to mean 'butter made from milk' because this meaning is totally predictable.[10]

To return to Downing, another important point of her experimental research in terms of predictability is her discussion of 'habitual' vs. 'temporary' relations. She – in contrast to Zimmer (see above) or Gleitman & Gleitman (1970) – does not accept the view that compounds – unlike relative-clause paraphrases – are always based on a habitual relation, and demonstrates – based on the results of her experiment – that the relation between the com-

pound constituents may sometimes be of temporary, fortuitous nature, in which case it is strongly bound to a context, but that there is a very strong tendency for compounds to be "created and interpreted on the basis of relationships perceived as permanent and habitual" (1977:836). She further points out that

> [i]f...the compound is based on a temporary or fortuitous relationship, it is unlikely that the hearer will be able correctly to deduce the nature of the class denoted by the compound, unless he was actually present at the moment the relationship occurred, or was specifically told of it (1977:837).

This was confirmed by her research in which informants showed strong *preference for stable relationships* when choosing from stable/fortuitous reading options. Moreover, the readings based on habitual relationships generally received higher ratings than temporary or fortuitous relationships, which were never rated as likely. These constraints thus contribute to keeping the number of readings which comply with the notion of predictability defined below at a manageable level. As pointed out by Aitchison (1987:155), "[t]hey subconsciously guide speakers, and also aid hearers in their interpretation, and so must be regarded as part of the lexical tool-kit."

The explanation of the dominance of habitual over temporary relations is, in my view, simple, and is related to the nature of word formation as an act of naming, of giving names to substances, actions, qualities, and circumstances. In this naming act, a coiner does not refer to any particular 'object' of extra-linguistic reality. Rather, (s)he gives names to a class of objects identifying common, general, that is to say, habitual features, and disregarding any fortuitous, individual, and thus, temporary features. This naming strategy may be supposed to also be taken over by an interpreter who also concentrates on what is permanent and (proto)typical, rather than on what is volatile and/or idiosyncratic.

And last but not least, Downing should be credited with introducing a method of evaluating *context-free interpretations of novel compounds*. In particular, her informants were asked to propose readings for possible primary compounds and, within a 'ranking task' "to evaluate the appropriateness of various interpretations proposed" (1977:817). Within the interpretation part of the task the informants were asked, among other things, to provide one or more interpretations for those compounds they judged to be possible. Within the ranking task they were asked to evaluate the individual interpretations as likely, possible, or impossible. Herewith she laid methodological foundations

for research into the predictability of readings of novel, context-free words in general.[11]

1.2.6 Allen

M. Allen (1978) rejects Roeper & Siegel's (1978) assumption that primary (root, non-verbal) compounds are all idiosyncratic, not rule-based, and thus only listed in the lexicon. Her *Variable R Condition* refers to the variability in primary compound meanings and employs one of the central principles of a line of later psycholinguistic works, accounting for conceptual combinations by the principle of *slot filling*. It establishes a range of possible (and consequently also impossible) meanings for a given compound. This range is specified in terms of semantic feature sets of the compound constituents: the semantic content of the first constituent may fill in any one of the available feature slots in the feature hierarchy of the head constituent. Certainly, the filling of a slot is conditioned by the semantic compatibility of the features in question (for instance, *water-mill* can mean 'mill powered by water', 'mill located near water', etc., but not 'mill which lives near water' (*mill* is [Inanimate] which is incompatible with the [Animate] feature of *live*), or 'mill which grinds water' (incompatibility of the features [Liquid] and [Grindable]).

Some of the acceptable meanings are more probable than others. This is related to the *hierarchy of semantic features* in the head constituent. The semantic content of the first constituent tends to 'plug in' to a semantic feature 'slot' which has a dominant position (for *mill*, the dominant semantic features are 'powered by' and 'production': *water mill, wind-mill, hand mill, steam mill vs. steel-mill, paper-mill, flour-mill, cotton-mill)*.

1.3 Basic psycholinguistic models

Three basic approaches to the interpretation of compound words (conceptual combinations) can be distinguished within the psycholinguistic framework.

First, an approach putting stress on the role of the *head noun* (head concept) functioning as a *schema* with a certain number of *slots* that are filled by the modifier values. This model is called a *feature model* or *schema model* (e.g., Allen 1978; Cohen & Murphy 1984; Murphy 1988, 1990; Smith & Osherson 1984; Smith, Osherson, Rips & Keane 1988), and is based on the *semantic representations* of the constituent nouns and associated *encyclopaedic knowledge*.[12] According to this approach, differences in the interpretability of

context-free primary compounds are related to the *relative salience* of particular meaning aspects (slots in a schema-based model, attributes in the feature-based model). Thus, 'high-interpretable' compounds are based on more salient meaning aspects of their motivating constituent and 'low-interpretable' compounds are based on less salient meaning aspects (Coolen, van Jaarsveld & Schreuder 1991).

Second, the *relation model* (Gagné & Shoben 1997; Gagné 2001) emphasises the central role of *thematic relations* between compound constituents and the language speaker's *linguistic knowledge* of the relative strength of the individual thematic relations bound to a particular *modifier concept* (in the characteristic binary modifier-head structure of English compounds). This knowledge facilitates the interpretation of compounds by preferring the interpretation based on a thematic relation which is more readily available to the modifier concept.

Third, the *analogy-based model* (Derwing & Skousen 1989; Skousen 1989; Ryder 1994) accounts for the interpretation of novel, context-free compounds primarily by lexicalised (i.e. established, institutionalised) compounds that serve as certain *interpretation patterns* or models. Thus, for example, *mud man* may be interpreted, among other possible readings, as 'a man who collects mud' if its interpretation follows the analogy with *garbage man*, or as 'a man who delivers mud' if it takes analogy with *milk man* as its pattern (van Jaarsveld et al. 1994: 113).

In the course of psycholinguistic research into the interpretation of compound words (concepts), a number of various modifications and/or combinations of these basic models have emerged. Some of the most fundamental are introduced below.

1.3.1 Slot-filling models

1.3.1.1 *Concept Specialization Model*

The idea of head-slot filling proposed by Allen (1978) as the Variable R condition reappears, among others, in Cohen and Murphy (1984) and Murphy (1988, 1990).

Cohen & Murphy (1984) and Murphy (1988) present a typical slot-filling approach to the interpretation of complex concepts (compound words) based on a *knowledge representation model of prototype theory*: the focal point in treating complex concepts underlying the N+N and A+N compounds is the idea that the interpretation of complex concepts is dependent on the interpreter's *world knowledge* without which it is not possible to capture the 'mediating

relation' between the component concepts. For example, the mediating relation between *morning* and *flight* in *morning flight* is that a flight occurs in the morning. Importantly, the mediating relation ('occurrence', in this case) does not follow directly from the two underlying concepts; it does not consist in the class of objects represented by the two concepts. Since the mediating relation does not obtain in the meaning of compound constituents, it must be, by definition, be inferred from the language user's knowledge.

Cohen & Murphy call their slot-filling model the *concept specialisation model* because the modifying concept fills a slot in the head concept's schema, thus specialising the head concept role with a modified value. For example, in *red apple*, RED modifies COLOUR of *apple*, in *ocean drive*, OCEAN modifies the LOCATION of *drive*, in *apartment dog*, APARTMENT modifies the HABITAT of *dog*, etc. The authors realise that the interrelation (matching) of the two motivating constituents (concepts) at this level provides too general interpretations. One cannot but agree with them. This level of meaning generalisation may indicate the *semantic compatibility* of the two underlying concepts; however, it has little to say about the predictability of the meaning(s) of the 'complex concept'. It is perhaps this fact that Cohen & Murphy have in mind when claiming that this mechanism makes it possible to generate atypical compounds like *virgin birth*. Their atypicality is explained by the fact that "the role that the value 'best fits' is usually filled by other more typical values; that is, the fit in this case is a relatively bad fit" (1984: 52).

Murphy (1988) discusses the above-mentioned idea of the *elaboration* of the basic interpretation identified by the slot-filling procedure. Elaboration, i.e. the use of one's knowledge of the world, appears to be an inherent part of any process of interpretation of novel, context-free complex naming units. This is demonstrated by Murphy's experiments (1988) showing that the feature-weighting approach (Smith & Osherson 1984; Smith et al. 1988) which 'computes' the meaning of a novel complex word merely from the information provided by the two motivating concepts – without using one's knowledge of the world – is untenable.

Murphy's model adapts case theory to represent the mediating relation between compound constituents by the possibility of filling the slots (roles) by features (values). Crucially, these values may be ordered by *typicality*.

Cohen & Murphy (1984: 47) refer to Barsalou's (1981) account of typicality reflecting "the ability of an object to fulfil the function typically associated with the category." This may be represented in the FUNCTION or the USED-FOR role, or any other role.

The idea of the 'inheritance of roles', employed, for example, in Finin (1980), reappears here, too. Thus, the concept of *piano* inherits its SOUND role from its superordinate concept *musical instrument*. For Cohen & Murphy the idea means that sub-concepts may specify more general concepts by identifying a particular value for an inherited role. For example, the general concept of *repair* is modified by filling the OBJECT role of *repair* by role value ENGINE to give *engine repair.*

The significance of language users' knowledge is illustrated by cases in which the typicality of an object, represented by a complex concept, is much lower than that of the individual constituent concepts. In other words, "the actual roles and role values of a complex concept need not be the same as those of its component concepts" (1984: 53). Cohen & Murphy give an example of *pet fish.* While pets are usually furry and cuddlesome, *pet fish* is not. This fact can only be explained by making use of 'domain knowledge' which makes it possible to substitute the prototypical role-values FURRY and CUDDLY by SCALES and SLIMY and by assigning the latter greater weight than the former role-values inherited from *pet*.

1.3.1.2 *Hampton's Model of Attribute Inheritance*

Cohen and Murphy's is a model based on the 'intensions' of the two concepts. (i.e., attributes that are typically shared by class members) – in contrast to the extensionally-based approach of fuzzy logic. An intensional approach is also taken up by Hampton (1983, 1987) who attempts to specify the rules which control the inheritance of attributes by a compound concept from its constituent concepts, i.e., the *inheritance rules*. Hampton demonstrates that only some attributes of each constituent of a compound are true of the compound concept, or in other words, the compound concept does not include all of the attributes characterising its constituent concepts. By implication a conjunction does not equate to the sum of the attributes of the constituent concepts. This explains why the typicality of, for example, a 'guppy' or 'goldfish', exemplars of *pet fish*, can make a perfect match with the conjunctive concept PET FISH in spite of the fact that it lacks some relevant attributes of the constituent concepts PET and FISH. It means that 'guppy' or 'goldfish' are better examples of the compound *pet fish* than they are of either constituent alone. To put it another way, its typicality is greater for the compound than for its constituents (Hampton 1987: 55, 56).

Hampton's model of attribute inheritance aptly predicts that the intension of a conjunction is formed as the *union* of the constituent attribute sets and that the importance of attributes for a conjunction may vary. This is no doubt true,

and is embodied in my model at the general level as a hierarchy of semes, and at the specific level, in the form of the dominant position of so-called prototypical semes with regard to the Predictability Rate.

In Hampton's view, the importance of an attribute for a conjunction can be significantly predicted from the importance of the attribute for the two constituents: "the importance of an attribute for a conjunction [is] more likely to resemble its importance for the dominant concept" (1987:64). Furthermore, the attributes necessary for any of the constituent concepts are necessary for the conjunction, and the attributes impossible for either constituent are impossible for the conjunction. This unequal status of constituent concepts in terms of the weight of their attributes is captured by the notion of *concept dominance*: if one concept of a conjunction has a greater number of important attributes than the other constituent concept the typicality in the conjunction is better predicted by the concept with the greater number of important attributes. Hampton's experimental results show that a conjunction inherits all attributes of its constituents except for those which are of low importance for the constituents.

1.3.1.3 *Selective Modification Model*

Another variant of the slot-filling model is the *Selective Modification Model* of Smith, Osherson, Rips & Keane (1988) developed within the framework of a prototype theory on the principles outlined in Smith and Osherson (1984). The notions of *typicality* and *prototype* are central to their conception because, *inter alia*, the "similarity-to-prototype plays some role in categorization, memory and communication" (1988:486).

Smith and Osherson (1984) provide experimental arguments against the notion of 'characteristic function' of the fuzzy-set theory specifying the degree to which an entity is a member of a prototype concept. Zadeh (1965) formulates a *minimum rule*, according to which the characteristic-function value of a combination of two prototype concepts (for example, PET FISH) is the minimum of the constituents' values (i.e. that of PET and FISH, respectively). Thus, if we evaluate the characteristic-function value of *guppy* in terms of concepts PET and FISH, the value of the constituent *pet* is lower than the value of *fish*, because *guppy* is more typical FISH than it is a typical PET. Therefore, the characteristic-function value of *guppy* with regard to the concept of PET FISH is given by the value of the constituent *pet*.

By means of experiments Smith and Osherson demonstrate the failure of the minimum rule of the fuzzy set theory and propose the basic principles of an alternative to a fuzzy-set theory on the principles of prototype representa-

tion which were later developed into a more comprehensive theory in Smith et al. (1988).

The main goal of the authors is to apply the principles of typicality computation from simple concepts to 'composite concepts' based on the knowledge about their constituents (i.e., simple concepts.). The general features of any model of prototype composition should, it is claimed, contain three basic components: (1) *attributes* of a given concept, (2) its *values*, (3) the *salience* of the individual values (i.e., the most likely features of each of them), (4) the *diagnosticity* of each of the attributes, i.e., their respective contribution to the discrimination of an instance of the concept (the rating of the typicality of an object with respect to a particular concept). For example, the concept of 'red apple' has attributes like colour, shape, texture, etc. Then, while the values of colour may be red, white, brown, etc., the salience (default) value is red.

These considerations underlie the model outlined in Smith et al. (1988), the purpose of which is to account for the way of identifying prototypes of composite concepts out of prototypes for simple concepts. This is relevant to my research because the model may contribute to the identification of the most characteristic (most predictable) reading of a complex word. The model proposed by Smith et al. includes three basic components: (1) a prototype representation for simple noun concepts, including the above mentioned attributes, values, salience, and diagnosticity; (2) procedures for modifying simple prototype to represent composite concepts; and (3) a means for determining the typicality of an instance in regard of a prototype. The authors apply their theory to A-N 'conjunctions', and show that there are two possible approaches to the treatment of A-N combinations: (1) a symmetric conception, postulating an equal contribution of the two constituents to the overall complex concept; (2) an asymmetric model, in which the noun (head or determinatum – to use the traditional terms) is the superordinate constituent which is modified by the adjective (modifier, determinant). Clearly, the latter of the two options is a typical Marchandean word formation syntagma based on the identification-and-specification scheme. It will follow from my discussion that I – like Smith et al., and also like Marchand (1960), Kastovsky (1982) and, importantly, Dokulil (1962) – prefer the latter option (even if not in its 'clean' form, with certain aspects of the symmetrical model being incorporated, too). The preference for the asymmetrical model is justified by Smith et al. by a striking change in meaning when the order of an adjective-noun combination is reversed (1988:492). In effect, Smith et al.'s is a typical slot-filling model: each attribute in the adjective concept selects the corresponding attribute in the noun concept, thus increasing the salience and diagnosticity. For instance,

when *red* modifies *fruit,* it selects the relevant attribute in the noun concept (colour), and increases the diagnosticity of the attribute 'colour' as well as the salience of the value expressed by the adjective (*red*).

1.3.2 Relation models

1.3.2.1 *Coolen et al.'s semantic representation model*

The slot-filling approach to compound interpretation is contrasted by Coolen et al. (1991) with the idea that "interpretation involves selection from a small set of frequent semantic relations, perhaps guided by the analogy of lexicalized compounds with the same nouns as the novel compounds" (1991:350). Coolen, van Jaarsveld, and Schreuder studied the interpretative processing of isolated novel compounds in a series of articles. In one of their experimental tests reported on in Coolen et al. (1991), informants were asked, for example, to determine the degree of 'interpretability', defined as "the plausibility of the object that is described by the compound" (Coolen et al. 1991:342), of both lexicalised (i.e., institutionalised, actual) primary compounds and novel (possible) compounds by using a 7-point scale, without, however, identifying the individual predictable readings. The results obtained suggest – in accordance with their expectations – that the interpretability of 'lexicalised' compounds is higher than that of the group of possible primary compounds labelled as 'low-interpretable' compounds prior to the experiment. In addition, the group of compounds labelled as 'high-interpretable' prior to the experiment are much more easily interpretable than that of 'low-interpretable' compounds.[13]

A disadvantage of this kind of experiment seems to be the fact that it disregards the semantic aspect of interpretation; in fact, we do not know which of the possible and/or predictable meanings are actually easy or difficult to interpret – different informants may well have assigned the same value to completely different readings of the same compound. There are usually a number of possible compound readings and, as will be demonstrated in my research, usually more than one predictable reading. By implication, the above-mentioned experiments and the resulting conclusions cannot bring relevant results for the determination of meaning predictability of naming units.

In reference to the aforementioned conceptions of basic semantic relations proposed by Levi (1978) and Li (1971), Coolen et al. formulate two semantically oriented predictions (1991:349). First, they postulate that the paraphrases of the meanings of particular 'high-interpretable' compounds are characterised by fewer basic semantic relations, which implies greater agreement between the informants in interpreting these compounds.

This claim raises certain doubts. First, one might ask what semantic relation can be classified as the 'basic' semantic relation. This is a rather vague notion, especially if we realise a considerable number of different proposals featuring considerable differences in the number of 'basic relations'. Let us compare, for example, Levi's (1978) nine fundamental Recoverably Deletable Predicates, Finin's (1980) eight interpretation rules, Szymanek's (1988) 25 fundamental cognitive categories, Beard's (1981) 44 universally available grammatical functions.[14]

Second, the results of my experiments, described in Chapter 4, do not favour their prediction: no direct proportionality between the number of 'basic' semantic readings, on the one hand, and the agreement between informants, and therefore, the predictability of readings, on the other hand, has been identified. In several cases a high number of various readings established a 'background' for one or two 'high-predictable' readings. In other cases the relatively low number of readings did not imply a high level of agreement between informants due to the existence of several 'competing' readings, a situation which renders high Predictability Rates unlikely. It should be, however, noted that the methodology of my research differs from that of Coolen et al., a factor which may have affected the results obtained.

The second prediction by Coolen et al. is that for high-interpretable and low-interpretable compounds there may be paraphrases that do not express one of these basic semantic relations. Importantly, "[t]he proportion of such idiosyncratic interpretations may be expected to be significantly larger for LI [low-interpretable – P. Š.] compounds" (1991:349). In other words, their experiment showed that the diversity of paraphrases of meanings of isolated novel primary compounds was much higher than that of easily interpretable ones, which may be attributed to the uncertainty of language users, especially in cases where there is no 'good' reading at hand for a novel compound. This conclusion has also been confirmed in my experiments.

Coolen et al. verified their hypothesis in an experiment in which informants were asked to say the first interpretation that came to their mind for each of 28 novel compounds. Then, the paraphrases were grouped according to Levi's classification of basic semantic relations (1978). As already indicated above, this type of classification is too rough to reflect subtle differences in various readings.[15]

Coolen et al. maintain that even if the slot-filling approach relies heavily on world-knowledge – and their approach is based primarily on semantic representations – the two models may be reconciled by the incorporation in the slot-filling model of the principles of the *interactive activation model*

postulating that the selection of particular slots is based on the relative dominance of meaning aspects of compound constituents (Coolen et al. 1993: 243). It is true, though, that the emphasis of the interactive model is on semantic representations rather than on the conceptual level. Both of these aspects, i.e., world-knowledge and semantic representations, are integrated in Meyer (1993), discussed below, and also in my approach to meaning predictability.

1.3.2.2 *Gagné & Shoben's thematic relation model*

Gagné & Shoben (1997) base their model of nominal compound interpretation on *thematic relations* between compound constituents. Inspired by Levi's relational model, based on a fixed number of Recoverably Deletable Predicates, they examine the role of Thematic relations in the *comprehension* of N+N compounds ('conceptual combinations', or 'combined concepts' to use their terminology). The central idea of their approach is that language users make use of their knowledge about both the meaning and the use of concepts. That is, they know which combinations are appropriate for a given concept. By implication, the interpretation of compound words rests on the knowledge of the interaction of the concepts involved in the conceptual combination. In their CARIN model (Competition Among Relations in Nominals), they use the term *competition* between various thematic relations to demonstrate that the thematic relation most readily available, i.e., the relation with greatest strength, is one which is preferred in the interpretation of combined concepts.

Gagné & Shoben assume (1997: 74) that "the availability of a specific thematic relation varies from constituent to constituent and that this difference in availability affects the ease with which two constituents are combined."[16] In other words, a compound with a more frequently used thematic relation is easier to interpret than a compound based on a relation that is not used frequently. For example, the main thematic relation of *mountain* is a 'Locative' relation (*mountain cabin*, *mountain stream*, *mountain resort*). On the other hand, there are only few 'made of' relations for *mountain* (*mountain range*). Consequently, language users tend to interpret its combinations as ones based on the Locative relation. By implication, the ease of interpretation of this kind of complex word is proportional to the degree of probability of a particular thematic relation.

Gagné & Shoben's experiments confirm this assumption to the effect that, in defiance of the general view of the dominating role of head in complex words, and in contrast with the conclusions of Hampton (1987) and Murphy (1988) concerning the dominant position of head noun in the interpretation of compound words, it is the *relational information* about the modifier which is crucial to the interpretation of conceptual combinations.[17] The authors assume

that the "information about thematic relations is stored directly with the modifier concept and can thereby guide the search for the correct interpretation of a combined concept" (1997: 83). The knowledge of the probability of the use of a particular thematic relation with a given modifier concept "is called a *relational distribution* and reflects a person's experience with the language and with combined concepts in particular" (Gagné 2001: 237).

These conclusions were further buttressed by a series of Gagné's experiments (2001). Gagné examined the influence of lexical priming and relation priming upon the interpretation of compounds, and arrived at the conclusion that – with regard to thematic relations between compound constituents – a previously presented compound word (prime) influences the interpretation of a current compound word (target) if the prime and target share the same modifier. The head noun priming has no such effects. This, as suggested by Gagné, argues against the schema-based theories according to which a particular slot in the head concept is filled with a corresponding modifier concept emphasising the central role of the head concept in the interpretation process.

Nevertheless, the head is not devoid of its role in the interpretation of combined concepts. In a brief note Gagné & Shoben (1997: 83–84) hypothesise that when the appropriate thematic relation is selected by the modifier the head noun may play a key role in elaborating the meaning of the combined concept.

Gagné & Shoben's idea of *competition* should be highly valued, because it points out the *gradeable* nature of the *acceptability* to language users of various possible meanings. In Štekauer (1998) the same term is used to account for and calculate the relative productivity of various Word Formation Types (WF Types) belonging to the same conceptually defined cluster (Agents, Instruments, Actions, etc.).

Since, as will follow from the account in Chapter 3, it is postulated that the word-*formation* and word-*interpretation* processes are mutually closely interrelated, the notion of competition is, naturally, re-introduced in my approach to meaning predictability. While in Gagné & Shoben's approach the competition pertains to thematic relations available for the interpretation of compound words, the onomasiologically founded theory of meaning predictability employs this notion to account for the unequal Predictability Rates of various potential readings of a novel naming unit. The concept of Predictability Rate thus reflects the competition between various potential readings of a naming unit, and makes it possible to identify the reading with the highest chances of being selected by a language user out of a number of possible candidates when such a naming unit is encountered outside context for the first time.

1.3.3 Analogy-based models

1.3.3.1 *Ryder's cognitive approach*

M. E. Ryder (1994) presents an interesting contribution to the theory of meaning predictability (of primary compounds). Her *analogy-based approach* draws on the principles of Langacker's *cognitive grammar*. Ryder assumes that

> to the extent that speakers and interpreters agree on the types of relationships considered relevant, there should be at least a limited amount of predictability for the meaning of novel compounds even when they are divorced from their presentation context,

and correctly adds that

> the relationship inherent in any morphological structure, including a compound, will be only a part of the meaning of the word. The rest of the meaning will result from the relationship between the word and the context or contexts in which it comes to be used (1994:10).

This seems to be a widely accepted view, also shared by lexicographers. An extreme variant of this position is held, for example, by J. Sinclair (1998) who argues against the word-based description of the meaning of lexical entries as is the established practice in dictionaries, because words occur in different 'cotexts' which give the final shape to the basic meaning of words, and significantly increase the number of possible meanings of a word. In Sinclair's view there is a significant difference between the finite number of meaningful items and infinite number of their applications (the number of meanings in texts).

To return to Ryder, she criticises the traditional generative approach based on defining certain rules and sweeping any deviations from these rules to the lexicon as a repository of irregularities and idiosyncrasies. In her view productivity is a cline and therefore it is not possible to put a clear-cut borderline between rules (i.e., what is regular) and lexicon (i.e., what is idiosyncratic) (1994:49).

It goes without saying that the rule-lexicon relation and the role of the lexicon in the system of grammar depends on one's theoretical framework. For example, in the onomasiological theory proposed in Štekauer (1998, 2001a) and briefly outlined below it is postulated that all naming units generated in the Word formation Component are regular (generated by productive WF Rules) and that any deviations (semantic shifts and/or formal modifications) are supposed to take place in the Lexical Component (which accommodates all simple and complex words, domestic and borrowed, as well as affixes). By implication, if one's attention is focussed on the predictability of *novel* (and therefore, reg-

ularly and productively coined) naming units (as is that of Ryder and myself) there is no need to bother about idiosyncrasies.

Since the traditional account of compounds has a number of flaws Ryder takes recourse in cognitive grammar, which provides her with the method of description and the terminology. Based on Langacker's (1987) ideas, Ryder takes two important postulates as her point of departure: (1) linguistic categories like other categories are *prototypes* rather than absolute sets, and (2) the human mind allows redundancy, and thus a lot of information can be recorded as both specific instances and as generalisations (1994: 63).

Her theory is based on the notion of schema. A *schema* is conceived of in accordance with Anderson and Pearson (1988: 42) as "an abstract knowledge structure." Schemas have variables which have a range of possible values, one of which may be a default value. They represent encyclopaedic information, and are dynamic processes, i.e., they can change. There are three basic schema types, notably, 'event schemas' (such as DOCTOR PERFORMING SURGICAL OPERATION), 'entity schemas' (DOCTOR: adult, habitually dressed in a white coat, having a stethoscope in his pocket, etc.), and 'feature schemas' (they can be abstracted from the entity schemas and event schemas for 'doctor').

The schemas used by Ryder for the interpretation of compounds are labelled as *linguistic templates*. The particular template used in constructing a new compound is called an *analogy base*. Linguistic templates represent a continuum of increasing abstractness and generality. An analogy base can be either an individual conventional expression such as *doghouse* or *birdcage*, or groups of compounds that share a common constituent, for example, *sea lion, seaman, sea cow, sea weed*, etc. Such a group underlies a more abstract linguistic template if there is a high-level correlation between the pairs of constituents, the so-called *cue-reliability*. This can be instantiated by *box* compounds (1994: 80):

(5) X + *box* = a box intended to contain/store X

Such groups of established words sharing a 'core word' are, according to Ryder, "likely to be favoured choices as analogy bases, since, on the one hand, they are fairly low-level and detailed, while, on the other, they are frequent enough to be fairly easy to find in memory" (1994: 89).

Still more abstract linguistic templates are based on similarities of meaning of the construction as a whole and the component structures, e.g. *cigar box, flour sack, water glass, flowerpot, saucepan, etc.*

Language users, in interpreting novel noun + noun compounds, bring to bear three types of knowledge, i.e., (1) the form that the speaker/writer chooses to use; (2) the context in which the form is presented (if any); (3) and certain

specific assumptions. One such assumption is the use of linguistic templates by speakers. Ryder gives the following example of a possible novel compound interpretation:

(6) *Water sock* (= something shaped roughly like a sock that water can flow through or by, that indicates the speed and direction of the flow. Based on *wind sock)*[18]

In her view, if the interpreter identifies *wind sock* as the linguistic template serving as an analogy base, (s)he will be able to come up with the 'correct' meaning for it. If this template does not come to his/her mind, (s)he might think of another conventional expression instead, one involving *water* rather than *sock*, such as *water balloon*, which, however, would be an erroneous interpretation (something like 'a sock filled with water and thrown at someone as a practical joke, usually used by children, often at outdoor parties').

There are also some other possible interpretations proposed by Ryder, such as that based on the linguistic template (7):

(7) Natural Element + Article of Clothing = clothing to be worn in contact with that element as protection from it (*raincoat*, *snowsuit*, etc.)

As suggested in the Introduction, the introduction of the term 'correct meaning' in the discussion of the meaning predictability of novel, context-free naming units should be eschewed. This is, in fact, related to the definition of the notion of meaning predictability of novel coinages. If this notion is defined as the degree to which the individual meanings of a novel naming unit can be predicted, i.e., the degree of their respective acceptability to language users, it is not possible to speak of a 'correct meaning'. A 'correct meaning' (i.e., the actual meaning of a novel naming unit) need not be the one with the highest predictability/acceptability to interpreters. From this it follows that the relation of correspondence between the meaning predictability and the 'correctness' of language speakers' judgements of the meaning(s) of novel naming units does not hold in all individual cases, and therefore cannot be a defining feature of meaning predictability.

Note that the Predictability Rate value in the example given above will be different for specialists, or at least people with some technical skills, and those who are technologically ignorant. In principle and in general, an experiment might demonstrate that the latter two of the above-mentioned possible interpretations of *water sock* might gain a higher Predictability Rate than the 'correct' one.

Ryder further maintains that the next assumption of the interpreter is the use of real-world knowledge, called *semantic information schemas*, i.e., event, entity, and feature schemas. The final step is *accommodation*: in the schema-based interpretation of the *valence* relations Ryder emphasises the importance of *correspondences* between the shared substructures and the component structure, which sometimes may only be achieved by accommodation. To put it another way, if the interpreter cannot find or create a common semantic information schema, he must accommodate a schema for one of the compound constituents to get the correspondence.

1.3.3.2 *Coolen, van Jaarsveld, and Schreuder*

The analogy-based model principles were widely examined by Coolen, van Jaarsveld, and Schreuder (1991), who assume that "the interpretability of isolated novel compounds may be determined by the availability of lexicalized compounds that can serve as a model for the interpretation" and the "[r]elations within these lexicalized compounds may be among the first ones that are considered in the interpretation process" (1991:350). This view was later experimentally verified by the same authors in van Jaarsveld et al. (1994). While their experiments show that analogous lexicalised compounds influence the decision-times of informants – suggesting that they are involved in processing novel isolated compounds – no influence upon the interpretability of the novel compounds has been observed. The experimental results lead the authors to the conclusion that "it is rather unlikely that the interpretative processing of novel compounds uses lexicalized compounds as models..." (1994:130). In addition, "[a]nalogous interpretative processing may not be a viable option for novel compounds, because of the semantic variation within sets of analogous compounds" (1994:131). Consequently, these authors believe the above-mentioned feature or schema models, which confine themselves to the semantic representations of compound constituents, to be more relevant to the interpretation of the compounds in question.

While my research seems to have confirmed these conclusions the analogy-based factor may not be excluded from consideration in individual cases, as follows from the respective readings of *dog spade*, *apple-juice seat*, and *garden whisky* discussed below.[19]

1.3.4 Combined and other models

1.3.4.1 *Finin's artificial intelligence framework*

Finin's work *The Semantic Interpretation of Compound Nominals* (1980) presents a theory designed within the framework of the science of artificial intelligence and implemented in the form of a computer program. Finin aptly notes that the selection of the most appropriate semantic relationship relies upon a host of semantic, pragmatic, and contextual factors. Finin's theory should be appreciated for placing emphasis on the conceptual basis of compound interpretation by developing a *frame-based representation system* to represent *concepts* and the *relationships* between them. Importantly, concepts are hierarchically organised, with the hierarchy representing different levels of abstraction in the order of 'frame' – 'slot' (role) – 'facet' – 'data'.[20] As a result, *attributes can be inherited*. An important part of Finin's system is a *concept matcher* which determines the mutual compatibility of the concept frames, a pattern concept and a target concept, underlying the compound. The *pattern concept* is a general description of a class of objects, while the *target object* provides a description of another class of objects. The matching operation is successful if the description provided by the pattern concept includes that provided by the target concept, that is to say, "if every object described by the target description is also described by the pattern description" (1980:74).

The semantic interpretation must first identify the underlying concepts and then identify the relationships between them. Compounds cannot be interpreted without employing our knowledge. Thus, for example, the compound *GM cars* (1980:4) in the meaning of 'cars made by the General Motor company' is primarily based on our knowledge of GM as a car manufacturer. Therefore, Finin's representational system is designed to capture intensional knowledge about the world as well as extensional knowledge, that is, that kind of knowledge which is not always evident in the surface form of a compound.

Within the interpretation part of his work Finin presents another attempt to develop a classification of basic semantic relationships between compound constituents by proposing general *structural interpretation rules*. The procedure is based on taking all rules which are applicable to a given compound. Each rule application yields one or more interpretations.[21]

The application of rules and the obtained semantic relations undergo the process of *role fitting* aimed at determining what role of frame Y can best accept X as its filler (i.e., value).[22] In principle, this is a matching process. Since there are generally several acceptable roles for which X can be a value, Finin proposes

a method of ranking these options, the details of which are, however, irrelevant to the present discussion.

The value of Finin's work follows from the fact that his theory integrates conceptual knowledge in the process of meaning interpretation and that it "can assign a sequence of two or more nominals a 'most likely' interpretation in a null context, given adequate representations for the concepts involved" (1980: 123).[23]

1.3.4.2 *Meyer's knowledge representation model*

Meyer (1993) is an example of a *knowledge representation model*. Meyer proposes a theory of meaning variability for novel root compounds in a verbal context which provides all kinds of knowledge necessary for novel root compound interpretation. While these compounds are ambiguous in isolation some of the possible meaning relations are more salient than others. In addition, their meanings are predictable in a particular context thanks to specific interactions of different knowledge systems.

As Meyer notes, ambiguity is not an inherent feature of novel root compounds (although it covers a major part of them). The compound *book fan*, for example, is hardly interpretable in a different way than 'fan of books' (1993: 4).[24] The ambiguity of root compounds is based on the existence of various conceptual relations. Meyer emphasises the importance of *conceptual (world) knowledge* and *semantic knowledge* for their interpretation, and explains their interaction in interpreting a compound. Importantly, the interpretation of compounds is based mainly on *prototypical features* of objects. What must be identified for a compound are the *plausible relations* based on prototypical concept properties.

While the centre of gravity of Meyer's book is the interpretation of novel NN-compounds in context, he also discusses relational ambiguity of isolated NN-compounds and explains why some relations are more salient or natural for a certain compound type than others. His approach is based on *Discourse Representation (DR) Theory*. Lexical meaning of single nouns is represented in lexical DR Structures. These are related to concepts in order to provide the set of possible domains a noun can denote. Compound interpretation is based on lexical meaning and conceptual representations. Crucial to NN-compound understanding is the knowledge of their respective concepts, the knowledge of the properties of the conceptualised extra-linguistic objects as well as the inference of the relation between the respective extra-linguistic objects.

I share Meyer's view that it is primarily the combination of features of concepts that triggers the interpretation of novel NN-compounds rather than any

compounding rule, and that "[t]he reason why a certain compound gets a certain number of possible interpretations is founded on conceptual knowledge rather than a fine-grained noun subcategorization" (1993:104).

Meyer distinguishes between two kinds of novel NN-compounds. *Sortal compounds* like *gold ship, screw table,* and *highway town* denote entities of a certain sort or concept, and for their interpretation they require conceptual knowledge of objects represented by the constituents (his term is 'extensions') and of possible relations between them. Since there are usually several such possible relations this type of compound is ambiguous.

Relational compounds like *cupboard side, book fan,* and *institute address* denote entities that are related to different entities, and are, in his view, interpretable by means of the relational property of the head noun, without access to the denoted concepts. In other words, the internal argument (θ–role) of the head noun is satisfied by the modifier. The internal argument may only be assigned if certain selectional restrictions are met. For example, while *soldier brother* is formed by the internal argument assignment, *computer brother* is not since the selectional restriction for *brother* is Animate; *computer* does not satisfy the selectional restrictions of *brother*. Therefore, the latter compound can only be understood by 'inference'; in particular, by inferring the relation between head concept and modifier concept (possibly giving the meaning 'the brother who owns a computer').

Meyer identifies three main hierarchical levels of the interpretation of context-free compounds (1993:110ff.) with respect to the saliency of relations in compounds: (1) Level of Theta-role assignment, (2) Level of Lexical representation, and (3) Level of Conceptual structures.

While the syntactic operation of internal argument satisfaction, i.e., theta-role assignment, leads to the most salient interpretation of primary compounds with a relational head, the most salient interpretation of primary compounds with sortal head is based on the lexical representation of the head noun. The interpretations based on a conceptual analysis are determined by object-specific properties and inferential operations.

In addition, the above-represented hierarchy implies that if selectional restrictions prevent an interpretation based on a higher level, the next preferred interpretation is based on the next lower level. In other words, the non-applicability of one relation leads to alternative interpretation possibilities. Access to a lower level is always possible in order to select alternative relations. Moreover, the conceptually determined relations are also hierarchically ordered, with unusual interpretations at the bottom of the hierarchy. These can be illustrated by the following example. Solids or pastes are sub-

domains of the domain of masses of substances. They may enter into mixing relations, one of them being 'made of'. Then, the following picture emerges:

> 'Made of' is a mapping from solids or pastes to non-decomposable physical objects. However, individual pieces have an inner space and a surface as possible places for locations. Thus, if the conditions for the made-of relation fail, there may be a switch to an interpretation by a location relation. Depending on the size of the individual pieces of substance and object as well as the other properties... location in the inner space of the substance or location on the surface of it is possible. Therefore, one arrives at the class of location relations. These are relations between a place and an object. This place may be a proper place or a place provided by a piece of substance or an object. If an object provides this place, it is conceptualized on the basis of spatial functions as supporting, containment or protection. (Meyer 1993: 147)

Meyer's work is extremely valuable for emphasising the role of prototypical features, and as an attempt to account for possible relations and for the unequal status of various interpretations in terms of their acceptability.

1.3.4.3 *Wisniewski's two-process model*

Wisniewski (1996) is a contribution to the theory of context-free word *interpretation strategies*. Within an extensive experiment he studied the strategies people use in interpreting novel, context-free N+N compounds, and demonstrated that one should avoid a one-sided approach to the interpretation of conceptual combinations. While previous models typically rely on a single interpretation strategy Wisniewski's model reflects various interpretation strategies of language users. He distinguishes two fundamental interpretation strategies, the *relation linking* and the *property mapping* from one compound constituent to the other, with *hybridisation* representing an extreme case of property mapping, i.e., the combination of properties of the constituents. Therefore, he proposes a two-process model: "One process would capture relations between the objects by filling a slot in the head concept with the modifier concept" and "[i]n another process, people would compare the modifier to the head concept, noting commonalities and differences related to those commonalities" (1996: 448–9). In its employment of two different interpretation procedures, Wisniewski's model resembles that proposed by van Lint (see Section 1.2.3). The relation-linking process produces two concepts linked by a relation while the property mapping process results in one concept containing part of another concept, or the two concepts form a hybrid.

For illustration, an interpretation-linking strategy for *zebra-box* gives the meaning 'a box with a zebra in it'. A property mapping strategy applied to *zebra*

horse yields the meaning 'a horse with a property very similar to one of a zebra', and a hybrid interpretation of the same compound refers to a kind of a new animal with properties of both zebra and horse.

These basic strategies have their modifications. For example, in construal strategies the referent of a compound constituent need not necessarily correspond to what one expects from the meaning of that constituent. For example, *moose pencil* may be interpreted as 'a pencil with a moose eraser'. This reading introduces the concept of 'eraser' as an actual referent of *moose*. It functions as a *representation* of the actual 'moose' (1996:436).

Wisniewski argues that property mapping is characteristic in cases of high similarity between the concept of head and the concept of modifier, while low-similarity pairs are mainly interpreted by the relation-linking principle. For instance, the reading 'a hatchet for pounding in nails' identified for *hammer hatchet* is very likely because

> the properties on which 'pounding in nails' depends are generally shared by the similar *hatchet*. A hatchet has the properties of being rigid and heavy and of having a solid, blunt end similar to the one that hammers have; thus mapping the additional property of 'used for pounding in nails' is plausible (1996:442).

Let us note that this sort of 'similarity' between the concept of head and the concept of modifier should not be mixed up with up with Downing's constraint upon the combinability due to semantic redundancy (see above for her above-mentioned examples *head hat*, *egg bird*, *book novel*, etc.). The *hammer hatchet* type of compounds is a different case. While, for example, *head* does not contribute to the basic meaning of *hat* (including its purpose of 'head protection') the *hammer* constituent of *hammer hatchet* identifies a purpose not typical of *hatchet*, and thus specifies its meaning. Therefore, while the *head hat* type of similarity between compound constituents is characterised by redundancy the *hammer hatchet* type of similarity illustrates the case of semantic compatibility.

I find the idea of the *similarity of concepts* and the *shared properties* the focal point of Wisniewski's theory, an idea corresponding to what can be called the semantic compatibility of (prototypical) semes of the two motivating constituents and what constitutes one of the cornerstones of my theory of meaning predictability. However, unlike Wisniewski I believe that the compatibility of (prototypical) semes underlies the interpretation of all novel, context-free complex words of various Predictability Rates (not only the strategy of property mapping).

Wisniewski extends the scope of Murphy's 'elaboration' operation. It not only refines the first-stage interpretation of a combined concept, but also – in

some cases – alters the typical referent of a constituent. In his example with *tiger chair*, this happens because of the (alleged) non-existence of a plausible relation between *tiger* and *chair*. Consequently, an elaboration based on world-knowledge replaces the concept of 'tiger' with its part, i.e., 'tiger skin', thus establishing a plausible relation-linking interpretation 'a chair MADE OF TIGER SKIN'.

1.3.4.4 *Libben's semantic transparency model*

Libben (1998) presents a model of compound representation and processing in which the crucial notion is that of semantic transparency. He distinguishes three levels of representation: (1) the stimulus level, (2) the lexical level, and (3) the conceptual level. The stimulus level makes it possible to identify novel compounds like redberry. But while its constituents can be activated in the same way as in, for example, blueberry, it cannot be comprehended through lexical representation because it is not stored in the lexicon, i.e., it is not institutionalised (there is no such word in the lexicon). The morphemes of redberry can only be identified by morphological parsing.[25] On the other hand, existing compounds are represented at the lexical level. Therefore, unlike redberry, blueberry and strawberry are represented at the lexical level.

In addition, Libben's model distinguishes between semantically transparent compounds (*blueberry*) and semantically lexicalised bimorphemic units which, as Libben assumes, are monomorphemic in the minds of language users (*strawberry*). To put it another way, native speakers realise that while *strawberry* can be analysed into *straw* and *berry*, *strawberry* does not contain the meaning of *straw*. A fact which is, in modern terminology, labelled as 'lexicalisation'. This difference in semantic transparency is captured at the *conceptual level*. Libben distinguishes two types of semantic transparency. *Constituency* pertains to the use of morphemes in their original/shifted meaning (in *shoehorn*, *shoe* is transparent because it is used in its original meaning, while *horn* is *opaque)*. *Componentiality* bears on the meaning of a compound as a whole: for example, *bighorn* is non-componential because the meaning of this word cannot be inferred from the meanings of its constituents even if these are related to independent morphemes. This makes it possible to inhibit, for example, the lexical representation of *boy* of the lexical unit *boycott*, and to inhibit the meaning of *straw* to interfere with the interpretation of *strawberry*.

By referring to these considerations in Libben (1998), Dressler (in press) distinguishes four fundamental degrees of morphosemantic transparency of compounds:

1. transparency of both members of the compound, e.g., *door-bell*;
2. transparency of the head member, opacity of the non-head member, e.g., *straw-berry*;
3. transparency of the non-head member, opacity of the head member, e.g., *jail-bird*;
4. opacity of both members of the compound, e.g., *hum-bug*.

It goes without saying that type 1 is the most appropriate and type 4 least appropriate in terms of meaning predictability. It will be shown that any non-established figurative meaning of a motivating constituent hampers the meaning-prediction process. While Dressler's scale of transparency gives the head transparency priority over the non-head transparency, the discussion in this chapter shows that the views of the role of the motivating constituents vary (compare, for example, Hampton (1987) and Murphy (1988), who emphasise the significance of the head constituent for the interpretation of compounds, and Gagné & Shoben (1997) assigning priority to the non-head constituent).

Of particular relevance to my discussion are the conclusions of Libben, Derwing & de Almeida (1999) and de Almeida & Libben (2002), demonstrating that there are no clear-cut boundaries between the individual stages of novel compound processing, namely (1) the identification of the constituent morphemes, (2) access to the representations and meanings of those morphemes, and (3) the interpretation of a compound on the basis of the headedness principle and the meanings of the individual constituent morphemes. That the process of interpretation of novel compounds is not so straightforward is illustrated by their findings, obtained through extensive experimental research, that novel ambiguous compounds like *clamprod*, which, in a listening test, admits twofold structure, i.e., *clamp* + *rod* and *clam* + *prod*, are parsed by the activation of all four morphemes rather than by taking the 'first possible parse' principle.

In my view, this kind of parsing (and, obviously, interpretation), "highly correlated with semantic plausibility" (Libben, Derwing & de Almeida 1999: 381), corresponds to one's intuition and also logical inference: since morphemes are bilateral units rather than mere forms, one may expect that any parsing and interpretation procedure takes into account the most plausible (i.e., the most predictable) combination of *meaningful* morphemes instead of being limited to more or less automatic identification of formal constituents. This leads the authors to conclude that "the primary function of the prelexical parser seems to be to supply all possible parses of a string. This conclusion is supported by the finding in this study that ambiguous compounds prime

semantic associates of all constituents" (1999:385).[26] In the light of these observations, it may be assumed that the three above-mentioned 'stages' more or less overlap.

1.3.5 Non-compound interpretation models

1.3.5.1 *The Clarks' 'contextuals'*

The previous brief outline indicates that the focal point of works directly or indirectly related to the meaning predictability and/or interpretation of novel, context-free naming units are compounds, in particular, root compounds. Still, there are several important exceptions to this general orientation, with Clark & Clark's *When Nouns Surface As Verbs* (1979) being of paramount importance. Although their article deals with *contextuals*, i.e., new noun-to-verb conversions which – as follows from the label – heavily depend for their interpretation on the specific context in which they occur, it presents highly valuable ideas some of which are of a general nature, because they (1) also apply to coinages resulting from other word formation processes (as admitted by Clark & Clark in their concluding remarks, and (2) are relevant to both context-free and context-dependent interpretation (and thus also predictability) of meaning(s). Their innovative denominal verb convention is a case in point:

In using an innovative denominal verb sincerely, the speaker means to denote

(a) the kind of situation
(b) that he has good reason to believe
(c) that on this occasion the interpreter can readily compute
(d) uniquely
(e) on the basis of their mutual knowledge
(f) in such a way that the parent noun denotes one role in the situation, and the remaining surface arguments of the denominal verb denote other roles in the situation (1979:787).

Disregarding points (a) and (f), and taking into consideration points (b) through (e), this can be translated for my purposes in such a way that only those meanings are predictable which are formed by a speaker with regard to their predictability, i.e., with regard to the mutual linguistic and non-linguistic knowledge of the coiner and interpreter and that only those semantic components of the motivating words are of relevance which can be inferred by an interpreter without intensive guesswork.

In addition, Clark & Clark use the term 'salience' to refer to those features which are "conspicuously unique, given our mutual knowledge" (1979:787). Thus, the sentence 'The boy porched the newspaper', including the 'contextual' *to porch*, is interpretable because the interpreter can identify the salient feature of 'porch' (a shelter adjacent to the main door into a house). In addition, (s)he is able to relate it to their extra-linguistic knowledge of how newspapers are delivered to subscribers in the U.S.A. From this it follows that the Clarks aptly assign a significant role to 'world knowledge', including the more or less generally shared 'generic knowledge' (what people know about space and time, the basic physical laws, natural kinds, artefacts and their functions, etc.) and 'particular knowledge' differing from individual to individual. Clark & Clark maintain that denominal verbs mostly rely on generic knowledge about concrete objects. However, this claim seems to be too strong a generalisation. Novel naming units in general may come into existence in speech communities of various size. A small group of friends is a sufficient community to justify the coining of a new naming unit which may fulfil its function within this small speech community by heavily relying on their group-bound, and therefore, more or less individualised, knowledge and/or experience. To re-introduce my example from Štekauer (1996:125), if, in such a small group, Peter is notorious for wasting a lot of money by gambling, then the meaning of conversion *to Peter* in a sentence like 'I Peter'd all my salary last week' is predictable thanks to the mutual knowledge of that small speech community. By implication, it may be concluded that the interpretation of novel naming units relies on the mutual world knowledge of speaker and interpreter, this ranging from particular knowledge of a minor speech community to the generic knowledge of the major speech community.

Another important observation made by Clark & Clark, which is of high importance for a theory of predictability, is the role played by *predominant features*, defined as features that "are more central to the characterization of the category than others" (1979:789). Here, in fact, the authors give support to the theory of prototypes: "A predominant feature of a category is one that tends to hold for most of its members – especially its typical members – but not for members of neighboring categories" (1979:790). For example, a red brick is more central to (or typical of) bricks than gold brick, wooden brick, glass bricks, bricks of cheese, etc. Importantly, a category may have more than one predominant feature. The notion of predominant features and their role in predicting the meaning of novel naming units is integrated in my below-described theory in the form of 'prototypical semes' and their place within the hierarchy of seme levels.

1.3.5.2 *Beard's transpositions*

Beard (1995) makes valuable observations concerning the meaning-predictability of conceptually recategorised naming units which, in his model of *Lexeme-Morpheme Base Morphology*, are treated as *transpositions*. Let us illustrate his approach with Noun → Verb transposition. The most common semantic functions in this type of transposed word are BECOME(XY) and CAUSE(XYZ).[27] These meanings are bound to the grammatical properties of verb: "If the output of a verbal transposition is marked [–Transitive], assign it the predicate structure [BECOME(XY)]..." (1995: 181) and "[i]f the output of a verbal derivation is marked [+Transitive], assign it the predicate structure [CAUSE(Xy_i[BECOME(Y_iZ))])]..." (1995: 182). Not all transposed verbs are of a piece with this common scheme. A semantic shift may result in more specific meanings which are more difficult to predict:

(8) a. He brushed his coat with his hand.
b. I paddled the canoe with a copy of the New York Times.
c. He combed his hair with his fingers.

These verbs are based on nouns which are lexically Instruments. Beard assumes the following:

> ... since the definition of a noun contains its natural function, all features denoting anything other than the natural function may be ignored when the noun is used in verbal contexts. The predicted meaning of the verbal derivation, then, is the natural function of the noun. Thus, (*to*) *hammer* will mean 'pound,' (*to*) *brush* will mean 'wipe (with the intent of removing,' and so on. Indeed, since the meaning of the derivation is the natural function of the base, no grammatical functional derivation can be involved (1995: 183).[28]

And therefore, "[t]he specific output of the derivation is most accurately predicted by the input, that is, the semantic representation of the base" (1995: 185). As follows from the phrase 'natural function', Beard clearly refers to prototypical features of objects as important reference points for meaning-predictability, determining, as demonstrated in Chapter 3 below, the most predictable readings.

1.3.5.3 *Kiparsky's principle of canonical features*

Kiparsky (1997) assumes that speakers of English are able to assign the appropriate meaning to a denominal verb even if they happen not to have heard it before. While it is not clear whether this assumption also concerns the context-free interpretation of such neologisms the explanation presented by Kiparsky

is highly relevant to my discussion. He maintains that "conceptual knowledge is essential to the formation of lexical meaning" (1997:477). With his account based on an analysis of Locatum vs. Location verbs, Kiparsky formulates a general principle highlighting the role of conceptual knowledge in the interpretation process:

(9) If an action is named after a thing, it involves a canonical [i.e., conventional, generic – P. Š.] use of the thing (1997:482).

For denominal instrumental verbs this principle means that such a verb reflects a characteristic use of the corresponding instrumental noun. For example, the general meaning of *to tape* must be 'to apply or use tape', and this general meaning may acquire its specific forms, for example, 'to fasten, tie, bind, cover, support, record, or measure with tape'. At the same time, principle (9) excludes "ad hoc uses of tape: e.g. using a roll of tape as a paperweight is not 'taping' the papers, using a piece of tape to strangle someone is not 'taping that person, etc." (1997:482). Kiparsky does not aim to evaluate the predictability of the individual specific competing readings (the purpose pursued in his study is different). However, inspired by the semantic theory of Bierwisch (1983, 1986), Bierwisch and Schreuder (1992) and Wunderlich (1997), Kiparsky formalises the conceptual knowledge at the level of Semantic Form. For illustration, the respective Semantic Forms of *saddle* as a denominal locatum verb and *corral* as a denominal location verb are given in (10):

(10) Locatum: $\lambda z \lambda y \lambda x$ [CAUSE (x, (HAVE-ON (y, z))) & SADDLE (z)]
Location: $\lambda z \lambda y \lambda x$ [CAUSE (x, (BE-IN (y, z))) & CORRAL (z)]
(Kiparsky 1997:484)

Two important conclusions follow for us from Kiparsky's ideas. First, Principle (9) specifies a general constraint on predictable readings. The Semantic Form, exemplified in (10), resembles Levi's Recoverable Deletable Predicates. By implication, the semantic level discussed by Kiparsky can 'only' account for the first, general step in the meaning-prediction process that requires further elaboration. This elaboration is based on our *conceptual knowledge*. Second, the *canonical* (conventional, generic) nature of the denominal verb meaning and the elimination of ad hoc interpretations corresponds to the abovementioned 'natural function' proposed by Beard, and gives further support to the idea of the crucial role of prototypical features of objects in the meaning-prediction process.

1.3.5.4 *Dokulil's meaning predictability*

Let me conclude this brief review with the ideas of M. Dokulil (1978). There are at least two reasons for doing this. First, Dokulil seems to have been the first to directly discuss some of the issues of *meaning predictability* of novel, context-free naming units, as also follows from the title of his short but seminal article *On the predictability of the lexical meaning of a word formation-motivated word* [my translation]. Second, he develops his ideas within the framework of an onomasiological approach to word formation, which is underlies my own discussion.

When examining the possibility of predicting the actual lexical meaning of a naming unit when first encountered by a language user, Dokulil points out the significance of a *general word formation (structural) meaning* (i.e., a general word formation rule/type/subtype, etc.). Dokulil (1978:247ff.) demonstrates that there can be different relations between general word formation meaning and a particular lexical meaning. A case of rough identity can be illustrated by Agent nouns, such as *teacher*, where the general WF Type 'a person performing an activity expressed by the verb' essentially indicates the specific lexical meaning. A very high level of such an identity is represented by, for example, feminine names derived from their masculine counterparts, the names of properties like Czech *chytrost* ('ingenuity') derived from *chytrý* ('ingenious'), verbal nouns, diminutives, and verbs expressing the repetition of action. Obviously, the predictability of these naming units is fairly high.

On the other hand, some naming units may hardly be predicted from the corresponding general word formation meaning. Thus, some nouns derived from colour adjectives acquire an unpredictable, highly specialised and idiosyncratic meaning. For example, the noun *zelenina* ('vegetable') derived from the adjective *zelený* ('green') is motivated by the green parts of plants used for consumption; *modřina* ('bruise') derived from *modrý* ('blue') is motivated by the (prevailing) blue colour of a bruise (at some temporal point of its existence); and *šedina* ('grey hair' sg.) derived from *šedý* ('grey') is motivated by hair colour. In each of these cases the lexical meaning is highly specialised with regard to the general word formation meaning. Consequently, such naming units are unpredictable.

A different kind of divergence between lexical meaning and word formation meaning obtains when the lexical meaning is broader than the latter. Thus *truhlář* ('cabinet-maker') does not only make cabinets but also other pieces of furniture, *sedlář* ('saddler') produces harnesses in general, etc.[29]

Dokulil concludes that while word formation meaning is integrated in the system of language, lexical meaning is a matter of norm. Not everything permitted by a system is implemented in the norm.

Dokulil identifies the individual factors affecting the meaning predictability of novel, context-free naming units. The first factor is the *onomasiological category*. Those naming units which belong to the so-called *modificational onomasiological category* (the basic concept is completed with a modifying element) are very well predictable thanks to the identity between the lexical and the word formation meanings. They include diminutives (*stromek* ← strom = 'little tree' ← 'tree'), feminine names derived from masculine nouns, mass nouns, verbs expressing repetitive action (*psávat* ← *psát* = 'to write habitually' ← 'to write'), etc.

The naming units falling within the *transpositional onomasiological category* (the meaning of the motivating word is transposed to a different phenomenal category, a different word class) are also highly predictable (*hoření* ← *hořet* = 'burning'$_{N}$ ← 'burn'$_{V}$, *plynulost*$_{N}$ ← *plynulý*$_{A}$ = 'continuity' ← 'continuous').

The problems connected with meaning predictability pertain to the fundamental onomasiological category called *mutational* (an element of one phenomenal category is determined by its relation to another element of the same or some other phenomenal category). Here the influential factors include the word-class of the motivating word, with substantives representing the most complex case because, as suggested by Dokulil, "rather than by one mark, substances are usually determined by a large set of marks which resist any reduction to a single mark as a motive of a naming unit" (1978:248) [my translation]. Therefore, it is here that word formation and lexical meanings may deviate from each other, which reduces the overall meaning predictability. Other influential factors include the implicitness and explicitness of a naming unit in close connection to the difference between compounding and affixation, the word-class of a naming unit, the word-class and the semantic class of the motivating word(s), the monosemantic vs. polysemantic (homonymic) nature of both the word formation structure and its constituents, and the productivity of the respective WF Type.

1.4 Summary

The various approaches, views, and positions presented in the précis in Chapter 1 can be summarised as follows:

1. Observations of generative morphology made about the semantic interpretation of compounds (like those cited from Lieber (2004) in Section 1.2.4 and Note 5) are very general to contribute to the predictability of meanings of novel naming units.
2. Classifications such as those by Lees (1960), Levi (1978), Li (1971) and Finin (1980) are also too general and do not exhaust all the possible thematic relations. Therefore they cannot be used as the exclusive basis for the meaning-prediction process. At best they can be understood as a step in the meaning-prediction process which must be completed with a subtler analysis, possibly in the sense of Murphy's (1988) conceptual 'elaboration' of the basic interpretation.
3. A crucial role for the meaning-prediction process is played by the conceptual level of coinage analysis, based on an interpreter's world knowledge, experiences, and pragmatic factors in general (Finin 1980; van Lint 1982; Cohen & Murphy 1984; Meyer 1993). Importantly, however, the prediction process also makes use of linguistic knowledge (e.g., the knowledge of the availability of thematic relations (Gagné & Shoben 1997; Gagné 2001) in combination with non-linguistic knowledge (Clark & Clark 1979).
4. The combination of two concepts constituting a new concept, that is to say, the semantic features of the motivating constituents, must be compatible (van Lint's 'necessary association of features', Zimmer's 'appropriately classificatory relationship', and various variants of the 'schema (feature) model', based on the slot-filling principle, such as Allen's (1978) 'Variable R' principle, Finin's approach (1980), the 'concept specialization model' of Cohen & Murphy (1984), and the 'selective modification model' of Smith, Osherson, Rips & Keane (1988).
5. Individual features of objects/their semantic representations are not of equal value, and represent a hierarchy. In other words, not all semantic features of the motivating words (features of concepts) are equally significant for a coinage interpretation. The most important are the 'predominant features' (Clark & Clark 1979), 'natural function' (Beard 1995), or the '(proto)typical features' (or, prototypes) (Murphy 1988; Hampton 1983, 1987; Smith, Osherson, Rips & Keane 1988).
6. The interpretation of combined concepts is significantly influenced by thematic relations connecting the combined concepts, in particular, the thematic relations stored with the modifier (Gagné & Shoben 1997; Gagné 2001).

7. The meaning-prediction process can, in some cases, rely on analogy as proposed by various forms of the analogy-based model (Derwing & Skousen 1989; Skousen 1989; Ryder 1994; and also Gagné 2001).
8. Meanings reflecting habitual, permanent, fixed relations are more predictable than meanings based on accidental, temporary, context-bound relations (Zimmer 1971, 1972; Gleitman & Gleitman 1970; Downing 1977). This is closely related to the idea that the interpretation of novel words is based mainly on prototypical features of objects (Meyer 1993).

Since the majority of the approaches outlined above are not *meaning prediction*-oriented some important issues have been left unaddressed, others have been answered insufficiently. These issues include:

1. While noun + noun compounds can have several potential meanings representing different relations between the compound constituents (combined concept) this fact is mostly disregarded, and it is only one of them which is usually taken into consideration. The existing models propose various approaches to the interpretation of 'conceptual combinations', but do not attempt to identify one or several readings of a combined concept that has (have) the highest chances to be picked up by a language user when encountered for the first time outside context. It will be demonstrated that the interpretation of context-free novel coinages in general cannot disregard the mutual *interpretation-conditioning relations* between the *multiplicity of possible readings*, and that the *strength of any reading depends on the number and the strength of the other competing readings*.
2. A number of theories outlined above account for the interpretation of compound words as either relation-based or slot-filling-based (property mapping) or, the former in some cases and the latter in some other cases. In addition, some models, such as those proposed by Finin (1980) and Murphy (1988), combine slot-filling with the representation of relations between the modifier and the head. It will be demonstrated that a more adequate view of the process of word *interpretation* can be obtained by examining the word-*formation* process. It will follow from the analysis in Chapters 2 and 3 that the basic structure that underlies the *act of naming per se* includes *prototypical features* of the motivating constituents (the concepts of which come to be combined into a complex concept), and that these features establish the *meaning-defining relation(s)* between the naming unit constituents. Thus, for example, while Wisniewski (1996: 428) distinguishes two different strategies for two different readings of *box clock*, a property mapping for 'square clock' and a relation-linking for 'clock

contained in the box', the present approach assumes that both can be interpreted as relations between compatible characteristic features (properties) of the objects involved: the former reading is based on the SIMILARITY (or, PATTERN) relation between 'box' and 'clock': the [Square Shape] of 'box' is its prototypical feature which is compatible with the feature [Solid] of the object 'clock'. The latter meaning is also based on the relation, this time that of LOCATION, enabled by the compatibility of the features [Container] for *box* and [Placeable] for *clock*.

3. In general, interpretation-oriented theories disregard a number of factors which should be reflected in any meaning prediction-oriented theory, such as the word formation factor, the relation between the morphological and the semantic structures, the underlying Morphological Type and Word Formation Type, and competition between the various possible readings.

Since one of the central claims of this work is that word-*formation* and word *interpretation* processes are closely interrelated, and that a model of the former facilitates comprehension of the latter, Chapter 2 establishes the necessary theoretical word formation framework for an account of a meaning-prediction theory.

Chapter 2

General word formation framework

2.1 An onomasiological model of word formation

A theory of predictability of complex naming units will be discussed against the background of an onomasiological approach to word formation (Štekauer 1996, 1998, 2001a). I share the view with T. van Lint (1982: 136) that "interpretation cannot be divorced from production", which means that "interpretation presupposes a production system which only produces interpretable strings". It will be assumed that the 'interpretable strings', i.e., new naming units are formed in accordance with the onomasiological model of word-*formation* and that their meanings are predicted by means of its partially mirror-like deployment in the process of word-*interpretation* of novel naming units.

There are two basic approaches to the study of word formation, those identified by M. Dokulil (1962, 1968b) as 'word formation' and 'word-formedness', by M. D. Stepanova (1973) as process and result, M. Aronoff (1976) as word formation and word-analysis, K. Hansen (1978) as '*Wortbildung*' and '*Wortbildungsanalyse*', etc. The onomasiological model of word formation accounts for the formation of new naming units, thus concentrating on the dynamic facet of this phenomenon.

The model serving as the basis for the present discussion has been developed as a reaction to what I consider three major deficiencies of the mainstream generative approaches to word formation: first, their prevailing formalism – (R. Beard 1995 is an important exception) – which for the most part disregards the semantic facet of word formation; second, the limitation of the discussion to purely linguistic aspects, without regard to extra-linguistic reality and speech community, i.e. two indispensable factors in the triad of any process of forming new naming units. New naming units do not come into existence in a vacuum or accidentally. There is always a demand on the part of speech community to give a name to a new object, action, quality, or circumstance. Each naming process is conditioned and determined by the knowledge and experience of a particular 'coiner'. Third, the binaristic principle underlying the generative approach results in a number of problems, like bracketing para-

doxes (Štekauer 1999), different results of morphological and syntactic analyses (Kastovsky 1982), unsuccessful attempts to draw a clear borderline between compounding and suffixation, etc. Bauer (1978: 214–15) speaks of an ideal of having one grammar in which among other things, "[t]he different types of word formation might not then be kept apart", which "might also be a desirable consequence, since the borderline between compounding and affixation, for example, is not always clear...".

The onomasiological theory outlined in Štekauer (1996, 1998, 2001a) and applied to some specific problems of word formation in a series of articles (e.g. 1992, 1997, 1999, 2001b, 2002, 2004, in press[1,2]), is an attempt to describe all productive word formation processes using one common mechanism. It emphasises the triadic aspect of word formation existing between extra-linguistic reality (object to be named), speech-community (coiner) and word formation, in order to emphasize the active role and cognitive capacity of a coiner. At the same time, it establishes a framework for the treatment of the individual word formation processes on a common basis. It assumes that naming units do not come into existence in isolation from factors such as human *knowledge*, human *cognitive abilities*, *experiences*, *discoveries* of new things, processes, and qualities, human *imagination*, etc. An object to be named is not named on its own but is envisaged and construed in relation to the existing objects. Thus, the structural relationships in the lexicon are preceded (or dominated) by a network of *objective relationships* which, by implication, should be taken into consideration in the process of naming.

In particular, the scheme in Figure 1, inspired by Horecký's model of linguistic sign (1983) and Dokulil's theory of onomasiological structure (1962), represents important interconnections between extra-linguistic reality, speech community, the conceptual level as a supralinguistic level, and the relations between the individual components of grammar as well as inside the Word-formation Component itself. In view of my research into the predictability of the meaning(s) of new naming units the following principles are of vital importance:

1. The model lays emphasis on an *active* role of language users in the process of giving names to objects instead of presenting word formation as an impersonal system of rules detached from the objects named and language users.
2. The naming act/process is not a purely linguistic act/process. Naming units do not come into existence in isolation from factors such as human knowledge, human cognitive abilities, experiences, discoveries of new things,

processes, and qualities, human imagination, etc. An object to be named is not named on its own but is envisaged in relation to the existing objects. Thus, any linguistic processes are necessarily preceded (or dominated) by a network of 'objectively' existing relationships. By implication, the naming act/process is a *cognitive process* relying on the intellectual capacities of a coiner.

3. The model stresses a close interconnection between linguistic and extra-linguistic phenomena.

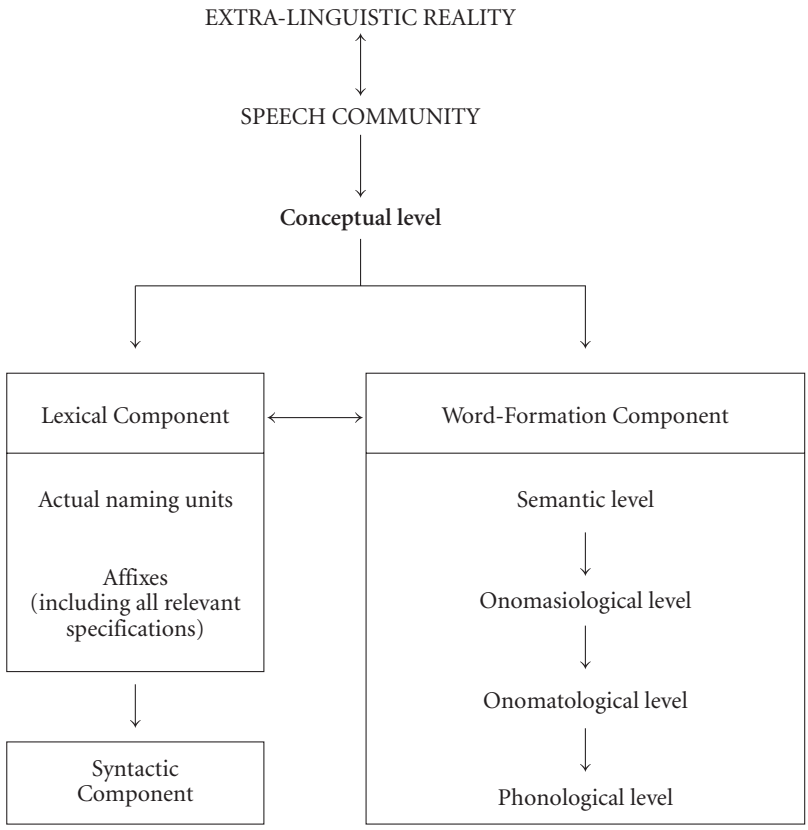

Figure 1

The onomasiological model of word formation includes the following levels:

(11) 1. Extra-linguistic reality
2. Speech community
3. Conceptual level
4. Semantic level
5. Onomasiological level
6. Onomatological level
7. Phonological level

The following theoretical account of the *act of naming* interprets the model graphically represented in Figure 1. For ease of understanding, the theory is illustrated with an example of the act of naming.

Extra-linguistic reality represents an *object to be named*. Each naming process responds to a specific demand of a speech community for assigning a name to an extra-linguistic object (in the broadest sense of the word). This is the reason why I find it necessary – in defiance of the mainstream theories – to shift the starting point of an account of word-*formation* beyond the limits of language as such, and include in it a *speech community* and its *linguistic demands*, i.e., (among other things) the need *to name* an object of extra-linguistic reality, and the level of the intellectual processing of an object to be named. By implication, through its manifold cognitive activities a speech-community selects what is there in extra-linguistic reality that deserves a name. This interrelation between extra-linguistic reality and a speech community predetermines all the subsequent steps.[1]

The notion of *speech community* should not be taken absolutely, i.e., there is hardly any word formation process which responds to the naming demand of all the speakers of a particular speech community. Rather, such a demand is closely connected with a limited number of 'first-contact' users, and a coinage may or may not subsequently find a wider use (i.e., become institutionalised).

To take up my example, one of thousands of 'objects' of extra-linguistic reality that were considered worth naming at some time in the past was 'a person whose job is to drive a vehicle designed for transportation of goods'.

The *conceptual level* as a supralinguistic level is independent of any particular language, and represents intellectual processing of the object to be named in a 'coiner's' consciousness by means of generalisation and abstraction processes. The primary task to be mastered is to analyse the object (in the broadest sense of the word) to be named (or better, a conceptual class of objects). A language user reflects the complexity of the object in the form of a *logical spectrum* delimiting the object by means of *logical predicates* (noemes), and by making use

of the most general *conceptual categories* (SUBSTANCE, ACTION (with internal subdivision into ACTION PROPER, PROCESS, and STATE), QUALITY, and CONCOMITANT CIRCUMSTANCE (for example, that of Place, Time, Manner, etc.)). The logical spectrum thus represents a knowledge structure reflecting the features of a class of objects, ranging from the most general features (shared with other classes of objects) to the most typical, the so-called *prototypical features*. Here, the notion of prototype is understood in accordance with Rosch (1978) and Rosch & Mervis (1975) as an 'average member' of the category featuring the most frequent attributes of the category members.

Furthermore, my understanding of the conceptual level is in accordance with Cohen & Murphy (1984: 31), who maintain the following: "When we offer a theory of concepts, we are not proposing a theory about how the external environment is structured, but rather about how people conceive of that environment."

Finally, it is important to note that any object in the sense of a conceptual class can have, in addition to the common, prototypical, features also features through which it differs from other members of the class, idiosyncratic features. The latter are latently present at the conceptual level.

In my example the logical spectrum can be represented as follows:

(12) The motivating Object 1 is SUBSTANCE$_1$.
SUBSTANCE$_1$ is Human.
The Human performs ACTION.
The ACTION is the Human's Profession (=Agent).
The Human is Agent.
The ACTION concerns SUBSTANCE$_2$ (=Object of Action).
The ACTION is based on Operation of SUBSTANCE$_2$.
SUBSTANCE$_2$ is a class of Vehicles.
SUBSTANCE$_2$ is an Object of ACTION performed by SUBSTANCE$_1$.
The Vehicles are designed for the Transportation of goods.
Etc.

It is postulated that the conceptual analysis is followed by *scanning* the Lexical Component. If a naming unit is found in the Lexical Component which can serve as a basis for semantic formation, it is the path of the Lexical Component which is preferred; otherwise, a naming process takes place in the Word-formation component (hence, two downward arrows from the Conceptual level in Figure 1).

The *Word-Formation Component* is considered to be an *independent* component of linguistic description. No natural language is a static system, fixed

once and forever. Rather, every language must be (and is) able to comply with an ever-changing extra-linguistic reality and the related language requirements of the particular speech community. From this it follows that every language is in a position to produce new naming units designating new 'objects', newly discovered phenomena, etc. It follows that every language needs a highly productive Word-formation Component. By implication, an independent Word-formation Component may qualify as a language universal.

All naming units coming into existence in the Word-formation Component are coined by *productive* and *regular* WF Rules.[2] Each new naming unit produced by a WF Rule is passed to the Lexical Component where it is stored. This approach makes it possible to simplify and regularise the Word-formation Component because any idiosyncratic changes take place in the Lexicon by way of *semantic formation* or *formal modification*. As a result, WF Rules are no less productive than Syntactic Rules or Inflectional Rules.

Since each act of naming responds to the immediate naming need of a speech community, the output of WF Rules is an *actual word*, i.e. a naming unit which was coined to satisfy a linguistic demand, be it the demand of a single member of a speech community, be it a single-act one-off demand. A word may only qualify for the status of an actual word if it has been coined. Whether its use will be spread over the whole speech community (implying frequent use), or whether it will be confined to a single use on the part of a single speaker, is insignificant. What is important is that the respective language has, by responding to the specific demand, manifested its capacity to provide a new, well-formed linguistic sign by its productive WF Rules whenever need arises. By implication, unlike, for example, Halle (1973) and Allen (1978), the present model does not overgenerate. By inclusion of extra-linguistic reality and the speech community this model only deals with those naming units which are required by a speech community, that is to say, with actual naming units.

I now proceed from the supralinguistic side to the description of the linguistic part of the word formation process. The structuralist approach to the linguistic sign emphasises its bilateral nature, including *signifiant* and *signifié*, i.e., form and meaning. The present model also follows this bilateral sign principle, and presents its subtler structure. The meaning facet of each new naming unit as a linguistic sign includes the semantic and the onomasiological levels, the formal side is composed of the onomatological and the phonological levels.

Individual logical predicates, specified by a 'coiner' at the conceptual level, are mapped onto the *semantic level* of a linguistic sign by means of *semes* (the notion of 'seme' is conceived of here in accordance with the notion of 'semantic marker' used in the theory of componential analysis) which constitute the

semantic structure (sememe) of the linguistic sign.[3] Those semes which represent the most typical, prototypical features of an object will be labelled herein as prototypical semes. It is postulated that semes are organised hierarchically, from the most general to the most specific. The meaning-prediction model described below distinguishes five levels of semes, with the prototypical semes representing level 4, and the idiosyncratic features of instances (tokens) of the conceptual class of objects representing level 5. For my word formation model a linear representation of semes will do. In my example the representation given in (13) is mapped onto the semantic level as follows:

(13) [+Material] [+Animate] [+Human] [+Adult] [+Profession] [+Agent];
[+Material] [–Animate] [+Vehicle] [+Transportation] [+Object of Operation] etc.

The name of the theory, the onomasiological theory, suggests that the *onomasiological level* is the central level of the model. The essentials of the concept of onomasiology as a dynamic approach to word formation were laid by M. Dokulil, a prominent representative of the Prague School of Linguistics, in his monograph (1962) and in a series of articles (1964, 1968a–d, 1997), and by Ján Horecký (1983, 1989, 1994). In Dokulil's view onomasiology deals with various types of conceptual structures resulting from a generalised reflection of objective reality in human consciousness and its processing in accordance with the naming means available in a particular language. Onomasiological categories are thus the fundamental conceptual structures underlying the process of naming. The phenomenon to be named is usually identified with a specific conceptual class, having its categorial expression in the particular language, and subsequently, within the limits of this class, it is determined by a mark. The conceptual class enters the onomasiological structure as a determined constituent – the onomasiological base, the mark as a determining constituent – the onomasiological mark (1962: 29).

Thus, at the onomasiological level we distinguish the *onomasiological base* ('head', 'determinatum') and the *onomasiological mark*. While onomasiological base is always simple, the onomasiological mark can be both simple and complex. In the latter case, it distinguishes the *determining constituent* (which sometimes can be analysed into the specifying and the specified elements) and the *determined constituent*. Both base and mark represent one of the abovementioned conceptual categories, SUBSTANCE, ACTION, QUALITY, and CIRCUMSTANCE. The base and the mark constituents represent an *onomasiological structure* reflecting the relations between the logico-semantic categories like Agent, Patient, Logical Object, Instrument, Time, Place, Factitiveness, Action, Process,

State, etc. The respective meanings of the individual logical-semantic categories (semantic case roles and predicates) as they are used in this book mostly rest on their use in Hansen et al. (1982).

To return to my example, it follows from the conceptual level analysis that a good candidate for the act of naming seems to be an onomasiological structure in which the onomasiological base stands for an Agent (the class of Humans performing the Action as their profession) of Action (the determined constituent of the onomasiological mark) aimed at its Object, i.e., the class of Vehicles (the determining constituent of the onomasiological mark):

(14) [(Logical) Object ← Action – Agent]

The term *onomatology* was introduced into linguistics by Vilém Mathesius, the founder of the Prague Linguistic Circle in 1926. Mathesius distinguished two levels of the description of language: (1) functional onomatology as a study of the naming units, and (2) functional syntax as a study of the means by which naming units are brought into mutual relation (1975: 16). The distinction between the terms 'onomasiology' and 'onomatology' in my conception of word formation follows from their reference to the levels belonging to two different facets of linguistic sign. Onomasiology zeros in on the meaning facet of linguistic sign, and identifies the semantic structure which conditions the 'surface form' of a naming unit.

Thus, at the *onomatological level* the onomasiological structure is assigned linguistic units based on the *Morpheme-to-Seme-Assignment Principle* (MSAP). Specifically, individual members of the onomasiological structure are linguistically expressed by word formation bases and possibly affixes, stored in the Lexicon.

In my example there are several options at this level. Thus, Agent can be expressed by *man, -er, -ist, -ant, -ian,* etc., because the meaning facet of each of them can be represented as 'Agent'. The Action of Operating the SUBSTANCE$_2$ can be expressed by WF bases of naming units *drive, steer, operate*, etc., because the meaning facet of each of them matches with the seme [Operation]. Finally, the (logical) Object can be represented by *truck*, *lorry*, and possibly some other WF base, the meaning of which is Vehicle. The selected options in my particular case are as follows:

(15) Object – Action – Agent
truck *drive* *-er*

There are at least two other possible representation types of the selected onomasiological structure. First, SUBSTANCE$_2$ may be backgrounded, in which case

the resulting naming unit may be, for example, *driver*; second, Action may be backgrounded, which may yield something like *truckist* or *truckman*.

The fact that all productively coined naming units are formed in the same way, by application of the same principle of Morpheme-to-Seme Assignment makes it possible to dispense with the traditional notions of word formation processes, including compounding, affixation, back-formation, or blending, and put the generation of all naming units on a uniform basis. While the traditional classification of word formation processes is exclusively based on formal criteria, such as extending vs. reducing the stem (word formation base), i.e., concatenative vs. non-concatenative processes; combination of two stems vs. stem + bound morpheme vs. internal stem modification; etc., the proposed approach emphasizes the bilateral nature of naming units as linguistic signs by reflecting their meaning facet. In addition to reducing the number of classificatory criteria to a single criterion (i.e., MSAP), the proposed approach makes it possible to reveal the naming strategies, and to show, for example, that formally different naming units may result from one and the same naming strategy. Thus, for instance, the naming process concerning 'a device designed to feed (machines with components)' may take the path represented by the onomasiological structure [Action – Instrument] and linguistically expressed by means of MSAP as either *feed unit* or *feed-er*. While the traditional classification does not reflect the common features of these two naming units and concentrates on their formal difference the proposed method views them as two different realisations of one and the same naming strategy, which is reflected in their falling within one and the same Onomasiological Type (for further discussion see 2.2 below).

At the phonological level, the new naming unit is 'shaped' in accordance with relevant phonological rules. In my example, it is the assignment of the corresponding stress pattern.

(16) ˈtruckˌdriver

2.2 Onomasiological Types

From the point of view of the final form of a naming unit it is important to determine what kind of onomasiological (semantic) structure is employed in the naming act. From this point of view five *Onomasiological Types* (OTs) can be distinguished. They are based on the criterion of which constituents of the onomasiological structure are linguistically expressed at the onomatological level. In general, the onomasiological structure includes three basic constituents:

(17) Determining constituent of the mark – Determined constituent of the mark – Onomasiological base

Let us recall that the onomasiological base identifies a whole class to which the object named belongs and corresponds to what is called *head* in generative linguistics, and the onomasiological mark 'entrenches' the object named with regard to all the other members of the class. The determined constituent of onomasiological mark generally stands for the concept of ACTION in its three different variants (PROCESS, ACTION PROPER, and STATE) and expresses the relation between the polar members of the onomasiological structure.

The example with *truck driver* illustrates *Onomasiological Type 1* in which all three onomasiological level constituents, i.e., the onomasiological base, and the determined and the determining constituents of the onomasiological mark are linguistically expressed. Other examples are *language teacher*, *brain-storming*, *air hostess*, *housing development*, *photo-sensor*, *sea-rover*, etc.

In *Onomasiological Type 2* the determined constituent of the onomasiological mark is expressed while the determining constituent is not (*teacher*, *lock nut*, *sensing electrode*, *stop button*, *stop-watch*, etc.). Importantly, this type is extendable to Onomasiological Type 1: *teacher* may be extended to include the determining constituent further specifying the meaning, for example, *teacher* may be extended to *language teacher*, *dance teacher*, *private teacher*, etc.; *lock nut* can be extended to *subassembly lock-nut*, *spring lock-nut*, etc. From this it follows that the determining constituent of the mark is present at the onomasiological level, and can anytime be 'activated' by MSAP.

In *Onomasiological Type 3*, it is the determined constituent of the onomasiological mark which is left unexpressed (*hatter*, *policeman*, *alpinist*, *honey bee*, *summer house*, *sun lamp, etc.)*.

In *Onomasiological Type 4*, the onomasiological mark cannot be analysed into the determining and the determined parts (*blue-eyed, unhappy*). Therefore, naming units falling within this Onomasiological Type, distinguish only two constituents, the onomasiological base and the onomasiological mark. For *blue-eyed*, the base is *-ed* and the mark is *blue eye*; for *unhappy*, the base is *un-* and the mark is *happy*, for *restart* the base is *re-* and the mark is *start*.[4]

Onomasiological Type 5 stands for *Onomasiological Recategorisation*, traditionally called conversion or zero-derivation. The onomasiological approach to conversion is based on the fact that each naming unit results from an intellectual analysis of an extra-linguistic object to be named. Within this analysis the object is classed with one of the four above-mentioned conceptual categories.

The individual aspects of extra-linguistic reality do not exist in isolation; on the contrary, they can be conceived of and subsequently linguistically expressed in various relationships, from different points of view. These different 'angles of reflection' of extra-linguistic reality can be cognitively brought into a close relation by re-evaluating the already existing logical spectrum and all the related lower levels. Then, the most striking feature of conversion is that it always linguistically expresses the conceptual recategorisation of extra-linguistic reality. Thus, for example, *databank* represents a SUBSTANCE. When conceptually recategorised, it becomes an ACTION; *experiment* expresses a PROCESS – after recategorisation it refers to an ACTION PROPER; *limit* is a CIRCUMSTANCE – after recategorisation it is an ACTION; *feature* is a QUALITY – its recategorisation yields an ACTION; *insert* is an ACTION – when recategorised it becomes a SUBSTANCE; *stand* belongs in a STATE – when recategorised it becomes a SUBSTANCE; etc.

What is the mechanism of these changes? As already mentioned, the individual logical predicates constitute a hierarchy. The recategorisation process consists in substituting the original dominating logical predicate which determines the conceptual category of a new extra-linguistic object to be named. The conceptual re-evaluation of extra-linguistic reality precedes the linguistic processes proper. It is the conceptual recategorisation which provides us with the evidence that conversion cannot be identified with zero suffixation: conceptual recategorisation is vital to conversion while only possible for suffixation.

Let us illustrate this point. The naming unit *milk* belongs to the conceptual category of SUBSTANCE. The conceptual level of Onomasiological Recategorisation with a hierarchy of logical predicates is given in (18). When the hierarchy within the logical spectrum in one of the converted meanings of *milk* ('to obtain milk from a female mammal') is changed, the recategorisation from SUBSTANCE to ACTION takes place. The central position within the hierarchy of logical predicates is assumed by a predicate focusing on the Actional aspect of the extra-linguistic object.

(18)

Original logical spectrum		New logical spectrum
SUBSTANCE	⟶	ACTION
		↓
{ It is material		GET {...}
It is inanimate		
It is liquid		
It comes from female mammals		
It is a foodstuff		
:		
...}		

As opposed to Onomasiological Types 1–4, Onomasiological Type 5 is characterised by an unstructured onomasiological level. There is no onomasiological base and there is no onomasiological mark. The original and the new dominating conceptual categories are related directly.

The following are some examples which also illustrate the way of classification of individual WF Types within the Onomasiological Recategorisation:

(19) a. *bond*$_{N}$ – *bond*$_{V}$:
[SUBSTANCE$^{\underline{\text{Result}}}$ACTION]
(in the meaning of a joint)
Interpretation: Substance as a Result of Action

b. *switch*$_{N}$ – *switch*$_{V}$:
[SUBSTANCE$^{\underline{\text{Instrument/Result}}}$ACTION]
(in the meaning of a device for completing or breaking an electric circuit)
Interpretation: Substance as an Instrument of Action

c. *insert*$_{V}$ – *insert*$_{N}$: [ACTION$^{\underline{\text{Object}}}$SUBSTANCE]
Interpretation: Substance as an Object of Action

d. *time*$_{N}$ – *time*$_{V}$:
[CIRCUMSTANCE$^{\underline{\text{Temporal}}}$ACTION]
Interpretation: Action in terms of Temporal dimension

e. *clear*$_{A}$ – *clear*$_{V}$:
[QUALITY$^{\underline{\text{Result}}}$ACTION]
Interpretation: Action Resulting in a certain Quality

Chapter 3

A theory of predictability

This chapter presents a general theory of predictability and its various aspects. Unlike Libben (1998: 31), who suggests that "models [of morphological processing] appropriate to one Morphological Type (e.g., prefixed words) may not be applicable to other Morphological Types (e.g., derived suffixed words)", I believe that the meaning predictability of all naming units falling within the first four Onomasiological Types is based on the same principles, irrespective of which of the traditional word formation processes is involved in a particular case. To put it another way, I believe that a unified theory of meaning predictability can be proposed, which applies equally to compounding and affixation. This assumption follows from the model of word formation outlined above which obliterates any differences between the traditional word formation processes. It will be shown, however, that Onomasiological Type 5, i.e., the type that corresponds to what has been traditionally called conversion or zero-derivation, requires certain modifications due to the absence of any onomasiological structure.

3.1 Why context-free meaning predictability?

A question may be raised as to why attention should be paid to context-free naming units if, in fact, novel naming units usually occur in some context – linguistic or situational. While this is true, the context may be said to establish the necessary preconditions for, as it were, the final tuning up of the meaning which follows most naturally from the relation between the motivating constituents of a novel naming unit, i.e., the meaning which most naturally results from the conceptual processing of the named object, in close relation to the objects captured by the meanings of the motivating constituents. This appears to be a consequence of the nature of the naming act which is context-independent, and is essentially of a cognitive nature. The core meaning (cognitive, denotative meaning) resulting from the act of naming is constant in each context. As suggested by Murphy (1988: 531), the semantic interpre-

tation of isolated complex words "avoids any contamination by discourse referents." Thus, context-free naming units, 'non-contaminated' by the context, represent the 'purest' result of the act of naming. Unlike various individual, unique, and idiosyncratic context-dependent meanings, context-free interpretation admits generalisation, this being a crucial condition for the meaning-prediction process.

A context-free interpretation of novel complex words is attributed importance by Coolen, van Jaarsveld, and Schreuder (1991: 341), who maintain that "empirical evidence about interpretation processes in isolation is essential for specifying the role of context in more detail." Another important argument supporting the relevance of my approach comes from Renouf and Bauer (2000), who – when trying to identify the contextual factors supporting the interpretation of novel compounds – arrive at a conclusion that

> the degree of support provided by the surrounding context is generally low: the individual kinds of support are diverse and sometimes indirect, and the interpretative process diffuses linguistically over many types of support... We have found that almost all our new compounds and derivations are semantically, if not referentially, interpretable by means of their internal components and that morphological processing is not just a vital back-up procedure to contextual analysis but that it is probably the single most effective starting-point for deducing the meaning of... neologisms... (2000: 256).

Wisniewski (1996) also points out several important reasons justifying the research into context-free interpretation of coinages. He notes that the contribution of context to understanding a novel word varies. Sometimes new complex words *are* interpreted with little context immediately present (newspaper headlines, the yellow pages). In addition, his experience resulting from experimental research indicates that in some cases "the context fails to 'spell out the meaning' or spells out the meaning sometimes *after* the initial occurrence of the combination" (1996: 450). Another important argument in favour of research into meaning-predictability of context-free naming units comes from research into various cognitive processes, indicating the interaction between context and prior knowledge. Therefore, Wisniewski aptly concludes that "it makes sense to first identify how the meanings of the constituents (i.e., prior knowledge) affect interpretation. Then the role of a discourse setting may be more meaningfully understood in light of these prior knowledge effects" (1996: 450–1).

One more point should be highlighted. Context-free meaning predictability is a *system-level* notion abstracting away from any particular speech-situation that can assume a multiplicity of forms each affecting the context-

dependent, speech-level meaning-predictability in a specific way and to a different degree. It follows that the context-free and the context-dependent meaning-prediction processes are different notions pertaining to different levels of language (langue vs. parole), requiring different methods of research, and serving different purposes (even if the ultimate objective is the same). It may be assumed that the most predictable reading, identified within a context-free meaning-prediction process, is a result of abstracting away from any particular speech situation, and, therefore, is a kind of *generalisation*. It is an *idealised* picture of what may be expected to be the meaning of a first-encountered context-free naming unit. This idealised meaning embodies the most characteristic constellation of predictability-boosting and predictability-reducing factors for a given naming unit.

This idealised picture need not be identical with the actual meaning of the naming unit in context and co-text, i.e., the meaning with which the naming unit is used in speech (parole). This has two major reasons. First, even if such an idealised meaning is the most probable from the interpretation point of view, it need not be the one with which a naming unit was *coined*. The reason is obvious. Both word formation and word interpretation are subject to a multiplicity of factors. The path from the object to be named to the name itself (linguistic sign) is not a highway with no digressions; rather than a straightforward process the naming act is complicated and is affected by a number of linguistic and extra-linguistic factors. The complexity of this process is captured in my concept of 'creativity within productivity constraints' (Štekauer in press[2], Štekauer et al. 2004) which rather than the rivalry of formal WF patterns (affixes) stresses the cognitive naming act performed by a particular language user who is aware of the available naming options. The same applies to the word interpretation process. There is no direct path from the form encountered to the identification of its referent. It is this diversity of factors that this monograph aims to identify.

Second, each naming unit is used in speech situations in which it is influenced by various *textual* and *co-textual factors* that modify the basic meaning.

Let us conclude answering the question asked in the heading of this section by resuming and elaborating on the tennis-player analogy mentioned in the Introduction. A tennis-player's position at the top of the ranking 'predicts' his/her chance to win the forthcoming tournament. This position is not haphazard; on the contrary, it is well justified. It is justified by the player's preceding results conditioned by the interplay of a number of internal and external factors. The internal factors include the player's talents, genetic characteristics, strong will, etc. The external factors include the overall conditions for train-

ing (economic aspects, availability of a suitable tennis-court, the qualities of his/her trainer, opportunity to participate in the tournaments), psychological aspects of his/her personal development, the influence of social environment, etc. If all these internal and external factors are in balance the 'Predictability Rate' with regard to the top-ranking player's winning the tournament is high. Certainly, there are also other players. The quality of the competition affects the top-ranker's chances of winning (the 'Objectified Predictability Rate'). So far, we are moving at the 'system level'. Each particular tournament takes place under specific circumstances that may influence the overall results by supporting the chances of the top-ranker or working against him/her. This is the practical context situation. But without regard to the actual results of one particular tournament the knowledge of the overall qualities and performance of a tennis-player are significant for tournament organisers and the player's overall position in the tennis world.

In the same vein we should view the importance of meaning predictability. One particular meaning of a novel complex word is most predictable, which follows from its linguistic characteristics, from the position of the object represented by this word in the system of the objects of extra-linguistic reality and, obviously, the relevant knowledge and experience of language users. Therefore its top-rank position is not haphazard. This meaning faces competition from other potential readings, and its position depends on the strength of the other readings (Objectified Predictability Rate). It may be that a coiner, under the influence of the specific circumstances accompanying the act of naming and the objectives pursued in the act of naming, chooses a particular formal representation for a different meaning than the most predictable one. One such case of 'mismatch', however, does not entail that the meaning of novel complex words is unpredictable – in the same way as one failure does not have disastrous effects upon the ranking position of a tennis player. The fact remains that it is the most predictable readings of novel complex words that tend to correspond with actual meanings of words as they are coined.

3.2 Predictability – lexical meaning – conceptualisation – extra-linguistic knowledge

In my discussion of the predictability of naming units I will use the term *seme* in the meaning 'semantic component', 'semantic feature', or 'semantic marker'. Semes are used here as a kind of shortcut to represent the meaning facet of naming units by mapping a conceptual structure in accordance with the ono-

masiological model (Figure 1). By implication, they are derived from the features of an object named. As such, they constitute what has been traditionally called 'lexical meaning'.[1]

In addition, I share Langacker's view calling into question the frequent attempts to draw a distinction between those specifications which are part of the conventional meaning of a linguistic expression and those which represent extra-linguistic knowledge leading to the distinction between 'semantics' and 'pragmatics'. As noted by Langacker (1988b: 57), "[t]he problem within this view is that the facts of language offer little basis for accepting its validity." Langacker refers to Haiman (1980), who surveyed the arguments usually presented in favour of drawing the line between the linguistic and extra-linguistic knowledge associated with a term, and found them all to be inadequate. He points out that the existence of a clear-cut boundary is only justified by methodological objectives (not facts). Therefore, he arrives at the following important conclusion:

> I see no a priori reason to accept the reality of the semantics/pragmatics dichotomy. Instead, I adopt an 'encyclopedic' conception of linguistic semantics. I posit no specific boundary between our linguistic and non-linguistic knowledge of the entity designated by a term, such that all those specifications on one side of the boundary clearly fall within the purview of semantics, while those on the other side are safely relegated to pragmatics. Far more realistic, I believe, is to posit a gradation of 'centrality' in the specifications constituting our encyclopedic knowledge of an entity: some domains and specifications are obviously more salient and linguistically important than others,... but the imposition of any precise or rigid boundary is considered arbitrary (1988b: 57–58).

A similar position is taken by Aitchison (1987: 195) when she assumes that it is "impossible to say where the 'meaning' of a word ends and general knowledge begins", and by Boguraev (1989: 8), who maintains that "it is difficult to pinpoint a boundary between the semantic knowledge that the use of a particular word (sense) implies and the expert background which prompts its use in a specific domain."

My position based on the onomasiological model is similar. The line connecting the conceptual level and extra-linguistic reality, including a speech-community, in my model (see Chapter 2) suggests that there is a direct connection between an object to be named and its conceptual processing by a naming person. The interconnection between extra-linguistic reality and the speech community is vital. It indicates that no naming process takes place in isolation from any and all other processes and relations in extra-linguistic reality,

on the one hand, and the experiences and knowledge of a speech-community, on the other. In other words, each act of naming reflects both the complexity of the relations in extra-linguistic reality and the complexity of perception of these relations by a speech community (a coiner). These circumstances of naming acts are inevitably reflected in the nature of the conceptual processing of the object named, and subsequently, in the mapping of the conceptual 'picture' to the meaning facet of a linguistic sign. From this point of view, the meaning of a naming unit can be conceived of as a representation of a conceptually processed class of objects of extra-linguistic reality. Since no features can exist without relations there can hardly be any linguistic meaning which does not reflect this (usually very complex) network of relations.

If, therefore, in my discussion of meaning predictability, I distinguish between the terms 'lexical meaning', 'conceptual structure', and 'extra-linguistic knowledge' (ranging from world knowledge to very specific knowledge, and including experiences, i. e., pragmatics), this distinction should be understood as reflecting different facets of our comprehension of extra-linguistic reality. Lexical meaning enables language users to *identify* a named object and to set it into a network of relations with other objects. These *relations* – so vital to the interpretation of novel naming units – can only be identified at the *conceptual level* by making use of one's extra-linguistic knowledge and previous experience with the relevant objects.

From this it follows that the approach presented here is one based on *knowledge representation*. It will be demonstrated that the (degree of) acceptability/non-acceptability and, consequently, meaning predictability/unpredictability (conceived of as a cline) depend on the interaction of linguistic knowledge (knowledge of the meanings of morphemes, including affixes, knowledge of productive Onomasiological, Word-Formation, and Morphological Types, knowledge of acceptable Onomasiological Structure Rules, etc.) and extra-linguistic knowledge (including knowledge of real and unreal, tangible and intangible objects of extra-linguistic reality, and their place in the narrower and broader system of relations and interactions). This view is in accordance with that expressed by Coseriu (1970: 116), who points out that linguistic knowledge is not sufficient to interpret a NN compound. It is the 'general knowledge of things' which makes it possible to meaningfully relate compound constituents.

With the preceding ideas in mind, it may be concluded that the semantic components of various abstraction levels represent the interpreter's knowledge of an object they stand for. It is for these reasons that a knowledge represen-

tation system is proposed here as the basis for the account of the meaning predictability of naming units.

3.3 Predictability and the native/non-native speaker factor

An important corollary of the previous considerations for the theory of the meaning predictability of novel naming units is its (relative) independence of the mother language of an interpreter. In particular, the process of meaning interpretation (meaning-prediction process) is based on the conceptual analysis of the phenomena of extra-linguistic reality, captured in naming units. Given the role of the conceptual level analysis, extra-linguistic knowledge and experience, and the onomasiological level as the conceptual basis for the naming process in the meaning-prediction process, it may be assumed that there is no principled difference in the ability of predicting the meaning of novel, context-free naming units between native speakers and non-native speakers provided that a non-native speaker has a standard command of the particular language (which implies that (s)he 'knows' the basic rules and principles of word formation, and understands the meanings of the morphemes constituting the particular naming unit) and his/her world knowledge and experiences are comparable to those of a common native speaker. At least, in the countries of Western civilisation the latter condition is met.[2]

For illustration, if a non-native speaker encounters a possible naming unit, say, *anthraxist*, and if (s)he knows the meaning of the lexical unit *anthrax* and the meaning of the suffix *-ist*, his/her position prior to the meaning-prediction process is in no way different from any English native speaker because the correct interpretation, i.e., the meaning-prediction process, is conditioned by the following knowledge:

1. both of them are sure to know that the substantival suffix *-ist* is an Agentive affix combined with nouns to denote persons who perform some activity related to the preceding noun constituent;
2. both of them may be supposed to know the meaning of the lexical unit *anthrax*;
3. both native and non-native speakers are expected to have the relevant extra-linguistic knowledge of 'anthrax' as a life-endangering substance;
4. both of them may be supposed to know that some persons in the U.S.A. received envelopes with anthrax from terrorists after September 11, 2001.

The first two conditions for mastering the meaning-prediction process are linguistic, the latter two require extra-linguistic knowledge, the knowledge of fairly recent events dominating world news for quite a long time.

It should be noted that the position of the native speaker and the non-native speaker is, in principle, the same. Based on the interrelation of linguistic and extra-linguistic knowledge, both of them can infer the possible meaning of *anthraxist*, which might be roughly 'a person mailing anthrax in order to murder/threaten other people'. Certainly, the absence of the determined constituent of the onomasiological mark at the onomatological level admits other readings as well (perhaps of lower Predictability Rate).

Given this situation, it is postulated that any experimental research may rely on both native speakers and non-native informants in the same way. The results of my experiments will demonstrate that this hypothesis is correct.

3.4 Predictability and seme level

When discussing root compounds R. Lieber maintains that

> ... lexical semantics fixes only so much of the interpretation of a newly coined compound, namely that the second stem determines the overall headedness of the compound, and that the compound as a whole has only a single referent. The rest is free and "involves context and encyclopedic knowledge" (2004: 53).

I will demonstrate in this and the following chapters that a context-free meaning-prediction process need not content itself with general interpretation statements like 'A is somehow associated with B'.

When Lipka (2002: 133) summarises the shortcomings of Feature Semantics he points out that it attributes an equal status to all features, without weighing them. One of the central hypotheses upon which I took up this project was that the meaning Predictability Rate of novel naming units is closely related to prototypical semes, that is to say, to the *prototypical features* of a particular conceptual class of objects represented by a novel naming unit. In particular, it is hypothesised that it is primarily the prototypical features of a particular class of objects which are first 'grasped' by a language user on encountering a novel naming unit, and therefore it is these features, and the corresponding semes, which underlie the most predictable readings. This view follows from the assumption that those meanings of naming units which are based on (the combination of) prototypical semantic components are more acceptable to language users, and therefore more predictable. This view accords, for

example, with the above-mentioned view of P. Downing that there is a very strong tendency for compounds to be "created and interpreted on the basis of relationships perceived as permanent and habitual" (1977:836).

Since the onomasiological base (head) determines the grammatical and the lexical features (word class, lexical class) of naming units, it may be assumed that it is the seme level of the base which is also crucial for the predictability of naming units.[3]

On the other hand, using a motivating constituent in a shifted (figurative, metaphorical) meaning reduces the predictability of a particular naming unit.[4] For example, Cohen & Murphy's (1984:52) example of atypical combination of the concepts in *virgin birth* can hardly be thought of as having a high Predictability Rate because the prototypical feature of *virgin* [–Having Child] and *birth* [+Bringing forth a Child] are mutually exclusive for an interpretation based on the semantic structure [Agent – Action – Result]. In addition, since any other possible interpretations seem to use *virgin* in a metaphorical sense their predictability is, by definition, low. My experimental research, using possible, non-institutionalised naming units, provided me with ample examples of this sort. Thus, the Predictability Rate of readings like 'a person who can be easily influenced' for *shape cloth*; 'one's life' for *game wheel*; 'a metal party' for *ball hammer*, approaches zero because of the metaphorical use of either of the constituents.

However, figurativeness (semantic shift) need not be an obstacle to a high Predictability Rate if a figurative, metaphorical meaning has become established, i.e., institutionalised, over time, as will be illustrated below by the personalised meaning of *star* in *hill star*.

Clark & Clark (1979) suggest that there are words whose senses depend entirely on the time, place, and circumstances in which they are uttered. As mentioned above, they are labelled as 'contextuals'. One of their features is claimed to be the non-denumerability of their meanings, i.e., "contextuals should possess not a small finite number of potential senses, but an indefinitely large number of senses" (1979:782). Let us mention Clark & Clark's *teapot* example in which the conversion *teapot*$_V$ does not serve its (proto)typical, inherent purpose; rather, it is used in an unpredictable way, and hence in an unpredictable meaning: 'to rub the back of leg with a teapot'.

H. Clark (1983) proposes a long list of 'contextuals' featuring non-denumerability of meanings, including indirect nouns (*one water* meaning 'one glass of water', 'one tub of water', 'one drop of water', 'one teaspoon of water', etc.), compound nouns, possessives (*John's dog* meaning 'the dog John possesses', 'the dog John is standing in front of', 'the dog John saw yesterday',

'the dog John always wanted', etc.), denominal nouns (*Nixonite, bicycler, saxophonist* whose interpretation depends on co-operation between speaker and addressee), eponymous verbs ('The photographer asked me to *do a Napoleon* for the camera'), pro-act verbs ('Alice *did* the lawn' where *did* can mean a number of activities), denominal adjectives (*Churchillian* meaning 'with a face like Churchill', 'smoking a cigar like Churchill', 'with a speaking style like Churchill', etc.), non-predicating adjectives, eponymous adjectives (*That is a very Picasso painting*).

This picture does not look very optimistic in terms of the meaning predictability of naming units. However, in this case appearances are deceptive. The Clarks' assumptions require several comments. First, the very notion of 'contextuality' as applied by the Clarks seems to be misleading, because – as admitted by Clark & Clark – the line between 'contextual expressions' and purely intensional expressions is vague: "A sense may be conventional within one community, as among newspaper reporters or computer users, but it may be a nonce sense for the people being addressed" (H. Clark 1983: 305).

This and other arguments against the exaggeration (and misuse) of the notion of 'context-dependence' are raised in Štekauer (2002) within a discussion of the status of nonce-formations and neologisms. To sum up these arguments, contextual dependence is a vague notion because (1) each naming unit, no matter how well it is integrated in a system, is used within its typical 'context', unless certain stylistic objectives require its use in the 'context' of a different register; (2) context-dependence is always a matter of speech (*parole*) and never that of system (*langue*). In the system every naming unit is accurately defined and has its distinct, context-independent meaning and function. This follows from the mechanism of coining new naming units, based on conceptual analysis. (3) One and the same context may have various implications for various groups of a speech community: one and the same naming unit may be context-free for a specific subset of a speech community (experts in a particular field) and fully context-dependent for another subset of a speech community.

The Clarks' assumption is even more misleading in terms of the predictability of meaning of novel coinages. In the vast majority of cases the number of possible combinations of semes of the motivating constituents is much larger (non-denumerable) than the number of actually predictable meanings. A meaning based on too general semes or too idiosyncratic semes cannot be obviously predicted, as follows from the numerous examples of the Clarks and from my experimental research discussed below.

Moreover, I do not share the Clarks' view that words like *Picasso* as in 'That is a very *Picasso* painting'; *Churchillian*; *apple-juice chair*, *Ferrari woman*,

Beethoven as used in 'Last night they played a *Beethoven*', the above-mentioned example of *water*, fully depend for their interpretation upon their context. Let us take, for instance, proper names. It has been demonstrated in Štekauer (1997) that proper names have sense, and function as polysemantic linguistic signs. The existence of recategorised (converted) proper names, which undoubtedly have their meaning, is the best evidence of this claim: meaning cannot be acquired out of thin air by such conversions. If we realise that our linguistic competence and performance are always conditioned by extra-linguistic knowledge and experience, the semantically shifted and converted proper names do not seem awkward any more. Thus, the above proper name meanings are predictable to a considerable degree thanks to our extra-linguistic knowledge (and unpredictable in the case of ignorance – as is the case with any other word where the sign–referent connection in one's mind is absent). If a language user is supposed to use and/or interpret a proper name properly (s)he must know the most characteristic features dominating the semantic structure of the respective person in the same way as (s)he must know the basic purpose, function, etc. of common words like *industrial robot*, *joystick*, *cricket*, *music*, etc. Each person is unique and can be captured by a set of characteristic and distinct features projected onto the semantic facet of a linguistic sign. For each such person there is a limited set of characteristic features, some of which are more salient and dominant than others, that can subsequently motivate a shifted (eponymic) or a recategorised meaning. To this extent, such naming units are fairly well predictable.

To sum up, each person may be viewed as a multi-feature object, which means that apart from the features common to all human beings (most general semes) (s)he has his/her unique personality features, behaviour features, extraordinary skills, ways of thinking, achievements, etc., that dominate the person(ality) and by which the person differs from any other individual. These salient, most conspicuous features of a 'conceptually' and 'linguistically' processed person(ality) set the limits to a range of possible – hence predictable – shifted or converted meanings.

Certainly, meaning predictability is restricted to a larger or smaller speech community whose members know the meaning of the person in question. Thus, the converted proper noun *to Havel* may be expected to be well-predictable in at least two speech communities, the Czech and the Slovak ones, because (1) the personality of Havel, his characteristic behaviour, views, and deeds are well known to people in these countries owing to his presidentship in Czechoslovakia and later in the Czech Republic, (2) close ties between the two nations, and (3) close affinity of the Czech and the Slovak languages. Several

striking features (semes) characterise this man, such as [Dissident], [Fighter for Human Rights], [Playwright], [Philosopher], [First Post-1989 President], [Leader of the Velvet Revolution], [Zealous Humanist], [Wrong Pronunciation of the Phoneme /r/], and perhaps one or two other salient characteristics. These features limit the possible, i.e., the predictable recategorisations of the proper name *Havel* for the Czech/Slovak speech community as a whole. Forming a recategorised naming unit outside this range of characteristics runs against the extra-linguistic knowledge and experiences of Czech/Slovak language users, and imposes excessive burden upon the context. To use the terminology of natural morphology, any meaning beyond the above specified range is unnatural.

Another example illustrates familiarity with a person's characteristic, salient features and/or behaviour within a small speech community, a group of friends all of whom know a person, let us call him *Peter*, who is notorious for wasting a lot of money by gambling. In this small speech community – as opposed to any outsider – the conversion *to Peter* may be expected to have a high Predictability Rate. For more detailed discussion on recategorised proper names see Section 3.5.1.7.

All in all, what can be predicted are meanings reflecting some characteristic features, and/or functions of the named objects. This assumption is based on my previous research into onomasiological recategorisation (= conversion, zero derivation). Štekauer (1996) demonstrates that it is the most characteristic, prototypical features that generally underlie the process of conversion, and condition the predictability of converted naming units.

Thus, for example, the word *hammer*$_{N}$ can be converted to the category of verb with the meanings that are intrinsically connected with the basic functions of 'hammer' as delimited by the logical spectrum at the conceptual level of the Word-formation Component, i.e., roughly, 'a manual tool for nailing'. Obviously, it would be highly improbable and against the basic principles of productive word formation and extra-linguistic knowledge of language users if the primary conversion process was based on a very general feature(s) of the tool called *hammer*, for example, in the meaning of a tool for drawing circles in sand, or for purposes similar to those of darts. None of these possibilities can be eliminated, though. In an appropriate context the sentences 'My children were hammering circles in sand', and 'Though they hammered the target for two hours they hit the bull's eye only once' should be perfectly comprehensible to a native speaker. In any case, these converted meanings are unexpected and unpredictable because they do not pertain to the object-specific features. Rather, it is the general features, such as SUBSTANCE and [Solid] which underlie these converted meanings, i.e., the features common to a large number

of objects. By implication, a naming unit standing for *any solid object* of appropriate size might be converted in the same sense ('drawing' and 'hitting', respectively). Thus, we might say 'My children were shoeing circles in sand', 'My children were stoning circles in sand', 'My children were cockleshelling circles in sand', etc., and ... they tea-potted the target, ... they shoed the target, ... they cockleshelled the target, etc. Since none of these readings of *shoe*$_V$, *stone*$_V$, *cockleshell*$_V$, *tea-pot*$_V$ correspond to the prototypical functions of the objects denoted by their corresponding nouns, none of them is predictable without relevant context.

The experimental research presented and analysed in Chapter 4 below provides a number of the readings of possible naming units, motivated by a too general seme(s) or seme combination, which are therefore unpredictable or barely predictable, for example 'the name of a bar, pub' for *garden whisky* (any compound, in fact, can be used for this purpose), 'all dresses which are sewn from cloth' for *shape cloth*, 'a wheel called Game' for *game wheel*, 'a good, pleasant seat where we feel well' for *apple-juice seat*, 'a hammer that can be thrown like a ball' for *ball hammer*; 'to exclaim 'boy" for *to boy*, 'to blossom' and 'to be interested in flowers' for *to tulip, etc.*

Štekauer (1996) proposes distinguishing three levels of generalisation of semantic components, with the more specific ones being the most probable candidates for motivating the process of conversion. The following structure was given for the naming unit *milk* (based on Ondrus, Horecký, & Furdík 1980: 37–38):

(20) a. classification semes: [–Proper Name] [+Concrete] [–Countable] [–Animate] [–Collective] [+Material]
b. identification semes: [+Liquid]
c. specification semes: [+White Colour] [Sweetish Taste], [+From Female Mammals] [+Beverage] [+Foodstuff]

As already suggested, it is not the most general (classification) semes that, in principle, become motivating elements because they only identify the referent with other similar objects in a large class of objects. Important are identification, and mainly specification semes. Thus, the combination of the semantic components [+From Female Mammals] and [+Liquid] motivate the converted meaning 'to obtain milk from cow', the combination of [+Beverage] and [+Foodstuff] underlies the obsolete meaning 'to feed with milk' (a ewe milks her lamb); the combination of [+Liquid], [+Beverage] and [+Foodstuff] motivates the meaning 'to add milk to (coffee)' (to milk the tea), etc.

These conclusions may be generalised and applied to any of the first four Onomasiological Types, although with onomasiological recategorisation (conversion) the situation appears to be more complicated. The reason is that while all the other Onomasiological Types are based on their respective characteristic onomasiological structures, the onomasiological recategorisation does not feature any such structure. Rather, it is based on the relation between the respective conceptual levels of the converting and the converted naming units as represented in (18). On the other hand, in discussing the meaning predictability of naming units with an onomasiological structure the interaction between the onomasiological base and mark must be taken into account.

For the present work a subtler classification of semes has been developed for the sake of evaluating the influence of seme level upon the meaning predictability of the individual readings of novel naming units. The above-mentioned three-level model has been completed at both extremes of generality. The introduction of level 5, including idiosyncratic semes, is important to demonstrate that the meanings motivated by idiosyncratic semes can hardly be predictable by being highly context-dependent. In addition to five levels of semantic components, the classification encompasses the level of word formation, reflecting the onomasiological structure which constitutes the very core of the naming act. This onomasiological structure, also identifying a WF Type, has proved to be a useful constraint upon the acceptability of certain meanings proposed.[5]

Level 1 includes the four most general conceptual categories capable of covering any object of extra-linguistic reality to be named, namely SUBSTANCE, ACTION, QUALITY, and CIRCUMSTANCE;

Level 2 is represented by 'classification semes';

Level 3 is represented by 'identification semes';

Level 4 is represented by 'prototypical semes';

Level 5 is represented by 'idiosyncratic semes'.

The different levels represent different levels of generalisation. Levels 4 and 5 differ from each other in their nature. The prototypical nature of level 4 semes means that they reflect the indispensable and fundamental features of the particular object which are supposed to be shared by all members of the class of similar objects, and by implication, they are the core semantic components of the meaning of a particular naming unit in terms of meaning predictability.

On the other hand, the idiosyncratic nature of level 5 semes refers to one pole of the 'general-individual' scale. A very specific seme of an object need not

be and usually is not indispensable and fundamental in terms of predictability. In principle, level 5 semes have an individualising effect and reflect either an acquired, non-prototypical, or unexpected, and therefore, idiosyncratic property which does not belong to the defining, prototypical characteristics of an object (for example, 'a spade used by dog for playing' in *dog spade*, 'a book in the shape of a baby' in *baby book*, 'a hammer whose one part is spherical' in *ball hammer*, 'a hat with the odour of flowers' in *flower hat*, 'a special seat filled up with apple-juice' in *apple-juice seat*, 'to wear a yellow sweater and green trousers' in *to tulip*, 'to run or move very fast' in *to river, etc.)*.

From the point of view of meaning predictability it is interesting that while the 'generalising' semes of level 4 contribute to and condition the meaning predictability of naming units, the individualising level 5 semes have an opposite effect. To put it another way, meaning predictability is supported by a combination of prototypical semes. The distinction between level 4 and level 5 semantic components more or less corresponds to the approach taken by Katz & Fodor who distinguish between 'semantic markers' and 'distinguishers':

> The semantic markers assigned to a lexical item in a dictionary entry are intended to reflect whatever systematic semantic relations hold between that item and the rest of the vocabulary of the language. On the other hand, the distinguishers assigned to a lexical item are intended to reflect what is idiosyncratic about the meaning of that item (1963: 187).

In other words:

> The distinction between markers and distinguishers is meant to coincide with the distinction between that part of the meaning of a lexical item which is systematic for the language and that part of the meaning of the item which is not (1963: 188).

While this distinction was subject to sharp criticism, an analogical distinction between prototypical and idiosyncratic semes (features) has proved very advantageous for the reasons mentioned in the preceding paragraphs.

For obvious reasons, it is not possible to give an exhaustive list of all the semantic components of Levels 3 and 4. In principle, the hierarchical relation of semantic components in (21) representing various kinship terms may be regarded as a pattern relation for 'filling in' the individual levels with their respective semantic components:

(21) Level 1 SUBSTANCE
Level 2 Animate
Level 3 Human

Level 4 Male/Female
Level 5 Adult/Middle-aged/Adolescent/Child/ Baby
or
Tall/Small
or
Red-haired/Blond(e)/Black-haired
or
Decent Person/Gentleman/Evil-doer/Killjoy
etc.

It follows from (21) that level 5 is reserved for those semantic components which implement various options offered by the next higher level. (22) extends the basic pattern of (21) with some more examples:

(22) Level 1 *Conceptual categories*:

SUBSTANCE, ACTION, QUALITY, CIRCUMSTANCE

Level 2 *Classification semes*:

Animate, Action, Process, State, Quality, Tangible, Abstract, Countable, Collective, Location, Direction, Time, Manner, etc.

Level 3 *Identification semes*:

Human, Animal, Plant, Material, Foodstuff, Artefact, Tool, Having Dimensions, Having Taste, Having Colour, etc.

Level 4 *Prototypical semes*:

Male, Female, Adult, Characteristic Material, Characteristic State of Matter (Solid, Liquid, Gaseous), Characteristic Colour, Characteristic Shape (Sphere, Triangle,...), Characteristic Taste, Characteristic Colour, Characteristic Application/Function/ Purpose of Tool, Characteristic Capacity – Ability – Skill, etc.

Level 5 *Idiosyncratic Features*, *Individualised Qualities*
Level of WF: *Onomasiological Structure Rules*

3.5 The meaning-prediction process

We may now proceed to an account of the interpretation of novel, context-free naming units which will be referred to hereinafter as the *meaning-prediction process*. It may be postulated that the novel naming unit meaning-predication process to be mastered by a member of a particular speech community travels roughly the direction opposite to that represented in the above-given model of

word formation. Thus, an interpreter first encounters the phonological form of a naming unit, which enables him/her to recognise its onomatological structure. This information is important for the identification of the individual linguistic units (morphemes) which were attached at the onomatological level to the conceptual constituents of the onomasiological level of word-*formation*, because – to use de Almeida & Libben's words –

> new compound forms [and it may be added that any new complex words] ... can only be understood in terms of their constituent morphemes. These morphemes must be isolated, accessed and understood in order for a compound interpretation to be formed (2002:98).

The identification of the individual morphological constituents is vital to the identification of the Morphological and the WF Types underlying the process of formation of a novel naming unit: the recognition of the particular Morphological Type and the WF Type (or WF Types if their morphological make-up is identical) is highly indicative of the possible range of general semantic relations between the constituents of the novel naming unit.[6] In principle, this step is based on an attempt to *reconstruct* the relation between the onomasiological and the onomatological levels, established by the application of the MSAP principle in the process of word formation.

This means that the next step consists in matching the meaning facet of the identified morphemes with the semes of the onomasiological structure. Based on mutual relations between these semantic constituents, it is possible to predict the possible onomasiological structure, or in other words, to employ one's competence with regard to the WF Rules and patterns to predict the particular lexical meaning of a novel complex naming unit on the basis of the identification of the *general word formation meaning* underlying the coinage.

This leads to the identification of one or (usually) several most probable readings of the word under interpretation. In other words, the language-user identifies the most possible onomasiological structures, and through them, the most possible WF Types. In addition, by interrelating the individual established meanings of the motivating constituents, a language user selects the most probable candidate, i.e., the most probable combination, for a closer predictability analysis. Thus, for example, in one of my sample naming units the informants faced the decision whether *star* in *hill star* is to be interpreted as a celestial body or a famous human being. Consequently, this step is based on important *decision-making operations*: (1) a language user selects one of several possible meanings in the case of a polysemantic unit which, in the process of word-*formation*, was used to represent the particular conceptual component

within the MSAP application; (2) a language user identifies possible WF Types (onomasiological structures).

The interpreter can now match the semantics bound to the individual WF Types with his *linguistic competence* (knowledge of *productive* WF Types, knowledge of the meaning of words and affixes) which informs him that, for example, a *nut-crack-er* can mean either an Instrument or an Agent. Based on his *extra-linguistic knowledge* (which encompasses knowledge of the world and experiences) the interpreter may come to a conclusion that *nutcracking* does not seem to be a professional activity, a job, and therefore the instrumental reading obtains a higher Predictability Rate.[7]

The process of interpretation and meaning prediction outlined above indicates that the interpretation process, especially in its first, general word formation-meaning-identification step, significantly benefits from a language-user's linguistic knowledge of productive WF Rules (Types), and that this knowledge functions as a filter through which only those possible readings of a novel word pass that comply with productive WF Rules. The next, fine-grained, filter employs a language user's extra-linguistic knowledge and experience. The two filters make it possible to reduce a considerable amount of more or less possible readings to the minimum number. My experimental research has shown that this number rarely exceeds two competing readings. Very frequently, however, one reading dominates all the other readings, and is therefore of the highest predictability value.

3.5.1 Predictability and the Onomasiological Type

One of the central assumptions of the model outlined is that the interpretation, and therefore, the predictability of novel, context-free coinages is strongly influenced by the Onomasiological Type to which the interpreted word belongs. It will be shown that the individual Onomasiological Types, constituted by the relation between the onomasiological and onomatological levels, differ in terms of predictability, which is closely connected with their general structural characteristics. In addition, the predictability of the individual Word-Formation Types and Morphological Types falling within a particular Onomasiological Type can significantly differ.[8] In general, the situation inside the individual Onomasiological, WF, and Morphological Types indicates a very close interrelation between the word-*formation* and the word-*interpretation* processes in the sense that the predictability of novel, context-free naming units heavily depends on the structural pattern employed in the process of for-

mation. The following is a brief and general characteristic of the individual Onomasiological Types in terms of meaning predictability.

3.5.1.1 *Onomasiological Type 1*

In the case of Onomasiological Type 1 the meaning-decoding step is not very demanding: it is not difficult to identify the semantic relations between the structural constituents thanks to the explicit expression of the Actional seme at the Onomatological level. To put it another way, each of the constituents of the semantic structure of the onomasiological level has its corresponding morpheme counterpart present in the particular naming unit. Since the Actional constituent is the key to understanding the meaning of naming units falling within OT1, because it is this Actional constituent which – in the function of the determined constituent of the onomasiological mark – relates the polar members of the onomasiological structure, the predication process seems to be, at first sight, simple. Therefore, it might be assumed that the Predictability Rate of the majority of OT1 naming units is very high.

For example, *piano-player* cannot but approach the maximum Predictability Rate because the Actional seme, onomatologically represented by the morpheme *play*, unambiguously relates the Instrument of Action (*piano*) and the Agent (*-er*). One of the central semantic components of *play* is [Human] because playing a piano requires a conscious and purposeful training which leaves very little space for an animal-based interpretation reflecting a language-user's experience with various animal performances (for instance, in circuses). Therefore, it may be predicted with almost 100% certainty that the intended meaning of *piano-player* is 'a person who (professionally) plays a piano'. Certainly, the situation-conditioned interpretations are available any time ('a person just now playing a piano', 'an animal trained for playing a piano'), however, their Predictability Rate may be supposed to be low due to the reasons mentioned in Chapter 1.

A lower Predictability Rate may be attached to words like *apple-eater* (to use Marchand's and Kastovsky's example) for the simple reason that this naming unit admits at least three tough-competing readings as a result of the ambiguous nature of the onomasiological base (certainly, from the point of view of an interpreter). It may be assumed that the process of interpretation of this naming unit pursues the following path:

Based on the identification of the onomatological structure, including three morphematic constituents *apple* – *eat* – and *-er* corresponding to the three constituents at the onomasiological level, a language user may employ his/her extra-linguistic knowledge to assign various degrees of predictability to

three competing readings closely related to the Agentive seme standing for the onomasiological base. Out of two non-figurative readings, the one identifying Agent with a person [Human] seems to have a higher Predictability Rate than that based on an [Animal] Agent. While both of them are acceptable in terms of the relation between [Eating] and Agent, our everyday experience seems to give preference to the former interpretation. Next in the hierarchy appears to be a figurative reading with a strong negative connotation (which is encoded neither in the onomasiological structure nor the onomatological structure), and therefore can only be inferred by having recourse to our extra-linguistic knowledge and experience: 'a kind of must or a pest playing havoc with one's crops of apples'.

This example demonstrates a close interaction between the onomatological, the onomasiological, and the conceptual (extra-linguistic) levels in interpreting novel naming units. It should be added that the existence of competing readings does not reduce the chances of a particular central reading to approach a 100% Predictability Rate, but it significantly reduces the Objectified Predictability Rate (see Section 3.9).

3.5.1.2 *Onomasiological Type 2*

In Onomasiological Type 2 the determined constituent of the onomasiological mark is expressed while the determining constituent is not (*lock pin*, *teacher*). From this it follows that the Actional constituent of the semantic structure is present and used for the specification of the Onomasiological Base in terms of what the object represented by the Onomasiological Base does or what happens with it. Again, the presence of this Actional constituent at the onomatological level facilitates the prediction process, and therefore the Predictability Rate of naming units belonging to this Onomasiological Type is expected to be fairly high: we are pretty sure that *teacher* is 'a person who professionally teaches (an unspecified subject)' (with the situation-bound interpretation in the sense of 'someone acting at the given moment as a teacher' possible, too).

Not all cases of this Onomasiological Type are that simple. The interpretation of naming units like *lock pin* is not so unequivocal. While we might predict that *lock pin* is 'a pin used for locking some other (unspecified) components or objects', it could just as clearly be 'a pin in a lock'. The obstacle to a high Predictability Rate of a single reading is the ambiguous status of the onomasiological mark – it may refer both to ACTION and SUBSTANCE, as is the case with, for example, *leakage* and *coverage*. Both of them admit at least two competing interpretations, the Process and the State/Result readings, which affect the Predictability Rate.

The question remains as to the range, the scope, and the relevance of information provided by this type of naming unit – that is, whether we expect any additional information concerning the actual use of *lock pin*, the subject(s) taught by a *teacher*, the place of *leak*ing, the object *cover*ed, etc., and whether this is at all a matter of predictability.

3.5.1.3 *Onomasiological Type 4*

I will skip Onomasiological Type 3 for a while because its coinages are the most difficult to predict and will be given special attention below, and proceed to Onomasiological Type 4, characterised by a simple structure (the onomasiological mark cannot be analysed into the determining and the determined parts). In general, the predictability of the relevant naming units seems to be very high, which, in my view, is related to the non-existence of the internal structure of the onomasiological mark. Therefore, *unhappy*, for example, is first identified as a naming unit composed of two morphemes of which *un-* is the onomasiological base (head) (exemplifying the general category of (contrary) [Negation]) and *happy* its onomasiological mark. The direct connection between the unstructured mark and the base at the onomasiological level gives no chances for multiple interpretations. Similarly, a 100 % Predictability Rate may be expected for naming units like *readable*, *readability*, *after-dinner*$_{A}$, etc.

The same applies to *lion-hearted* in which the onomasiological mark (*lion-heart*) specifies the base (exemplifying the general category of [Feature]) and makes the interpretation of 'featuring lion-heart' the only possible interpretation.

Here, too, some more complicated cases occur, such as *parliamentary*, which out of context may be interpreted either as an adjective 'concerning the Parliament' (OT4), or a converted, recategorised, Agent noun 'a member of the Parliament' (OT5). Knowledge of the word-class in cases like this boosts the Predictability Rate.

Thus, it seems that the previous three Onomasiological Types are good candidates for a fairly high Predictability Rate. Unfortunately, the situation is not so simple. Some problematic examples were already adduced above. More of them will be discussed in the course of my analysis of the experimental results. All the same, the three Onomasiological Types establish much better general conditions for the prediction process than the following one which lacks the expression of the Actional constituent at the onomatological level.

3.5.1.4 *Onomasiological Type 3*

In the case of naming units without the determined constituent of the onomasiological mark, the situation is more complicated because the interpreter cannot unambiguously identify the logical-semantic relation between the two polar members of the onomasiological structure. The (theoretical) multiplicity of compatible semantic relations between the polar members is usually very high. From a purely linguistic point of view their reduction to one or two acceptable meanings is hardly possible. In other words, this is the point where *extra-linguistic factors* come to play a central role in the meaning-prediction process – no doubt in close co-operation with the interpreter's linguistic competence and intuition.

For example, when faced with a (possible) compound naming unit *baby book* the interpreter is able to identify the morphemes attached to the onomasiological structure. (S)he also knows the meaning(s) and therefore the semantic structure(s) of the determining constituent of the onomasiological mark (*baby*) as well as the meaning(s) and the semantic structure(s) of the onomasiological base (*book*). But not only that. (S)he also intuitively knows the respective hierarchical structures of the semantic components of *baby* and *book*, respectively, in terms of the general-individual scale. It should be kept in mind that this hierarchy maps the logical predicates of the conceptual level.

What is postulated here is that this knowledge enables the interpreter to trigger the *matching process* which starts at the level of prototypical semes. To put it another way, what is matched first is the *prototypical semes* of the motivating constituents, and therefore the prototypical features of the motivating referents. This view corresponds with Miller & Johnson-Laird (1976: 291), who maintain that perceptual identification procedures are related to the perceptually salient features of the core of a concept, defined as "an organized representation of general knowledge and beliefs about whatever objects or events the words denote – about what they are and do, what can be done with them, how they are related, what they relate to" (which seems to correspond to my conceptual structure) as well as that presented by Meyer (1993: 5), who assumes that "interpreting novel compounds is based mainly on prototypical features of objects and of certain domains."

The point of departure for the matching process is the onomasiological base. In my example, [For Reading] seems to be such a prototypical semantic component for *book* as an onomasiological base of *baby book*. This semantic component therefore may be supposed to be taken as a primary reference point in scanning the hierarchical semantic structure of the onomasiological mark *baby*; the result of the scanning operation is the identification of [–Reading

Capacity] as a relevant semantic component for the matching operation. In this way, the Agentive interpretation is automatically eliminated.

Since the first matching operation failed, the matching process continues; now, the subsequent search appears to branch because there seem to be two matching operations as the next best candidates for evaluation. The interpreter may resume matching from the same semantic component of the Onomasiological Base, i.e. [For Reading], and search for a semantic component in *baby* that is compatible with it, thus identifying, for example, [+Perceptual Capacity]. In the relation of these two dominating semantic components the latter assumes the logical-semantic function of Target of the overtly unexpressed Action:

(23)	Target	←	(Action)	–	Theme
	baby	[Perception by listening]	[Reading]	[For reading]	*book*

After checking all the possible semantic components at the lowest level the matching operation resumes at the next higher level, yielding the combination (24):

(24)	Topic	–	(State)	–	Patient
	baby	[Class of babies]	[Containing information of taking care of babies]		*book*

and its individualised variant (25):

(25)	Topic	–	(State)	–	Patient
	baby	[Single baby]	[Containing records of one's baby]		*book*

The other line of the matching process takes another central, prototypical feature of *book* as a reference point, that is to say, [+Contains Drawings/Pictures], which yields reading (26):

(26)	Agent	–	(Action)	→	Theme
	baby	[Perception by viewing]	[Viewing]		*book*

Another matching operation combines [+Dimension] as a fairly general semantic component of *book* and one of the central and most specific semantic components of *baby* [+Small Size], this giving reading (27):

(27)	Quality	–	(State)	–	Patient
	baby	[Small size]			*book* [Substance]

After completing the matching process (which may be conceived of as a trial-and-error process, and therefore also covers a number of other possible combinations) the interpreter faces one or more readings complying with the semantic compatibility principle. However, it has been pointed out that the notion of semantic compatibility cannot be mixed with the notion of predictability. The former is much wider and also encompasses those combinations which are not easily predictable. At this point it is extra-linguistic knowledge, including a language user's experiences, that is involved in the decision-making process aimed at the identification of the most plausible reading(s).

An even more demanding situation for the interpreter is represented by the case traditionally labelled as blending, for example, *smog* (smoke + fog), *autocide* (automobile + suicide), *molecism* (molecule + organism), *pornotopia* (pornography + utopia), *stagflation* (stagnation + inflation), *fixidity* (fixation + rigidity), *seavacuation* (sea + evacuation), *transciever* (transmitter + receiver).[9]

Štekauer (1998) suggests that such naming units are generated within a two-step process. First, 'full versions', i.e. compound naming units, are generated in the Word-Formation Component by a productive WF Rule. Then the naming unit is directly moved to and stored in the Lexicon which is the place where the naming units generated in the Word-Formation Component can undergo various formal and semantic changes. In this particular case an unpredictable form reduction takes place. This formation-related unpredictability cannot but pose serious problems to the 'first-encounter' meaning-prediction process. Although this type of novel, context-free naming unit has not been tested within my experiments described below it may be expected that, in principle, the predictability of this sort of coinage is rather low. Still, this is not a hard and fast rule. (Un)predictability depends on the degree of recognisability of the individual constituents in the blend, which is conditioned by an interplay of several factors, including the degree of transparency of such naming units (the 'visibility', i.e., the possibility to identify the individual constituents in a blend), one's linguistic competence (the extent of one's vocabulary), and the register to which a word belongs (common words vs. highly specialised terms vs. puns, etc.), which goes hand in hand with the speaker's expertise in and linguistic experiences with the field of the given register. Thus, the predictability of blends can, it may be surmised, oscillate to some extent. Given the important role of the language competence factor it may be expected that the differences between native speakers and non-native speakers will be most striking in this group of coinages.

This group of coinages does not fall within the scope of my research, which is aimed at new naming units generated by the productive rules of the Word-Formation Component. Since blends receive their final shape in the Lexicon they represent a special category of coinages. Similar conclusions apply to acronyms and clippings which, in fact, rather than new naming units are 'mere' form-reductions of the already existing naming units, once coined in the Word-formation Component and subsequently stored and modified in the Lexicon.[10]

3.5.1.5 *Summary*

To summarise, it follows from the previous account that, in three-constituent onomasiological structures (determining constituent of OM + determined constituent of OM + onomasiological base), the Actional constituent is vital from the interpretation point of view because it relates the onomasiological base and the determining constituent of onomasiological mark, thus expressing the general Actional orientation of the base. By identifying the presumed semantic structure an important step has been made towards predicting the meaning of a naming unit: the range of possible meanings of a naming unit has been considerably narrowed and/or limited.

It follows that the above-described process of meaning-prediction maps the word formation model in the reverse sequence of steps. This mapping is not perfect, though. The conceptual factor, drawing on one's knowledge of the world and experiences, is involved in meaning identification as early as the onomasiological level and participates in all subsequent steps and related decisions. In this respect my model significantly differs from that proposed by Smith and Osherson (1984) and Smith et al. (1988), who maintain that "even in cases where general knowledge is used, it may not come into play until after the procedures specified in the selective modification model" (1988:525), which means that (1) extra-linguistic knowledge is not an indispensable, intrinsic factor of the interpretation of complex naming units, and (2) if this kind of knowledge is included in the interpretation process, it happens no sooner than the second stage of a two stage process: the first stage, based on a rapid composition process, is sometimes followed by a slower composition process employing one's general knowledge. A criticism of this position may be found, for example, in Murphy (1988).

In addition, the picture of the meaning-prediction process outlined above suggests that the process of interpretation of novel naming units is not totally arbitrary. Rather, each novel naming unit generated by a productive WF Type, is predictable in the widest sense of the word: the regularity of the particular

productive WF Type establishes the necessary compatibility between the semantic components of the combined morphemes, reflecting extra-linguistic affinity of the related objects. Under these compatibility conditions certain semantic component combinations are more acceptable because they reflect more realistic relations in extra-linguistic reality. This is the reason that justifies developing a predictability model in which the meaning predictability of naming units is conceived of as a gradeable phenomenon, as a continuum, which as such, can be computed by statistical and mathematical methods.

This view corresponds, in one of its aspects, with the conclusions drawn by Coolen et al. (1993: 243) from their experimental research. The conclusions concern the relation between dominant meanings and dominant aspects of meanings, on the one hand, and the interpretability of novel primary compounds, on the other. Coolen et al. propose that dominant meanings may become available sooner than subordinate meanings, and the same applies to dominant and nondominant aspects within a particular meaning. Dominant aspects may become available sooner than nondominant ones. Therefore,

> [t]he selection of a meaningful semantic relationship between compound members may depend on the dominance of compatible meaning aspects. Dominant meaning aspects will be considered earlier in the interpretation process than will subordinate ones, and meaningful relations involving dominant meaning aspects will be computed before semantic relationships in which subordinate meaning aspects are involved (1993: 243).

To put this approach in a broader theoretical framework it will be useful to refer to the discussion of two basic types of the interpretation of meanings of naming units as presented by Coolen, van Jaarsveld, and Schreuder (1993). According to one hypothesis, semantic representations of the constituent nouns are available either simultaneously or in rapid succession. Because of temporal overlap in the access of the semantic representations, *activation* of these representations may be *interactive*. This approach hypothesises that

> common or compatible meaning aspects may be reinforced, and the activation of mutually inconsistent meaning aspects may be inhibited... Meaning aspects for which the activation is enhanced may be taken up more readily in the interpretative process, whereas meaning aspects that are inhibited may not be considered at all. (Coolen et al. 1993: 236)

An alternative hypothesis speaks of *independent activation*: the individual aspects of semantic representations become available regardless of their relevance for interpretation. Consequently, "[i]nterpretative processes will have to consider meaning aspects of the constituent nouns more elaborately be-

cause activation of the semantic representations itself does not provide cues for interpretation" (ibid.: 236).

In view of these two basic conceptions the above-outlined meaning-prediction process may be treated as an *interactive model* by outlining the procedure of the individual matching steps. This procedure starts at the level of the most characteristic, prototypical features of 'motivating' objects (reflected in the semantic components of level 4). It may be surmised that, if there is 'close competition' between two or more naming unit interpretations of roughly equal Predictability Rates, the corresponding matching operations take place simultaneously.

3.5.1.6 *Idiosyncrasies of Onomasiological Type 5*

The previous considerations apply to the first four Onomasiological Types. A few additional comments are required for Type 5 (Onomasiological Re-categorisation, traditionally discussed under the label of conversion or zero-derivation), which is peculiar by having no onomasiological structure. As a result the connection between the converted naming unit and the motivating, converting one is very close, or better, direct. While the first four Onomasiological Types are based on the interrelation of two (or more) motivating extra-linguistic objects, with the new naming unit representing a new quality resulting from these horizontal (level-internal) and vertical (inter-level) relations (Figure 1), the process of conversion results from the direct interaction between the motivating and the motivated objects (example (18)). This close relation characterises all the levels. Due to the nature of this naming process one cannot distinguish between the onomasiological base and the mark. Consequently, rather than by the relation between the base and the mark, the semantic relations follow from the direct relation between the most general conceptual categories of SUBSTANCE, ACTION, QUALITY, and CIRCUMSTANCE, which in the other four Onomasiological Types establish a framework for the semantic structure. In the case of Onomasiological Type 5 they are simply mapped down from the conceptual level and put into direct semantic relation (i.e., without any mediating semantic structure).

I assume that the prediction process of converted naming units takes a similar strategy in terms of the sequence and interrelation of the individual steps. Its idiosyncrasies are related to the unique features of the WF process of conversion. Since there is no onomasiological structure, no matching process takes place. Rather, the emphasis is laid on the possible 'radiation' of the central meanings of the converting (motivating) naming unit. In other words, prediction is closely related to the identification of the *established* (institutionalised)

meanings that seem to be the best candidates for the conceptual recategorisation underlying conversion, and – equally important – the *direction* in which they can be recategorised. Consequently, the prediction of the meaning(s) of converted naming units appears to be a two-step process: within the first step, the possible general semantic relation(s) between the concepts of the motivating and the motivated naming units is (are) proposed; within the next step, these general relations are specified.

Thus, for example, the general semantic relation of [SUBSTANCE$^{\underline{\text{Manner/Pattern}}}$ACTION] identified in the first step as the best candidate for the conversion of *boy* primarily on the basis of its prototypical semes [Male] [–Adult] and [Characteristic Behaviour] is, in the second step, specified (apart from other low-predictable proposals) as 'to act or behave the way boys do (immature)' and 'to (try to) look/behave like a boy (clothes, haircut, motions, walking) – of girls' on the basis of prototypical features, and therefore prototypical semantic components [Male] and [Adult], respectively.

To take another example, there are two central semantic structures available for *cableway* following from the prototypical and the specification (level 3) features/semes like [Means of Transportation] and [Constructable]: [SUBSTANCE$^{\underline{\text{Instrument}}}$ACTION] and [SUBSTANCE$^{\underline{\text{Result}}}$ACTION]. In the second prediction step they are specified as 'to travel by cable' and 'to construct a cableway', respectively.

3.5.1.7 *Predictability of recategorised proper names*

Štekauer (1997) discusses the semiotics of proper names and their conversion. Some of the ideas are relevant to the discussion of predictability, and therefore it will be worth summing up the basic ideas from the present point of view. In principle it is claimed that there is no difference between common names and proper converted names. In the same way as common names, proper names, too, can be delimited by a set of logical predicates at the conceptual level (logical spectrum). In the same way as for common names, these logical predicates specify the most characteristic features of the object name, in this case, a person(ality), both physical and mental features, their behaviour, achievements, etc.

Still, there are two differences between proper and common names. The first one bears on the fact that while the logical spectrum of any 'common object' identifies the prototypical characteristics of a class of objects through the process of generalisation and prototype identification, in the case of persons as objects of the naming process, this delimitation is individualised. This individualisation does not mean that the respective semes mapping the logi-

cal spectrum are 'distinguishers'. Even at the individualised level we do justice to the general, characteristic properties, behaviour features, etc., of the person(ality). By distinguishers we can refer to their non-typical, idiosyncratic features or behaviour.

The second difference concerns the fact that the conceptual structure (and hence the semantic structure) of a proper name reflects by necessity gradual and/or sudden modifications or changes in the development of a particular person and the specific life-shaping events. We all develop and act under the influence of both our age and the social environment in which we live. These changes cannot but find their reflection in conceptual and semantic structure. In the case of common names such changes are possible, though not necessary.

These differences, however, do not call into question the above-mentioned postulate. The logical spectrum of a proper name is, in the same way as with common names, projected onto the semantic level of a linguistic sign in the form of hierarchically ordered semes constituting a semantic structure.

Importantly, the conceptual analysis of proper names is objective in the same sense as that of common names. It is independent of individual assessments and positive or negative attitudes of a language user. In this sense, the *meaning* of a proper name is not what I or you think of, or how I or you perceive the person(ality) named by a proper name.

It follows from the previous considerations that the prototypical semes of level 4 capturing the most characteristic features of a person(ality) named are the best candidates for the process of recategorisation (conversion). In the process of recategorisation any of level 4 semes can be activated and thus determine the direction of conversion.

It was shown that the process of WF within the framework of Onomasiological Type 5 does not, in principle, distinguish between common and proper names. What about the meaning-prediction process?

It will be argued that there are no principled differences in the prediction process either. The principle of world knowledge and experience is vital for the interpretation of a recategorised proper name, too. We cannot predict the meaning of, for example, the converted naming unit *display*$_{V}$ if we do not understand the meaning, function, purpose, etc., of the converting *display*$_{V}$. In a similar vein, we cannot predict the meaning of the recategorised naming unit *to Havel*$_{V}$ (a famous personality of the Czechoslovak velvet revolution) if we do not know who Havel is and what the characteristic features of this personality are (were). The knowledge of and/or the experience with the object named is the fundamental predictability condition for both recategorised common and

proper names. It suggests that, as with common names, the predictability of recategorised proper names is not language dependent.

Štekauer et al. (1997) report on an experimental piece of research, one part of which examined the ability of language users to predict the meaning of recategorised proper names. The research included 75 undergraduates studying at the Department of British and American Studies, Faculty of Arts, Prešov University. The research took the form of an anonymous test consisting of six parts, one of which was focused on recategorised proper names. The subjects were given the following task:

(28) 1. Explain in detail the meaning of converted proper names in the following sentences.
2. Suggest their Slovak translation.[11]
Both in 1. and 2. avoid using the phrase 'He did it like...". Try to be more explicit.

The fact that the recategorised proper names were presented in context does not seem to be of high relevance, because the context could not be of much use if the subjects did not know the objects named. This can be illustrated by the following examples:

(29) a. He *Joseph Hellered* the situation in that company.
b. He *Ben Johnsoned* his promising swimming career.

In the test ten such names were offered; nine of them were well-known international personalities at that time, *Joseph Heller*, *Stephen Spielberg*, *Ben Johnson* (a Canadian 100 m world-record holder who was later disqualified for doping), *Richard Nixon*, *Salieri, Maradona, Havel, Madonna, G. Bruno*, and one well-known Slovak political satirist, *Milan Markovič*.

Their answers were evaluated as 'correct' (if they reflected the characteristic features of the personalities) and 'false'. The findings indicated serious gaps in the extra-linguistic knowledge of our students who frequently missed the point.

Some readings made no sense, others were too general, and mostly activated level 3 semes. So, for example, readings like 'to organize a meeting like R. Nixon' cannot be accepted as 'correct' for a speech community. People are hardly aware of the way Nixon organised meetings, with the exception of a small speech community surrounding Richard Nixon at the time of his presidency, which, obviously, was not the case with our informants. Rather he is known, at least in our speech community, for the Watergate affair.

For similar reasons the readings for *Heller*$_V$, such as, 'to criticize', 'to analyze in detail' are too general, in the same way as 'to be successful', or 'to direct' for *S. Spielberg*$_V$. There are many novelists who 'criticise' and/or 'analyse in detail', and there are many directors who are 'successful', and each director 'directs'. A very low number of our informants were able to predict the meaning of *Salieri*$_V$, and quite a lot of them confused the fate of G. Bruno with that of G. Galilei.

The test showed that those subjects who had the necessary world knowledge managed to predict the readings of the recategorised proper names, the others gave no proposal, or fell back on too general interpretations. Not surprisingly that there were many 'correct' answers for *Madonna*$_V$ and *Maradona*$_V$.

It was mentioned above that the recategorisation-related prediction process takes two steps. To illustrate this point with recategorised proper names, let us take, the example, *Ben Johnson*, a 100 m world-record holder. The knowledge of the object to be interpreted, the fact that he used drugs, and consequently spoiled his promising career, enabled our informants to identify one of the most characteristic 'behavioural' features of Johnson in the form of a general semantic relation [SUBSTANCE$^{\underline{\text{Manner/Pattern}}}$ACTION]. The next step identified the Action as, roughly, 'to spoil one's career as a result of doping'.

To summarise the previous discussion, I argue that the word formation and the meaning-prediction processes in terms of recategorised common and proper names are, in principle, the same; they take place in two steps, and that they heavily depend on the world knowledge and experiences of language users.

3.6 Onomasiological Structure Rules

Meaning predictability is also conditioned by an *acceptable onomasiological structure*. This can be articulated as the following principle:

(30) The meaning prediction capacity of a language user is conditioned by his knowledge (subconsciously performed competence) of *Onomasiological Structure Rules* that function as constraints on the interpretation of naming units.

The experiments discussed in Chapter 4 provided me with ample examples of readings which are unpredictable because the onomasiological structure (WF Type) of a reading proposed does not correspond to the onomasiological structure of the naming unit. Examples are given in (31):

(31)

Naming unit	Proposed reading	Appropriate naming unit
flower hat	'a flower in the shape of a hat'	*hat flower*
hill star	'a hill in the shape of a star'	*star hill*
	'a very high hill that 'touches' the sky'	*star hill*
	'many stars at one place'	*star hill*
garden whisky	'vegetable meal with a little whisky'	*whisky garden*
	'very good fruit for somebody who likes this fruit as much as whisky'	*whisky garden*
age bag	'a period when bags were/are popular'	*bag age*
dog spade	'a dog for watching spades'	*spade dog*
	'the shape of a dog in the ground which is made with a spade'	*spade dog*
shape cloth	'a figurine for cloth makers'	*cloth shape*
blondesjoker	'a blonde haired woman who is good for a laugh'	*joke-blonde*
feather-dialer	'a person who decorates dials with feathers'	*dial-featherer*
anthraxist	'a person who fell ill due to anthrax'	*anthraxee*

The inadequacy of the readings proposed can be judged from simple Onomasiological Structure Rules imposing constraints on the internal structure of WF Types. Examples of these rules are given in (32):

(32)
a. The onomasiological base is on the right in English compounds.
b. The Pattern seme is left of the State seme.
c. The Quality seme is always left of the Patient seme.
d. The Source feature is left of the Action seme.
e. The Purpose seme is left-oriented.
f. If the structure contains the Agent seme the Object seme is left of the Action seme (or, Action directed at Object is left-oriented in a structure with the Agent seme).
g. If the structure contains the Agent seme the Instrument seme is left of Action.
h. If the structure contains the Instrument seme and the Result seme, then the Result seme is the right-hand neighbour to the Action seme, and the Instrument seme is the left-hand neighbour to the Action seme.
i. If the structure contains both the Material seme and the Object seme, the material seme is a right-hand neighbour of Object.

A set of such rules makes it possible to decide whether a particular interpretation is admissible for the underlying WF Type. For example, the reading 'a period when bags were/are popular' is not admissible for *age bag*, because it is excluded by Rule (32a.) above. According to this rule, it is 'period' of the proposed reading that must be placed on the right-hand side of the naming unit corresponding to the proposed reading: it follows from the definition of onomasiological base in Chapter 2 that it stands for a general class of objects; the onomasiological mark provides more specific information ('when bags are/were popular'). Hence, the proposed reading corresponds to the naming unit *bag age*.

The same result for this reading is obtained by applying rule (32c.). The general onomasiological structure of 'mark – base' can be represented as in (33):

(33) Quality – Patient
bag *age*

where *bag* stands for the characteristic feature (Quality) of a particular period, and as such it must be on the left-hand side to the Patient seme.

Rule (32b.) prohibits the reading 'a flower (Patient) in the shape of a hat (Pattern)' for *flower hat*. The naming unit corresponding to the proposed reading is *hat flower* because it has the corresponding onomasiological structure (34):

(34) [Pattern – (State) – Patient]

The same rule is applicable to the unpredictable reading 'very good fruit for somebody who likes this fruit as much as whisky' of *garden whisky*. Here, *garden* (Patient) stands metonymically for 'fruit' and 'whisky' serves as a Pattern (of quality). Therefore, the required naming unit for the proposed reading is *whisky garden*.

Rule (32f.) predicts that, in English, the Object of Action is, by default, a left-hand neighbour of the Actional constituent. Therefore, while

(35) [Object ← Action – Agent]

is admissible for the reading 'a person (Agent) who tells/makes (Action) blonde (Object) jokes', the reading 'a blonde haired woman who is good for laugh' is controlled by the onomasiological structure

(36) [Quality – Patient]

and, for this reason, the corresponding naming unit must be *joke-blonde*.

The reading 'a person who decorates dials with feathers' for feather-dialler would correspond to structure (37):

(37) [Material – Object ← (Action) – Agent]

However, Rule (32i.), requires the reverse order of the Material and Object semes. For this reason, an admissible naming unit for the proposed reading would be *dial-featherer*.

The reading 'a dog (Agent) for watching (Action) spades (Object)' has the onomasiological structure (38):

(38) [Object ← (Action) – Agent]

Rule (32f.) prohibits the proposed reading for the naming unit *dog spade*, because, in this case, the structure would be as in (39):

(39) *[Agent – (Action) → Object]

Rule (32h.) must be held responsible for the unpredictability of the reading 'the shape of a dog (Result) in the ground which is made (Action) with a spade (Instrument)'. A naming unit corresponding to the proposed reading is therefore *spade dog*.

3.7 Predictability and productivity

At first sight it might seem that the predictability of novel, context-free naming units is in direct proportion to the productivity of WF and Morphological Types underlying these coinages, that is to say, that those naming units coined by the most productive WF Rules will be the most predictable. The situation is, however, much more complicated, and it will be shown later in this work that there are a number of factors which overshadow the productivity factor. There are at least two crucial reasons preventing productivity from becoming a central predictability-influencing factor.

The first reason does not require any in-depth comment: while productivity is a matter of *speaker/writer* predictability is that of any subsequent *interpreter*. While the former coins a new naming unit with a *single general* meaning, the number of possible compatibility-based combinations available to the latter abounds in many cases.

The second reason concerns the crucial difference which any treatment of the predictability-productivity relation must take into account. While predictability concerns the *meanings* of naming units, productivity usually per-

tains to *WF Rules* (WF Types), on the basis of which new naming units are generated.

By implication, while productivity is about the *general*, predictability is about the *individual*. In other words, this relation exemplifies the classical philosophical dichotomy of the universal and the particular. Productivity does not (and cannot) refer to individual meanings. Like any other system of rules, Onomasiological Types, WF Types, and Morphological Types also represent a certain level of generalisation. This generalisation means that all the naming units generated by a particular Type – although differing in their subtleties – can be subsumed under one general meaning. It is for this reason that the notion of productivity necessarily assumes a certain amount of abstraction from individual idiosyncrasies.

On the other hand, the predictability of naming units is based on the identification of specific, frequently idiosyncratic meanings. Rather than being interested in what a particular novel naming unit has in common with other, already existing naming units, predictability zeroes in on what makes a coinage unique and different from any other naming unit.

The experimental research outlined below provided us with numerous examples of this dichotomy. The general, the productivity-related, is, in my approach, represented by semantic structure, standing for the internal, structural variants of various conceptual categories underlying the various Word-Formation Type Clusters.[12]

Let us illustrate the point by one of the sample naming units. In the case of *baby book*, the predictable readings 'a book about babies and how to take care of them' and 'a book with photos of one's baby(ies)/album; with records of baby's development (first steps, first word, . . .)' fall within one and the same conceptual category of Patients, here represented by the semantic structure of [Stative (=Theme) – (State) – Patient].[13] The latter, necessarily being a generalisation, cannot discern the subtle, but vital, difference between these two distinct readings of *baby book*.

Similarly, both of the predictable readings for *flower hat* 'a hat with flowers on it' and 'a hat made of flowers' belong to a productive WF Type represented by the semantic structure [Stative (=Material) – (State) – Patient]; the predictable readings of *game wheel* 'a wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games' and 'a wheel which is a part of a game equipment, a wheel with which a game is played' belong in a productive WF Type represented by the semantic structure of 'Process – Instrument'; the predictable readings of *to boy* 'to act or behave the way boys do (immature)' and 'to (try to) look/behave like a boy (clothes, haircut, mo-

tions, walking) – of girls' fall within a productive WF Type represented by the semantic structure of [SUBSTANCE$^{\underline{\text{Manner/Pattern}}}$ACTION]; the predictable readings of *to morning* 'to get up early in the morning' and 'to do one's morning routine' belong to a productive WF Type represented by the semantic structure of [CIRCUMSTANCE$^{\underline{\text{Temporal}}}$ACTION]; and the predictable readings of *to triangle* 'to draw a triangle' and 'to make a triangle from something/to make triangular things' belong to a productive WF Type represented by the semantic structure of [QUALITY$^{\underline{\text{Result}}}$ACTION].

In any case, the role of the relation between productivity and predictability in the meaning-prediction process cannot be excluded. As already indicated in the outline of the predictability theory, the process of predicting the meaning(s) starts at the phonological level and proceeds upwards through the onomasiological WF model, with the interventions of the interpreter's extralinguistic knowledge and experience. The identification of a specific morphological structure at the onomatological level is a precondition for the identification of a possible semantic structure, with the latter subsequently showing a language user the path to more specific readings. It may be assumed that an interpreter identifies the possible semantic structures on the basis of his/her competence, that is to say, his/her knowledge of the productive rules of WF. It is here where the connection between WF productivity, on the one hand, and novel coinage meaning predictability, on the other, may be sought.

As already indicated, this connection does not mean that more productive WF Types/Morphological Types are automatically more predictable. One of a number of factors that can work counter to the direct proportion between morphological productivity and predictability is the *competition* of productive (!) Types. Thus, for instance, a number of coinages with *-er* (*driver*, *reader*, *printer, manager*, etc.) in the function of the Onomasiological Base may have both Agentive and Instrumental meanings. As suggested in connection with the Objectified Predictability Rate (see Section 3.10 below), the competition of two readings, both in terms of their number and a close Predictability Rate Gap, is an important obstacle to predictability, significantly reducing the Objectified Predictability Rate. Thus, in fact, the existence of more than one highly productive WF Type is a factor having negative rather than positive, boosting effects on meaning predictability.

3.8 Predictability and typicality

Apart from the fallacious assumption of the direct relation between predictability and productivity there is one more fallacy which is equally tempting, and concerns the relation between predictability and typicality, the latter being one of the central notions of the prototype theory.

An important lesson is learned from the following example, which demonstrates that the typicality of a naming unit as a whole cannot be tied to only one of its constituents. Thus, it is not correct to speak of the typicality of *lion pet* exclusively in terms of typicality, that is to say, in terms of the defining, prototypical characteristics of 'pet' (domesticated, small-size, non-dangerous animal, a frequent object of cuddling, etc.). Each of the constituents underlying (motivating) a naming unit has its typical features and prototypical representatives, exemplars. The resulting naming unit, building upon two underlying 'typicalities', i.e., the prototypical features of two motivating objects, represents a new, conceptually processed object with its new typical features which cannot be reduced to one of the motivating constituents.

The interpretation of a new object relies on the combination of prototypical features of the two (or more) 'motivating' objects. Therefore, the prototype of *lion-pet* cannot be inferred merely from *pet*. It is a new 'quality' resulting from the combination of selected prototypical features of 'lion' and those of 'pet'. From this point of view, while the claim that *lion-pet* is an atypical representative of 'pets' is true, it is of little relevance to our discussion. What matters is that, based on and thanks to the prototypical features of 'lion' and those of 'pet', the new naming unit *lion-pet* is fairly predictable in the reading 'a lion which is a pet'. It is within this meaning, standing for the particular mental concept, that we should seek a typical representative, a prototype of 'lion-pets': perhaps, a lion that is perfectly tamed and trained to obey and which is kept in a space attached to a house'. A wild lion can hardly be considered a *lion-pet*.

The most important conclusion in this respect is that the Predictability Rate of *lion-pet* in the above-mentioned reading may be fairly high, even if *lion* is not a typical *pet*!

This apparent contradiction may be accounted for as follows:

The main reason for the above-mentioned hampering effect of level 5 semes is that while the notion of *meaning predictability* refers to the mental concept of a *class* of objects (to be) named, level 5 usually refers to the individual members of the class whose features need not fully correspond to the prototype. On the other hand, *typicality* refers to fuzziness of conceptual

categories implying *individual differences*. Individual members of a cognitive category (i.e. a class of objects covered by a particular naming unit) may share all the prototypical features or only some of them, or different degrees of these features. Thus, a particular member of the class may approach the 'ideal' prototype to varying degrees. This is captured by cognitive linguistics in the notion of *typicality scale*.

Now, let us return once more to the example of *lion-pet*. There are usually a number of interpretations of a novel, context-free naming unit which can assume various positions on the typicality scale. Some of them share only few (prototypical) features with the rest of the class to which they belong on the principle of family resemblance. This is also the case with *lion pet*. Still, even if at the periphery, or, at the bottom of the typicality scale with regard to the concept of 'pet', its reading 'a lion that is a pet' appears to be at the top of the predictability scale when compared to any other possible reading of *lion-pet*. As already indicated, this object of extra-linguistic reality is *conceivable* thanks to the compatibility of semes of the motivating constituents reflecting the features (attributes) of the 'underlying' objects. This conceivability ensures it a place within the field of pets, even though at a low typicality scale level.

Conceivability is also important for predictability considerations, but from a different point of view. When we are faced with the task of interpreting the combination of *lion-pet* the above considerations based on the interpretation 'a pet that is a lion' concern only one of several possible interpretations. Other interpretation options (matching operations) are also available and evaluated in terms of their respective acceptability, for example, 'a pet that guards lions', 'a pet that eats lions', 'a pet that is used for lion hunting', etc. It is our knowledge of the world that helps us overcome the low typicality handicap of the former interpretation by comparing it with all the other mentioned (and still other, unmentioned) readings. This evaluation process indicates that even if the discussed reading is not typical of the class of pets, it is the most predictable reading with regard to the range of possible readings resulting from the combination of *lion* and *pet*, because it is the best conceivable reading.

Thus, while the typicality rating process primarily relies on comparison *external* to the particular naming unit, i.e. its comparison with the other naming units belonging to the same lexical class (the class of pets, in our example), or, in other words, the comparison of a particular object with other objects of the same class, the predictability rating is primarily based on a comparison *internal* to the particular naming unit, the comparison of the possible meanings within one and the same naming unit, and thus within one and the same object (itself representing a class of 'lion pets').

Interestingly, each of the possible readings brings *lion pet* to a new network of typicality relations. Thus, when the reading 'a pet that guards lions' is evaluated in terms of typical 'utilisation' of pets it follows that it is only one kind of pet, the dog, that comes into question for this purpose. Therefore, rather than a hyperonym class name of 'pet', the name of hyponym 'dog' would be expected if – unlike our experiences – a dog were used for guarding lions, i.e., *lion dog*. In this respect, a pet as a guard of wild animals in general seems to rank even lower at the typicality scale than the former, more predictable interpretation.

By the same token, when considering the possible reading 'a pet that is used for lion hunting', it is only dogs from among the class of typical pets that are used in wildlife hunting.[14] On top of it, dogs are not used for the hunting of lions because other techniques of lion hunting appear to be more effective, which is given by the environment in which lions live. Using dogs for lion hunting would probably doom the dogs to die. Therefore, the typicality of this 'function' of pets approaches zero.

Comparison of the individual typicalities brings us to the conclusion that it is the former reading which has the highest relative typicality, or, in other words, that it is best acceptable to language users based on their knowledge and experiences. Since it is the most acceptable of all the readings conceivable its Predictability Rate is the highest of all.

This example illustrates the basic idea of this section, that is to say, that *the level of typicality of features need not necessarily imply the level of meaning predictability* (and vice versa). The point is that a new object named is not a simple combination of two prototypes. It is only one or several, but not all, of the prototypical features that are combined in a new naming unit. Otherwise many naming units would be impossible due to the incompatibility of the motivating constituents. To elaborate on the *pet* example, while the 'relative' smallness (compare the size difference between a kitten and a St. Bernard), domestication, and perhaps (relative) obedience (compare a well-trained Alsatian vs. cats with their tendency to roam), (relative) attachment (compare a cocker spaniel and a tortoise), and (relative) peacefulness (compare again a cocker spaniel vs. bullterrier and pitbull)' are the prototypical features of 'pet' they at the same time indicate that a prototype of 'pet' is a difficult-to-define 'mixture' of these 'ingredients'. In addition, a prototypical feature of pet is also 'an animal kept for pleasure'. It appears that it is primarily this feature which is activated in establishing a concept of a *lion pet* – lions can be kept for pleasure – as our knowledge and experience confirm, and they can be (relatively) domesticated (tamed) and trained.

In this way I complete my account of *predictability as a series of matching operations activating various prototypical features/semes in close connection with extra-linguistic knowledge*. The activation of prototypical semes and their matching provides us with a series of possible readings which are then evaluated in terms of acceptability within their respective interpretation classes.

3.9 Predictability Rate

This section proposes a method of calculating the *Predictability Rate* (PR). The introduction of this notion is based on the postulate that the meaning predictability of novel naming units can be quantified, computed, and mutually compared. Dressler & Ladányi (2000: 127, 128) maintain that "[g]radualness, instead of discreteness, is an essential property of any model which works with prototypes" and since "productivity is a prototypical property of rules" of word formation, productivity is gradual. Given the close interconnection between word formation and word-interpretation, the same may be assumed about meaning predictability, that is to say, meaning predictability of new coinages is a prototypical feature of word formation, and therefore it is gradual. And like productivity, meaning predictability is not an all-or-nothing notion.

The calculation of Predictability Rate is based on the following postulates:

1. The predictability of meanings of naming units correlates with the acceptability of these meanings to interpreters. It may be proposed that *acceptability* is a system-level analogue to Labov's speech-level term 'consistency' "with which a given sample of speakers does in fact apply the term" (Labov 1973: 353).
2. Since there is no clear-cut boundary between acceptable and unacceptable meanings the predictability of the meanings of naming units is a *cline*.

Then, the Predictability Rate of a particular reading of a novel, context-free naming unit can be calculated as its frequency of occurrence weighted for the scores assigned:

(40) $$PR = \frac{r}{R_{max}} \times \frac{p}{P_{max}}$$

where

r = the number of informants identifying a particular meaning as acceptable

R_{max} = the total number of informants
p = the sum total of the points assigned to a given meaning by all informants (on a scale from 1 to 10, where 10 stands for the highest acceptability of the meaning)
P_{max} = the maximum possible number of points assignable by all informants

For example, the reading 'a book for babies (fairy tales, rhymes, pictures; drawings)' of the naming unit *baby book* (see Section 4.2.2.1) was proposed by 38 out of 40 informants, i.e., the frequency of occurrence of this reading is 38/40 = 0.95. The scores assigned to this reading is 306 points of the total of 400 assignable points, which is 0.765. The resulting PR of this particular reading is therefore 0.727. It is much higher than the PR of, for example, 'a naive, babyish book', also proposed for this sample naming unit, because it was only proposed by 16 out of 40 informants (16/40 = 0.40), and its frequency of occurrence is merely 0.213, which gives the PR of 0.085.

By implication, this method of calculation of PR makes it possible to evaluate the strength of various readings proposed for a novel, context-free word, and thus determine the degree of their predictability.

3.10 Objectified Predictability Rate

While the PR value is sufficient to compare the meaning predictability of readings within one and the same naming unit it does not allow for comparing the predictability of readings of various naming units. As follows from Section 3.8, Predictability Rates calculated for the individual predictable readings take into consideration two significant variables: first, the number of informants who adduce a particular naming unit reading, which indicates that the reading is acceptable to them; second, the assigned/assignable scores proportion for the individual readings, determined by the informants' rating activity. The Predictability Rate is therefore directly proportional to the number of informants who identify the reading and the points assigned to this reading.

The Predictability Rate calculated in this way considers, however, a particular reading of a naming unit in isolation, regardless of the meaning predictability of the other possible readings of the same naming unit. It may be assumed that the strength of the individual readings, i.e., their Predictability Rates affect each other, which is reflected in their respective positions on the meaning predictability scale.

It may be assumed that the individual readings *compete* with one another. I first used the idea of 'competition-rather-than-blocking' in Štekauer (1998) in connection with the evaluation of the word formation productivity of various WF Types belonging to the same conceptually determined cluster (WF Types of Agent, Patient, Instrument, Negation, Action, Location, etc.). As already proposed in the preceding sections, the word formation and the word-interpretation processes are closely interrelated. Consequently, the notion of competition can also be advantageously employed to account for mutual relations between various readings of one and the same word, featuring various PRs.

Thus, the concept of *Objectified Predictability Rate* (OPR) builds upon the notion of Predictability Rate which is taken as a point of departure for subsequent calculations reflecting the differences in PRs of a certain number of most predictable readings of a naming unit. It follows from the *Competition Principle* that the greater the *Predictability Rate Gap* between the most predictable reading of a naming unit and the next lower PRs of the same naming unit (and at the same time, the greater the Reading 1/Reading 2 and Reading 1/Reading 3 ratios) the higher the OPR. The corollary of introducing this variable can be illustrated by the following example:

Let us suppose that there are two naming units X and Y. Their three most predictable readings are X_1, X_2 and X_3, and Y_1, Y_2 and Y_2, respectively. Let us further suppose that X_1 and Y_1 are the top PR readings of their respective naming units and happen to have identical PRs of, let's say, 0.486. Furthermore, let us suppose that the PR of X_2 is 0.194 and that the PR of $Y_2$0.362. Finally, let us assume that the third-rank readings' PRs are identical, for example, X_3 = 0.088, and also Y_3 = 0.088. This situation is given in (41):

(41)

Naming unit X		Naming unit Y	
	PR		PR
X_1	0.486	Y_1	0.486
X_2	0.194	Y_2	0.362
X_3	0.088	Y3	0.088

Since the competition of predictable readings in the case of the naming unit Y is much tougher than in the case of X, intuitively the actual (objectified) predictability of X_1 is higher than that of Y_1. This fact is captured by the proposed Objectified Predictability Rate.

This type of relations may be advantageously calculated using Luce's (1959) choice rule which makes it possible to weigh (in my case) the strength (PR) of the most predictable reading against the strength (PRs) of any number of other

competing readings. This method was applied, among others, by Gagné & Shoben (1997) for the calculation of the strength of the thematic relation which is the best candidate for the interpretation of a particular complex word.[15]

The formula adapted for the calculation of the Objectified Predictability Rate, is as follows:

$$(42) \quad OPR = \frac{PR^{top}}{PR^{top} + PR^{top-1} + PR^{top-2}}$$

If formula (42) is now applied to (41), we get the $OPR_X = 0.633$ and $OPR_Y = 0.519$. By implication, with other values identical it is the higher $PRG_{X1\text{-}X2}$ value compared to the $PRG_{Y1\text{-}Y2}$ value which is responsible for the higher OPR of X1. This result confirms our intuition according to which reading Y_1 faces much 'tougher competition' on the part of reading Y_2 than X_1 on the part of X_2. Consequently, the predictability of X_1 is much better than that of Y_1 in spite of these two having identical PR values.

From this it follows that a high absolute PR does not guarantee a high OPR: a naming unit reading of lower PR may be comparably more predictable than a reading of another naming unit of a higher PR, if the former can take advantage of a considerable PRG. This postulate has been confirmed in my research on a number of occasions. For illustration, within a conversion-oriented Experiment 2, the top PR reading 'to take part in a conference' of a possible naming unit *to conference* has by far the highest PR (0.427) among ten potential conversions in Experiment 2. However, it ranks as low as ninth in terms of OPR. This is caused by tough competition on the part of the other two readings, especially the extraordinarily high PR of the rank 3 reading (0.165).

It goes without saying that, unlike the simplified example (41), the OPR does not depend on only one PR Gap value; that is, it not only pertains to the relation between the first and the second-rank readings. Generally, the higher the number of competing predictable readings, i.e., the higher the number of close PR Gaps, the lower the OPR.

3.11 Hypotheses

The discussion in Chapter 3 can be summarised in the following hypotheses which were examined and evaluated in an extensive experimental research project presented and commented on in detail in Chapter 4.

1. The predictability of meanings of naming units correlates with the acceptability of such meanings to interpreters.[16]

2. Since there is no clear-cut boundary between acceptable and unacceptable meanings the predictability of meanings of naming units is a cline.
3. An important condition for meaning predictability is a combination of prototypical semantic components (= as linguistic representations of logical predicates reflecting the prototypical features of the objects named).
4. The predictability of the meanings of any naming unit heavily relies on the conceptual level analysis, on the cognitive abilities of language users (i.e. supralinguistic level) the principles of which are identical to mankind as a whole. From this point of view, the experimental results for native speakers should not significantly differ from those for second language speakers.[17]
5. The meaning-prediction process is significantly influenced by extralinguistic knowledge and experiences of language users. Given a more or less homogeneous group of my informants (university education in the field of humanities), living in *similar cultural settings*, the results should not be negatively influenced by their belonging to different speech communities.[18]
6. These postulates can be experimentally verified by computing the Predictability Rate and the Objectified Predictability Rate. While the former reflects the meaning predictability relations within one and the same naming unit, the latter variable makes it possible to compare the predictability strength of top readings of various naming units.
7. For each novel, context-free naming unit there is *one or two central meanings* that are most acceptable to language users and hence most predictable.

Chapter 4

The Experiments

4.1 Method

I now turn to the application and verification of the theoretical considerations presented in the preceding chapter, by means of four experiments. The first two of them make use of non-established (i.e., either possible, non-existing, or non-institutionalised) naming units, belonging to the selected Onomasiological Types, in particular to Onomasiological Type 3 (it is this type of naming unit which has been the focus of the majority of predictability research projects, i.e., it includes *inter alia* primary compounds) and Onomasiological Type 5 (in particular, Noun → Verb conversions). The primary goal of these experiments was to test the proposed method of computing the Predictability Rate and the Objectified Predictability Rate, and through the calculated PR values to evaluate the correctness of the prediction concerning the key role of prototypical semes of level 4. Equally important was an assessment of the hypothesis according to which there are no principled differences between native and non-native speakers in terms of their meaning-prediction capacity. The role of semantic structure (underlying my notion of WF Type) will be examined. I will also refer to some cases directly supporting the analogy-based model, but also will refer to some cases in which analogy runs counter to the predictability of a reading. Additionally, I will trace the interconnection between the linguistic and extra-linguistic factors in the various predictable readings of sample naming units. The proportion between the number of possible readings and the actually predictable readings will clearly follow from the experimental data. The application of two differently designed methods of data collection (requiring the informants to propose all acceptable readings vs. only the most acceptable reading) will make it possible to find out whether the meaning-prediction process is affected by the time factor (the time available to the informants for proposing the most acceptable reading(s)).

The third experiment examines possible naming units belonging to the five different Onomasiological Types, each of which is represented by two possible naming units. The primary objective is to apply the proposed method of cal-

culation of the Objectified Predictability Rate to various types of naming units, and, in this manner, to demonstrate the viability of this method of comparing the predictability of meanings of various naming units.

The fourth experiment is a brand new type in terms of the method applied. It evaluates the Predictability Rates of, as it were, 'impossible' naming units, i.e., those naming units which do not comply with the principles of productive word formation and belong to the 'types' specified by various morphologists as 'unacceptable' from the synchronic WF point of view because of violating various restrictions on productivity. The main objective is to relate the notion of predictability and productivity, and to examine whether or not there is any inter-relation and/or influence of productive Word-Formation/Morphological Types upon the Predictability Rate value.

The basic method of the experimental part of my research is as follows. Each experiment included ten possible, non-established naming units (with the exception of *refusnik* – see Section 4.4.1). The reason for using possible, non-established rather than actual naming units is obvious: to get a realistic picture of the way of interpreting novel naming units, informants should not be influenced by the established meanings of existing naming units. The motivating words whose WF bases are present in the sample naming units belong to the core vocabulary. In this way I wanted to remove any linguistic advantage of native speakers over non-native speakers. Given identical linguistic conditions I could evaluate the other factors affecting the meaning-prediction process.

Experiments 1 and 2 included twenty native speakers and twenty non-native speakers each, Experiment 3 included 90 native speakers, and 50 native speakers took part in Experiment 4. The common task of the informants was formulated as follows:

(43) The following are potential English words. Propose as many possible meanings for each of the words as you can think of, and assign a score to each of the meanings proposed, with 10 points indicating the highest probability of occurrence in the language of the meaning proposed (and therefore most predictable), and 1 point the minimum chance of occurrence in the language of such a meaning.

Since the experimental naming units used in Experiment 4 violate the constraints on productivity, this wording was modified for this experiment by omitting the word *potential*. This instruction was followed by an example. Thus, for instance, Experiment 3 was introduced by the following illustration:

(44) *water-mill*
– a mill driven by water 10 pts.
– a mill near water 4 pts.
– a mill having the colour of water 1 pt.
etc.

The example was followed by ten 'test' words. The informants were asked to do the test in writing. Since I could not take part in testing all the informants for objective reasons (see below) I imposed no time limit for responses. The informants were asked to do the test in their own time and deliver it to me via e-mail, snail-mail, or personally. In this way all of them were provided similar conditions for completing the test.

The informants in the first two experiments fall within three basic groups. There was a sample of 20 native speakers and 20 Slovak informants (non-native speakers) who completed the test as described above. The purpose of including in the experiments both native and non-native speakers was to give support to the above-mentioned hypothesis of no principled differences in their meaning-prediction capacity. Most of the native speakers were contacted by friends of mine abroad or by my students who have friends in various English speaking countries. A number of the questionnaires were also completed by native speakers who teach at a language school in Košice, Slovakia. All in all the major portion of native speakers were university undergraduates, and a smaller part university graduates. All Slovak informants were my students at the Department of British and American Studies, attending either *An Introduction to English Linguistics* or *English Lexicology* course. The experiments were carried out in two consecutive years, in the order from Experiment 1 to Experiment 3.

In addition there was another group of 25 non-native speakers, the so-called check-group of Polish undergraduates (my students in Rzeszow who studied English) whose task was a little different. Rather than proposing the greatest possible number of meanings, they were asked to propose only one single meaning for each of the possible naming units which they found most acceptable, i.e., the best candidate for integration in the lexicon of existing words. In addition, they did the experiment under timed conditions, having only 60 seconds per naming unit. The instructions were given verbally, including an example, in which the most predictable reading of, for instance, *water-mill* was highlighted. The next procedure was as follows: a naming unit was read by myself twice, and then written on a blackboard. From this point, I counted 60 seconds. Then I uttered another possible word, and the cycle

continued through all ten possible words. At the end, I immediately collected the 'tests'.

The results were processed in the form of tables representing the individual readings and their respective scores assigned by the individual informants. In order to assess my hypothesis concerning equal predicting capacity of both native and non-native language users the results obtained from native speakers and non-native speakers in the first two experiments were first analysed separately. Then they were joined in order to obtain more statistical power.

As already indicated, all 'main group' informants were presented the sample naming units in written form. While this method of research was 'enforced' by my ability to reach native speakers, and the necessity to establish identical conditions for both native and non-native informants, this circumstance cannot be assessed negatively. I surmise that the major part of novel naming units is encountered by language users in a written form, in newspapers, professional and popular journals, on the Internet, documentation to inventions, projects, designs, etc., and therefore the conclusions of the research seem to be relevant in terms of the objectives pursued for this research.

It should be noted that the conditions of interpretation of contextualised and established words during natural-language comprehension differ from the comprehension of novel, context-free naming units. While the processing and intrerpretation of individual words, substantially facilitated by (linguistic and/or situational) context takes few hundred milliseconds the time required for the comprehension of a novel, context-free naming unit is usually much more demanding and time-consuming. This explains much longer time provided to my informants for fulfilling their tasks. As postulated in the Introduction, context-free meaning predictability provides a general, objective, and unbiased view of the interpretation of novel naming units, undistorted by the infinite number of potential linguistic and extra-linguistic contexts, including time pressure, that may influence and condition a specific act of using a naming unit.

4.2 Experiment 1

This section gives a detailed discussion of both predictable and unpredictable readings as proposed by both groups of informants in order to demonstrate the method of my approach to the analysis of the individual non-established naming units, their respective readings, and the multiplicity of factors that affect the meaning-prediction process. The analysis of naming units of Experiment 2

can therefore be confined to the predictable readings with highest Predictability Rates. An analysis of each experimental naming unit is introduced by two tables, the first presenting the data for native speakers, and the second the data for non-native speakers. For obvious reasons, the respective lists of readings in these tables needn't coincide as the proposals of native and non-native informants may differ in details. For space reasons, the tables do not present single-occurrence readings. This kind of readings is of little value for our discussion due to a high degree of their accidentalness. In spite of this fact, some of single-occurrence readings are discussed in the analysis of the individual naming units in an effort to identify an informant's reasons for proposing such a reading and/or show the irrelevance of such a proposal. Columns 1–20 refer to the individual informants, and the numeral values in each of the columns identify the degree of acceptability of the individual readings to the respective informants. Empty boxes mean that a given reading was not proposed by a particular informant. For technical reasons, the maximum acceptability value (10 points) is represented as 'X' throughout the tables.

4.2.1 Sample naming units

The test included ten possible and/or non-established (non-institutionalised) primary compounds falling within the scope of Onomasiological Type 3. With one exception they were devised by myself.[1] As the basic criterion, the onomasiological base was defined by the semes [Inanimate] and [Tangible]. Then, the onomasiological marks of the individual possible compounds were specified according to the criteria specified in (45):

(45) The scheme underlying the selection of primary compound constituents functioning as onomasiological marks:

[Animate]	– [Human]			*baby book*
	– [Animal]			*dog spade*
	– [Plant]			*flower hat*
[Inanimate]	– [Tangible]	– [Solid]	– [Movable]	*ball hammer*
			– [Immovable]	*hill star*
		– [Liquid]		*apple-juice seat*
[Process]				*game whee*
[Form]				*shape cloth*
[Location]				*garden whisky*
[Time]				*age bag*

4.2.2 Experimental data and their analysis

4.2.2.1 *baby book*

Table 1. baby book – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a book for babies (fairy tales, rhymes, pictures; drawings)	7	X	X	X		4	X	7	8	X	X	7	X	X	8	7	7	X	8	5
– a book about babies and how to take care of them		8	8		X			9	8	8		8	5				6	8	X	7
– a book with photos of one's baby(ies)/album; with records of baby's development (first steps, first word,. . .)	X	9								X	9	X	9	7		X	6			X
– a (very) small book	4	2	3	2	6	2	1			5	8	2		8		1	1	2		2
– a naive, babyish book						8				5	8	1	3		6	5		2		
– a book written/drawn by babies/children				1			1					2		1						

Table 2. baby book – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a book for babies (fairy tales, rhymes, pictures; drawings)	9	7	8	X	6	8	X	9	9	9	5	5	8	6	6	X	5	9		9
– a book about babies and how to take care of them	7		5				2	8	4	X	7	7	2	8	9	X			8	X
– a (very) small book	5				2	4	X	6	6	7			3	7			2			
– a book with photos of one's baby(ies)/album; with records of baby's development (first steps, first word,. . .)				4	8	6					9	X	6	X		8	7		7	8
– a naive, babyish book		4				2	8	7	X				7		5				4	
– a book with a cover of a 'baby skin colour'		1					3									5				
– a new book (new release)				7			3	8	5											
- a book for girls							5		1											
– a book written/drawn by babies/children							1					2					8			6

Type: Animate Human SUBSTANCE – Inanimate SUBSTANCE
baby
Seme Level 1 – SUBSTANCE
2 – [Animate]
3 – [Human]
4 – [±Male] [–Adult] [Listening Capacity] [Perception by Watching] [Object of Parental Love] [–Reading Skill] [–Writing Skill] [Limited Intellectual Capacity] [Very Small Size] [Period after Birth] [±Drawing Skill][2]
book
Seme Level 1 – SUBSTANCE
2 – [Inanimate] [Tangible]
3 – [Artefact]
4 – [Rectangular Shape] [For Reading/Listening/Perception by Watching] [Having Some Content] [Limited Size Range] [±With Photos/Pictures] [±Containing Data] [–Smell]

The most predictable reading:
'A book for babies (fairy tales, rhymes, pictures; drawings)'

WF Type:		[Stative (=Goal) – (State) – Patient]	
Seme Level Combination:		4–4	
	NS	*NNS*	*Total*
FO:	19/20	19/20	38/40
Scores:	158/200	148/200	306/400
PR:	0.751	0.703	0.727
OPR:			0.574

The next three readings:
'A book about babies and how to take care of them'

PR:	0.285	0.340	0.312

'A book with photos of one's baby(ies)/album; with records of baby's development (first steps, first word, . . .)'

PR:	0.225	0.228	0.227

'A (very) small book'

PR:	0.184	0.130	0.157

Comments
The most frequently suggested meanings are those based on the WF Type of [Stative – State – Patient], with Stative being either Goal (the highest PR), Theme, or Quality. The 'Goal' reading is primarily based on the activation of the following semes: [Listening Capacity], [Perception by Watching] and [For Reading/Listening/Perception by Watching], i.e., the semes of level 4. The PR

value of the top reading (0.727) is the highest PR of all the readings that were identified in this experiment. At the same time, it is the only reading in Experiment 1 whose PR exceeds the value of 0.500. What strikes one are the relatively high PRs of the next few readings of *baby book*. There seem to be several factors contributing to the existence of a fairly high number of predictable readings in this case: first, the SUBSTANCE – SUBSTANCE combination with an [Animate] [Human] onomasiological mark; second, pragmatic circumstances favouring the combination of the two motivating words; third, a good combinability of the semes involved in the meanings proposed.

'A book about babies and how to take care of them' (rank 2) is a 4–4 Seme Level Combination. While, in general, anything can become a topic of a book (SUBSTANCES, ACTIONS, QUALITIES, and CIRCUMSTANCES), in this particular case it is [–Adult] and [+Period after Birth]. In the case of *book*, the activated semes are [±With Photos/Pictures] and [±Containing Data], i.e., level 4 semes.

This and the next readings belong to the identical WF Type and have the same Seme Level Combinations. By implication, the motivations underlying the two readings are closely related. 'A book with photos of one's baby(ies)/album; with records of baby's development (first steps, first word, . . .)' activates level 4 semes for the *baby* constituent which constitute the meaning of the object of photographing/picturing (the book contains pictures of one particular human being) and [+With Photos/Pictures/Records] and [+Containing Data] for the *book* constituent.

The Word-Formation Types with Quality/Pattern feature a range of scores and frequencies of occurrence. At the top of them there is the reading 'a (very) small book'. This level 4 combination activates the seme [Very Small Size], on the one hand, and [Limited Size Range], on the other. The prototypical features of 'book', including [For Reading/Listening/ Perception by Watching] and [Having Some Content] are backgrounded in this case. The [Very Small Size] seme introduced by *baby* causes a kind of tension in regard to the standard [Limited Size Range] seme of 'book' and, consequently, shifts the size-related seme to level 5 (making it idiosyncratic). The resulting combination is therefore level 4 and level 5. The negative effect of the level 5 semes seems to be outweighed by analogy with the institutionalised naming units such as *baby car, baby ladder,* or *baby grand*.

From among other readings, 'a naive babyish book' activates the semes [Low Intellectual Capacity] and [Having Some Content]. The ideal 4–4 Seme Level Combination is not reflected in the PR that is much lower than the PRs of the top readings for *baby book* (0.085). The explanation seems to be related

to the principle of *blocking*, which keeps this reading in the unpredictability range: the existence of *babyish* and *infantile* blocks the use of *baby* in this particular meaning.

Behind the two closely related readings 'a new book (new release)' and 'an author's first book', there is one and the same semantic motivation, in particular, the combination of [Period after Birth] and [Artefact], i.e., level 4 and level 3, respectively. The former seme activates the [Age] seme of *book* which, however, is not its prototypical seme. Hence, level 3 [Artefact] is here completed with level 5 [Age]. This fact does not establish a favourite condition for a high PR. Another factor which seems to reduce the predictability is the *figurative* use of *baby* (SUBSTANCE) in the [Temporal] CIRCUMSTANCE function.

The reading 'a clever baby' in which 'baby' is compared to a 'book' is difficult to accept because of the central position of the [Limited Intellectual Capacity] seme. The development of intellectual capacities of babies is roughly the same, and any differences become apparent at a more advanced age. This comparison might be applicable to 'child' rather than to 'baby'

'A book with a cover of a baby skin colour' could not gain a higher PR because skin colour does not belong to inherent, or most conspicuous, characteristics of 'baby'. Such a naming unit would perhaps require the identification of colour reference by means of a specified element of the determining constituent of the onomasiological mark, such as *baby-colour book*. The same is true of the reading 'a book that smells like a baby'. Its single occurrence can be accounted for by the [Smell] seme not being a prototypical feature of 'book' (definitely, it is level 5) and by the figurative use of the first constituent.

Out of the remaining interpretations, 'a book written/drawn by babies/children', bound to the WF Type [Agent – Action – Result], is of interest as one of only two Actional WF Types. There seems to be something awkward about this interpretation, which is based on the combination of the activated semes of level 4 [±Drawing Skill] and level 4 [Having Some Content]. These two semes are not compatible due to the semes [Very Low Age], [–Writing Skill], and [Limited Intellectual Capacity] characterising *baby*. The pitfall of this reading is that it results from the failure to distinguish between the skills and intellectual capacities of 'baby' and 'child'.

'One's most favourite book' activates the level 4 seme [Object of Parental Love] of *baby*, functioning here – strangely enough – as an Actional constituent in the meaning of 'to love', thus transforming the typical primary compound structure of Onomasiological Type 3 into that of Onomasiological Type 2. In other words, the determining constituent of the onomasiological mark is re-evaluated as its determined constituent. The single occurrence and its lowest

rating seem to have several causes. First, it is a figurative use of *baby*. In this sense, it could be used in combination with any SUBSTANCE-bound onomasiological base. Second, and of greater importance, this WF Type *does not represent a productive pattern* of WF in English. There is a natural tendency in English to combine the Actional constituent with the Agent or the Instrument of Action. In English, Agent is usually a right-hand member of the onomasiological structure (unlike this case: [(Agent) – Action → Object]). Hence, it violates the relevant *Onomasiological Structure Rule*.

The reading 'a book a child is currently holding' is heavily *context-dependent*, and as such, can hardly be predictable. This kind of meaning is usually expressed by a syntactic possessive structure. Similarly, the readings 'an insurance book for a child' and 'a cheque book for a child' are context-bound, and therefore, unpredictable.

'A book shaped like a baby' suffers from *pragmatic restrictions*. The level 4 [Rectangular Shape] of the second compound constituent (*book*) seems to be a predictability-limiting factor: while [Topic] generally varies from book to book, the rectangular shape tends to be (almost) universal. Other shapes are unexpected (level 5), and therefore, unpredictable.

NS – NNS comparison

A comparison of the two groups of informants shows minimal difference in the assessment of the acceptability, and thus, the predictability of the 'central' readings of *baby book*. The results obtained from these two groups do not show significant differences in the majority of cases (the top reading difference: 0.048; rank 2 reading difference: 0.055; rank 3 reading difference: 0.003; rank 4 reading difference: 0.054). This fact supports my hypothesis according to which the predictability of meanings heavily relies on the conceptual processing and extra-linguistic knowledge, and by implication, it is not conditioned by the status of being a native speaker.

Check-group results

Readings	*FO*
A book for babies (fairy tales, rhymes, pictures; drawings)	15/25
A book for mothers, including instructions of taking care of babies	5/25
A book with photos of one's baby(ies)/album; with records of baby's development (first steps, first word, . . .)	3/25
A (very) small book	2/25

The check-group results indicate that the most predictable readings of the main group also dominate under completely different experimental conditions. They

support both the hypothesis given above and the results of the former two groups. The reading 'a book for babies' clearly dominates both the main groups and the check-group. The rankings of the other readings also correspond to the main group results. The reading 'a naive, babyish book' does not occur in the check-group, which may be due to the possibility of proposing only one reading by an informant.

4.2.2.2 *dog spade*

Table 3. dog spade – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a spade in the shape of dog	3		3	1			5	3		2	4	7	3					7		5
– a spade used for scooping up a dog's excrement	8	7			8			5	8			4	9		9	8	1	5	5	8
– a spade with a picture of a dog on it					6									5						
– a spade in a bad condition/of poor quality					1							1								
– a bad hand in a card game/worthless spade card											2		3	5						
– a playing card with animal pictures											3	1								
– a spade for dogs to play with			1				1		2				X	2	2	1				

Table 4. dog spade – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a spade in the shape of dog	8						2	3	2				2	3		2	2	5		
– a spade used for scooping-up a dog's excrement							7	7	9					2			1	4		
– a spade with a picture of a dog on it	9																3			
– a spade whose blade is as sharp as dog's teeth		8													5					
– a spade that barks and bites – a toy				7									1							
– a spade in a bad condition/of poor quality				6			3		5						6	1				
– a lazy person					6	6														
– someone who doesn't like working in a garden					8	4														
– a spade used for burying dogs							3		4					7						
– a faithful dog							4	5	1											
– a spade for dogs to play with											3	5			3				4	

Type: Animate Animal SUBSTANCE – Inanimate SUBSTANCE
dog
Seme Leve1 1 – SUBSTANCE
2 – [Animate]
3 – [Animal] [Canine]
4 – [Characteristic Shape] [±Watching Skill] [±Domesticated] [±Human Partner] [Pet] [–Using a Toilet] [–Intellectual Capacity] [Ability to Play]
spade
Seme Level 1 – SUBSTANCE
2 – [Inanimate] [Tangible]
3 – [Artefact] [Tool]
4 – [For Digging] [Long Handle] [Pointed Shape]

The most predictable reading:
'A spade used for scooping-up a dog's excrement'
WF Type: [Object – (Action) $\underline{\text{Purpose}}$ Instrument]
Seme Level Combination: 4–5

	NS	*NNS*	*Total*
FO:	13/20	6/20	19/40
Scores:	85/200	30/200	115/400
PR:	0.276	0.045	0.137
OPR:			0.548

The next three readings:
'A spade in the shape of a dog'
PR: 0.118 0.065 0.090
'A spade for dogs to play with'
PR: 0.033 0.015 0.023
'A spade in a bad condition/of poor quality'
PR: 0.001 0.026 0.010

Comments
The results reveal confusion among the informants. The prevalence of single occurrence readings indicates that there is something in the relation between *dog* and *spade* which the informants found awkward. The major limiting factor appears to be related to the *incompatibility* of the dominating seme of *spade*, i.e., [An Instrument for Digging] requiring [Human] seme in the other compound constituent (implying hand–foot–eye co-ordination), and the [Animal] seme which implies such inherent semes of *dog* as [–Intellectual Capacity] [–Hands] which hampers using a spade for digging by dogs.

On the other hand, a typical feature of a domesticated dog [Human Partner] and [Pet] is incompatible with the feature [Inanimate] of spade. The seme [For Watching] for *dog* is, on the other hand, compatible with *spade*. Here, however, the semantic compatibility is eliminated by a word formation factor, the ordering of the primary compound constituents. The interpretation of 'a dog for watching spades' requires a reversed sequence of the compound constituents, that is to say, a different WF Type: [Object ← Action/(Action) – Agent] which yields *spade-dog*. In the absence of any 'strong' reading – the PR of the most predictable reading is only 0.137 (less than the rank 4 reading of *baby book*) – the informants had to take pains to propose at least partly acceptable readings, which is reflected in a number of questionable interpretations.

The most predictable reading 'a spade used for scooping-up a dog's excrement' should be interpreted as 'an Instrument used for Action concerning dogs'. Remarkably, the 'dog' itself is involved in this reading only in an indirect way – through its excrement! This reading clearly demonstrates the role of what Murphy (1998) calls 'conceptual elaboration' reflecting our knowledge of the world, and what Wisniewski (1996) illustrates with construal strategies applied to those cases in which the referent of a compound constituent does not correspond to what one expects from the meaning of that constituent (cf. Section 1.3.4.3 and his example of *moose pencil* interpreted as 'a pencil with a moose eraser').

What seems to be another hampering factor related to this reading are the shape and the function of a spade. This is a matter of *pragmatics*: it is a shovel that fits the proposed purpose better, and therefore is generally used for this function. Hence, using a spade for shovelling implies idiosyncrasy, and therefore, level 5.

Along with this conclusion, the proposed interpretation is hampered by an extra-linguistic factor, in particular, by the existence of a special-purpose instrument in English-speaking countries. The corresponding naming unit *pooper-scooper* used for this kind of Instrument supports the blocking of this reading from the linguistic side: there is no need to have two synonymous units.[3] Thus, there are two blocking factors at play: the *pragmatic factor* (the inadequate shape of 'spade') and the *avoid-synonymy principle* (the existence of *pooper-scooper*).

Consequently, rather than relying on a prototypical semantic feature, this reading is based on the [Can Be Used as a Shovel] seme which can hardly be considered to be a prototypical feature of spade. Rather, its use as a shovel is motivated by a level 5 seme. If – despite these circumstances – the PR of this reading dominates the range of proposed readings the reasons for this fact

should be sought in the 'weakness' of the other readings proposed. The lack of any significant competition from the other readings boosts the Objectified PR of this reading, which is higher than that of the most predictable reading of *baby book* in spite of the considerable difference between their PRs in favour of the *baby book*'s top reading.

The reading 'a spade for burying dogs' is closely related to the top one. Its WF Type is the same [Object – (Action) – Instrument]; in this case, 'dog' is directly involved in the Action. There being neither word formation obstacles nor those related to semantic incompatibility (the combination of semes: level 2 [Animate] → [Mortal] and level 4 [For Digging]) nor pragmatic blocking of the preceding type, one might wonder why this type of interpretation was not assigned a higher Predictability Rate. It may be surmised that another pragmatic blocking factor is at work, i.e., with few exceptions, there are no cemeteries for dogs and there are, therefore, no professional dog-gravediggers; by implication, there is no need to have a special-purpose instrument for burying dogs.

'A spade for dog's food' has also the same WF Type, with the indirect involvement of 'dog'. The atypical motivating seme functions as an obstacle to this kind of interpretation due to the relevant *extra-linguistic knowledge*.

'A spade by which a dog was killed' – the past tense in the proposal makes this reading too much context-bound and therefore unpredictable. But even if the reading were more general (paraphrased in the present tense) the predictability would hardly be higher. Spades are not designed for the killing of dogs. One could equally propose a number of other similar relations between 'dog' and 'spade', such as feeding (see above), carrying dogs from one place to another, training, etc. And, even more important, any [Tangible] SUBSTANCE of an appropriate size admits the 'killing'-motivated reading. From this it follows that the semantic component [Tangible] is a *too general seme* (level 2) to permit a higher PR. A high *context-dependence* is also a serious drawback to 'a spade that a dog is sitting next to'.

The reading 'a spade in the shape of dog' is identical to that for *baby book* ('a book shaped like baby'). Both are [Pattern – (State) – Patient], and the same comment is applicable to both of them. In fact, the slot reserved for the [Characteristic Shape] (level 4 seme) of spade is filled with an atypical shape of a different [Tangible] SUBSTANCE (implying level 5). This leads us to a conclusion similar to that drawn in the preceding paragraph: any [Tangible] SUBSTANCE may become a Pattern for shaping any other [Artefact]. A dog-like shape of a spade is not its prototypical feature (this being a pointed shape plus a long

handle). As indicated in the case of similar interpretation for *baby-book*, other than typical spade-like shapes are unexpected, and therefore, unpredictable.

By the same token, the low PR of 'a spade for dogs to play with' follows from the fact that almost any [Tangible] SUBSTANCE of a corresponding size may become a toy for dogs. This reading does not follow from the prototypical features of spade functioning as an onomasiological mark.

In the group of non-native informants the reading indicating poor quality of a spade ('a spade in a bad condition/of poor quality') gained support from 5 informants. On the other hand, only two native speakers proposed this reading and assigned it the lowest predictability value (1 point). This may be accounted for by a language interference factor, in particular, a Slovak collocation *pod psa*' (literally 'under dog' = it's bloody awful; it's of very poor quality; in poor condition). While in English there is a structurally identical and semantically similar expression *underdog*, the latter does not seem to be so closely related with the 'quality' feature. This case thus demonstrates the possibility of an interpretation to be influenced by a linguistic factor, notably, the *unequal structuring of the world* by different languages and different *linguistic connotations*.

'A special spade for dog which helps it to dig its own garden' suffers from the *incompatibility of the semes* as accounted for above. The same applies to 'a special spade which helps a dog to protect himself from other dogs' and 'a faithful spade'.

'The shape of dog in the ground which is made with spade' is excluded due to the violation of an *Onomasiological Structure Rule*; for this heavily context-bound reading a reversed sequence of compound constituents is required: *spade-dog* corresponding to the WF Type of [Instrument – (Action) → Result]. Similarly, the reversed ordering of the constituents is required for the reading 'a dog for watching spades' because it is determined by the WF Type [Agent – (Action) → Object] at the onomasiological level. A reversed order also conditions the reading 'a dog with an ugly face': [Pattern – State – Patient].

'An ugly spade' is awkward because the majority of (pet) dogs are lovely and the general attitude to dogs seems to be more positive than negative (even if pitbulls and rottweilers are also dogs!). This contrasts with negative shifted meanings, including 'an ugly woman' and 'a mean, contemptible fellow').

Some other proposals are of a figurative nature, and consequently unpredictable. Thus, 'someone who has to know everything' is perhaps inspired by a 'tracker' dog and the sharpness of the spade tip.

NS – NNS comparison

The results are not so straightforward as in the earlier example of *baby book*. On the one hand, the first two readings swap their positions in the respective groups of the informants. On the other hand, the NS – NNS differences between the PRs of two top readings are significant. What strikes one at first sight is that the PR of 'a spade used for scooping-up a dog's excrement' is higher by 0.231 in the NS group. This in spite of the fact that native speakers may be expected to be aware of the existence of an established synonymous word. By implication, this reading should be blocked for them. It appears, however, that the experience of native speakers with this kind of instrument may have had the opposite effect; it facilitated the interpretation of the newly encountered compound. The lesson to be taken from this case concerns a close interplay between extra-linguistic (cultural) and linguistic factors. In Slovakia there is no tradition for dog-owners to clean excrement left by their dogs in public. Given the absence of the relevant law and/or traditions, there is no need for a special-purpose instrument, and – therefore – no need for a naming unit. Therefore, the above claim about the blocking of the particular interpretation might be modified in terms of culture-related blocking of the NNS groups' interpretation.

Thus, it may be surmised that the native speaker informants – when facing the problematic combination of *dog* and *spade* – clutched at a straw by resorting to their extra-linguistic knowledge and experience – and used a 'pooper-scooper' as a pattern object. The non-native speakers lacked this referential analogy.[4]

Check-group results

Readings	*FO*
A spade used for scooping-up a dog's excrement	9/25
A spade in a bad condition/of poor quality	5/25
A spade in the shape of dog	3/25
A spade for burying a dog	3/25
A species of a dog which looks like a spade	1/25
A tool which digs holes like dogs do	1/25
A spade for a dog to play with	1/25
Dog's tongue	1/25
Dog's paw	1/25

The main group results are buttressed by the check-group. The only deviation from the main group is represented by the 'poor quality' reading. The explanation of this fact is proposed above.

4.2.2.3 *flower hat*

Table 5. flower hat – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a hat with a flower(s) on it	7	9	X		8		8	5	9				9	X	7	8			X	X
– a hat with a flower pattern/design/ornaments				7		4				4	X	9			6					
– a hat made of flowers	8	7	3	8	8		6	6	7	6	9	9	9		6	X	6		6	5
– a hat with a picture of a flower on it																	1	7		
– a hat in the shape of a flower	7										9		6			7		5		
– a special hat for protection of flowers											4	1			3					
– flowers in the garden growing in a way that suggests a hat	6											2								
– a hat to wear when you work in the garden (compare, e.g. 'shopping hat')		4						4		2			5			3	5	6		

Table 6. flower hat – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a hat with a flower(s) on it	X		6	8	8	6		7	X	9	7	9			8	X	5	9		9
– a hat with a flower pattern/design/ornaments	9		4				X	8	X							X		X	7	8
– a hat full of flowers	7						8						8							
– a hat made of flowers		X	9				8		8		5	6		9	9		6		8	8
– a hat that smells good		6													5					
– a hat that has a lot of different colours						4			4	6			5			5				
– a very big hat						5										7				
– curly or funny hair							4		1											
– a hat in the shape of a flower		4			6					6			5				1			
– a flower in the shape of hat										6					6					2
– a hat that has been forgotten in the garden and flowers have grown and blossom on it											2	6								
– a special hat for protection of flowers											1	4								

Type: Animate Plant SUBSTANCE – Inanimate SUBSTANCE
flower
Seme Level 1 – SUBSTANCE
2 – [Animate]
3 – [Plant] [Colour]
4 – [For Decoration] [Light Weight] [Short] [Material] [Immovable] [Grows] [Fragile] [Characteristic Shape] [±Fragrance] [±Used as a Symbol]
hat
Seme Level 1 – SUBSTANCE
2 – [Inanimate] [Tangible]
3 – [Artefact]
4 – [±For Decoration] [±For Protection] [Covering] [Characteristic Shape] [Characteristic Material]

The most predictable reading:
'A hat with flowers on it'

WF Type:		[Stative (=Material) – (State) – Patient]	
Seme Level Combination:		4–4	
	NS	*NNS*	*Total*
FO:	13/20	15/20	28/40
Scores:	110/200	121/200	231/400
PR:	0.358	0.454	0.404
OPR:			0.463

The next three readings:

'A hat made of flowers'			
PR:	0.506	0.237	0.359
'A hat with flower design/pattern/ornaments'			
PR:	0.060	0.171	0.109
'A hat in the shape of flower'			
PR:	0.043	0.028	0.035

Comments
The most predictable readings come down, in principle, to a single WF Type based on the [State – Patient] relation in which *flower* functions as Stative or, more specifically in some cases, as Pattern. Actional readings are rare. This follows from the [Inanimate] seme of *hat* as an onomasiological base (which thus cannot initiate any Action) and from the [Immovable] seme of the Animate onomasiological mark.

The reading 'a hat with flowers on it' activates the seme [+For Decoration] for *hat* and the seme [For Decoration] for *flower*. In this way, we obtain a perfect *match of prototypical semes* which favours predictability. It goes without saying that the [For Decoration] seme of *flower* subsumes other prototypical semes such as [Specific Fragrance] and [Specific Shape] of the particular flowers.

'A hat made of flowers' is motivated by a level 4 seme [For Decoration] and an idiosyncratic level 5 [For Making Hat], on the one hand, and level 4 semes [Characteristic Shape] and [+For Decoration] and a level 5 seme [Idiosyncratic Material], indicating an unusual material used to make a hat, on the other. Consequently, the activated semes indicating the [Purpose] of 'flower' and the [Material] of 'hat' are not their respective prototypical features. This unfavourable situation in terms of predictability is, however, compensated by the *perfect match of the level 4 seme* [For Decoration] which belongs to the prototypical semes of both *flower* and *hat*. An equally important role, boosting the predictability, seems to be played by *analogy* with *straw hat*. In principle, this interpretation is very close to 'a hat with flowers on it'.

'A hat with flower design/pattern/ornaments' is another reading of *flower hat* represented by the [Stative – (State) – Patient] WF Type. In this case, the Stative takes the form of Pattern because – unlike the first reading – this interpretation does not involve actual flowers. A close relation between this and the previous readings is also demonstrated by the identity of activated semes: [+For Decoration] and [For Decoration].

The reading 'a hat that has a lot of different colours' and perhaps also 'a crazy hat' (a mixture of weird colours?) appear to be motivated by a level 3 seme [Colour] in the sense of a range of different colours characterising the class of flowers. In relation to *hat* this seme is activated as 'a multicoloured' hat.[5] The reason for the unpredictability of the colour-motivated readings can be sought in their being secondary in regard of the former three 'main' readings because any of them can entail this latter one.

The same explanation can be applied to 'a hat which smells good' which is secondary to the readings 'a hat made of flowers' and 'a hat with flowers on it'.

In the native speaker group, a relatively high frequency of occurrence is featured by 'a hat to wear when you work in the garden (compare, for example, 'shopping hat')' which has zero occurrence in the NNS group. Its motivation is indicated by an explanation on the part of one of the informants, saying that the interpretation was inspired by *analogy* with *shopping hat*. Thus, the *pragmatics* may explain the non-existence of similar proposals in the NNS group: there is no tradition of wearing 'shopping hats' in Slovakia. Thus, we encounter

an interesting case: an extra-linguistic factor, i.e., *different cultural traditions and habits*, precludes the working of a linguistic phenomenon of analogy. This case further supports the assumption of the interconnection of the multiplicity of factors affecting the meaning-prediction process. It should be, however, noted that the indicated analogy is indirect because, unlike *shopping* there is no activity specified in *flower hat*. This fact may account for the relatively low PR of this reading.

'A hat in the shape of a flower' is activated by a too specific seme (level 5); a hat may be designed in any shape (this being a piece of one's extra-linguistic knowledge). What was said about *baby book* and *dog spade* can be re-iterated here: a flower-like shape of hat is not its prototypical feature (level 4 is 'a covering with a brim and a crown'). By implication, other than typical hat-like shapes are unexpected, and therefore, unpredictable.

'A hat forgotten on the standing water full of water-flowers' and 'a hat that has been forgotten in the garden and flowers have grown and blossomed on it' are too much *context-bound*, and therefore, unpredictable.

Some other readings are used *figuratively*, with the same consequence as entailed by context-boundness, for example, 'a gardener' and 'an exhibitionist'. Still others are excluded due to the violation of an *Onomasiological Srtucture Rule*, for instance, 'a flower in the shape of hat'. 'A hat worn by a girl called Flower' is – like in all other naming and ownership cases – unacceptable from the meaning-predictability point of view. For a similar instance, see below 'a ball with the label (the picture) of hammer' for *ball hammer*.

Interestingly, 'a special hat for flowers in hot weather which protects them' is figurative, however, acceptable because it activates prototypical features of hat, i.e., [For Protection] and [Covering].

NS – NNS comparison

The two topmost readings dominate in both groups, although in reverse order. The next two reading rankings are identical in the two groups.

Check-group results

Readings	*FO*
A hat with flowers on it	9/25
A hat made of flowers	5/25
A hat in the shape of flower	4/25
A 'hat' protecting flowers against winter	2/25
A colourful hat	2/25
A hat worn during gardening	1/25
A group of flowers growing together in the shape of a hat	1/25
A hat full of flowers	1/25

Also in this case, the check-group confirmed the main group results, in terms of both the top position of the two dominating readings and the occurrence of other readings proposed by the informants. A relatively strong position is taken by 'a hat in the shape of flower' which maps the results of the NS group.

4.2.2.4 *ball hammer*

Table 7. ball hammer – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a hammer (the top of) which has the form of a ball	8	7	5	6	8	8	5	7		9	X	X	9	9	4	8		X	8	X
– a (ball + hammer) toy for children to play with; a game				3		4				6	7	6					7		4	
– a (special shaped) hammer used for ball-like components	5							5							6	1			5	
– a hammer like a stick for baseball, cricket, croquet. . .(for sport)				8	6		1						4		6					
– a hammer for hitting a heavy ball in a game of strength		4															3			

Table 8. ball hammer – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a hammer (the top of) which has the form of a ball	X	6	7		8	2	8		7		2		3			5	5	7		6
– a (ball + hammer) toy for children to play with; a game		2	5	5	4	4	6	5	8	1		8		2			1	2		
– a (special shaped) hammer used for ball-like components	6	7								9	8	5			7	2		4	X	X
– a hammer like a stick for baseball or cricket. . .(for sport)				9										4						
– an aggressive person					6	3														
– a soft hammer which does not hurt you when you hit your fingers						5		3	5				2				3			
– a clumsy person with poor dancing skills							8	6	7											
– a climax of a program/show at the ball									5											8
– complete disaster/ embarrassment that happened at the ball							3		6											

Type: Tangible Solid Movable SUBSTANCE – Inanimate SUBSTANCE
ball
Seme Level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Artefact]
4 – [Sphere] [Solid]
hammer
Seme Level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Artefact] [Tool]
4 – [For Pounding] [For Hammering] [Hard] [Characteristic Shape]

The most predictable reading:
'A hammer (the top of) which has the form of a ball'

WF Type:		[Pattern – (State) – Patient]	
Seme Level Combination:		4–5	
	NS	*NNS*	*Total*
FO:	18/20	13/20	31/40
Scores:	141/200	76/200	217/400
PR:	0.635	0.247	0.420
OPR:			0.681

The next three readings:

'A (ball + hammer) toy for children to play with; a game'			
PR	0.065	0.172	0.113
'A (special shaped) hammer used for ball-like components'			
PR	0.028	0.170	0.084
'A hammer like a stick for baseball or cricket. . . (for sport)'			
PR	0.031	0.007	0.017

Comments
The crucial seme of the most predictable reading 'a hammer (the top of) which has the form of a ball' appears to be the level 4 [Sphere] of *ball* because it interacts with (and contradicts) *hammer*'s level 4 seme [Characteristic Shape]. The shape specified at level 4 is vital to this meaning and the same is true of the function (purpose) of this tool. The form of the top of the tool characterises the (range of) possible applications of this particular tool rather than *any* tool. Being spherical is not a prototypical feature of 'hammer'. This means that the level 4 seme [Characteristic Shape] is replaced due to an idiosyncratic shape of hammer (level 5).

The PRs of the other readings are much lower. As a result, the OPR of the top reading is high.

Closely related to this meaning is 'a (special shaped) hammer used for ball-like components'. This reading might be considered an extension of the dominant reading. Also here the shape is of vital importance, one which is somehow adapted to the shape of the objects for which hammer is designed. However, the Purpose (Function) of 'hammer' is foregrounded, which is reflected in the logical-semantic category of Instrument. Therefore, the Object of Action is also included: [Object – (Action) – Instrument]. Actually, the two readings appear to represent two sides of the same coin, which is not to say that they should be taken as a single reading. Their mutual relation can, however, explain the disproportional distribution of the PRs in the two groups of informants with respect to the two readings in question.

The *extra-linguistic knowledge-motivated* train of thought leading to the reading 'a (ball + hammer) toy for children to play with; a game' (PR = 0.113) seems to be the following: hammer-tops are hardly ever round. Therefore, such a hammer is somewhat suspicious in terms of form and application. By implication, some substitute application should be looked for, and the most readily available one is that of a ball-top hammer functioning as a toy. Thus, while in the former two readings, the prototypical features of form and function are vital to the respective readings, here it is the *pragmatics* which significantly co-establishes the reading. This has its impact upon (or follows from) the seme-level of at least the *hammer* constituent. Its function in the game seems to be vague, and perhaps the [Tool] seme rather than [For Nailing] is the motivating one.

Quite a number of the readings proposed are of a *metaphorical nature*, with hammer referring to the special conduct of humans beings or to the specific impact of some Action, for example, 'a clumsy person with poor dancing skills', 'an aggressive person', 'complete disaster/embarrassment that happened at the ball', 'a metal party'. In the former two readings, the *hammer* constituent represents an Agent and the Manner of a specific Action (dancing; violence). The figurativeness and context-dependence makes these readings, in principle, unpredictable.

The meanings with single occurrence include, for example, 'a hammer of low quality', 'a small hammer', 'a hammer of low quality', 'a jumping hammer'. The motivations behind these interpretations are unclear to me.

'A flexible hammer' seems to be a *self-contradiction*. One of the central characteristics of 'hammer' is its hardness.

'A hammer that can be thrown like a ball': since any [Tangible] SUBSTANCE of appropriate size and weight can be thrown (level 2), the unpredictability of this reading appears to be obvious. Its unpredictability is strengthened by the closely related *extra-linguistic factor*: 'hammers' are not designed for throwing.

'A help for women in case of being endangered by men' is, despite its single occurrence, an interesting proposal although *hammer* is used metaphorically referring to the hardness of the tool and the effect of the instrument rather than to the shape. Its unpredictability is increased by the *low register* and *slang* use of *ball*.

'A hammer used for crashing the Christmas balls to powder in order to decorate postcards' is extremely context-bound, and consequently unpredictable at the system level.

'A soft hammer which does not hurt you when you hit your fingers' is an interesting proposal because it, surprisingly, came to mind for several informants. One can only guess that they refer to a round shape of one hammer end, possibly rubber-coated. All in all, this proposal is a *self-contradiction* in the same way as the above-mentioned 'flexible hammer', because hardness is an indispensable feature of hammer.

The reading 'a hammer used for doing something to balls' is too general, and says nothing more than there is some relation between ball and hammer. For a naming unit to be predictable it must somehow 'imply' the Actional relation between the members of the onomasiological structure.

The final proposal to be mentioned is 'a ball destroyed by hammer' which is excluded by an *Onomasiological Structure Rule*. This doubtful meaning may theoretically be attached to the reverse order of ball and hammer. Only the latter order makes it possible to conceive of 'ball' as representing a class of objects, a subset of which is characterised by being destroyed by 'hammer'.

NS – NNS comparison

The points assigned might give an impression that this compound is dominated by one strong meaning. Yet as suggested above, this is not quite so, because the 'Action-focused' meaning of 'a (special shaped) hammer used for ball-like components' can be viewed as a superstructure upon the basic 'Form-focused' meaning of 'a hammer, a part of which has the form of a ball'. This relation follows from the fact that [Characteristic Shape] plays a crucial role in the Actional meaning as well. If we realise this circumstance, the differences between the NS and NNS groups of speakers (0.635 vs. 0.247; 0.028 vs. 0.170) do not seem to be that serious. It may be assumed that while the NS group placed emphasis on the [Characteristic Shape] seme the NNS group inclined to stress

the semes [For Nailing] and [For Hammering], i.e., the application facet entailed by the [Characteristic Shape]. If the frequencies of these two readings are added up, we obtain identical numbers (23:23) for both groups of informants! Given, however, two different WF Types underlying the two readings in question, I decided to treat these two readings separately.

Check-group results

Readings	*FO*
A hammer (the top of) which has the form of a ball	11/25
A hammer like a stick for baseball or cricket . . . (for sport)	6/25
A hammer used to break/crush balls	2/25
A device in bowling to retrieve the ball after the pins have been knocked down	1/25
A tool for hitting a metal part in a system, with a ball 'shot upwards' indicating one's physical strength	1/25
A footballer, or a handball/volleyball, or basketball player scoring a lot of points	1/25
A hammer used for throwing, e.g. in sport	1/25
A soft hammer (not for hammering nails)	1/25
Something very hard that cannot be broken	1/25

The central reading is confirmed unambiguously. The second most frequent proposal occurs twice in the NNS group and does not occur in the NS group. What strikes one is the contrasting connotation of the *ball* constituent. It indicates both 'softness', or better, a lesser degree of hardness compared to a sharp-edged hammer, and a high degree of hardness when the emphasis is laid on the material (metal balls are usually very hard and heavy). As already discussed, the 'soft' hammer reading can hardly be predictable, because it is [Hardness] which is a prototypical feature of 'hammer'.

4.2.2.5 *hill star*

Table 9. hill star – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a person who is brilliant at hill climbing/running/ cycling	8	2	3		6	6					9	5	5	5		6		4		5
– a star that can be seen beyond/above the hill	7	5			4	1					1	3			8	6	1			
– a star shaped like a hill/a star that belongs to a constellation in the shape of a hill	7				1			6				1				1			5	
– an object in the shape of a star situated on a hill	5			2			1	4										6		
– a hill from which the stars are visible very well/from which stars are observed				3					6	2										
– a hill where the famous stars live										1	6			8						
– a hill in the shape of a star				2						1										
– the name of a star			2																7	8

Table 10. hill star – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– an object in the shape of a star situated on a hill	X																		7	
– the brightest star	6							8		6										
– a person who is brilliant at hill climbing/running/ cycling				X	8	5	8		4	7			4	5		1				
– a star that can be seen beyond/above the hill				6			X	6	7				5		8	5	5	7		5
– a star shaped like a hill/a star that belongs to a constellation in the shape of a hill							8		1								2	1		
– a hill in the shape of a star			7	3							3	3								
– a hill which is very high			7											2			3		4	1
– the sun					6	4														
– the first star on the sky					4	3														
– a star that belongs to a constellation in the shape of a hill							8		1								2	1		
– a very famous hill where only the most famous stars can appear											6	7								

Type: Inanimate SUBSTANCE (Location) – Inanimate SUBSTANCE;
Inanimate SUBSTANCE – Animate SUBSTANCE (shifted meaning)

hill

Seme Level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Natural Object]
4 – [Pyramidal Shape] [Elevation] [Natural Point of Orientation] [Covered with Trees, Grass, Shrubs, Rocks, etc.] [Mighty Appearance] [Natural Border]

star$_1$	*star*$_2$
Seme Level 1 – SUBSTANCE	1 – SUBSTANCE
2 – [Inanimate]	2 – [Animate]
3 – [Celestial Body] [±Visible]	3 – [Human]
4 – [Visible at Night]	4 – [Celebrity]
[Characteristic Shape]	[Excellence]

The most predictable reading:
'A person who is brilliant at hill climbing/running/cycling'
WF Type: [Pattern – (State) – Patient]
Seme Level Combination: 4–4

	NS	*NNS*	*Total*
FO:	12/20	9/20	19/40
Scores:	64/200	52/200	100/400
PR:	0.192	0.117	0.152
OPR:			0.521

The next three readings:
'A star that can be seen beyond/above the hill'

PR	0.081	0.160	0.119

'A star shaped like a hill/a star that belongs to a constellation in the shape of a hill'

PR	0.032	0.012	0.021

'An object in the shape of a star situated on a hill'

PR	0.023	0.009	0.015

Comments

The pair of the motivating words *hill* and *star* represent a combination for which it is not easy to find points of unambiguous semantic combinability. The above-given table demonstrates that the informants were at a loss when asked to propose a predictable meaning for this specific combination. Therefore, the PR values of the most predictable readings of *hill star* are rather low.

What might seem at first sight a surprise, that is, the top position of a reading ('a person who is brilliant at hill climbing/running/cycling) motivated by a figurative interpretation of one compound constituents ('star' in the shifted meaning 'a famous and popular human beings'), is not so surprising if one realises that this transferred meaning is frequent in everyday speech. Movie stars, pop stars, sport stars, and other 'stars' and 'starlets' make headlines of newspapers, journals, and other mass-media. This is also reflected in the relatively high number of proposals, ranging from most specific, such as 'Sean Connery' (started his career in a film called *The Hill*), which is a highly specialised context (level 5) to have a higher Predictability Rate, to more general readings like 'a famous actor who lives in Beverly Hills'.

It is important to note that the top position of a reading based on a semantic shift does not impair our predictions concerning the negative influence of figurativeness upon the predictability. The shifted meaning of *star* is a *well-established*, *institutionalised meaning*, and as such, it enters a new naming unit. Such a case must be distinguished from those in which a naming unit is motivated by a shifted meaning which is not established and becomes activated in a new naming unit itself.

In general, it may be assumed that the individual PRs of this kind of reading would have been even higher if the informants had been explicitly asked to concentrate on the shifted meaning of *star*. The strength of this reading of *star* in reference to the compound *hill star* is buttressed by the scarcity of options provided for the semantic compatibility of *star* (original meaning) and *hill*. The weakness of the other, non-figurative, readings is manifested by their very low PRs. The PR of the strongest among them, the second-rank 'a star that can be seen beyond/above the hill' is merely 0.119. The combination of motivating semes for this reading is as follows: for *hill*, it is the level 3 [Natural Object] and the level 4 [Overtops the Surrounding]; the latter semantic component seems to imply another level 4 seme, [Natural Point of Orientation]. In the case of *star* it is unambiguously the [Celestial Body] seme (level 3).

The PRs of the remaining readings are very low (well below 0.100) for various reasons, including the non-established *figurativeness*, *personification*, *context-boundness* ('in tales – the star that sleeps on a hill'; 'a celebrity as big as the 'hill"; 'a lighthouse'; 'a person who had to pass a very hard route to become popular'; 'a famous actor who lives in Beverly Hills' – in this latter reading, the meanings of both *star* and *hill* are shifted); a violation of an *Onomasiological Structure Rule* ('a hill in the shape of star'; 'a very high hill that 'touches' the sky'; 'many stars at one place'); a *combination of reasons* ('a well known hill because of evidence of life in the past' – *star* in the transferred meaning of pop-

ularity is commonly applied to human beings rather than to inanimate objects + too much context-bound + reversed WF Type; 'the only hill in the surroundings' – *star* in the transferred meaning of popularity is commonly applied to human beings + reversed WF Type; 'Noah' – too specialised and personified). Other readings are excluded because *any* word could be used in the sense proposed ('a sportswear trademark'; 'an actress popular in the city of Hill'; 'first and last name of a person').

NS – NNS comparison

While the perception of the readings, based upon the shifted meaning of *star*, was fairly strong among the native speakers, particularly in reference to the meaning 'a person who is brilliant at hill climbing/running/cycling' (PR = 0.192 vs. NNS PR = 0.117), the focus of the NNS group was mainly on the non-figurative reading 'a star that can be seen beyond/above the hill' (NS PR = 0.081 vs. NNS PR = 0.160). Interestingly, the sum of the PRs for the top two readings is almost identical (0.273 vs. 0.277) which means that while these two readings play a central role in both NS and NNS groups, the former preferred the figurative and the latter the literal *star*-based motivation. Under these circumstances, the results of the check-group were anticipated with special interest.

Check-group results

Readings	*FO*
A star that can be seen above a hill	8/25
A very famous person (pop, movie, sport)	7/25
A star-shaped monument/structure on a hill	3/25
A person famous only in her community (village)	2/25
A star which is visible from a hill	2/25
A person very good at climbing	1/25
The Sun	1/25
A beacon on a hill to warn airplanes	1/25
The topmost person in a hierarchy	1/25

These results give support to both of the above-identified meaning-prediction directions and confirm a very strong awareness of the personified meaning of *star*, this time in a more general framework, in which both star and hill are used figuratively: the figurativeness of *hill* means 'a considerable degree'.

The comparison of this reading with the 'two-point' reading of 'a person famous only in her community' exemplifies the possibility of foregrounding two different semes of the same constituent, yielding sometimes opposite meanings. While the seme [Mighty Appearance] is foregrounded in the dominant reading, the feature of 'hill' [Natural Border] is emphasised in the latter.

Consequently, while the former makes the space for *star* almost unlimited, the latter restricts it in an exact way.

4.2.2.6 *apple-juice seat*

Table 11. apple-juice seat – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a seat for drinking apple-juice in a restaurant, bar, etc.	6	5	7				9	3	6	5	4	7	5	8	6	6		7	3	
– a seat in the colour of apple-juice				2	6				5		7		1		5					
– a seat with apple juice spilled on it	5			4			7			X		8				1	4	6		4
– a seat smelling like apple-juice						6														2
– a special seat filled up with apple-juice			1							4	1	2								
– a seat in the shape of an apple								4											5	
– the origin/source of apple juice		2			2															
– a seat on which a carton of apple-juice is placed/a seat with apple-juice next to it			1	1	9	2														

Table 12. apple-juice seat – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a seat for drinking apple-juice in a restaurant, bar, etc.	6	X		9		5	8	8	7	8	7	X	8	9	9	X	5	9	X	X
– a seat in the colour of apple-juice	X	7	6	4		5		5		7	8			9			X		5	
– a seat smelling like apple-juice	6			8				1		2		8			7					
– a seat with apple juice spilled on it		8	3	8		6	3		5	3		X	6			5	9			3
– the place where it's banned to drink alcohol; non-alcohol area			7									4	4							
– a chair in a shape of an apple in sweet shop used by children just for having fun; apple-like shape where one drinks only apple-juice					2			6				5	1		2	2				
– a name of a bar serving all kinds of apple-juice (made of all sorts of apples); a seat in the Apple-Juice Club																	3			3
– a special seat filled up with apple-juice			5						6								1		6	
– a seat on which a carton of apple-juice is placed						3						7				4				

Type: Inanimate Liquid SUBSTANCE – Inanimate SUBSTANCE
apple-juice
Seme Level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Foodstuff] [Having Colour] [Having Taste] [Having Smell]
4 – [Liquid] [Temperance Drink] [Sticky] [Characteristic Yellowish Colour] [Pleasant Smell] [Sweet] [Causing Stains on Cloth]
seat
Seme Level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Tangible] [Artefact] [Furniture]
4 – [For Sitting]

The most predictable reading:
'A seat for drinking apple juice in a restaurant, bar, etc.'

WF Type:		[Object – (Action) – Purpose – Location]	
Seme Level Combination:		4–4	
	NS	*NNS*	*Total*
FO:	15/20	18/20	33/20
Scores:	87/200	148/200	235/400
PR:	0.326	0.666	0.485
OPR:			0.648

The next three readings:

'A seat with apple-juice spilled on it'			
PR:	0.110	0.207	0.155
'A seat in the colour of apple juice'			
PR:	0.039	0.209	0.108
'A seat smelling like apple-juice'			
PR:	0.004	0.048	0.020

Comments
The most predictable reading of the *apple-juice* + *seat* combination is a *level 4* + *level 4* seme combination 'a seat for drinking apple juice in a restaurant, bar, etc.'. The central seme of *seat* is not the level 3 [Furniture]; rather, it is 'a special piece of furniture [For Sitting] designed for a restaurant, bar, café, etc., which implies level 4. The central seme of *apple juice* in the most predictable reading is also a level 4. It is not a mere [Foodstuff] of level 3. The emphasis is on [Temperance Drink]. Interestingly, the *extra-linguistic knowledge* does not work counter to the predictability of this possible naming unit despite the non-existence of apple-juice seats in either European or American cultural setting.

The explanation seems to consist in the compatibility of the two meanings, and, ironically, in the pragmatics. There are a number of soft-drink establishments – so why not one serving apple-juice at a particular seat. This is confirmed by a relatively high PR of this reading.

The level 4 [Liquid], [Sticky], and [Causing Stains on Cloth] and the level 4 [For Sitting] determine the reading 'a seat with apple-juice spilled on it'. One might object that any SUBSTANCE might have apple-juice on it. However, this particular reading conceals a negative *pragmatic implication* which seems to underlie a relatively high predictability of this reading: the fact that apple-juice is sticky and causes stains on cloth. This foregrounds the function of 'seat' [For Sitting] which, in connection with spilled juice, may have unpropitious consequences.

In this case, rather than Action (unlike the topmost reading) the State interpretation seems to be central. Something happened with apple-juice which resulted in an occurrence at a particular [Location].

What is of considerable interest is the seemingly situation-bound meaning: the spilled juice does not remain on a seat forever (hopefully). This fact should work against the predictability of this reading. On the other hand, the semes [Liquid] of *apple-juice* and [For Sitting] of *seat* are perfectly compatible, and imply that any *apple-juice* can be spilled on any *seat* (thus causing inconvenience), which therefore favours the *generalisation* required for a high PR.

'A seat in the colour of apple juice' is a level 4 [Characteristic Colour] + level 4 [For Sitting] combination. As to the extra-linguistic knowledge, there is a strong tendency to relate colours to [Pattern] SUBSTANCES (lemon, orange, navy-blue, violet, raven black, coal black, etc.). Unlike the two more predictable cases, the [Location] seme is backgrounded.

The reading 'a good, pleasant seat where we feel well' is, in fact, the opposite of the negative interpretation concerning spilled juice. It was mentioned by a single informant, and thus its predictability approaches zero. Little wonder, because there is hardly anything in the semantic structure of 'apple juice' which would contribute to the idea of comfortable sitting.

'A seat smelling like apple-juice' (NS PR = 0.004; NNS PR = 0.048; Total = 0.020) is a combination of level 4 [Characteristic Smell] and level 4 [For Sitting], which indicates that the mere level 4 combination need not be sufficient for a high predictability, for example, in cases when the competitive readings are 'stronger'. Moreover, it must be admitted that the 6/20 frequency of occurrence of this reading in the NNS group seems to me surprisingly high because of my extra-linguistic knowledge saying that smell is not the most striking feature of apple-juice among its prototypical features. I realise, however, that sense

perception differs from person to person, and that the sense of smell may be more developed in some people.

'A special seat filled up with apple juice' was proposed by four native and four non-native informants. Apparently, this reading has been inspired by a template, an *analogy base* (to use Ryder's term) *water-bed*. While *water-seat* in the meaning of a seat filled up with water would, no doubt, receive a more significant support, the pragmatics seems to be a crucial obstacle, apparently stronger than the favourable Seme Level Combination.

The other readings have poor results. 'A seat from which you can see a shop where apple-juice is sold' is too much context-bound. The reading 'a seat on which a carton of apple-juice is placed' seemingly resembles that with apple-juice spilled. The reality is, though, different. While the sticky and stain-causing qualities of spilled apple-juice have far-reaching consequences and are not unusual, thus allowing for generalisation, the case with apple-juice cartons seems to be highly *situation-bound*.

The idea of having seats with apple-juice drawn was espoused by only a single informant. In my view, the reason consists in this reading roughly falling within the scope of the top-most interpretation, i.e., the function of objects tends to be symbolised by their respective drawings.

The reading 'a seat in the shape of an apple' apparently missed the point because the 'juice' constituent is paid no regard to, and (as aptly remarked by L. Lipka – personal communication) juice has no [Shape]!

All in all, *apple-juice seat* features one reading with clearly dominant predictability, and two other readings above 0.1. All of them are level 4–4 combinations. The low predictable or next-to-zero predictable combinations are mostly characterised by unfavourable Seme Level Combinations. The extra-linguistic factors seem to have strong PR-reducing effects in some cases which eliminate the 'advantage' of a favourable Seme Level Combination.

In conclusion, I would like to return to P. Downing who introduced this compound in her seminal research (1977). *Apple-juice seat* was presented in a situation-bound meaning of a 'seat in front of which a glass of apple-juice' [is] placed' (1977: 818) as a personal experience of a friend of hers. It is paradoxical that no such reading has been proposed by any of my native or non-native informants! Downing's meaning of *apple-juice seat* might be perhaps subsumed in the most productive reading of my research, that is, 'a seat for drinking apple juice in a restaurant, bar, etc.'. Nevertheless, there is a significant difference between these two readings. Downing's one is *context-bound*, *deictic*, while 'my' reading is *generalised*, a *system-level* reading.

The scope of the readings proposed by the informants, and primarily, the predictable ones, calls into question Downing's assumption that

> the lexicalization potential of such a compound [i.e., in the deictic meaning referred to by Downing'] seems to be quite low, since it is based on a relationship of a very temporary, fortuitous nature, and such states are generally not considered name-worthy by the community at large (1977: 822).

While it is true – as stated in Chapter 1 – that naming units tend to express permanent rather than context-free relations, there does not seem to exist any pragmatic, user-oriented obstacle which would prevent a generalisation of the 'deictic', context-bound circumstances and relations. This, in fact, has been demonstrated by the most productive reading proposed by my informants: 'a seat for drinking apple juice in a restaurant, bar, etc.' is in fact the generalised variant of Downing's reading, that is to say, 'seat in front of which a glass of apple-juice is *usually* placed'.

NS – NNS comparison

In both of the groups, the reading 'a seat for drinking apple-juice in a restaurant, bar, etc.' clearly dominates, and is very strong especially in the NNS group (0.666).

While 'a seat with apple-juice spilled on it' and 'a seat in the colour of apple juice' are almost equally acceptable to the non-native speakers (0.207 and 0.209, respectively) and their PRs are relatively high for a 2nd and 3rd rank readings, the values of these readings in the NS group are lower (0.110 and 0.039, respectively).

Compared to the other naming units of this Onomasiological Type, the number of proposed readings for this compound is relatively low (15 : 17), indicating that the informants did not face serious problems in identifying the meaning of this possible compound.

Check-group results

Readings	*FO*
A seat in a bar/restaurant/party where apple-juice seat is served	12/25
A seat which is of the apple-juice colour	5/25
A seat where apple-juice was spilt	2/25
A very comfortable/cozy seat	2/25
One of the seats in a pub/restaurant reserved for children	2/25
A very soft seat	1/25
A place where apple juice is produced	1/25

The check-group bears out the dominant position of the reading 'a seat for drinking apple-juice in a restaurant, bar, etc.'. With regard to the other two readings, it favours the way non-native speakers perceive this compound.

4.2.2.7 *game wheel*

Table 13. game wheel – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a wheel for playing roulette and casino games; a wheel in the Wheel of Fortune Type games	5	6	X	8		6			9		8		8	X	9	X	5	5		9
– a wheel which is a part of a game equipment, a wheel with which a game is played	5	1			X		2	X		8	X	9	9	X		3			X	
– a toy for babies; a wheel for children to play with										2	9		9						7	
– a never-ending game/a cycle of several games						5				3	5									
– a turntable with a selection of games	9									5										
– a wheel made of games										3									4	

Table 14. game wheel – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games		6	9	9		8				8	6	X		8		8			5	8
– a toy for babies; a wheel for children to play with		5	3						8		8	9	3	9				7	6	
– a wheel which is a part of a game equipment, a wheel with which a game is played					7	4		5	9	9	8		4	9	5	7	6		3	
– a never-ending game/a cycle of several games	X			7												5				
– a wheel you need to play the computer games; a part of TV game console	9	7																		
– a game played in the circular motion; a game in which children change their positions in clock direction; a game in which the players stand in the circle				6			7									4				
– a game wheel – one's life; circle of life		3											1							

Type: PROCESS – Inanimate SUBSTANCE

The classification of this naming unit is ambiguous. It may be classed with the other primary compounds falling within Onomasiological Type 3 because of the functional and structural analogy of *game* with cases of Onomasiological Type 1, such as [Object ← Action – Instrument], for example, *game-indicating wheel* or *game-selection wheel*, where *game* assumes the same polar position in the onomasiological structure as in the *game wheel*. In the examples given, it functions as an Object of Action, with the Object being Process (a game is a Process). In this case *game* functions as a determining element of the onomasiological mark.

On the other hand, an Onomasiological Type 2 of *game wheel* can also be understood as a structural and functional analogy with an Onomasiological Type 1 in which *game* stands for the specified element of the determined constituent of mark, as in the case of *baby game wheel*.

game

Seme Level 1 – ACTION
2 – [Process]
3 – [Amusement] [Competition] [Game]
4 – [Specific Game], [Instruments Used], [Rules], [Players]

wheel

Seme Level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Artefact]
4 – [Round] [Turning Motion]

The most predictable reading:

'A wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games'

WF Type:		[Process – Instrument]	
Seme Level Combination:		4–4/5	
	NS	*NNS*	*Total*
FO:	14/20	11/20	25/40
Scores:	108/200	85/200	193/400
PR:	0.378	0.234	0.302
OPR:			0.490

The next three readings:

'A wheel which is a part of a game equipment, a wheel with which a game is played'

PR:	0.261	0.228	0.245

'A toy for babies; a wheel for children to play with'

PR:	0.027	0.131	0.069

'A never-ending game/a cycle of several games'

PR:	0.010	0.017	0.013

Comments

The Seme Level Combination in the top reading is a *level 4 + level 4* combination because a specific type of game is identified and the feature of the turning motion of 'wheel' is vital to this game. A wheel used in this type of game requires specific marking, such as a pointer, i.e., certain individualising features, which point to level 5. This type of reading owes its (for primary compounds) relatively high PR to some *extra-linguistic factors* – a number of (so popular) TV programmes of the Wheel of Fortune type and the large number of gambling houses.

'A wheel which is a part of a game equipment, a wheel with which a game is played' is a more general variant of the preceding, more specific, reading, that is to say, no specific game is referred to in this reading. As such, it is a 3–4 Seme Level Combination (the turning motion as a prototypical semantic component of *wheel* is of primary importance in this reading, too).

The reading 'a toy for babies; a wheel for children to play with' is another variant of the WF Type [Process – Instrument']. While the second-rank reading provides no details of the game to be played, and the top-rank reading restricts the meaning in terms of a specific group of games, this reading restricts the meaning of *game wheel* in terms of 'players' of the game: [(Agent) – Action – Instrument]. This specification of 'game' implies activation of a level 4 seme.

As far as the very low PR readings are concerned, 'a game wheel – one's life; a circle of life', 'a never-ending game; game that is played again and again', and 'regularly repeating events' are *too metaphorical* to claim a higher predictability level.

Two readings, 'a deciding point in a game' and 'an action to entertain someone or to 'break the ice' missed the point because 'wheel' and its rotary motion imply continuity while the two readings proposed imply an abrupt change in the course of things, i.e., the meaning of 'turning point'.

'An exercise-equipment used by gymnasts' does not refer to a game – gymnastics can hardly be labelled a game. The same holds of 'a training machine for future car drivers'. Driving a car is not, in its principle, a game.

As with the other sample compounds, the 'nick-name' and 'name' interpretations are eliminated due to their universal applicability to almost any primary compound and even any substantival naming unit.

It may be concluded that this primary compound does not offer a wide range of options, which may be considered to be another factor which strengthens the position of the top two readings of *game wheel*.

NS – NNS comparison

This compound favours the initial hypothesis. The results of the two groups are fairly similar, especially in terms of the identical ranking of the first four readings. The differences between the individual PRs are not big either.

The limited number of possible readings is manifested in the two groups differently: as a very low number of proposed readings in the NS group, and as a fairly high number of weird single-occurrence readings in the NNS group.

Check-group results

Readings	*FO*
A wheel used in a game (e.g. for choosing a player to take turn; throw-and-catch, etc.)	11/25
A wheel in the game of roulette/fortune's wheel, etc.	8/25
A wheel used in a game instead of dice to decide the number of moves, etc.	2/25
A kind of a 'merry-go-round' for grown-ups	1/25
A circle formed by game participants	1/25
A game played without interruption	1/25
A baby toy	1/25

The results indicate that the two main groups more or less coincide in their respective assessments of the predictability of the two most predictable readings. Their respective positions in the check-group are exchanged. Still, the two readings clearly dominate.

4.2.2.8 *shape cloth*

Table 15. shape cloth – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a cloth cut into a particular shape	6		1	3	2	2			5		7	1	1			4			8	
– elastic cloth shaping woman's figure, very tight clothes shaped by the form of the body				8		5						6		7	8					
– a stencil/pattern for making shapes on material			6								6									
– a cloth for using when cutting out shapes to make clothes with (like a cloth template)																	1		6	
– a cloth with different shapes on it		7			8		5													
– an unstretchable cloth – does not loose its shape											3		4							

Table 16. shape cloth – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– a cloth cut into a particular shape	X								8		8		7		1	7			5	
– elastic cloth shaping woman's figure, very tight clothes shaped by the form of the body					7	X	8				9	2	3	X		5	8			8
– homeless person, poor man – poor dressed; shabby dressed		4	7												1					
– universal cloth												X	4						3	
– a mummy; cloth used for mummification		2										3								
– something in abnormal, unusual shape or extraordinary cloth				8										3				1		
– a contemporary fashion hit – skirts, dresses which have an unusual shape					9			1											1	
– a cloth with different shapes on it	6							9												
– item with a shape of cloth									6							4				

Type: Intangible Form – Inanimate SUBSTANCE
shape
Seme Level 1 – QUALITY
2 – [Geometric Feature] [Form]
3 – —
4 – —
cloth
Seme level 1 – SUBSTANCE
2 – [Inanimate] [Tangible]
3 – [Artefact] [Material]
4 – [±Elastic] [±For Clothes]

Most predictable reading:
'Elastic cloth shaping woman's figure, very tight clothes shaped by the form of the body'

WF Type:		[Stative (=Quality) – (State) – Patient]	
Seme Level Combination:		2-4	
	NS	*NNS*	*Total*
FO:	5/20	10/20	15/40
Scores:	34/200	70/200	104/400
PR:	0.043	0.175	0.098
OPR:			0.476

The next three readings:
'A cloth cut into a particular shape'
PR: 0.110 0.081 0.097
'A cloth with different shapes on it'
PR: 0.015 0.008 0.011
'A contemporary fashion hit – skirts, dresses which have an unusual shape'
PR: 0.0003 0.008 0.003

Comments
The PRs of *shape cloth* readings are in general very low as a result of a *limited combinability of the semantic structures* of the two motivating words. Moreover, the top readings combine the meanings of *cloth* and *clothes*. Since this happened in both NS and NNS groups of informants one may surmise that this 'misinterpretation' did not result from a lack of linguistic knowledge; rather, it was a way out of the situation they faced. In fact, *cloth* and *clothes* are in the relation of 'material-product'.

Out of the readings based on the non-shifted meaning of the motivating words, the most predictable is 'a cloth cut into a particular shape'. With one exception the scoring assigned to this reading by the non-native informants is fairly high (the average score is 6.6 pts) and rather scattered in the NS group (ranging from 1 to 8 pts.). Quite surprisingly, the PR of this reading is rather low in both of the groups. In the non-native group the advantage of high average score is reduced by a small number of occurrences. In the NS group the situation is quite the opposite: a fairly high number of occurrences (11) are played down by the low scores assigned. All in all, its PR is 0.110 % for the NS and 0.081 for the NNS.

There is one more reading which gained – contrary to my expectations – very low support: 'a cloth with different shapes on it'.[6]

The apparently *limited combinability* of the two semantic structures is also demonstrated by a fairly high number of proposals of vague motivation. Furthermore, there are quite a lot of *figurative/personified readings*, such as 'a person who can be easily influenced, persuaded and abused', 'something in abnormal, unusual shape, or extraordinary cloth', 'somebody who tries to be perfect in everything', 'somebody very eccentric in dressing', 'a place where you can shape cloths', and 'an instrument used for shaping clothes'. Others, like 'a table cloth which can be adjusted to different shapes of table', are much too *context-bound*. The reading 'cloth manufactured by 'Shape' company' is unacceptable for reasons previously discussed. Finally, 'a figurine for cloth makers' is excluded for a combination of reasons, including *violation of an Onomasiological Structure Rule* and *transferred meaning*.

NS – NNS comparison:
The first two readings swap their ranking in the two groups of informants. The generally low PRs of the top readings do not favour significant NS-NNS differences. In general, the number of readings proposed in the NS group is low, and an extremely high number of questionable proposals and single-occurrence readings in the NNS group. All of these facts indicate the problems the informants must have faced when trying to come up with some reasonable readings for this compound.

Check-group results

Readings	*FO*
Elastic cloth shaping woman's figure	7/25
A cloth cut into a particular shape	5/25
A piece of cloth used as a template	4/25
Something like scissors	2/25
A fashionable cloth	2/25
A machine of the future to adjust a kind of clothing for any person so that it may fit perfectly	1/25
A cloth with various-shaped ornaments on it	1/25
A piece of material ready-cut for sewing a shirt, dress, etc.	1/25
A kind of cloth which keeps its shape for a long time	1/25
A fabric from which cloths of different shapes are made	1/25

In principle, the check-group bears out the main groups' results. Interestingly, while the 'template' reading occurred four times in the check-group it only occurred twice in the NS and did not appear at all in the NNS group.

4.2.2.9 *garden whisky*

Table 17. garden whisky – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– whisky served in the garden (on various occasions for example at parties, barbecues, etc.)		7	6	8	X	4	3		8		X	8	2	6	8	5		5	4	5
– cheaper, low-quality, home-made whisky	8	2								2	7		6	X		8	1	4		7
– whisky made from garden products	4				9			3		6		3	6		5	6	8		9	
– bad advice about gardening										2										
– whisky made, produced, distilled in the garden				2		2													7	
– a brand name of whisky		5	5																	

Table 18. garden whisky – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– whisky served in the garden (on various occasions for example at parties, barbecues, etc.)				6	X	5	8		7		8	9	4	7		7	7		X	
– cheaper, low quality, home-made whisky	X	7	8				7									5		6	5	
– whisky made from garden products		6			4	7		X	5	3	4	X	2	6	X	3	2	8		9
– whisky made, produced, distilled in the garden						1		1		7		8	1	3			5			
– the name of a bar, pub		8	3																	
– it's something like tea time but it is time for whisky				3							5									
– a person who prefers drinking whisky in the garden				8									3		3					
– whisky stored in the garden					4								3							
– garden full of whisky bottles		3														6				
– whisky stolen from, found, hidden in the garden						1						7				8				
– a brand name of whisky									9	X										
– whisky buried in the garden to be better	3														6					

Type: CONCOMITANT CIRCUMSTANCE of Location – Inanimate SUBSTANCE

garden

Seme Level 1 – SUBSTANCE
2 – [Inanimate] [Location]
3 – [Plot]
4 – [±For Gardening] [±For Relaxation] [±Fruit Trees] [±Vegetable] [±Flowers] [±Shrubs]

whisky

Seme Level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Artefact] [Drink]
4 – [Strong Liquor] [Made of Grain]

Most predictable reading:

'A whisky served in the garden (on various occasions, for example at parties, barbecues, etc.)'

WF Type:		[Location – (Process) → Theme]	
Seme Level Combination:		4–4	
	NS	*NNS*	*Total*
FO:	16/20	12/20	28/40
Scores:	99/200	88/200	187/400
PR:	0.396	0.264	0.327
OPR:			0.490

The next three readings:
'A whisky made from garden products'
PR: 0.148 0.334 0.231
'A cheaper, low quality, home-made whisky'
PR: 0.138 0.084 0.109
'A whisky made, produced, distilled in the garden'
PR: 0.008 0.046 0.023

Comments
The most central seme of *garden*, activated in the top reading, is the *level 4* [For Relaxation]. The central semes of *whisky* seem to be all the *prototypical features* of 'whisky': it is not *any* (alcoholic) [Drink] that is referred to. The PRs are relatively high for Onomasiological Type 3. It may be assumed that the *extra-linguistic knowledge* of garden parties may have contributed to the predictability of the most productive reading. There is one more reading which, perhaps, might be included here, in particular, 'whisky drunk by epicures'. This reading heavily relies on the seme [For Relaxation]. The single occurrence of the latter reading may be explained by its too specialised meaning which the former, much more predictable meaning avoids.

Garden as an onomasiological mark in 'whisky made from garden products' is used metonymically, referring primarily to 'garden' as a [Source] of fruit used for the production of 'whisky' rather than to [Location]. Therefore, its conceptual category of CIRCUMSTANCE is replaced by SUBSTANCE ([Material]). In other words, rather than on a [Plot] located next to a house or in a 'garden area', the informants put emphasis on [Fruit Trees] and [Shrubs].

While these level 4 semes of *garden* match the level 3 seme [Drink] of *whisky*, I am not sure whether they fit the defining semantic components at level 4, i.e., [Strong Liquor] and [Made of Grain]. Rather, they appear to contradict the prototypical features of whisky. Put differently, the question is whether whisky can be made from garden fruit. The deviation from the expected raw-material (rye, wheat, corn, or barley), which is not grown in gardens, may indicate that the informants (no doubt, the vast majority of them being laymen in terms of whisky production) were tempted by a perfect match of the combination of [Raw-material] and the final [Product].

While *analogy-base templates* usually have positive effects upon the predictability of naming units, this case seems to provide an opposite result: here we face *negative interference*, perhaps under the influence of a semantic analogy base represented by wine made from garden fruit. A fairly high frequency of this reading indicates that quite a lot of the informants succumbed to this

temptation. This reading is a typical example of the role played by the (*lack*) *of extra-linguistic knowledge* in the predictability of naming units.

The same comment may also accompany the reading 'cheaper, low quality, home-made whisky'. The only difference consists in the fact that the informants went one step further in their conceptual analysis (indicated in my previous remarks), once again under the influence of a negative interfering analogy base (a cheap wine made from garden products). Thus, the activated semes [Fruit Trees] [Shrubs] simultaneously point out the quality of whisky. [Strong Liquor] does not seem to be activated in this reading as suggested by the attributes 'cheaper', 'low quality'. The [Location] seme representing [Source] simultaneously entails the [Quality] seme, in particular [Low Quality]. This reading might be perhaps completed with the closely related 'soft whisky – with lower percentage of alcohol' which gained two fives. Interestingly, for one non-native informant the quality of a garden-fruit-made drink is associated with positive associations ('whisky of a very high quality')!

Out of the remaining, fairly numerous proposals, some are excluded due to an *inappropriate onomasiological structure*, for example, 'vegetable meal with a little whisky', and 'very good fruit for somebody who likes this fruit as much as whisky'. These readings seem to require a *reversed ordering of the constituents*.

The reading referring to a brand-name of whisky, to the producer of whisky, a garden called *Whisky*, 'whisky made in the city of Garden – future city', and the 'nickname' reading can be applied to any primary compound. They represent *disadvantageous Seme Level Combinations*, which automatically reduces the PR to a minimum. This also applies to 'a bottle of whisky given to someone for taking care of your garden' which combines level 3 conceptual categories (one can be remunerated for his/her work in any way, by both concrete and abstract SUBSTANCES) and (Work on a [Plot]).

Finally, one should mention a series of proposals relating whisky to the [Location] seme of *garden*: whisky stored, made, produced, distilled, found, hidden in, stolen from the garden. While all of them are possible and acceptable in terms of word formation compatibility of the meanings of the two constituents they were overridden by the above mentioned 'whisky served in the garden', possibly as a result of the *experience* of the informants.

NS-NNS comparison

While in the NS group the PRs of the second and the third readings are almost identical (0.148 and 0.138), in the group of non-native speakers a similar agreement can be found between the first and the second readings (0.264 and 0.334). The biggest difference in the PRs concerns the reading 'a whisky made

from garden products) which is rank 1 in the NNS group and rank 2 in the NS group. If the total PRs are taken into consideration the first four readings are almost evenly graduated by steps per about 0.1.

Check-group results

Readings	*FU*
Whisky served in a garden (during parties, etc.); such a party	20/25
Whisky produced from fruit growing in a garden	3/25
Whisky of low quality	2/25

Clearly, the number of different readings proposed by the check-group for this compound is lowest of all sample naming units in Experiment 1, and corresponds with the results of the main groups. Moreover, the dominance of the first of the above readings is unprecedented. Such a result indicates a very good meaning predictability, and maps the results obtained from the NS group.

4.2.2.10 *age bag*

Table 19. age bag – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– an old bag, a bag that looks old			7	9		6	8		7	8		5					5			
– a bag with the number on it indicating a person's age	5				6			4	4		6									
– a bag under eyes that signifies tiredness, exhaustion and possibly age; circles underneath the eyes caused by age											8		7							8
– something which hides age (cosmetics)										1				7					8	
– a bag documenting the story of one's life – a bag full of things indicating the age of the owner; a file with age-relevant data		5														6			2	
– a collection of anything for a specific age group												4								6
– a special bag designed for a specific age category	6										5		1		2	1				
– an (grumpy) old woman/a group of old women								5		1			3							

Table 20. age bag – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– an old bag, a bag that looks old	9	8	3	9	9	4			7	4			1	X	4	4	6	5	3	
– a special bag designed for a specific age category		4				9	6			5	4		8				7	1	2	
– a group of people of the same age, same generation/ group of elderly people			9	9			7		3			5			3	2				
– a bag made of good material which survives ages; long lasting								5					6	9				3	4	
– a bag documenting the story of one's life – a bag full of things indicating the age of the owner; a file with age-relevant data		4							3				1		2		5			
– a man who thinks in old way, old opinions, old-fashioned views		3		8											2					
– a bag under eyes that signifies tiredness, exhaustion and possibly age; circles underneath the eyes caused by age							8					9				6				8
– a sum of all properties which characterize old people	X										8									
– a very fashionable bag								5								5				

Type: CONCOMITANT CIRCUMSTANCE of Time – Inanimate SUBSTANCE

age

Seme Level 1 – CIRCUMSTANCE
2 – [Time]
3 – [Stage in Life] [Impact of Time]
4 – [±Ugliness] [±Shabbiness] [Old-Fashioned]

bag

Seme Level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Artefact] [Appearance]
4 – [±For Decoration] [±For Shopping] [Container] [±High-Quality]

The most predictable reading:
'An old bag; a bag that looks old'

WF Type:		[Temporal/Quality – (State) – Patient]	
Seme Level Combination:		4–3	
	NS	*NNS*	*Total*
FO:	8/20	15/20	23/40
Scores:	55/200	86/200	141/400
PR:	0.110	0.323	0.203
OPR:			0.744

The next three readings:
'A special bag designed for a specific age category'

PR:	0.019	0.104	0.053

'A group of people of the same age, same generation/group of elderly people'

PR:	–	0.067	0.017

'A bag documenting the story of one's life – a bag full of things indicating the age of the owner; a file with age-relevant data'

PR:	0.010	0.019	0.014

Comments

Although *old bag* is a slang expression for an unattractive woman, in the case of the most predictable reading of *age bag* the interpretation 'old bag' may be supposed to be meant literally rather than figuratively. This is borne out by the second part of the paraphrase ('a bag that looks old').

While the *age* constituent clearly falls within the conceptual category of CIRCUMSTANCE, in this particular meaning the [Temporal] seme rather implies [Quality] characterising the [Patient], i.e., 'bag'.

The Seme Level Combination is 4–3; the level 3 [Impact of Time] is negatively specified at level 4 by [Ugliness], [Shabbiness] and/or [Old-Fashioned] for *age* while the [Artefact] seme of *bag* does not need any further specification. The indicated *age* semes comply with the [Appearance] of *bag*.

A not very high PR of the top reading, and the extremely low PRs of all the next readings (0.053; 0.017; 0.014) might suggest that the combination of [Temporal] CIRCUMSTANCE and SUBSTANCE bodes ill for meaning predictability. However, this is not true. The Predictability Rates of primary compounds like *morning tea* or *evening paper* would be, no doubt, much higher.

The second rank among the NS informants is taken by 'a bag with the number on it indicating a person's age' (PR = 0.031); among the non-native speakers it is 'a special bag designed for a specific age category' (PR = 0.104). In these two readings the [Temporal] category *age* is personified, and refers,

respectively, to individuals and a group of people of a particular age. The activating seme of *bag* in these readings does not seem to be any of the level 4 semes. The [Purpose] is clearly backgrounded, and the idea of [Artefact] in general is sufficient for this interpretation.

One additional proposal falls semantically within this reading, i.e., 'a bag for all generations'. It was, however, treated separately, because the same informant assigned the same value to both of the readings in question.

While in the preceding cases it was only the left-hand constituent that was interpreted metaphorically, the reading 'a group of people of the same age, same generation/group of elderly people' is an instance of a semantic shift of both constituents: *age* is perceived as 'people of a particular age' and *bag* is interpreted figuratively as 'a group' of people. Therefore, the very low PR of this reading (0.017) is not surprising.

The WF Type of the reading 'a bag made of good material which survives ages; long lasting' is the same as in the case of the 'old bag' reading. In both of them, the [Temporal] seme has some Quality-related semantic consequences. While the maximum possible SLC activates the semes [Impact of Time] and [High-Quality], the figurative use of *age* appears to work against a higher PR.

One more group of readings deserves comment: 'a sum of all properties which characterize old people' and 'a bag documenting the story of one's life; a bag full of things indicating the age of the owner; a file with age-relevant data'. In principle, these meanings refer to a single underlying idea of 'all the features and stuff characterising one's life'. This kind of reading is a genuine [Temporal] CIRCUMSTANCE + SUBSTANCE reading. The *metaphoric* nature of both of the constituents, however, seems to work against the predictability of this reading.

Out of the numerous single/double occurrences, I would like to draw the reader's attention to the 'hunch' meaning and the 'fat on one's body' meaning. The former seems to me to be *pragmatically* much more relevant because elderly people tend to walk hump-backed. Fatness is not age-bound.

The reading 'a period when bags were/are popular' must be rejected due to a relevant *Onomasiological Structure Rule*. This reading corresponds to the reversed order of the two constituents, i.e., *bag-age*.

That figurativeness may bring contradictory interpretations is borne out by accounting for *age bag* as either 'a disgusting bag' and 'an old bag, a bag which looks old' or 'a very fashionable bag'. Furthermore, these two readings are examples of *foregrounding* and *backgrounding* (cf. Cruse 1986) processes at the system level. The notion of 'age' is a [Temporal] category of a very wide comprehension. Thus, while in the former reading the emphasis is laid on one of the negative effects of the flow of time the latter reading figuratively puts

emphasis on one point in the flow of time. Moreover, while 'a disgusting bag' highlights the negative effects of time flow, 'a bag made of good material which survives ages' foregrounds its positive effects.

NS – NNS comparison
In both of the groups there is one central reading, although the 'centrality' of the NNS's top reading is much stronger.

Check-group results

Readings	*FO*
One's life experience	10/25
A very old, worn-out bag	6/25
Collection of souvenirs/old items concerning one's life	5/25
(Health and other) problems accompanying elderly people	2/25
A bag suitable for people of a particular age	1/25
A bag that resembles a bag of the past	1/25

In this case the results do not fully comply with those of the main group. The top reading of the main group takes the second position in the check-group. The metaphorical meaning 'one's life experience', which found almost zero support in the main group of informants, roughly corresponds to the reading 'sum of all properties which characterize old people'. This reading has three occurrences in the main group.

4.2.3 Summary 1

The results obtained in Experiment 1 are summarised in Tables 21 to 25.

4.2.3.1 *Native speakers vs. non-native speakers*
Table 21 summarises the data concerning the PR similarities/differences between the native and non-native groups of speakers, and Table 22 compares the rankings of the most predictable readings in the native and non-native groups of informants (including the check group).

Table 21. Differences in PRs between NS and NNS groups

Naming unit/reading	*Difference in PRs*
baby book	
1. A book for babies (fairy tales, rhymes, pictures, etc.)	0.048
2. A book about babies and how to take care of them	0.055
3. A book with photos of one's baby(ies)/album; with records of baby's development	0.003
4. A (very) small book	0.054
dog spade	
1. A spade used for scooping-up a dog's excrement	0.231
2. A spade in the shape of a dog	0.053
3. A spade for dogs to play with	0.018
4. A spade in a bad condition/of poor quality	0.025
flower hat	
1. A hat with flowers on it	0.096
2. A hat made of flowers	0.269
3. A hat with flower design/pattern/ornaments	0.111
4. A hat in the shape of flower	0.015
ball hammer	
1. A hammer, a part of which has the form of a ball	0.388
2. A (ball + hammer) toy for children to play with; a game	0.107
3. A (special shaped) hammer used for ball-like components	2.142
4. A hammer like a stick for baseball or cricket…(for sport)	0.024
hill star	
1. A star that can be seen beyond/above the hill/that touches the hill	0.079
2. A person who is brilliant at hill climbing/running/cycling	0.191
3. An object in the shape of a star situated on a hill	0.020
4. A star shaped like a hill/a star that belongs to a constellation in the shape of a hill	0.014
apple-juice seat	
1. A seat for drinking apple juice in a restaurant, bar, etc.	0.340
2. A seat with apple-juice spilled on it	0.097
3. A seat in the colour of apple juice	0.170
4. A seat smelling like-apples	0.044
game wheel	
1. A wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games	0.144
2. A wheel which is a part of a game equipment, a wheel with which a game is played	0.033
3. A toy for babies; a wheel for children to play with	0.104
4. A never-ending dame/a cycle of several games	0.007
shape cloth	
1. Elastic cloth shaping woman's figure, very tight clothes shaped by the form of the body	0.132
2. A cloth cut into a particular shape	0.029
3. A cloth with different shapes on it	0.007
4. A contemporary fashion hit – skirts, dresses which have an unusual shape	0.008

Table 21. (*continued*)

garden whisky	
1. A whisky served in the garden (on various occasions, for example at parties, barbecues, etc.)	0.132
2. A whisky made from garden products	0.186
3. A cheaper, low quality, home-made whisky	0.054
4. A whisky made, produced, distilled in the garden	0.038
age bag	
1. An old bag; a bag that looks old	0.213
2. A special bag designed for a specific age category	0.085
3. A group of people of the same age, same generation/group of elderly people	0.067
4. A bag documenting the story of one's life – a bag full of things indicating the age of the owner; a file with age-relevant data	0.018

Table 22. Comparison of the top rankings of NS and NNS

	Ranking		
	NS	*NNS*	*Check-group*
baby book			
'A book for babies (fairy tales, rhymes, pictures, etc.)'	1	1	1
dog spade			
'A spade used for scooping-up a dog's excrement'	1	2	1
flower hat			
'A hat with flowers on it'	2	1	1
ball hammer			
'A hammer, a part of which has the form of a ball'	1	1	1
hill star			
'A star that can be seen beyond/above the hill/that 'touches' the hill'	2	1	1
apple-juice seat			
'A seat for drinking apple juice in a restaurant, bar, etc.'	1	1	1
game wheel			
'A wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games'	1	1	2
shape cloth			
'Elastic cloth shaping woman's figure, very tight clothes shaped by the form of the body'	2	1	1
garden whisky			
A whisky served in the garden (on various occasions, for example at parties, barbecues, etc.)'	1	2	1
age bag			
'An old bag; a bag that looks old'	1	1	2

Tables 21 and 22 indicate that the hypothesis postulating similar acceptability of the individual readings to the NS and NNS groups of informants has been borne out:

1. In general, the two groups of informants *coincide in identifying the same predictable readings*. This agreement is striking especially when we compare the four most predictable readings for each of the sample naming units.
2. Even more striking, these readings are assigned *the same PR ranking* in a large number of cases. In 25 out of 40 possible cases, including four most predictable rankings of each sample naming unit, there is an absolute match between the groups. With three naming units (*baby book*, *ball hammer*, and *game wheel*) the agreement is complete: the ranking of all the four most predictable readings is identical in both groups of informants. With other naming units the individual positions are swapped (single-rank difference). The top reading ranks agree in five cases; in the other cases, the difference is one rank. The 'inconsistencies' are mostly insignificant, resulting from very slight PR differences. Thus, for example, the ranking of the second and the third rank readings for *apple-juice seat* is reversed, however, the PR Gap between the rank 2 and rank 3 readings in the NNS group is merely 0.002.
3. The PR differences for the top rank readings are significant in two cases (they exceed the value of 0.25); what is, however, interesting is that in neither of these cases does a significant PR difference have any effect upon the ranking agreement: with *ball hammer* the difference of 0.388 does not have any influence upon the top position of the reading 'a hammer (the top of) which has the form of a ball' in both of these groups of informants; and the 0.340 PR difference in the case of *apple-juice seat* concerns the rank 1 reading ('a seat for drinking apple juice in a restaurant, bar, etc.') in both groups, too. In six of ten cases the PR difference between the top readings is under 0.15. There seem to be only two cases in which the *differences in cultural habits affect the language system and language use*: first, the existence of a special tool for 'scooping-up a dog's excrement' in the NS countries and its non-existence in the NNS country may be held responsible for the 0.231 PR difference in the reading 'a spade used for scooping-up dog's excrement'; second, using 'shopping hats' may have served as an analogy base for some of the native speaker informants in proposing the reading 'a hat to wear when you work in the garden'.
4. As far as the second-rank readings are concerned, the PR differences are generally smaller (which is logically related to the generally lower PR val-

ues for second-rank readings), with six cases under 0.1 and three other cases under 0.2. The only significant difference pertains to *flower hat*, which may be accounted for by the fact that the top two readings swapped their positions in the two groups of informants.

5. Finally, let us take into account the results obtained from the check-group. The absolute match with the total PR of the main group has been obtained in 8 out of 10 top rank readings; in the remaining two cases the difference is one rank (*game wheel* and *age bag*). However, in neither of these cases are the differences dramatic.
6. To summarise, *the two groups of informants and the check-group significantly agree in identifying the predictable and unpredictable readings* for the sample naming units despite the fact that one main group was composed of native speakers (Americans and Englishmen), the other one of Slovaks, and the check-group of Poles. Given the above-mentioned idea that the differences in habits affect the language system and use, it may be surmised that the minimum impact of this factor upon the sample naming unit interpretation follows from the more or less shared European culture (in the widest sense of the word).
7. Based on these facts, it may be concluded that the predictability of the individual readings of novel, context-free naming units is a matter of both *linguistic competence*, the level of *conceptual processing*, and *extra-linguistic knowledge* and *experience*. Under the conditions of similar cultural traditions, habits, value systems, and ways of life, the extra-linguistic factors play a minimum role in establishing different conditions for native and non-native speakers in a meaning-prediction process. This is to say that non-native language users are equally 'qualified' for the assessment of the predictability of meaning of a novel naming unit as any native speaker under the condition of a relevant degree of proficiency in the particular language (English in our case).

A remarkable overall agreement between the groups of native and non-native speakers bears out the assumption that predictability is not just a vacuous term, and gives support to hypothesis 4. (Section 3.11), justified in Section 3.3, assuming that the results for native speakers should not differ significantly from those for non-native speakers.

4.2.3.2 *Predictability Rates*

Table 23 gives an overview of the Predictability Rates of the four most predictable readings for each sample naming unit.

Table 23. An overview of four most predictable readings for each naming unit

Naming unit	*PR*	*Naming unit*	*PR*
baby book		*apple-juice seat*	
	0.727		0.485
	0.312		0.155
	0.227		0.108
	0.157		0.020
dog spade		*game wheel*	
	0.137		0.302
	0.090		0.245
	0.023		0.069
	0.010		0.013
flower hat		*shape cloth*	
	0.404		0.098
	0.359		0.097
	0.109		0.011
	0.035		0.003
ball hammer		*garden whisky*	
	0.420		0.327
	0.113		0.231
	0.084		0.109
	0.017		0.023
hill star		*age bag*	
	0.119		0.203
	0.057		0.053
	0.021		0.017
	0.015		0.014

The following conclusions can be drawn from Table 23.

1. The Predictability Rate values of the top readings differ significantly and range from the poor 0.098 (*shape cloth*) to an extremely high value 0.727 (*baby book*). This might indicate that the Predictability Rate does not depend on the Onomasiological Type – all the sample naming units in Experiment 1 belong to Onomasiological Type 3. However, given the fact that the PR of nine of ten top readings is (mostly significantly) below the value of 0.5, it may be assumed that Onomasiological Type 3 does not favour high Predictability Rates: only three top readings have the PR above 0.4, two above 0.3, one above 0.2, two above 0.1, and one top reading is even below 0.1 (*shape cloth*). The average PR value of the ten top readings is 0.322. If we disregard the extreme value of *baby book*, then it is only 0.277. It may therefore be assumed that the *general tendency for the PR of primary compounds*, falling within Onomasiological Type 3, *is the PR value about 0.3*. Let us divide the predictability scale – for the sake of evaluation – into four predictability levels as follows:

(46) 0–0.25 unpredictability level
0.26–0.50 low predictability level
0.51–0.75 medium predictability level
0.76–1.00 high predictability level

Then, we may assume that *the tendency for the meaning-predictability of novel, context-free naming units belonging to Onomasiological Type 3 is the lower range of the low-predictability level*. This conclusion bears out my considerations presented in 3.5.1.4 concerning the difficulties in predicting the meaning of those naming units in which the Actional seme does not find its onomatological representation. Obviously, this general trend does not preclude the existence of naming units with a much higher Predictability Rate, as evidenced by the case of *baby book*, which is almost at the level of high predictability.

2. Given the division proposed in Point 1., four top readings are in the unpredictability range (*dog spade, hill star, shape cloth,* and *age bag*). This indicates that the semantic compatibility of some motivating words is poor, which is determined by the limited possibilities of relating their corresponding extra-linguistic objects in a logical way. While this poor compatibility is reflected in a low Predictability Rate, it has hardly any influence upon the Objectified PR (see Table 23 and the related comments).

4.2.3.3 *Meaning predictability and dominant readings*

The top readings with the highest PR in the sample, i.e., *baby book*, *ball hammer*, and *apple-juice seat* do not face tough competition, i.e., the PR Gap between these readings and rank 2 readings are significant (0.415; 0.307; and 0.330). The only exception among the ten top readings is that of *flower hat*, with the PR Gap a mere 0.045.

The average PR Gap between the rank 1 and rank 2 readings of the ten sample naming units is 0.151. The average PR Gap between the rank 1 and rank 3 readings is immense: 0.244.

An even clearer picture of the dominant position of the most predictable readings is obtained if the rank 1 : rank 2 (hereinafter R1/R2) and the rank 1 : rank 3 (R1/R3) ratios are calculated. The average PR value of the rank 1 readings is almost twice (exactly, 1.88 times) higher than that of the rank 2 readings. The gap between the most predictable reading and the third rank reading is huge: the average PR of the former is 4.14 times higher. It follows from the formula for the OPR calculation that even more than the absolute

PR Gap value it is the closely related *R1/R2 and R1/R3 ratios that are crucial* in terms of the dominant position of rank 1 readings.

Almost universally (with the exception of *baby book*), *the PR values of rank 4 readings are insignificant*, mostly below 0.1. The average value of the ten rank 4 PRs is as low as 0.031. If the extraordinary case of *baby book* is disregarded the average value approaches zero (0.017). For this reason the effect of rank 4 PRs upon the Objectified PR values is negligible. Consequently, these values were not included in the calculation of the OPR.

All these statistical data provide an unambiguous support to hypothesis 7. (Section 3.11) assuming one (rarely two) dominant reading(s) for each novel, context-free naming unit. This conclusion has far-reaching consequences for meaning predictability.

4.2.3.4 *Objectified Predictability Rate*

The Objectified Predictability Rates for the individual sample naming units and their comparison with the Predictability Rates are given in Table 24.

Table 24. Top reading ranking, PR Gaps, Objectified Predictability Rates

Most predictable reading for each naming unit	*PR*	*PR-based rank*	*PR Gap*		*OPR*	*OPR-based rank*
			absolute value	*R1/R2 ratio*		
baby book	0.727	1	0.415	2.330	0.574	5
dog spade	0.137	8	0.047	1.522	0.548	6
flower hat	0.404	4	0.045	1.125	0.463	10
ball hammer	0.420	3	0.307	3.717	0.681	2
hill star	0.119	9	0.062	2.088	0.604	4
apple-juice seat	0.485	2	0.330	3.129	0.648	3
game wheel	0.302	6	0.057	1.232	0.490	7–8
shape cloth	0.098	10	0.001	1.010	0.476	9
garden whisky	0.327	5	0.096	1.416	0.490	7–8
age bag	0.203	7	0.150	3.830	0.744	1

A comparison of the PR values of the most predictable readings of the individual sample naming units and the corresponding OPR values leads to an important conclusion: there is *no direct proportionality relation between the Predictability Rate and the Objectified Predictability Rate*. While a particular reading may be the most predictable one for a given naming unit, its objectified value, that is, its predictability value relativized with regard to the other readings of this naming unit may be significantly lower.

For illustration, the top reading of *age bag* has a very low PR (0.203), and ranks as low as seventh. At the same time, it has the highest OPR in the sample

(0.744). Similarly, the OPR of *hill star* is relatively high (0.604) in spite of a very low PR value (0.119). On the other hand, while *flower hat* ranks fourth in terms of its PR value it has the lowest OPR from among the evaluated naming units. These 'discrepancies' can be explained by the respective PR Gaps: low PR Gaps present a serious obstacle to a high OPR. Thus, the PR Gap between the two top readings of *age bag* (0.150) is the fourth largest; what is even more important, the PR of the rank 2 reading of this naming unit is almost four times smaller than that of the top reading. Moreover, the PR of the rank 3 reading of *age bag* approaches zero (0.017).

A *crucial role of the PR Gap-based R1/R2 ratio* in determining the OPR can be best seen from the comparison of the respective rankings of *dog spade* and *flower hat*. The absolute values of their PR Gaps are almost identical (0.047 vs. 0.045). However, the R1/R2 ratio values differ considerably: 1.522 vs. 1.125. As a result, the OPR of the top reading of *flower hat* is much lower than that of the top reading of *dog spade*. This in spite of the fact that the PR value of the former is higher than the PR value of *flower hat*. This is reflected in their respective rankings. PR-based ranking: *dog spade* 8 vs. *flower tree* 4; OPR-based ranking: *dog spade* 6 vs. *flower tree* 10.

In *ball hammer*, the big PR Gap of 0.307 is supported by an extremely big R1/R2 ratio of 3.717. This contributes to the overall second rank of its top reading.

Similar factors account for a significant difference between the PR and the OPR rankings of *hill star*. The top reading's PR is more than twice higher than that of rank 2, and the rank 3 reading approaches zero.

The PR Gap of *shape cloth* is as low as 0.001, which gives the R1/R2 ratio of 1.010. In addition, its PR value is the lowest of all the ten top readings. One would expect that the top reading of this naming unit cannot but be the least predictable in terms of OPR. The reality is different – it ranks ninth, because the rank 3 reading of *shape cloth* is much lower than that of *flower hat* (ranking tenth).

An interesting case is represented by *baby book*. Its most predictable reading (PR = 0.727) approaches the high predictability level. However, its relatively low PR Gap and, mainly, the strength of the competing predictable readings (whose values are higher than the values of some of the top readings in the sample: rank 2 = 0.312 and rank 3 = 0.227) strikingly reduce its OPR. As a result, the reading 'a book for babies (fairy tales, rhymes, pictures; drawings)' that has the highest PR among the top readings ranks fourth in terms of OPR.

This brings us to a paradoxical conclusion: the existence of a relatively large number of semantically fairly well compatible, possible, and predictable seme

combinations, implying several acceptable meanings of a naming unit, appears to be an obstacle to the overall predictability due to tough competition between the acceptable (predictable) readings. On the other hand, poor compatibility of the semes of the motivating units need not entail a poor OPR owing to the absence of any competition (rival readings) for an otherwise low PR reading.

As a result, novel naming units with a single, low-predictable reading like *dog spade* may be objectively more predictable in context-free linguistic situations than novel naming units which offer a range of possible, semantically well acceptable readings, such as *baby book* and *flower hat.*

With the exception of the top two readings, in particular *age bag* (0.744) and *ball hammer* (0.681), the majority (8 out of 10 cases) of the top readings are concentrated within a narrow range between 0.463 (*flower hat*) and 0.648 (*apple-juice seat*), which might indicate a certain tendency of OPRs for this Onomasiological Type of naming unit.

4.2.3.5 *Seme level*

Table 25. Seme Level Combinations of top readings

Naming unit/ the most predictable reading	*Seme Level Combination*
baby book	
'A book for babies (fairy tales, rhymes, pictures, etc.)'	4–4
dog spade	
'A spade used for scooping-up a dog's excrement'	4–5
flower hat	
'A hat with flowers on it'	4–4
ball hammer	
'A hammer, a part of which has the form of a ball'	4–5
hill star	
'A star that can be seen beyond/above the hill/that 'touches' the hill'	3/4–3
apple-juice seat	
'A seat for drinking apple juice in a restaurant, bar, etc.'	4–4
game wheel	
'A wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games'	4–4/5
shape cloth	
'Elastic cloth shaping woman's figure, very tight clothes shaped by the form of the body'	2–4
garden whisky	
'A whisky served in the garden (on various occasions, for example at parties, barbecues, etc.)'	4–4
age bag	
'An old bag; a bag that looks old'	4–3

This overview appears to confirm hypothesis 3. (Section 3.11) that the *combination of prototypical* (level 4) *semes, reflecting the prototypical features of relevant objects, is one of the conditions for good predictability*. The readings motivated by other than prototypical semes tend to have a lower PR. In any case, there seems to be a condition imposed on a reading to become a top one in terms of meaning predictability, according to which *at least one of the two dominantly activated semes should be of level 4*. On the other hand, *the 4–4 Seme-Level Combination itself does not guarantee a high Predictability Rate*, as indicated by a number of unpredictable readings in the sample. This factor works in close co-operation with the other factors affecting the predictability of novel, context-free naming units.

4.2.3.6 *Tendencies*

The following tendencies emerge from the previous analysis:

Tendency 1: One dominant reading.

Tendency 2: Its PR is in the lower range of the low predictability level (0.26–0.50). The average PR value for the ten top readings in the sample is 0.322).

Tendency 3: The OPR-conditioning average R1/R 2 ratio of 1.88 is fairly high and the R1/R3 ratio is as high as 4.14, which gives support to Tendency 1. The closely related PR Gap between the first and the second readings usually exceeds the value of 0.1 (the average PR Gap for the ten top readings in the sample is 0.151). The PR values of the rank 3 and lower rank readings are usually insignificant and negligible: the average PR for rank 3 readings in the sample is 0.078.

Tendency 4: An important (not sufficient, though) condition for meaning predictability is level 4 of the motivating semes. In other words, a 'good' PR tends to be conditioned by prototypical semes.

It may be assumed that meaning predictability is conditioned by the inherence of features, their prototypical nature; any individualisation, the trespassing on the prototypical feature limits, appears to become a serious obstacle to predictability. Figurative meanings rarely exceed the predictability level (exceptions in this research are the readings based on the well-established metaphorical meaning of 'star'). It should be, however, emphasised that the activation of level 4 semes itself does not guarantee a high PR.

Tendency 5: Meaning predictability is influenced by extra-linguistic factors, such as world knowledge and experience.

Tendency 6: In culturally similar communities, the meaning-prediction process tends to bring similar results, no matter whether the informants are native speakers or non-native speakers of a language. In other words, the degree of their agreement in identifying the most predictable readings of novel, context-free converted naming units is very high.

4.3 Experiment 2[7]

4.3.1 Sample naming units

The test included ten possible conversions (conceptual-recategorisations) devised by myself according to the following scheme. It maps the principles used of Experiment 1 (see (45) in Section 4.2):

(47)	[Animate]	– [Human]			*boy*
		– [Animal]			*lion*
		– [Plant]			*tulip*
	[Inanimate]	– [Tangible]	– [Solid]	– [Movable]	*cableway*
				– [Immovable]	*planet*
			– [Liquid]		*river*
		– [Intangible]	– [Process]		*conference*
			– [Form]		*triangle*
	[Location]				*courtyard*
	[Time]				*morning*

4.3.2 Experimental data and their analysis

4.3.2.1 *to boy*

Table 26. to boy – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to act or behave the way boys do (immature)			5		3	4		X	8	X	X	8	3	X		2		6	8	5
– to (try to) look/behave like a boy (clothes, haircut, motions, walking) – of girls						2					8				X					X
– to give birth to a boy						6						4			3	6	X			7
– to (try to) look/behave like a boy (clothes, haircut, motions, walking) – of girls		5		8													9			
– to 'staff' with boys												X						X		
– to treat like a boy	2				2								5							

Table 27. to boy – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to act or behave the way boys do (immature)		9	X			X		9		X	X	X		X	X	9		5	X	9
– to (try to) look/behave like a boy (clothes, haircut, motions, walking) – of girls		5	3			8		8	5	1	8	7	9	4		7		7		3
– to bring up a male child				X					X	8										
– to be foolish/to fool around					5													9		
– to be childish						9											10			
– to change sex from girl to boy						3		X	8			5							5	
– to treat somebody as a boy instead of taking him as an adult	3								2											
– to joke like boys do										7							8			

Type: Animate Human
boy
Seme level 1 – SUBSTANCE
2 – [Animate]
3 – [Human]
4 – [Male] [–Adult] [Characteristic Behaviour]

The most predictable reading:
'to act or behave the way boys do (immature)'
WF Type: [SUBSTANCE$^{\underline{\text{Manner/Pattern}}}$ACTION]
Seme Level: 4

	NS	*NNS*	*Total*
FO:	13/20	13/20	26/40
Scores:	92/200	121/200	213/400
PR:	0.322	0.393	0.358
OPR:			0.601

The next two readings:[8]
'to (try to) look/behave like a boy (clothes, haircut, motions, walking) – of girls'

PR:	0.162	0.244	0.203

'to change sex from female to male'

PR	0.030	0.039	0.035

Comments
The major part of the readings proposed are based on a WF Type in which Action, in particular, the [Characteristic Behaviour] of boys, is imitated. Exceptions to this rule are represented by readings like 'to organise a team or a group where there are only boys' ([SUBSTANCE$^{\underline{\text{Object}}}$ACTION]), 'to bring up

a boy child' ([SUBSTANCE$^{\underline{\text{Object}}}$ACTION]), and 'to change sex from girl to boy' ([SUBSTANCE$^{\underline{\text{Result}}}$ACTION]). Interestingly, the reading based on the analogy with the central meaning of *to man*, i.e., 'to furnish with staff', gained minimum support. Where this 'link' was effective the reading gained two tenths in the native speaker group.

The two most predictable readings are based on a WF Type in which 'boy' as SUBSTANCE serves as a Pattern indicating the Manner of ACTION (i.e., [SUBSTANCE$^{\underline{\text{Manner/Pattern}}}$ACTION]). The Manner refers to the characteristic behaviour of 'boy'. This behaviour is mostly conceived of as a whole, and there are only a few readings in which a more specific aspect of behaviour is salient (for instance, the characteristic way of speaking, dressing, walking, haircut, tendency to attract attention of girls by force and/or by showing off). Since none of these individual features dominate, the activation of any of them usually requires a contextualised situation. We might say that, for example, 'Peter likes to boy when he meets a nice girl' and – knowing him well – refer to his inclination to show off.

The two dominating readings emphasise two different facets of opposition into which 'boy' as [–Adult] and [+Male] enters. In particular, reference is made to the characteristic way of boys' behaviour or acting that makes them different from (1) adults, especially mature adult men, and (2) girls. While the most predictable reading is based on the [+Adult] vs. [–Adult] contrast, the second reading is based on the contrast between [+Male] and [–Male].[9] Rather than on boys themselves (first reading) the second reading lays emphasis on imitation by girls of a boy-like behaviour. The relatively high PR of this latter reading is not surprising in the age of emancipation of females in the countries of Western civilisation, which is manifested in the way of dressing and hairstyle (unisex), in the growing social and economic opportunities of women and also girls, in their growing chances to excel and to become independent.

An important conclusion that can be drawn from the *boy* conversion is that the process of conversion can hardly be accounted for by sticking to the linguistic data. These must go hand in hand with the listener's/reader's extra-linguistic knowledge. There is nothing in the semantic structure of 'boy' that would reflect, for example, the present-day 'unisex' trends in fashion and/or the emancipation trends.

The reading 'to change sex from female to male' also makes reference to *extra-linguistic knowledge*, in particular, to the developments in medicine, in surgery and in genetic engineering. Its WF Type is [SUBSTANCE$^{\underline{\text{Result}}}$ACTION]. The reading, 'to hire someone to carry luggage' is too much context-bound, and therefore, unpredictable. Carrying luggage is not a prototypical feature of

'boy' in its basic, cognitive, meaning. 'To behave like a mother's little boy' is an interesting modification of the most predictable reading: it is its negation. To behave like a mother's boy means not to behave like a 'true' boy. It may be this fact that caused only a single 1-point occurrence of this reading. Readings like 'to exclaim 'boy'' are 'automatically' doomed to unpredictability because all verbs resulting from conversion might be interpreted this way.

Check-group results

Readings	*FO*
To act or behave the way boys do (immature)	19/25
To behave like a boy (of a girl in a company of boys; to wear boy clothes, to speak that way, etc.)	4/25
To be interested in things that typically boys are fond of (e.g., cars)	1/25
To treat sb. as if he were a boy	1/25

The check-group maps the results of the main group of informants, and confirms the order of the two predictable readings, and the dominating position of the reading based on the contrast [–Adult] vs. [+Adult], which is closer to the native group's results.

4.3.2.2 *to lion*

Table 28. to lion – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to have the power and rule as a lion			8	7				6				2	8			X		7		5
– to roar like a lion						2		4					4	8	5	3	X	4	7	X
– to behave/act like a lion (proud, lordly and/or brave)	6				2		6		X	X	X	2								
– to catch/hunt/kill a lion										3	9		7	5			5		X	
– to populate with lions					4				2											
– to treat someone with respect									4					3						
– to have a hair-cut that resembles a lion mane												2			8					

Table 29. to lion – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to roar like a lion	5		8	X		1		8	X	5	9	5	4	5	2	4		5		7
– to have the power and rule/to act as a leader		X			X			9			4		8	X		5	6			2
– to behave like a lion (proud, lordly and/or brave)					9		4		8			X	X							
– to protect somebody as lioness protects her cubs									2				6				X	X		
– to look like a lion										8	5			9			8			
– to dress up as a lion			9								1								X	
– to act like a predator, hunter						X														9
– to kill in order to survive						9									8					
– to eat a bar of chocolate called 'Lion'		5				2														
– to be self-confident/ proud							2			6		3								
– to be selfish							2				2									

Type: Animate Animal
lion
Seme level 1 – SUBSTANCE
2 – [Animate]
3 – [Animal]
4 – [Predator] [Extreme Physical Strength] [Strong Voice] [Majestic Appearance] [Courage]
All of the proposed readings take lion's features and/or behaviour as Pattern and Manner. One can hardly think of any other onomasiological structure.

The most predictable reading:
'to be aggressive; to roar like a lion'[10]

WF Type:		[SUBSTANCE$^{\underline{\text{Manner/Pattern}}}$ACTION]	
Seme Level:		4	
	NS	*NNS*	*Total*
FO:	10/20	15/20	25/40
Scores:	57/200	88/200	145/400
PR:	0.143	0.330	0,237
OPR:			0.581

The next two readings:
'to have the power and rule/to act as a leader'

PR	0.106	0.144	0.124

'to behave/act like a lion (proud, lordly and/or brave)'

PR	0.081	0.051	0.065

Comments

What makes this reading interesting is the different treatment by informants of the prototypical characteristics of 'lion' compared to 'boy'. While there does not seem to be an absolutely salient feature for the latter, and 'boyhood' is understood rather as a sum of more or less equally significant features, the notion of *lion* is dominated by such features as [Predator] and [Strong Voice] and, obviously, [Extreme Physical Strength], implying the ruling position in the world of animals (the king!). It may be for this reason that the reading 'to behave like a lion' – as opposed to the reading 'to behave like a boy' – has not acquired wider support. The informants apparently preferred to point out various striking properties of 'lion', such as beautiful mane, courage, strength, 'to devour like a lion', etc.

The reading 'to treat someone with respect' seems to miss the point by reversing the view. Rather, it is the other animals that treat the lion with respect. Surprisingly, while six native speaker informants proposed a reading in which 'lion' is an Object of Action, 'to catch/hunt/kill a lion', no such proposal occurred in the non-native group! One would expect a higher Predictability Rate for this reading on the basis of analogy with *to fish*.

Check-group results

Readings	*FO*
To be aggressive; to roar like a lion	10/25
To behave bravely, courageously	9/25
To have the power and rule/ to act like a leader	3/25
To give oneself airs	2/25
To take the biggest and/or best part of sth for yourself	1/25

The check-group confirmed the highest predictability of the [Predator]-based reading; however, the frequency of occurrence of the [Courage]-based reading is less by one only. This is different from the main group results where the PR of a partly corresponding reading (rank 3) is very low (0.065).

It is my subjective opinion that the results in both the main group and the check-group are rather surprising. To me, 'lion' symbolises majesty, the king-status in the animal empire rather than aggressiveness, the latter being attributable – as opposed to the former characteristic – to a number of other predators.

4.3.2.3 *to tulip*

Table 30. to tulip – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to grow tulips		7	5		2	3			X	4	9	X	X	9		9	9	5		8
– to hit someone with a tulip	4					6							2							
– to pick tulips											X	8		X		X				
– to shape something as a tulip	7				1		2		X											
– to make colourful				3													2			
– to have sexual or intimate relations with sb (euphemism); to perform sexual favours						3								6						
– to decorate (a room) with tulips							4			8									5	

Table 31. to tulip – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to grow tulips			X	9		8		X	X						9		X	8	9	X
– to pick tulips		5								9					8		9	8		
– to put tulips into a vase			X																X	
– to look beautiful						7		8							6	7				
– to turn red in face	1	5																7		
– to blossom						X									7					
– to decorate a room with tulips									8				X							
– to make a bunch of tulips										X										5
– to smell like a tulip										7										7
– to keep secrets																	2		8	
– to give somebody a tulip	2												4							
– to go to Holland													8				5			

Type: Animate Plant
tulip
Seme level 1 – SUBSTANCE
2 – [Animate]
3 – [Plant] [Flower] [Growable]
4 – [Characteristic Features of Tulip]

The most predictable reading:
'to grow tulips'

WF Type:		[SUBSTANCE$^{\underline{\text{Result}}}$ACTION]	
Seme Level:		3	
	NS	*NNS*	*Total*
FO:	14/20	10/20	24/40
Scores:	100/200	93/200	193/400
PR:	0.350	0.233	0.290
OPR:			0.831
The next two readings:			
'to pick tulips'			
PR:	0.038	0.049	0.043
'to decorate with tulips'			
PR:	0.024	0.009	0.016

Comments

The Seme Level of the top reading is 3. The Result of Action follows from the prototypical feature of the next higher level: any plant is [Growable]. As with any other plant that has been a common object of gardening or agricultural activities, the results confirm the expectations of high predictability of the 'growing'-based reading. In the non-native group the dominating position of this reading owes a lot to an extremely high average score (approaching the maximum value (!)) rather than to the frequency of occurrence (10/20). In contrast to it, the roughly same PR in the native speaker group is primarily achieved by a higher frequency of occurrence.

The PRs of the other readings are far below the predictability level. The dominating position of the top reading is highlighted by the fact that out of the remaining 34 readings in the non-native group 32 gained support from three or fewer informants (mostly only one). A similar situation is in the native group. As far as the two remaining readings are concerned, 'to pick tulips' [(SUBSTANCE$^{\underline{\text{Object}}}$ACTION]) has also a very high average score (9.5 and 8.2 pts., respectively), but only a 25% frequency of occurrence.

'To shape something as a tulip' is based on the 'Manner/Pattern' type of relation, a very productive WF Type in the Noun → Verb conversion class; however, this particular reading did not find wider support.

'To look beautiful' is an example of a *metaphorical meaning*. It is understandable that it has not got over the predictability level. There are many [Tangible] SUBSTANCES which are beautiful, and more specifically, many flowers which are beautiful. Being beautiful is a matter of *subjective assessment*, and

therefore, rather than an intrinsic, prototypical feature, it is an extrinsic feature of any such Object.

Some readings are over-specific, for example, 'to have a yellow sweater and green trousers' – tulips can have a number of different colours. Therefore, the identification of colour is an *individualising feature of level 5*. Furthermore, the predictability of this particular reading is hampered by its *figurativeness*.

The reading 'to become beautiful in spring' is at the other pole of the specific – general scale. Any flower can motivate this reading. And the same applies to the reading 'to be interested in flowers' – this reading fits better the conversion of *flower*$_N$.

A considerable meaning-predictability potential is demonstrated by two tens for 'to put flowers into a vase'.

In principle, any activity whose Object/Result are tulips is acceptable; the results indicate, however, that only one of them is really predictable.

Check-group results

Readings	*FO*
To grow tulips	4/25
To be nice, friendly, polite, gentle to people	4/25
To be shy, to blush with shame	3/25
To flourish, look beautifully like a tulip	2/25
To be innocent	2/25
To flatter someone	2/25
To open sb's heart by flowers	1/25
To look beautiful	1/25
To give flowers to a girl	1/25
To be in love with oneself	1/25
To apologise to sb. by giving him/her a tulip	1/25
To act as a playboy, a womanizer	1/25
To blossom, to flourish	1/25
To be as sensitive and fragile as tulip	1/25

The results of the check-group both confirm those of the main group and differ from them. They confirm the former by placing the reading 'to grow tulips' at the top and by a large number of single-occurrence readings. The number of the readings proposed by the main group is extraordinarily high – unlike the other converted naming units. The difference consists in a considerable gap between the most predictable reading and the other readings of the main group, on the one hand, and the minimum differences in the check-group.

The check-group also confirmed that 'tulip', like a number of other flowers, is frequently used as a symbol (mostly symbol of beauty) and, importantly, for different language users it evokes various connotations, which is reflected in a considerable scattering of the readings. The reading 'to be in love with oneself' seems to result from confusing *tulip* with *narcissus*.

4.3.2.4 *to cableway*

Table 32. to cableway – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to travel by cableway				8	8		6		X	X	8	8		X	8			7	X	X
– to build/construct a cableway	6	4	8			5	8			8	X		8		X	9	X	7		5
– to lay cables												8	9	7	2					
– to transport items via a cableway					4			7								6	8			

Table 33. to cableway – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to travel by cableway	8		X	X		X		9			8	9	4		7	9	9			X
– to build/construct a cableway		6				8			X	X	7			5	X	7				8
– to work one's way up				3		9		8										3		
– to make progress	2																		5	
– to climb the mountains		7																	8	

Type: Inanimate – Tangible – Artefact – Movable
cableway
Seme level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Artefact] [Equipment] [Constructable]
4 – [Means of Transportation]

The most predictable reading:
'to travel by cableway'

WF Type:		[SUBSTANCE$^{\underline{\text{Instrument}}}$ACTION]	
Seme Level:		4	
	NS	*NNS*	*Total*
FO:	12/20	12/20	24/40
Scores:	103/200	103/200	206/400
PR:	0.309	0.309	0.309
OPR:			0.564

The next two readings:
'to construct/ build a cableway'

PR:	0.319	0.160	0.232

'to lay cables'

PR:	0.026	–	0.007

Comments
Almost unbelievable: the results of the two groups are absolutely identical in both the frequency of occurrence and the scores. An extra-linguistic factor is vital to this result: the Instrumental reading for *cableway* is a very natural and expected one because – like any other vehicle – it is a *means* of transportation. Since the transportation function is the prototypical feature and function of *cableway* the Seme Level is 4.

The results obtained for the second-rank reading 'to construct/build a cableway' are not surprising. In principle, this reading is analogical to 'to grow tulips', the only predictable reading of *tulip* featuring the semes [Animate] [Plant]. Both of them reflect the Result of Action. In both of these cases the motivation comes from level 3: with *tulip*, the reading follows naturally from the level 3 seme [Growing Capacity] characterising a plant in general. The *cableway* reading is activated by the level 3 seme [Artefact]; and each artefact is [Constructable]. These more general, level 3, semes are specified in the converted meanings bound to the particular flower and equipment, respectively.

These two readings, as it were, 'saturate' the possible WF Types, and consequently, the predictable options. Their predictability-related dominance is so impressive that only a few more readings were proposed, which primarily holds for native speakers (only 6 readings in total). Out of the remaining readings, the 'four-occurrence' reading 'to work one's way up' is clearly figurative, 'to lay cables' is a less successful construction-related variant of 'to construct/build a cableway', and 'to transport items via a cableway' a less successful transportation-related variant of 'to travel by cableway'.

Check-group results

Readings	*FO*
To travel by cableway	13/25
To move up, to progress and advance	3/25
To haul/help somebody during a climb	3/25
To help sb make progress in a particular activity	2/25
To use sb's help while trying to make a career	1/25
To pull sth up	1/25
To build a cableway	1/25
To risk, to act hazardously	1/25

While the top reading of the main group also dominates very clearly in the check-group the second predictable reading was given only one vote from the check-group. This is partly accountable by the 'one-only principle' in the check-group and the tough competition on the part of the main reading.

4.3.2.5 *to planet*

Table 34. to planet – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to fly to other planets; to land on a planet	8	6	3		7	4		6	8	9	X	7		9	5	X	5	4	X	
– to populate a planet			9	X		8	7							8			X			
– to do something to save the planet											4									6
– to export goods all over the world																6				
– to observe planets through a telescope											5					9				
– to discover a planet					3					4					3					

Table 35. to planet – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to travel to other planets/to land on a planet	7	X		7	X	X		9	7		X		9	8				8		9
– to populate/colonize another planet		7		8			X					X				7	9	6		
– to dream about something	1															3				
– to be interested in planets and space	5					6				X					2					6
– to create a new planet									3						8					
– to explore or describe a planet		9			9											9			9	
– to bring someone back to reality			4														4			
– to name a planet					7														7	
– to discover a planet						7														7

Type: Inanimate – Tangible – Natural – Moving

Planet

Seme level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Celestial Body]
4 – [Resolves About its Star] [±Colonisable]

The most predictable reading:
The two predictable meanings are clearly based on and conditioned by extra-linguistic knowledge of the informants, especially that concerning scientific discussions and sci-fi treatments of human penetration into space and possible future colonisation of planets.

'to travel to other planets/to land on a planet'

WF Type:		[SUBSTANCE$^{\text{Target/Direction}}$ACTION]	
Seme Level:		4	
	NS	*NNS*	*Total*
FO:	16/20	12/20	28/40
Scores:	111/200	104/20	215/400
PR:	0.444	0.312	0.376
OPR:			0.795

The next two readings:

'to populate/colonize another planet'			
PR:	0.078	0.100	0.089
'to discover a planet'			
PR:	0.008	0.007	0.008

Comments
The most predictable reading represents two closely related aspects, or better, the first stage of space colonisation: travelling to and landing on a planet, which is reflected by the Target/Direction category. Since travelling to planets is closely related to the idea of landing on them, I decided to classify these two semantic aspects as one and the same reading.

The Seme Level is 4: it is not any celestial body which can be (at the present stage of technology) the target of travelling and landing. Only (some) planets are potentially reachable.

The reading 'to populate/colonize another planet' reflects the next stage of human space travel, aimed at the settlement of planets. I may only hypothesise that the higher PR of the former follows from the *logic of extra-linguistic reality*: travelling to other planets is technologically viable nowadays while their colonisation is a matter of the remote future.

The other readings, with very low Predictability Rates, reflect the possibility of a planet becoming an Object of human interest and exploration, some are motivated by extreme space distances ('to travel long distances'), or identify 'planet' with our Earth.

Check group results

Readings	*FO*
To travel to/land on a pl anet	8/25
To populate/colonize/another planet	7/25
To think/dream about travelling to another planet	2/25
To behave as if being the centre of the Universe	2/25
To navigate one's journey according to planets	1/25
To orbit round a planet	1/25
Not to stay in one place, move all the time	1/25
To think globally	1/25
To move in a circle	1/25

The ranking of the top readings is the same as in the main group, although the gap between them is much smaller in the check-group. In the main group the most predictable reading clearly dominates.

4.3.2.6 *to river*

Table 36. to river – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to swim/bathe in a river		9		6		6			8	4				X				6	5	X
– to travel down a river (in a canoe, kayak, etc.)	7				6		4				X		7		7		7		X	
– to direct a river into an area in order to irrigate it		7					6	2		4								5		
– to urinate excessively												5						7		
– to be/act like river: patient, slow but succeed in life			3				2													
– to flow like a river						4					1									

Table 37. to river – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to swim/bathe in a river		5		X		8		X	X	8		9	8			X		8	X	
– to travel down a river (in a canoe, kayak, etc.)	6	8									6	8	9					7		X
– to cry very much			7					3	8					9						
– to swim with the stream					8										4					
– to water a garden					4				4											
– to flow like a river						7									6	6				
– to build an additional way for river to float through some territory									5	X										
– to catch fish which live only in rivers													6						6	
– to live by the river																1			9	

Type: Liquid
river
Seme level 1 – SUBSTANCE
2 – [Inanimate]
3 – [Natural Object]
4 – [Liquid] [Water] [Stream] [Flow] [Level]

The most predictable reading:
'to swim/bathe in a river'

WF Type:		[SUBSTANCE$^{\underline{\text{Location}}}$ACTION]	
Seme Level:		4	
	NS	*NNS*	*Total*
FO:	9/20	11/20	20/40
Scores:	64/200	96/200	160/400
PR:	0.144	0.264	0.200
OPR:			0.635

The next two readings:
'to travel down a river (e.g., in a canoe)'
PR: 0.116 0.098 0.107
'to direct a river into an area in order to irrigate it'
PR: 0.030 0.008

Comments
Swimming, bathing, boating, and similar types of relaxation are activities which are bound to the inherent and prototypical feature of river, in particular [Liquid] and [Flow]. Since water pools, lakes, and rivers, unless polluted (it seems that [+Pollution] has been working up to the prototypical characteristics of 'river' in the technological era) have been traditionally connected with the notion of a pleasant way of spending summer time (extra-linguistic experience); this reading is based on the relevant level 4 semes.

It may be that the preference of this reading over the next one, 'to travel down a river', results from the knowledge and experience that bathing and swimming in rivers is a much more common phenomenon than sailing on a river.

Out of the other readings, the greatest support was given to 'to direct a river into an area in order to irrigate it', a reading not mentioned in the non-native group. An interesting pair is represented by 'to act like a river – patient, slow but succeed in life' and 'to run or move very fast'. One may surmise that this contradictory evaluation of 'river', which occurs both in the native and the

non-native groups as well as the check-group, is based, once again, on one's *subjective connotation*.

The reading 'to swim with the stream', no matter how much it reflects the reality (herd-behaviour of the major part of mankind) cannot be predictable because of its *figurativeness*.

The figurative meaning also blocks the predictability of the reading 'to stay afloat (in spite of difficulties)'.

Check-group results

Readings	*FO*
To swim/bathe in a river	7/25
To sail on a river	4/25
To move quickly, to hurry	3/25
To flow like a river (e.g. blood)	2/25
To talk a lot – a flood of words	2/25
To change oneself because of the influence of time; to pass	2/25
To move slowly and majestically like a river	1/25
To be very nosy/curious about sth.	1/25
To feel free	1/25
To be unpredictable	1/25
Not to stay in one place	1/25

The two most predictable readings found analogical support in the check-group. Furthermore, as mentioned above, 'the 'slow' and the 'fast' connotations of river also occur in the check-group, highlighting the role of subjective perception.

4.3.2.7 *to conference*

Table 38. to conference – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to take part in a conference					X		4	9		X	X	X		X	9	X	7	X	8	X
– to organize a conference	X		7		X		8	4				8	8		9	8	7		5	
– to (meet) to talk about/ consult/discuss a subject		8	3	6	8	9	2	3	X		6		7	2		5	X			5
– to take a conference call on the phone											8				X					

Table 39. to conference – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to take part in a conference	6	X		X	8	X	4	X	8	X		9		X		8	X		8	6
– to (meet) to talk about/consult/discuss a subject		8	X			X	6			8		X			8			X		
– to organize a conference							X		9		8		7	X	X				9	
– to lead a meeting			8	9																
– to ask somebody several questions (like journalists do)				8			2													
– to make an arrangement						9	2													
– to report about a conference													5							3
– to talk in a long and boring way																	5		2	

Type: Process
Conference
Seme level 1 – ACTION
2 – [Human Beings]
3 – [Meeting] [Preparation]
4 – [Official Event] [Presentation of Information] [Exchange of information]

The most predictable reading:
'to take part in a conference'
WF Type: [PROCESS$^{\underline{\text{Purpose}}}$ACTION]
Seme Level: 4

	NS	*NNS*	*Total*
FO:	13/20	15/20	28/40
Scores:	117/200	127/200	244/400
PR:	0.380	0.476	0.427
OPR:			0.531

The next two readings:
'to meet to talk about/consult/discuss a subject'
PR: 0.294 0.140 0.212
'to organize a conference'
PR: 0.231 0.110 0.165

Comments
Conversion turns *conference*$_{\text{N}}$, conceived of as a Process, into the Action of taking part in a conference, i.e., a more general and abstract concept is turned into a more active and individualised concept. The two concepts are related by

the Purpose of Action (presentation and exchange of important information, data, views).

'To meet to talk about/consult/discuss a subject' is a generalised variant of the previous interpretation. The platform for discussion and exchange of views need not be an official event; it may be any official and unofficial meeting of people. It is the much broader scope of this reading, resulting form the activation of level 3 [Meeting] (used in combination with level 4 [Presentation of Information] and [Exchange of Information]), that made me treat the two readings separately. At the same time, the partial shift towards level 3 explains an almost double PR value of the former reading, based exclusively on the activation of level 4 semes.

The WF Type of 'to organize a conference' differs from that of the former two readings by focusing on the Result of Action, which is a Process of holding an official event at which information and views are presented and exchanged (hence level 4). Compared to the most predictable reading this one views 'conference' from the opposite perspective. While the former activates its 'participant' aspect, this reading activates the 'organiser' aspect. The reasons for the much higher PR of the 'participant aspect' may be sought in its direct connection to the process of a conference itself, i.e., to the exchange of views, while the latter puts emphasis on the preparation stage, i.e., a much more general characteristic, common to a number of other events which are not aimed at the exchange of views within an official event.

Check-group results

Readings	*FO*
To take part in a conference	14/25
To discuss problems/ideas with other people	8/25
To organize a conference	2/25
To gossip on social matter	1/25

This is an ideal case – the results of the check-group faithfully copy those of the main group.

4.3.2.8 *to triangle*

Table 40. to triangle – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to make a triangle from something/to make triangular things	2			4	X	6				X	X		X			7	X			7
– to draw a triangle		7		8			8	5			X	9			8		5	X		
– to be a part of a triangle of lovers								2	6			6	6			4				
– to move between three points											7			X		6				
– to enter into a three-way interaction										6		X								
– to play a triangle																3			4	

Table 41. to triangle – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to make a triangle from something/to make triangular things	8			9		X		8	X	8			8				X			8
– to draw a triangle	5	X	X	X			6	7								8	7	9	X	X
– to be a part of a triangle of lovers				8	8		4								5					

Type: Form
Triangle
Seme level 1 – QUALITY
2 – [Geometric Feature] [Form]
3 – [Two Dimensions]
4 – [Three Angles]

The most predictable reading:
'to draw a triangle'
WF Type: [QUALITY$^{\underline{\text{Result}}}$ACTION]
Seme Level: 4

	NS	*NNS*	*Total*
OS:	9/20	11/20	20/40
Scores:	70/200	92/200	162/400
PR:	0.158	0.253	0.203
OPR:			0.489

The next two readings:
'to make a triangle from something/to make triangular things'
PR: 0.190 0.178 0.184
'to be a part of a triangle of lovers'
PR: 0.030 0.025 0.028

Comments
The conversion of *triangle*$_{N}$ provides us with two, almost equally predictable readings. The close results make the reversed ranking of these two readings in the two groups of informants almost insignificant. Moreover, since the Predictability Rate Gap is very small (0.019), this significantly reduces the Objectified PR. On closer inspection the two readings are variants of the same 'motive', that is to say, that of giving an object, a configuration a triangular shape. In the former case the purpose is to make a geometric drawing, in the latter, it is any product which, by its form, resembles a triangle. I was hesitant to include these two readings in one group, the former being just a variant of a more general notion of producing – by any technique – anything which is of triangular shape. Both of them are based on the Factitive link (Result of Action) between the conceptual categories of QUALITY and ACTION. Eventually I decided to split them into two different readings due to the different nature of the Results of Action, i.e., those which can be manually handled and those which cannot. Another reason was that eight informants proposed both the drawing-based reading and the 'tangible' product reading.

Moreover, this is another example demonstrating how difficult it is to set borderlines between individual readings due to the *fuzziness* phenomenon.

As to the other readings, only 'to be a part of a triangle of lovers' gained any significant support (nine votes in total).

Almost all the other readings are single occurrences, thus contributing to a very strong position of the two central, closely related readings of *triangle*. It is worth mentioning those like 'to have three tasks', 'to give birth to three babies', and 'to arrange items into groups of three', which seem to miss the point. While the concept of 'triangle' implies 'relations' these readings lack any relational nature. Number 'three' is not sufficient (while being indispensable) with regard to the notion of 'triangle'.

The shifted meaning of *triangle*, referring to a musical instrument, occurs in two different readings: 'to make the musical instrument called a triangle' and to 'play a triangle'.

The reading 'to visit the Bermuda triangle' activates level 5 seme, an individualised feature of triangle. In this particular case, it is defined geographically.

Check-group results

Readings	*FO*
To give sth the shape of a triangle	6/25
To draw a triangle	4/25
To be a part of a triangle of lovers/relationship, business, etc.	4/25
To move on a triangular route	4/25
To look at a problem, to delve into a problem, to consider all of its aspects	1/25
To solve puzzles	1/25
To do an exercise in a triangle-like body position	1/25
To "encircle sb./sth.	1/25
To triple something	1/25
To provide three solutions to a problem	1/25
To have three different faces	1/25

The data show an interesting situation with three different readings taking the second rank. The two highest PR readings of the main group swap their positions, and copy the results of the native informants. They are accompanied with the reading which is based on a well-established shift of meaning of *triangle*, i.e. 'ménage à trois', and the reading 'to move on a triangular route', which also occurred in the main group.

4.3.2.9 *to courtyard*

Table 42. to courtyard – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to play/spend time in the courtyard			8	7		5	4	9	X	6	7			9	3	8	X			7
– to build a courtyard/to make an area into a courtyard	8	5	5		X	5	8		5			X	7		5	6		2	X	
– to enclose a space									2				X				6	8	5	
–to make something look like/in the style of courtyard				5			6													
– to pave a courtyard											8	7								
– for a prisoner to be punished in a courtyard (e.g. Smith was courtyarded for 3 days in the stocks)													4		1					

Table 43. to courtyard – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to play/spend time in the courtyard (with friends)	3	6		7	9	X	8					X		5		7	3	7	6	X
– to build a courtyard	2						5		X	X				X	4		7			8
– to keep something in the courtyard						8													9	

Type: Location
Courtyard
Seme level 1 – SUBSTANCE
2 – [Inanimate] [Location]
3 – [Lot]
4 – [For Relaxation] [For Play] [Specific Facilities] [±Next to a Block of Flats] [±Next to a House]

The most predictable reading:
'to play/spend time in the courtyard (with friends)'
WF Type: [SUBSTANCE$^{\underline{\text{Location}}}$ACTION]
Seme Level: 4

	NS	*NNS*	*Total*
FO:	13/20	13/20	26/40
Scores:	93/200	91/200	184/400
PR:	0.302	0.296	0.299
OPR:			0.604

The next two readings:
'to build a courtyard'
PR: 0.280 0.112 0.186
'to enclose a space'
PR: 0.039 – 0.010

Comments
The top reading 'to play/spend time in the courtyard (with friends)' offers another important piece of evidence that there are no significant differences between native and non-native speakers in terms of their meaning-prediction capacity. The PRs of the two groups of informants are almost identical: the difference is a mere 0.006!

The reading 'to play/spend time in the courtyard (with friends)' goes back to quite common, very vivid recollections of language users of their childhood: to spend one's leisure at a courtyard by playing football, badminton, volley-

ball and/or other games, or simply by meeting and talking to friends. Thus, an *extra-linguistic factor* – one's *experiences* – plays an important role in proposing this reading. The reading is level 4, the emphasis being put on the Location of courtyard, especially with regard to blocks of flats, because it is usually children of the same block that *courtyard*$_V$.

The PR Gap between this reading and the reading 'to build a courtyard' is much smaller – in fact, negligible – in the native group of informants. While with some other conversions the Factitive reading may be chosen as a last resort, the fairly high PR for a second-rank reading in the native group indicates that it does not apply to this case. The less so that there is an 'attractive' top-rank reading.

A large number of readings occur just once. 'To be a homeless/pavement dweller' better applies to the concept of 'pavement' or 'street' rather than courtyard. 'To play tennis' is *too specific*. It activates a level 5 seme by individualising the level 4 [For Play]. The same applies to 'to use a courtyard as a place for concerts'.

The reading 'to buy a courtyard' could be, in theory, applied to any Noun →Verb conversion by being *too general*. Therefore, the 'Change-of-Possession' reading does not seem to be a productive type of conversion.

Check-group results

Readings	*FO*
To play/spend time in the courtyard (with friends)	19/25
To waste time, to spend time by doing nothing, to hang around	1/25
To stay outside one's house, one's family	1/25
To enclose sth.	1/25
To have fun in a crazy way	1/25
To adapt accordingly a space between blocks of flats	1/25
To keep sth/sb safe	1/25

The topmost reading was given an extraordinary level of support from the check-group, with the other readings occurring only once. Interestingly, the reading 'to build a courtyard' does not occur among the readings proposed by the check-group. A possible explanation follows from the *Conclusions* in point 4.3.3 below: the Factitive reading is usually second in rating, following a reading with a higher acceptability. It may have been that when the check-group informants had to select one of the two or more options, the Factitive reading was overshadowed by a more predictable one.

4.3.2.10 *to morning*

Table 44. to morning – native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to get up (early) in the morning			5		4		7	6	4		7	9		X		5	9	6		X
– to do one's morning routine/to perform usual morning activities				8		4					9	X			7	8	8		X	
– to greet somebody in the morning/to say 'good morning'		6				6				8							X		5	
– to have breakfast														8			7			
– to work in the morning	3										8				6		4			
– to be sleepy														2	1					

Table 45. to morning – non-native speakers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
– to get up early in the morning	3	X			7	9	6	7			X			7	X		9		X	X
– to do one's morning routine/to perform usual morning activities	X			6					6	X		8	6				8			
– to greet somebody in the morning/to say 'good morning'	2			X		X		9	X						7	8				
– to make/have breakfast or morning coffee	8			7	6								6	8						
– to wake up somebody in the morning				8		4			8								5	X		
– to start a new day							4							9						
– to refresh something												5		9						

Type: Time
Courtyard
Seme level 1 – CIRCUMSTANCE
2 – [Time]
3 – [Part of Day]
4 – [±Starting from Midnight] [±Starting from Dawn]

The most predictable reading:
'to get up early in the morning'

WF Type:		[CIRCUMSTANCE$^{\underline{\text{Temporal}}}$ACTION]	
Seme Level:		4	
	NS	*NNS*	*Total*
FO:	12/20	12/20	24/40
Scores:	82/200	98/200	180/400
PR:	0.246	0.294	0.270
OPR:			0.601

The next two readings:

'to do one's morning routine; to perform usual morning activities'

PR:	0.128	0,095	0.111

'to greet somebody in the morning/to say 'good morning"

PR:	0.044	0.098	0.068

Comments

This conversion also demonstrates the *closeness of the meaning-prediction results* obtained from the two groups. This time the PR difference for the top reading is only 0.048.

The results reflect two different definitions of 'morning'. For some, the concept of morning is connected primarily with early hours, for others, 'morning' starts with dawn. The latter notion is relative, because it depends on the season of the year and constantly changes. Moreover, the notion of morning seems to be culture-bound. It would be interesting to make a public survey focused on the perception of 'morning' in different countries. Since in the U.S.A. and Britain business and school hours usually begin later than in the Central European region the notion of 'morning' is necessarily affected by this *pragmatic circumstance*: people in Slovakia are early risers compared to those in the USA and Britain. This indicates a significant role of the *extra-linguistic factor* affecting one's *experience*.

The reading with the highest PR seems to prefer the 'don't-wait-for-dawn' interpretation. As indicated above, 'early' may mean something different to people coming from different cultural settings. Furthermore, the *morning*$_{\text{V}}$ readings cannot be arrived at without certain *experience* with what is usually done in the morning.

The second-rank reading does not solve the temporal delimitation of 'morning'. Rather, it concentrates on the first activities performed when getting up, which – at least in Western civilisation – include hygiene-related activities. While these are more or less general some other activities like drinking coffee, applying make-up, or doing morning exercises are less common, and therefore treated separately by some informants. All the same, this entails the previously

discussed problem of a *fuzzy borderline* between the individual readings and their classification.

Check-group results

Readings	*FO*
To get up early in the morning	10/25
To do all the standard morning activities	7/25
To greet sb. in the morning	5/25
To do sth. in early morning hours	1/25
To participate in an all-night party	1/25
To 'wake up' after day-dreaming; to return to reality	1/25

The results of the check-group map those obtained from the main group's aggregate results, and, at the same time point out the above mentioned classification problems.

4.3.3 Summary 2

The results obtained in Experiment 2 are summarised in Tables to 46–48. Table 46 summarises the most predictable readings for all of the converted naming units used in Experiment 2.

Table 46. An overview of the most predictable readings

Naming Unit	*Predictability Rate*		
	NS	*NNS*	*Total*
boy			
'to act or behave the way boys do (immature)'	0.322	0.393	0.358
'to (try to) look/behave like a boy (clothes, haircut, walk) – of girls'	0.162	0.244	0.203
'to change the sex from female into male'	0.030	0.039	0.035
lion			
'to roar like a lion; to be aggressive'	0.143	0.330	0.237
'to have the power and rule as a lion'	0.106	0.144	0.124
'to behave/act like a lion'	0.066	0.031	0.047
tulip			
'to grow tulips'	0.350	0.233	0.290
'to pick tulips'	0.038	0.049	0.043
'to decorate with tulips'	0.024	0.009	0.016
cableway			
'to travel by cableway'	0.309	0.309	0.309!
'to build/construct a cableway'	0.319	0.160	0.232
'to lay cables'	0.026	–	0.007

Table 46. (*continued*)

planet			
'to travel to/land on other planets'	0.444	0.312	0.376
'to populate/colonize a planet'	0.078	0.100	0.089
'to discover a planet'	0.008	0.007	0.008
river			
'to swim/bathe in a river'	0.144	0.264	0.200
'to travel down a river (e.g., in a canoe)'	0.116	0.098	0.107
'to direct a river into an area in order to irrigate it'	0.030	–	0.008
conference			
'to take part in a conference'	0.380	0.476	0.427
'to (meet) to talk about/consult/discuss a subject'	0.294	0.140	0.212
'to organize a conference'	0.231	0.110	0.165
triangle			
'to draw a triangle'	0.158	0.253	0.203
'to make a triangle from something/ to make triangular things'	0.190	0.178	0.184
'to be a part of a triangle of lovers'	0.030	0.025	0.028
courtyard			
'to play/spend time in the courtyard'	0.302	0.296	0.299
'to build a courtyard/ to make an area into a courtyard'	0.280	0.112	0.186
'to enclose a space'	0.039	–	0.010
morning			
'to get up early in the morning'	0.246	0.294	0.270
'to do one's morning routine/ to perform usual morning activities'	0.128	0,095	0.111
'to greet somebody in the morning/ to say 'good morning"	0.044	0.098	0.068

4.3.3.1 *Native speakers vs. non-native speakers*

Let us first compare the data obtained from native and non-native speakers. The most valuable acknowledgement of hypothesis 4. (Section 3.11) (and confirmed in Experiment 1), saying that the results for native speakers should not significantly differ from those for non-native speakers, comes from almost identical proposals of most predictable readings for the individual converted naming units (in principle, the differences bear on the number and the range of very low-occurrence, usually single-occurrence, readings), and, even more important, from the almost total agreement in their ranking. Thus, the two groups identified the same rank 1 reading for 8 of 10 conversions. In one (*cableway*) of the two exceptions, the difference between rank 1 and rank 2 in the NS group is as low as 0.010; in the other case (*triangle*), it is 0.042.

Even more impressive is the fact that in eight cases the two groups agree in the identification of the ranking of two topmost readings. Furthermore, the differences between the two groups in terms of the PR value are insignificant: the average difference for rank 1 readings is as low as 0.087, for rank 2 readings it is 0.077, and for rank 3 readings 0.034.

In addition, the absence of any major differences in the results of the two groups of informants is an important piece of evidence for the assumption that both native and non-native speakers make the same use of extra-linguistic knowledge and experience. Certainly, differences in culturally similar communities may arise if an object is named that is unknown and/or not used in one of the speech-communities, as it follows from the discussion on *pooper-scooper* in Section 4.2.2.2. No similar case has occurred in my conversion-oriented experiment.

4.3.3.2 *Meaning predictability and dominant readings*

An analysis of the data shows that there is only one case in which the PR is above 0.4 ('to take part in a conference'). Compared to Experiment 1, where the PR values were scattered and ranged from 0.098 to 0.727, the PR values of recategorised naming units seem to be more consistent, with the PR range being much narrower: from 0.200 (*river*) to 0.427 (*conference*). As to the distribution, one reading is above 0.4, three above 0.3, and 6 above 0.2. The lowest PR of a top reading is exactly 0.200 ('to swim/bathe in a river'). None of the readings reached the medium predictability level. Seven readings are above the predictability level (0.25) three top readings fall below it. In spite of the different structure of the PR values in Experiments 1 and 2, their average Predictability Rates do not differ markedly: 0.297 for Onomasiological Type 5 vs. 0.322 for Onomasiological Type 3.

The majority of non-top readings fall below 0.2. The PRs of the third-rank readings are well below 0.1, the only exception being *conference*, whose rank 3 reading's PR is 0.165.

In seven out of the ten experimental conversions, the top reading surpasses the rank 2 reading by more than 0.1. In this respect the 'strongest' predictability position can be attributed to 'to grow tulips'. Its PR is higher by 0.257 than that of the rank 2 reading for *tulip*$_V$. What strikes one is that the PRs of the third-rank readings are well below 0.1, the only exception being *conference* whose rank 3 reading's PR is 0.165.

The average PR Gap between the rank 1 and rank 2 readings of the ten sample naming units is 0.148. The average PR Gap between the rank 1 and rank 3 readings is immense: 0.258.

An even clearer picture of the dominant position of the most predictable readings is obtained if the R1/R2 and the R1/R3 ratios are calculated. The rank 1 PR value is twice (exactly, 1.99 times) higher than that of the rank 2 reading. The average gap between the most predictable readings and the third rank readings is abysmal: the average PR of the former is 7.57 times higher.

All these statistical data provide an unambiguous support to hypothesis 7. (Section 3.11) assuming one (rarely two) dominant reading for each novel, context-free naming unit. This conclusion has far-reaching consequences for meaning predictability. In particular, it may be assumed that the difference between the PR of the top reading and the PRs of the lower-rank readings influences the predictability strength of the most predictable reading. Therefore, this variable should be taken into consideration in evaluating the meaning-predictability, and is reflected in the Objectified Predictability Rate. The relevant results are commented on in Section 4.3.3.5.

Let us have a closer look at the 'strongest' top readings in the sample.

The WF Type underlying the reading 'to take part in a conference' (0.427) is [PROCESS$^{\underline{\text{Purpose}}}$ACTION]; the WF Type of 'to travel to/land on other planets' (0.376) is [SUBSTANCE$^{\underline{\text{Target/Direction}}}$ACTION]; the WF Type of 'to act or behave the way boys do (immature)' (0.358) is [SUBSTANCE$^{\underline{\text{Manner/Pattern}}}$ACTION]; and finally, that of 'to travel by cableway' is [SUBSTANCE$^{\underline{\text{Instrument}}}$ACTION]. The diversity of WF Types underlying these readings might indicate that their 'success' is not bound to any specific WF Type. Rather, the WF productivity of each of these WF Types establishes general, but necessary, conditions for the interpretation of the respective naming units in the manner they are interpreted.

The relation between WF Types and meaning predictability features certain tendencies. In particular, the nature of an object to be named appears to influence the selection of a WF Type. Based on the experimental data, it may be supposed that language speakers make use this intuitive knowledge in the meaning-prediction process. For illustration, artefacts are preferably converted in the Factitive-mode (*cableway*, *triangle*) and/or Instrumental mode (*cableway*). An overview of the relations observed in my research is summarised as Tendency 7 in 4.3.3.6. It is only in this sense that Ryder's (1994) templates, designed for the field of primary compounds, may perhaps be of some (rather limited) relevance here.

Making any more conclusive generalisations on the relation between the WF Type and the PR is, however, preconditioned by implementation of a large-scale research focussed on this particular issue. Such a research should answer several basic questions: Do particular nouns (or better, lexically determined groups of nouns) prefer a particular direction of conversion (i.e., WF Type)?

If yes, is there any correlation between this preference and a particular level of meaning predictability? Is the preference for a particular WF Type characteristic only of the most predictable reading of such a group of nouns, or is there any broader pattern identifying a typical WF Type for each rank within the group?

For more specific reasons behind the dominance of the above-mentioned readings, let us examine the role played by prototypical semes in the interpretation process. Thus the reading 'to take part in a conference' is motivated by the level 4 semes [Presentation of Information] and [Exchange of Information] that specify the general semantic category of Purpose. While motivated by this top reading, the rank 2 reading of *conference*, i.e., 'to (meet) to talk about/consult/discuss a subject', is semantically shifted and generalised, and hence not bound to the prototypical features of conference (concerning special ways of organisation, topic-selection, course, etc.). In this respect it is not motivated by a level 4 seme; rather, the motivation is bound to a higher level – level 3. Even if the PR of this reading is above the predictability level, the PR Gap between this and the top *conference* reading is more than significant, i.e., 0.215.

The reading 'to travel to/land on other planets' outscores the remaining readings of *planet* in a most significant way. Its basic semantics expressed by the category of Direction is specified further by the motivating level 4 semes [Solid] and [±Reachable]. The '±' mark should be interpreted in a relative and conditional way: while only few planets can be reached by mankind at this level of technological development many others *can* be reached in human dreams, science-fiction, or long-term visions. In this respect, the 'reachability', including the meaning of 'travelling to and landing on a planet', is both a level 4 seme and, at the same time, a very strong extra-linguistic reading-motivation factor, contributing to the relatively high PR of this reading, especially in the era of closely monitored space flights.

It appears that the considerable PR Gap between this and the next reading for *planet* (highlighting the colonisation of planets) appears to bear on the aspect of implementability. Language users, when interpreting the naming unit *planet*, may be guided by the 'closer-to-reality' principle, thus preferring the more realistic option and/or assigning it much higher scores.

The reading 'to act or behave the way boys do (immature)' is motivated by the level 4 [–Adult] and [Characteristic Behaviour] which 'detail' the category of Pattern/Manner. Here, however, one cannot account for the significant gap of 0.155 between this and the next reading 'to (try to) look/behave like a boy (clothes, haircut, walk) – of girls' by a different seme level motivation. The latter reading is also motivated by two level 4 semes, i.e., [Male] and

[Characteristic Behaviour]. The explanation must be sought in extra-linguistic reality, our world knowledge and experience. Out of the two contrasting relations [+Adult] vs. [–Adult] and [+Male] and [–Male] the latter still appears to be based on a more solid basis, resting upon much longer traditions and the role of the two genders in society. In spite of what has been mentioned above in connection with the emancipation process, the gender distinctions are perceived to be crucial, including the male-female role opposition in everyday life. Therefore, it may be surmised that while the favourable predictability conditions (productive WF Type, level 4 seme motivation) helped this reading to rise above the predictability level (0.203) it is clearly outscored by the other reading with equally favourable predictability conditions, further supported by relevant extra-linguistic factors.

The gap between the two top readings of *cableway* is relatively small. The top reading 'to travel by cableway' corresponds with my expectations: the general semantic category Instrument is supported by the level 4 seme [Means of Transportation]. The second-rank reading 'to build/construct a cableway' is motivated by level 3 [Equipment] which implies the seme [Constructable]. However, all [Artefacts] (level 3) are constructable and producible. Therefore, it comes as no surprise that the reading motivated by this general seme played a relatively important role in the meaning-prediction process of the native speaker group.

An important role of world knowledge and experience in the interpretation process can be further illustrated by the dominant position of 'to grow tulips' (0.290). The PR of the next lower reading ('to pick tulips') is merely 0.043, which gives the PR Gap of as much as 0.247. Interestingly, both of these readings are motivated by level 3 [Growable] and [Pickable], respectively. The small PR Gap and, simultaneously, the top position of the level 3 seme motivated readings can be accounted for by the features of tulip and their perception by language users. Like many other flowers, 'tulip' arouses a number of disparate connotations, and functions as a symbol (as do many other flowers). Due to this speaker-bound figurativeness and 'connotability', the number of its possible readings (motivated by level 4 semes) is very high, which results in an extreme scattering of the readings proposed, enabling one central reading (motivated by a level 3 seme) to take an absolutely dominant position.

Let us finally comment on the smallest PR Gap between two top readings. This is the case of 'to draw a triangle' (0.203) and 'to make a triangle from something/to make triangular things' (0.184). The PR Gap of only 0.019 contradicts the above-indicated tendency of one clearly dominating reading for each first-encountered converted naming unit. This finding cannot be ac-

counted for by the identical WF Type [QUALITY$^{\underline{\text{Result}}}$ACTION], based on the Factitive relation, because, for example, the above-discussed difference between 'to act or behave like boys do (immature)' and 'to (try to) look/behave like a boy (clothes, haircut, walk) – of girls', also characterised by the same WFT, is more than one predictability level (0.155). With *triangle* the decisive factor appears to be the competition of more or less equally probable readings.

4.3.3.3 *Seme level*

Nine of ten most predictable readings are motivated by a level 4 seme. This seems to justify the hypothesis of the important role of level 4 semes in the meaning-prediction process.

The only top reading motivated by level 3 is 'to grow tulips'. The Result (tulip) of Action (growing) follows from the feature [Growable]. Certainly, this feature is not typical of tulips only. It may be assumed that the reading based on level 3 seme [Growable] would be a preferred direction for the majority of conversions in the lexical group of plants.

4.3.3.4 *Main group vs. check group*

Table 47 compares the results obtained from the main group and the check group of informants.

Table 47. Comparison of the two main groups and the check-group: rankings of the most predictable readings

Converted NU/reading	*Ranks*	
	Main group	*Check-group*
boy		
1. 'to act or behave like boys do (immature)'	1	1
2. 'to (try to) look/behave like a boy (clothes, haircut, walk) – of girls'	2	2
lion		
1. 'to roar/shout; to be aggressive'	1	1
2. 'to have the power and rule/to act as a leader'	2	3
tulip		
1. 'to grow tulips'	1	1
2. 'to pick tulips'	2	–
cableway		
1. 'to travel by cableway'	1	1
2. 'to build/ construct a cableway'	2	5–8
planet		
1. 'to travel to other planets/to land on a planet'	1	1
2. 'to populate/colonize another planet'	2	2

Table 47. (*continued*)

river		
1. 'to swim in a river'	1	1
2. 'to travel down a river (e.g., in a canoe)'	2	2
conference		
1. 'to take part in a conference'	1	1
2. 'to meet to talk about/consult/discuss a subject'	2	2
triangle		
1. 'to draw a triangle'	1	2–4
2. 'to make a triangle from something/to make triangle-shaped things'	2	1
courtyard		
1. 'to play/spend time in the courtyard (with friends)'	1	1
2. 'to build a courtyard/to make an area into a courtyard'	2	–
morning		
1. 'to get up early in the morning'	1	1
2. 'to do one's morning routine/to perform usual morning' activities	2	2

In spite of the fact that the research techniques applied to these two groups differed substantially, Table 47 reveals a significant ranking identity of the most predictable readings in the main and the check groups. As for the most predictable readings, agreement between the two groups obtains in nine out of ten cases. The only exception is *triangle* where the two most predictable readings have swapped their positions. This is not surprising, as the difference in the PRs of the two readings in the main group is very small (0.019). The agreement is especially remarkable with *conference*, where the respective rankings of the top three predictable readings coincide fully.

A big gap between the main group of informants and the check-group seems to exist in the case of 'to build/ construct a cableway' (rank 2 in the main group, rank 5–8 in the check-group); however, a closer view reveals that all the readings in the check-group, except for the unambiguously dominating 'to travel by cableway', feature minimum frequency of occurrence: the readings 'to move up, to progress and advance' and 'to haul/help somebody during a climb', which share the second position, occurred three times, and the rank 4 'to help sb make progress in a particular activity' occurred only twice.

4.3.3.5 *Objectified Predictability Rate*

The Objectified Predictability Rates for the individual sample naming units and their comparison with the Predictability Rates are given in Table 48.

Table 48. Top reading ranking, PR Gaps, Objectified Predictability Rates

Most predictable reading for each naming unit	*PR*	*PR-based ranking*	*PR Gap*		*OPR*	*OPR-based ranking*
			Absolute value	*R1/R2*		
boy	0.358	3	0.155	1.764	0.601	5/6
lion	0.237	8	0.113	1.911	0.581	7
tulip	0.290	6	0.247	6.744	0.831	1
cableway	0.309	4	0.077	1.332	0.564	8
planet	0.376	2	0.287	4.224	0.795	2
river	0.200	10	0.093	1.869	0.634	3
conference	0.427	1	0.215	2.014	0.531	9
triangle	0.203	9	0.019	1.103	0.489	10
courtyard	0.299	5	0.113	1.608	0.604	4
morning	0.270	7	0.111	2.432	0.601	5/6

What was stated in Section 4.2.3.4 when commenting on the Objectified Predictability Rates in terms of the results of Experiment 1 is borne out in Experiment 2. In particular, the OPR values depend on the PR Gap, and primarily on the related R1/R2 ratio. This means that the OPR-based ranking need not, and actually does not, correspond with the PR-based ranking. The most striking examples are the naming units *tulip* and *river*. An unusually high OPR value (0.831) of the top reading of *tulip* (whose PR value is relatively low – rank 6) results from an extremely high R1/R2 ratio (6.744), supported by a high PR Gap (0.247). With *planet*, the situation is different. Its 2nd PR rank is 'confirmed' by the same OPR rank thanks to a very high R1/R2 ratio as well as the highest PR Gap (0.287) from among the ten sample naming units. On the other pole, we can observe a big difference between the PR-rank (1) and OPR-rank (9) of the top reading of *conference*. This cannot be accounted for by the R1/R2 ratio as it is fairly high (2.014), and the same is true of the PR Gap (0.215). The reason for the 'poor' OPR-position of the top reading of *conference* is connected with an extremely high PR value of the 3rd rank reading of this naming unit. Consequently, the PR Gap between rank 1 and rank 3 readings is relatively very small, and the R1/R3 ratio is several times smaller than an average R1/R3 value for the naming units of this Onomasiological Type. Further details of the interrelation between the PR and OPR values, and the influence of the R1/R2 ratios as well as the PR Gap are given in Table 48.

If we now compare the OPRs of the top readings in Experiment 1 and Experiment 2, the average is 0.572 for Onomasiological Type 3 and 0.623 for Onomasiological Type 5, which means that while the average PR is higher in Onomasiological Type 3 (let us recall that the average PR of Onomasiological Type 5 is 0.297, for Onomasiological Type 3 it is 0.322), the OPR ranking is

reversed. While the difference is not significant (0.051) in this case either, the reversed proportion might indicate that the strength of the competing readings in the Type 5 sample of naming units is smaller, which contributes to the strength of the top readings. A comparison of the second- and third-rank readings reveals that the average PR of rank 3 readings is 0.032 for Onomasiological Type 5 and 0.078 for Onomasiological Type 3; the average PR of rank 2 readings is 1.491 for Onomasiological Type 5 and 1.712 for Onomasiological Type 3. These data suggest that the responsibility for the higher average OPR value of Onomasiological Type 5 can be attributed primarily to the 'weakness' of second-rank readings of Onomasiological Type 5. In other words, the PR Gaps between top readings of OT5 naming units and the next two readings (and especially between rank 1 and rank 2) are higher than those in Onomasiological Type 3. This is further supported by the higher average R1/R2 ratio (1.99 for Onomasiological Type 5 vs. 1.88 for Onomasiological Type 3) as well as R1/R3 ratio (7.57 for Onomasiological Type 5 vs. 4.14 for Onomasiological Type 3).

Consequently, it may be assumed that there is a *tendency for novel, context-free naming units of Onomasiological Type 5 to be slightly more predictable than those of Onomasiological Type 3*.

4.3.3.6 *Tendencies*

The previous discussion may be summarised in the following tendencies for the meaning predictability of novel, context-free conversions, that is naming units falling within Onomasiological Type 5:

Tendency 1: One dominant reading.

Tendency 2: Its PR is above 0.250 (the average value for the ten top readings in the sample is 0.280).

Tendency 3: The PR Gap between the first and the second readings usually exceeds the value of 0.1 (the average PR Gap for the ten top readings in the sample is 0.148). The average R1/R2 ratio of 1.99 is significantly high and gives support to Tendency 1. The PR values of the rank 3 and lower rank readings are insignificant and negligible: the average PR for rank 3 readings in the sample is 0.027, and the R1/R3 ratio is as high as 7.57.

Tendency 4: An important, though not a sufficient condition for meaning predictability is level 4 of the motivating seme; in other words, a 'good' PR tends to be conditioned by prototypical semes.

Figurative meanings rarely exceed the predictability level (exceptions in my research are, e.g., 'to roar/shout; to be aggressive', 'to have the power and rule/to act as a leader' for *lion*). It should be, however, emphasised that the activation of a level 4 seme itself does not guarantee a high PR.

Tendency 5: Predictability is influenced by extra-linguistic factors, such as world knowledge and experience.

Tendency 6: In culturally similar communities, the meaning-prediction process tends to bring similar results, no matter whether the informants are native speakers or non-native speakers of a language. In other words, the degree of their agreement in identifying the most predictable readings of novel, context-free converted naming units is very high.

Tendency 7: The relation between WF Types and meaning predictability reveals certain tendencies:

(a) Converting nouns denoting objects with behaviour tend to convert in the Manner/Pattern mode, i.e., the converting unit functions as a Pattern and determines the Manner of Action of the converted naming unit. This is the case of the predictable readings of *boy* and *lion*.

(b) Conversion of naming units denoting natural objects (*planet, river*) and artefacts of relevant size (*courtyard*) usually rests upon the Location/Direction mode.

(c) Artefacts, in general, are also convenient for the Factitive-mode conversion (*cableway*, *triangle*) and the Instrumental mode (*cableway*).

(d) Events are naturally bound to the Purpose mode (*conference*).

(e) Time-related naming units usually convert via the Temporal mode (*morning*).

4.4 Experiment 3

4.4.1 Sample naming units

In addition to the exploration of the predictability-conditioning factors discussed and assessed in Experiments 1 and 2, Experiment 3 was primarily designed to apply the computation of the Objectified Predictability Rate to non-established naming units covering all five Onomasiological Types. The sample of informants included 90 native speakers of English, the vast majority of them university undergraduates from Britain and the U.S.A. The underlying principle of this experiment was the same as in the previous two experiments, i.e., the informants were asked to propose as many readings for each of ten possible naming units as they could think of and to assign them a score indicating the degree of acceptability.

A group of 25 non-native speakers, Polish undergraduates, functioned as a check-group, under the same conditions as the check-groups in the first two experiments. This time each Onomasiological Type was represented by two naming units with the aim of including in this experiment a range of various WF Types and Morphological Types. The possible naming units selected for the research are given in (48):

(48)

OT1	Word-Formation Type
blondesjoker	[Theme – Action – Agent]
feather-dialler	[Instrument – Action – Agent]
OT2	
contactee	[Action – Patient]
refusnik	[Action – Agent]
OT3	
anthraxist	[Instrument – (Action) – Agent]
smile-man	[Stative – (State) – Patient][11]
OT4	
removage	[Action – Process]
befoot	[Action – Object]
OT5	
obsess$_{N}$	[STATE$^{\underline{\text{Patient}}}$SUBSTANCE]
leave-behind$_{N}$	[ACTION$^{\underline{\text{Result}}}$SUBSTANCE]

In fact, all ten naming units were proposed with a certain specific meaning in mind. This meaning determined the WF Type and Morphological Type of each of these naming units. Thus, the experiment was also designed to verify

the degree of coincidence between the informants' readings and the respective intended meanings of the coiner. In this way, the word formation and the word-interpretation processes were interrelated.

Four of these naming units fall within the Word-Formation Type Cluster of *Agents*. The intended readings with which I proposed these naming units classify them to different WF and Morphological Types:

(49)		*Word-formation Type*	*Morphological Type*
	blondesjoker	[Object ← Action – Agent]	$[N\ V\ \text{-er}]_N$
	feather-dialler	[Instrument – Action – Agent]	$[N\ V\ \text{-er}]_N$
	refusnik	[Action – Agent]	$[\theta\ V\ \text{-er}]_N$
	anthraxist	[Instrument – (Action) – Agent]	$[N\ \theta\ \text{-ist}]_N$

In addition, I smuggled into the list of possible naming units one actual naming unit, *refusnik* 'a Soviet citizen, especially a Jew, who is refused permission to emigrate' (Webster New World Dictionary. Third College Edition 1988). The purpose was to test (1) the meaning predictability of a naming unit which does not belong in the core vocabulary (and for this reason, needn't be a part of the vocabulary of the informants); (2) the effect of the foreign suffix *-nik* (with the meaning 'a person who rejects standard social values' (Bauer 1983:87)) upon the meaning-prediction process, that is to say, to what degree native speakers are aware of its meaning if they do not know the actual meaning of a naming unit (*refusnik* in this case); (3) how native speakers dispose of the meaning-prediction task if the highly productive WF Type of [Action – Agent] is assigned a 'peripheral' morphological realisation.

4.4.2 Experimental data and their analysis

Given the extensive research, including 90 informants, this section does not present space-consuming tables, including the proposals of the informants. Therefore, the evaluation of the meaning predictability of the individual naming units is presented, as it were, *in medias res*.

4.4.2.1 *blondesjoker*

'A person who tells/makes blonde jokes'

FO:	76/90
Scores:	452/900
PR:	0.422
OPR:	0.630

'A blonde-haired humorist, comedian, joker'
FO: 58/90
Scores: 341/900
PR: 0.242

'A joker in a card pack'
FO: 13/90
Scores: 36/900
PR: 0.006

Comments
It was assumed in Section 3.5.1.1 that the meaning predictability of coinages falling within Onomasiological Type 1 is very high and may approach 100%, because the relation between the polar members of the onomasiological structure, that is to say, between the onomasiological base and the determining constituent of the onomasiological mark, is (usually) unambiguously specified by the Actional constituent of the onomasiological structure. This is true in cases like *piano-player* and *novel writer*, where the relation between the Animate base, represented by the suffix *-er*, and the Inanimate Object (*piano*) and the Result of Action (*novel*), respectively, is explicitly specified by *play* and *write*, respectively. However, it was also indicated that meaning predictability may be affected by the ambiguity of the onomasiological base, as in the case of *apple-eater* where the base may stand for a [Human] Agent, an [Animal] Agent, or a shifted meaning ('must').

The present instance illustrates another type of ambiguity that impairs the Predictability Rate – the ambiguity of the determining constituent of the onomasiological mark. Thus, *blonde* in this particular case may be either an Object of joking (the 'intended' meaning) or Quality (i.e., the colour of the Agent's hair). It is interesting that a large number of informants proposed both of these readings. In addition, and this seems to be a *PR-reducing factor*, out of 15 informants who did not propose the former reading 13 proposed the latter competitive reading. In any case, the data indicate that rather than the Predictability Rate it was the Objectified PR that has been significantly reduced by the *competition* of the two readings.

Of the remaining readings there is only one that deserves attention in terms of the frequency of occurrence (thirteen instances), but not in terms of the scores assigned (mostly 1 or 2 points): 'a joker in a card pack'. This reading results from the activation of a *shifted*, and importantly, *institutionalised*, meaning of *joker*. Its low PR can be accounted for by *joker*'s poor compatibility with *blonde*. There are hardly any card packs with a blonde joker. Thus, the

world knowledge and *experiences* of the informants seem to have played down this interpretation.

Check-group results

A person who tells/makes blonde jokes	21/25
A blonde-haired humorist, comedian, joker	4/25

The check group results coincide with those of the main group. It is remarkable that only these two readings were proposed.

4.4.2.2 *feather-dialler*

'A feather(-like instrument) used for dialling the phone'

FO:	62/90
Scores:	277/900
PR:	0.212
OPR:	0.752

'A type of telephone that requires a very light touch to dial'

FO:	25/90
Scores:	147/900
PR:	0.045

'A person who dials a telephone with a feather'

FO:	20/90
Scores:	101/900
PR:	0.025

Comments

This naming unit was coined with the onomasiological base representing an Agent, expressed by the suffix *-er*. To my surprise, sixty-two native speakers interpreted this suffix as an Instrument. This interpretation of *feather dialler* has the following onomasiological structure:

(50) [Instrument_2 – Action – Instrument_1]

in which the determining constituent of the mark functions as a part, or a closer specification, of the base. In other words, feather is a part of the Instrument for dialling telephone numbers, or the Instrument *is* (made of) a feather. It may be surmised, however, that the latter reading would be – by most coiners – more probably realised by a more productive WF Rule [Action – Agent], onomatologically realised as *dial(ling) feather*.

As a result, the 'intended' meaning has not received support sufficient to get above the predictability level. Only 20 informants voted for this interpre-

tation, and even those who proposed this reading valued it rather low. The average score is 4.9 pts., which gives a poor PR of 0.025. What strikes me is the fact that the Agent-based reading seems to be more acceptable from the point of view of extra-linguistic reality: a freak dialling with a feather appears to be much closer to the reality than a feather designed specifically for dialling a phone.

The *ambiguity* is also present at the other pole of the onomasiological structure. Feather may be conceived of as an Instrument (as in the 'intended' Agentive-based WF Type, or in the dominating reading – an Instrument of an Instrument), as Material ('a dial made of feathers'), or Pattern/Manner ('a type of telephone that requires a very light touch to dial'). This ambiguity can also be held responsible for stealing several percentage points from the most predictable reading.

Of the other Agentive readings the most significant support was given to 'a person who dials very softly', in which case the prototypical feature [Light] of 'feather' matches with the Action of dialling that can be executed in the proposed 'soft' way. This reading is an Agentive variant of the second-rank reading.

Check-group results

An extreme scattering of readings is illustrated by the following review of the proposed readings.

A person who dials phone numbers by a feather	3/25
Sth. in the shape of feather used to dial numbers	2/25
A person who always looks for details in everything	1/25
A dialler made of feather	1/25
A very light-weight dialler	1/25
A person who has no problems with keeping in touch with other people	1/25
A person who collects feather	1/25
A special device for removing feather from ducks in a gentle way	1/25
A person who selects appropriate feathers for pillows	1/25
A person who collects feathers by phoning his suppliers	1/25
A producer of duvets	1/25
A person who deals with feathers of killed birds	1/25
A young person dialling a phone number	1/25

Nine informants were unable to propose any reading for this naming unit. This might indicate that meaning predictability may, in more demanding cases, be partly affected by the time factor.

4.4.2.3 *contactee*

'A contacted person, a person that can be contacted (for various purposes)'

FO:	89/90
Scores:	807/900
PR:	0.897
OPR:	0.988

'A person who wears contact lenses'

FO:	13/90
Scores:	32/900
PR:	0.006

'A person who does the contacting'

FO:	7/90
Scores:	54/900
PR:	0.005

Comments

The top reading of this naming unit features the highest OPR of all the naming units in my four experiments, and the frequency of occurrence is almost at the maximum. Only one informant did not propose this reading (the Agentive reading was preferred by him/her). This may be related to the high productivity of Onomasiological Type 2, the WF Type [(Agent) – Action → Object] (with the general meaning of 'a person who is affected by Action') as well as the Morphological Type [Verb + *-ee*]$_N$. The competing Agentive reading ('a person who performs Action') was only proposed by five informants. The 'unintended' reading based on 'contact lenses' occurs slightly more frequently (even if with very low scoring in general).

Thus, an extremely high harmony of the described circumstances contributed to the extremely high Predictability Rate of this coinage.

Check-group results

A contacted person, a mediator, one we can reach if necessary (to provide information, etc.)	15/25
A person whose task is to contact other people (various purposes, including a spy, business, etc.)	6/25
A person from the underground providing information to police	2/25
A rival in the contact kickbox	1/25
A person who gets information from others	1/25

Like in the case of *blondesjoker*, the main and the check-group results coincide in the most predictable reading. No reading concerning contact-lenses

occurred, which can be explained by there being only one really acceptable option for the informants.

4.4.2.4 *refusnik*

'A person who usually refuses to do something, a stubborn person, a denier'

FO: 57/90
Scores: 402/900
PR: 0.283
OPR: 0.823

'A political dissident (from the former USSR, or from a totalitarian regime)'

FO: 21/90
Scores: 204/900
PR: 0.053

'A person who rejects obedience to conventional rules, laws, orders, and/or to be loyal'

FO: 9/90
Scores: 71/900
PR: 0.008

Comments

This is one of the most interesting cases in my research. As indicated above, this is the only actual naming unit in the sample. Not all actual naming units are of the same 'value' in terms of the degree and range of their institutionalisation in a speech community. 'Institutionalisation' and 'speech community' are vague and relative notions. A particular naming unit may be accepted by and therefore institutionalised in an (almost) complete community of native speakers – this is the case of core (central) lexical units like *father*, *boy*, *table*, *bread*, *sky*, *good*, *go*, *run*, etc. Some other lexical units may be only known to (understood/used by) experts in a particular field (highly specialised terms in various branches of science). Still another one may be a 'property' of a small group of friends, classmates, etc. (various slang words). In general, the notion of partial or *minor speech community* is determined by a range of sociolinguistic factors, including age, profession, education, social status, etc. Crucially, in all of these cases the respective naming units fulfil their basic communication function within the specific (partial) speech communities in which they have been institutionalised. For the remaining part of the speech community they function as non-existing words. In our present case the naming unit *refusnik* is somewhere at the periphery of the lexicon, institutionalised in a limited group of speakers.

Clearly, the reasons for the 'unpredictable' status of the actual meaning may be sought in a *sociolinguistic factor*, in particular the *age of informants*. My informants were, with few exceptions, young university students who do not seem to have had any opportunity to learn about Soviet (Jewish) citizens and/or political dissidents not allowed to emigrate from the former USSR. More than a decade ago, when the Soviet Union broke down into a number of independent countries, my informants were children. The results indicate that the bulk of the informants do not know this naming unit, which is only familiar to those who are keen on politics and/or the history of the Soviet Union. One of the informants even commented on this naming unit with words "not an English word"! This explains why the PR of the 'institutionalised' reading is as low as 0.053. On the other hand, it is not surprising that its average score is extremely high (9.2 pts.): those who knew this naming unit (21 informants) had no reason to assign it less than 10 points. From this point of view, assigning less than 10 points to this reading by three informants is surprising.

What may be regarded as a *linguistic factor* is the 'feeling' of some informants for the Russian origin of the suffix *-nik*. This follows from proposals like 'a mock name for Russian', 'a Russian-sounding way of saying refuse', '(jokingly) a recent immigrant in the US of Russian, Ukrainian or other former Soviet extraction', 'Russian foreign policy of refusing US demands', and 'a (Russian) pacifist'.

The circumstances just described made the majority of informants look for a non-actual interpretation. Two basic paths of search were pursued. The first one resulted in the dominant position of the reading '*a person who usually refuses to do something*, *a stubborn person*, *a denier*' which gives witness to the correct Agentive perception of *-nik* and corresponds to the general meaning of this suffix 'a person who usually refuses to do something, a stubborn person, a denier'. For this reason, this reading is not a surprise.

The next path brought the informants to various low-frequency readings connected with 'rubbish', 'refuse', 'garbage', for example, 'garbage man', 'a person who goes through the garbage', 'a person who produces lots of garbage', 'a special type of refuse or trash', 'a trash can'.

Check-group results

A person who always/usually refuses (various circumstances)	16/25
A person who rejects obedience to conventional rules, law, and/or to be loyal	4/25
A device used to inform of one's refusal	1/25
A pessimist	1/25

A meeting which ends with a refusal of the offer	1/25
A keyboard button for cancelling a function	1/25
An act of refusal	1/25

The check-group confirmed the top position of 'a person who usually refuses to do something, a stubborn person, a denier' as well as the peripheral position of the institutionalised reading in the English word-stock. In addition to the above-mentioned reasons for the low frequency of the latter reading, one more reason can be assumed for the group of Polish informants. There was hardly any opportunity to learn this word at language classes in a non-English speaking country.

4.4.2.5 *anthraxist*

'A terrorist who uses anthrax as a weapon (e.g. by mailing it in envelopes)'

FO:	76/90
Scores:	608/900
PR:	0.570
OPR:	0.659

'A person who works with/produces/studies anthrax'

FO:	58/90
Scores:	410/900
PR:	0.293

'A fan of a heavy-metal group 'Anthrax"

FO:	5/90
Scores:	33/900
PR:	0.002

Comments

Expectations may sometimes be deceptive and rather different from reality. I included this naming unit in the sample because I was sure that its Predictability Rate would be high. Who does not know about the events of September 11, 2001? Now, if one looks at the OPR (0.659) of the expected reading ('a terrorist who uses anthrax as a weapon') (s)he can be surprised at the value not being extremely high. The answer is both simple and complicated, and must be sought in both *linguistic* and *extra-linguistic factors*.

Given the impact of the tragic event of September 11 upon mankind as a whole and the expected general awareness of the above-mentioned tragedy (informants' *world knowledge*) both the average score (6.76 pts) and the number

of occurrences (76/90) seem to be rather low: it follows from the data that 14 of 90 informants (16%!) did not propose this reading.[12]

On the other hand, a significant *predictability-reducing role* is played by Onomasiological Type 3 representing an onomasiological structure *without any Actional constituent*, which allows for multiple relations between the polar members of the onomasiological structure – in the same way as with primary compounds. This ambiguity is aggravated by a strong *competition pressure* on the part of a productive WF Type [Object ← (Action) – Agent] as well as Morphological Type [N + *-ist*]$_N$. By implication, the interpretations 'a person who works with/produces/studies anthrax' are fully justified and, admittedly, expectable.

As a result, two contradictory tendencies have affected the results for this naming unit: *world knowledge* supporting the 'terrorist'-oriented reading, and the *linguistic pressure* admitting productive coining of naming units with the multiplicity of relations between the polar members of an onomasiological structure based on Onomasiological Type 3.

Check-group results

A terrorist mailing anthrax in envelopes	15/25
A person who tests/produces/distributes anthrax	8/25
A person who fell ill due to anthrax	1/25
A life-threatening bacteria	1/25

The proposals of the check-group result in a perfect match between the two groups of informants.

4.4.2.6 *smile-man*

'A person with a smile; a person who smiles a lot, an optimist'

FO:	77/90
Scores:	593/900
PR:	0.564
OPR:	0.851

'A person whose job is to make people smile, comedian, clown'

FO:	30/90
Scores:	191/900
PR:	0.071

'A yellow-faced symbol on T-shorts, a smiley-face icon (Happy Face)'

FO:	19/90
Scores:	120/900
PR:	0.028

This naming unit admits two semantic structures: either the 'static' reading, i.e., [Stative – (State) – Patient], which corresponds to the interpretation 'a man with (i.e., 'who has' → State) a (constant, frequent) smile in his/her face', and which fits a productive Morphological Type [N – (State) – *man*]$_N$ (*oilman, policeman, milkman, salesman, craftsman*); or, a dynamic reading which corresponds to the interpretation 'a man who (constantly, frequently) smiles', i.e., [Action – Agent], demonstrating the Morphological Type [V – *man*]$_N$. These two readings belong in two different Onomasiological Types, which should have serious repercussions upon their meaning predictability.

As assumed in Section 3.5.1, the general predictability level of Onomasiological Type 2 is much higher than that of Onomasiological Type 3 owing to the ambiguity of the latter (due to the absence of an Actional constituent at the onomatological level) resulting in a number of competing readings. The reason why the two readings, falling within two different Onomasiological Types, were classified as a single reading is that a large number of informants gave a combination of these interpretations within one and the same reading proposal.

A fairly high frequency of occurrence contributes to the medium predictability level of this reading. This suggests that rather than by Onomasiological Type 3 the informants were motivated by a much more productive OT2-based WF Type functioning as a *predictability-boosting factor* thanks to the *Actional nature* of the *smile* constituent. One way or another, the strength of this reading is extraordinary (OPR: 0.851) as a result of a high PR value, an unusually big PR Gap (0.493), and an extremely high R1/R2 ratio (7.94). Consequently, the reading 'a person with a smile; a person who smiles a lot, an optimist' for *smile-man* ranks among the most predictable readings in the sample of 40 naming units used in my four experiments.

At the same time, this naming unit gives support to the view that it is primarily *permanent*, *stable*, and *constant relations* that are worth naming: there was no reading with a paraphrase in the present progressive (...who is smiling...).

An interesting comment came from an informant who interpreted *smile-man* as a 'politically correct' synonym with *comedian*![13]

The reading 'a yellow-faced symbol on T-shorts, a smiley-face icon (Happy Face)' occurred eighteen times, and seems to be bound to the specific extralinguistic context because of its zero occurrence in the check-group.

Readings like 'a false hypocritical person', 'a mentally retarded person who smiles for no reason', 'a dentist who fixes people's teeth so their smile looks better' give additional support to the crucial role of *world-knowledge* and *experiences* in the meaning-prediction process. None of these readings could result from 'mere' linguistic competence.

Check-group results

An optimistic, good-natured person often smiling	18/25
A person whose task is to cheer somebody up	5/25
A person who makes people smile when photographed	2/25

The check-group results map those of the main group. The third-rank reading occurred three times in the main group, with low scores assigned.

4.4.2.7 *Removage*

'Something that has been/needs to be removed, waste material, garbage'

FO:	59/90
Scores:	441/900
PR:	0.321
OPR:	0.662

'The act/process of removing a person or object(s), goods, waste, etc.'

FO:	43/90
Scores:	299/900
PR:	0.159

'The act of moving from one place to another'

FO:	8/90
Scores:	46/900
PR:	0.005

Comments

The morphological structure of this naming unit corresponds to several Morphological Types. It is for this reason that *removage* was selected for my experiment. The purpose was to test the influence of a *multiple subcategorisation of a suffix/the existence of several polysemous suffixes* (in this particular case the suffix *-age*) upon the meaning-prediction process.[14] Thus, *-age* can combine with nouns, verbs, and adjectives, and may refer to Process, the Result of Process, State, Material, Collectivity, Location, etc.[15] No matter which of the theoretical positions concerning the polysemy/homonymy opposition is ac-

cepted the one-to-many relation between form and meaning can be expected to have considerable *predictability-reducing effects*.

The picture 'created' by the informants corresponds with my expectations. There are two competing readings. In fact, the top reading 'something that has been/needs to be removed, waste material, garbage' combines two different WF Types:

(51) a. [Action/Process → Result]
b. [Action → Object].

A number of informants combined these two possible readings into a single reading in which *-age* carries the meaning of Material. On the other hand the next reading 'the act/process of removing a person or object(s), goods, waste', places the focus on 'Action/Process'.

Strangely, a different meaning-prediction path brought some informants to a different onomatological analysis and structure: *re* + *movage*, as occurring in the readings, such as 'the act of moving from one place to another', 'the act of moving sth more than once', 'the fee you pay someone to move equipment for you'.[16] Interestingly, some informants proposed readings resulting from both of these prediction paths, which means than they are not disjunctive.

The OPR of the topmost reading is not very high: while the PR of the 3rd rank reading approaches zero, the competition of the second-rank reading is significant.

Check-group results

An act of removing something, e.g. rubbish, hair, etc.	14/25
A change of place	4/25
All unnecessary things left after removing a place	1/25
Things, stuff that need to be removed from some place	1/25
An act of moving sth once more	1/25

Four informants failed to propose any meaning. The top reading of the check-group corresponds with the rank 2 reading of the main group. The rank 1 reading of the main group was proposed by a single check-group informant. Thus, unlike the native speakers, the non-native group preferred readings other than the 'intended' one. It is difficult to find the reasons for this discrepancy.

4.4.2.8 *befoot*

'To walk'

FO:	29/90
Scores:	191/900
PR:	0.068
OPR:	0.382

'To (surgically) remove a person's or animal's foot/paw/hoof'

FO:	16/90
Scores:	130/900
PR:	0.026

'To kick someone/something'

FO:	19/90
Scores:	87/900
PR:	0.020

Comments

This possible naming unit proposed with the intended WF Type [Action – Object] proved to be a disaster for the informants. Their total confusion is reflected in the absolutely lowest PR of the top reading ('to walk', PR = 0.068) from among all the 30 sample naming units used in Experiments 1 to 3. The top reading is followed by the 'intended' reading 'to (surgically) remove a person's or animal's foot/paw/hoof' (PR = 0.026). The disorientation of the informants is best manifested by the total number of the proposed readings that is by far the highest among the thirty naming units in the three experiments: 55 different readings! Let us therefore look at this case in detail.

The first peculiarity of the prefix *be-* is that it can function as both a class-changing and class-maintaining affix (*behead*, *becloud*, *belittle* vs. *bemoan, besmear*, *bemock*). Second, it can express a range of various meanings. Third, Marchand (1960: 100) maintains that "[i]t is only with the shade of overloadedness, disparagement, or ridicule that *be-* is a productive morpheme in PE [Present-day English – P.Š.]." Furthermore, Bauer (1983: 217) assumes that this prefix "is probably no longer productive except in the sense seen in *be-jewelled*, where the past participle of the verb is used adjectivally..." In spite of these circumstances, I proposed *befoot* with the meaning 'to remove N' because I believed (erroneously, by all accounts) that this meaning could be inferred by analogy with *behead* (cf. Ryder's 'analogy-based formation' – Section 1.3.3): 'to remove a part of the body'. Given the above-mentioned unproductivity of this suffix it was postulated that analogy would outweigh the unproductivity as a predictability-reducing factor.

Check-group results

To cut sb's foot	5/25
To put on shoes	3/25
To take off shoes	3/25
To walk	3/25
To enter a place for a short time	1/25
To be settled somewhere for good	1/25
To kick	1/25
To step on sth	1/25

Seven informants failed to suggest any reading. The scattering of the proposed readings in the main group due to the unpredictability of this naming unit is also present in the check-group. None of the readings gained a more substantial degree of support. The readings 'to put on shoes' and 'to take off shoes' are not the first instance of contradictory interpretation of one and the same naming unit in my experimental sample.

4.4.2.9 *obsess*$_{N}$

'A person who is obsessed'

FO:	39/90
Scores:	319/900
PR:	0.153
OPR:	0.539

'An object which a person is obsessed by'

FO:	29/90
Scores:	203/900
PR:	0.073

'Persistent preoccupation with an idea or emotion/obsession'

FO:	25/90
Scores:	189/900
PR:	0.058

Comments

The WF Type underlying the intended reading is given in (52):

(52) [STATE$^{\underline{\text{Patient}}}$SUBSTANCE]

While my expectations of the top reading were correct, it is not strong enough (PR = 0.153) to dominate the other two readings, which are a variation on the same theme: the Object of obsession (0.073) and the State of obsession (0.058), respectively. The latter interpretation is suspicious because it actually

substitutes for the actual word *obsession* and, as in the case of *dialler* (used by some informants in interpreting the above-discussed *feather-dialler* in the Instrumental meaning reserved for the actual naming unit *dial*), it violates the principle of blocking. From a different viewpoint, since these three readings are 'readily' available, the number of proposed readings is extremely low (11).

Check-group results

An obsessed person	12/25
A state of mind – constant thinking of sth.	6/25
A thought one cannot get rid of	5/25
Obsession lasting only a very short time	1/25
A particular form of obsession	1/25

Apart from the identity of the most predictable readings, one more interpretation deserves attention, in particular, 'obsession lasting only a very short time': a reduced morphological form (*obsess* vs. *obsession*) seems to have implied 'reduced' interpretation for one of the informants!

4.4.2.10 *leave-behind*$_{N}$

'Sth that has been/usually is (e.g. umbrella) left behind/forgotten (e.g. after an event)'

FO:	58/90
Scores:	485/900
PR:	0.347
OPR:	0.878

'A person who was left behind'

FO:	18/90
Scores:	110/900
PR:	0.024

'Anything unwanted, disposable, ignored (both material and immaterial)'

FO:	17/90
Scores:	114/900
PR:	0.024

Comments

The intention behind this naming unit was to test a converted phrasal verb. Similar to the previous converted naming unit, there are several plausible conversion directions, also reflected in the informants' proposals. In addition to the predictable reading there are three [Human] [Animate] readings: 'a person who falls behind, unable to keep up with the rest of a group', 'a person who is

forgetful/who often leaves things behind', and 'a person who was left behind'. The reading 'anything unwanted, disposable, ignored (both material and immaterial)' has the same PR as the reading 'a person who was left behind'. A completely different path brought some informants to the reading 'people who were not taken to heaven in rapture'. None of these readings gained, however, a sufficient level of support. As a result the OPR of the most predictable reading is extremely high (0.906). This is connected with a very high PR Gap (0.323) and a huge R1/R2 ratio (14.46).

Check-group results

Sth. that we have gone through in the past, our (unpleasant) experience	7/25
An easy-going person who does not worry about anything very much, who pushes worries out of his mind	3/25
A person who falls behind, unable to keep up with the rest of a group (in education, etc.)	3/25
Ath. not actual any more	1/25
A person one wants to forget about	1/25
A person who starts a new life and tries to forget about his past	1/25
An act of separation from a person you loved	1/25
A person deserted (by a partner)	1/25
A backlog	1/25
An object/belonging that was left somewhere/forgotten about	1/25
An act of forgetting about sth	1/25

This is an unusually long list of proposed interpretations. The main group's No. 1 gained support from only one informant. The non-native speakers preferred the interpretation based on 'unpleasant experiences'. They occur in three different variants.

4.4.3 Summary 3

The results obtained in Experiment 3 are summarised in Table 49.

Table 49. Onomasiological Type, Predictability Rate (top reading), and Objectified Predictability Rate

	OT	*PR*	*OPR*
blondesjoker 'A person who tells/makes blonde jokes'	1	0.422	0.630
feather-dialler 'A feather(-like instrument) used for dialling'	1	0.212	0.752
the phone contactee 'A contacted person, a person that can be contacted'	2	0.897	0.988
refusnik 'A person who usually refuses to do something, a stubborn person, a denier'	2	0.283	0.823
anthraxist 'A terrorist who uses anthrax as a weapon'	3	0.570	0.659
smile-man 'A person with a smile; a person who smiles a lot, an optimist'	3	0.564	0.851
removage 'Something that has been/needs to be removed, waste material, garbage'	4	0.321	0.662
befoot 'To walk'	4	0.068	0.382
obsess$_{N}$ 'A person who is obsessed'	5	0.153	0.539
leave-behind$_{N}$ 'Sth. that has been/usually is left behind/forgotten'	5	0.347	0.878

This table indicates the complexity of factors that influence the meaning predictability of novel, context-free naming units. In interpreting the data it is necessary to keep in mind that they are based on only two naming units per Onomasiological Type. In addition, the representatives of the individual Onomasiological Types are – in a way – extreme cases which were intentionally proposed in order to demonstrate that there is no single factor that conditions the meaning predictability of novel, context-free naming units.

1. In Section 3.5.1 it was claimed that Onomasiological Type may influence the Predictability Rate. From this point of view Experiment 3 does not provide unambiguous and conclusive evidence. On the one hand, the highest OPR (0.988) of *contactee* (Onomasiological Type 2) and a very high OPR of another Onomasiological Type 2 naming unit *refusnik* (0.823) favour my hypothesis. The OPR of *anthraxist* (0.659) (OT3) more or less fits the typical OPR range of Type 3 naming units, identified in Experiment 1. It

should be noted that while belonging to Onomasiological Type 3, *anthrax-ist* represents a different general Morphological Type ([WF base + suffix]$_N$) compared to those tested in Experiment 1 ([WF base + WF base]$_N$).
On the other hand, there are several cases which call the above-mentioned hypothesis into question. Thus, *smile-man* (Onomasiological Type 3) has the third highest OPR value (0.851), and the OPRs of both of the Onomasiological Type 1 representatives are fairly low. In addition, there is no predictable reading for *befoot* (Onomasiological Type 4). An evaluation of these results in isolation may bring about highly distorted and misleading conclusions.

2. The reasons for the discrepancies between the general theory of the role of Onomasiological Types and the actual results were already indicated in the comments on the individual naming units above. Let us summarise the *predictability-reducing factors* in order to get a more lucid picture of the situation:

   a. Perhaps the most negative factor which reduces the OPR is the *number of competing readings*. The PR value of a 'competing reading' cannot be fixed within any range. It depends on the value of the top reading, the PR Gap, and the R1/R2 (R1/R3) ratio. It follows from Summaries 1 and 2 that a tough competition on the part of one or more readings may significantly reduce the OPR of a top reading with a (relatively) high PR value. There are no extreme cases of competition in Experiment 3. Examples of partial competition include 'a person who tells/makes blonde jokes' (PR = 0.422) vs. 'a blonde-haired humorist, comedian, joker' (PR = 0.242), and 'a terrorist who uses anthrax as a weapon (e.g. by mailing it in envelopes)' (PR = 0.570) vs. 'a person who works with/produces/studies anthrax' (PR = 0.293).

   b. The *ambiguity of one or even both polar constituents* of a WF Type, which translates into the existence of several WF Types available to a language user, reduces OPR. If some of these WF Types are productive, an overpressure of acceptable readings develops, which reduces the Predictability Rate Gap and the R1/R2 (R1/R3) ratio, and consequently also the OPR. This is the case of the *homonymous suffix -er* which occurs in two highly productive Morphological Types representing Agents and Instruments. In the sample case represented by *feather dialler* the Agentive interpretation is very weak (perhaps due to the pragmatic experience of the informants), which reduces the adverse effects of this factor upon meaning predictability. In general, however,

cases like this suggest that, *apart from positive effects*, *productivity may also have reducing effects upon the predictability of meaning of a novel, context-free naming unit*.

c. The *foreignness/low productivity of an affix*. Such an affix assumes a peripheral position in the system, and – logically – is not known to a large number of language users. If such an affix occurs in a naming unit its awkwardness transfers to the naming unit as a whole. In that case the advantage of an affix as a predictability-supporting constituent turns into a disadvantage, which reduces the Predictability Rate. For this reason the PR value of the top reading of *refusnik*, including a peripheral suffix *-nik*, is low (0.283). In spite of this fact, the OPR value of 'a person who usually refuses to do something, a stubborn person, a denier' is extremely high, because the number of competing readings is practically zero: the other readings proposed by the informants are so weak that they boosted the OPR of the top reading to as much as 0.823.

d. The object named requires *highly specialised world knowledge and/or experience* which is not readily available to the majority of language users. *Refusnik* is a case in point. The established reading gained minimum support exactly due to the fact that the reading 'a political dissident (from the former USSR, or from a totalitarian regime)' was not a part of the mental lexicon of the majority of the informants. This resulted in a paradox, the PR of an established reading is much lower than that of a potential, non-existing reading. Hence, this predictability-reducing factor seems to be very powerful (see also for our discussion of the meaning predictability of converted proper names in Section 3.5.1.7).

e. The *absence of a morphological expression of an Actional seme* is a serious obstacle to meaning predictability. On the other hand, the *Actional nature* of the determining constituent of the onomasiological mark boosts the predictability of Onomasiological Type 3 because it compensates for the absence of the determined constituent which is a typical bearer of Actional meaning, relating the polar members of the onomasiological structure. This is exemplified by *smile-man* in the static Onomasiological Type 3 interpretation of the top reading ('a man with a smile'). However, as noted above, this interpretation cannot be discussed separately from the Onomasiological Type 2 interpretation. Thus, the fact that this ambiguous case of Onomasiological Type 3 shares the features of Onomasiological Type 2 accounts for its

exceeding the maximum OPR obtained for the Onomasiological Type 3 naming units in Experiment 1 (0.744 for *age bag*).

3. In this sample, four naming units were generated in the meaning belonging to the *Word-Formation Type Cluster of Agents*. None of them reached an expected high level of PR, each of them for some of the reasons already mentioned in the previous points. Given the peculiar nature of these naming units one can hardly draw relevant conclusions concerning the meaning-predictability of Agent nouns; the more so that four cases are too small a sample for this purpose. In spite of this fact, I believe that the notion of Word-Formation Type Cluster (covering the names of Agents, Instruments, Actions, etc.) lends itself to the examination and evaluation (under the condition of a sufficiently large sample) of the relation between the productivity of the individual Onomasiological, WF, and Morphological Types, and their respective meaning-predictability. Taking into consideration the necessary extent of such a research project this remains a task for future projects.

4.5 Experiment 4

4.5.1 Sample naming units

Experiment 4 differs from the preceding three experiments in its focus. Its aim was to examine the relation between meaning-predictability and word formation productivity. The basic idea underlying this experiment was to evaluate the meaning predictability of various naming units which were coined in defiance of productive WF Rules, that is to say, each naming unit in the sample violates some of the constraints upon productivity. The purpose was to evaluate – in so far as the limited sample permits it – the influence of 'unproductivity' on the meaning predictability of such naming units, and the degree to which various constraints affect (if at all) the meaning predictability of novel, context-free naming units. All sample naming units are potential naming units (in Bauer's (2001) sense) and are morphologically transparent. The sample of unproductively coined naming units, including the relevant constraints, is given in Table 50:

Table 50. Unproductively coined naming units and the relevant constraints violated

Naming unit	*Relevant constraint*
evilen$_V$	The inchoative suffix *-en* only attaches to monosyllabic stems and, moreover, only if they end in an obstruent, optionally preceded by a sonorant (Halle 1973) – *structural-phonological constraint*
sillily	Adverbial *-ly* should not be added to adjectives which already end in *-ly* (Bauer 2001) – *phonological-morphological constraint*
unill$_V$	Negative affixes are *not* used with simple adjectival stems that have a 'negative' value on evaluative scales such as 'good-bad', 'desirable-undesirable' (Zimmer 1964) – *semantic constraint*
two-computered (friend)	An inalienable possession as a condition for *-ed* compound adjectives (Bauer 2001) – *semantic constraint*
stealer	Kiparsky's Elsewhere Condition (1982), for example, permits only one rule to be applied, that is to say, the one which is more specific (*thief* in this particular case) – *level-ordering constraint* ('blocking by synonymy' constraint)
sittee	Blocking by a more productive WF Rule – *'blocking by synonymy' constraint*
concentrationhood	Uninterpretable because there are two abstract suffixes (Plag 1999) – *structural-semantic constraint;* moreover, *-hood* only combines with nouns denoting human beings – violation of the *subcategorisation constraint*
horribility	*-ity* is only productive in combination with the productive *-able* function, which maps transitive verbs to adjectives. By implication, the domain of the *-ity* function f_{ity} is the function f_{able}, and its range is the composed function $f_{ability}$ (Raffelsiefen 1992) – *morphosyntactic constraint*
nearth	(A + *-th*)$_N$ is an unproductive WF Rule (Bauer 2001) – *WFR-related constraint*
becomable	*-able* is used on transitive but not intransitive verbs (Bauer 2001) – *morphosyntactic constraint*

The informants' views were obtained in a manner identical to the former three experiments. All 50 informants were native speakers, mostly university undergraduates, and were approached by my students personally or via e-mail messages. They replied either by e-mail or by snail mail. From this it follows that the 'unproductive' coinages were presented to them in writing.

4.5.2 Experimental data and their analysis

Evilen
'To do something wicked/(more)evil'
FO: 32/50
Scores: 133/500
PR: 0.178
OPR: 0.973

'To increase the evilness of something'
FO: 5/50
Scores: 17/500
PR: 0.003

'To make someone/something look evil'
FO: 3/50
Scores: 18/500
PR: 0.002

sillily
'In a silly manner, foolishly'
FO: 46/50
Scores: 339/500
PR: 0.624
OPR: 0.998

'A lily placed upon a window sill'
FO: 2/50
Scores: 16/500
PR: 0.001

'Childish'
FO: 1/50
Scores: 10/500
PR: 0.0004

unill
'To cure, to become healthy'
FO: 34/50
Scores: 156/500
PR: 0.212
OPR: 0.977

'To make null and void, to obliterate'
FO: 5/50
Scores: 21/500
PR: 0.004

'To unify or make one in purpose'
FO: 2/50
Scores: 11/500
PR: 0.001

Two-computered (friend)
'A friend with two computers'
FO: 39/50
Scores: 285/500
PR: 0.445
OPR: 0.987

'A friend who spends all their time on the computer'
FO: 5/50
Scores: 27/500
PR: 0.005

'An extremely calculating, thoughtful smart friend'
FO: 3/50
Scores: 14/500
PR: 0.001

Stealer
'A thief'
FO: 48/50
Scores: 396/500
PR: 0.760
OPR: 1.000

'A physically handicapped person unable to walk'

FO: 1/50
Scores: 5/500
PR: 0.0002

'Not a get-away driver'
FO: 1/50
Scores: 2/500
PR: 0.0001

Sittee
'Someone who sits (a lot) on something – as opposed to standee'
FO: 23/50
Scores: 93/500
PR: 0.086
OPR: 0.887

'A sofa, a piece of furniture for sitting'
FO: 8/50
Scores: 27/500
PR: 0.009

'A location/place of some thing'
FO: 3/50
Scores: 18/500
PR: 0.002

Concentrationhood
'The state of (deep) concentration'
FO: 28/50
Scores: 130/500
PR: 0.146
OPR: 0.936

'A larger group of persons concentrating on the same thing, having similar interest'
FO: 5/50
Scores: 27/500
PR: 0.005

'The ability to concentrate'
FO: 5/50
Scores: 23/500
PR: 0.005

Horribility
'The degree of horror, awfulness, disgusting condition'

FO: 27/50
Scores: 180/500
PR: 0.194
OPR: 0.770

'The ability to scare someone/the ability to be horrible'
FO: 16/50
Scores: 85/500
PR: 0.054

'(An attribute of) horribleness'
FO: 4/50
Scores: 25/500
PR: 0.004

nearth
'The state of being near to; proximity; opposite of 'farth"
FO: 22/50
Scores: 87/500
PR: 0.077
OPR: 0.852

'Something near the earth'
FO: 10/50
Scores: 32/500
PR: 0.013

'A small mound of dirt in which mites live'
FO: 1/50
Scores: 9/500
PR: 0.0004

becomable
'possible; that can become'
FO: 41/50
Scores: 252/500
PR: 0.413
OPR: 0.983

'Able to be made beautiful or attractive'
FO: 6/50
Scores: 27/500
PR: 0.006

'That can be changed'

FO: 2/50
Scores: 10/500
PR: 0.001

4.5.3 Discussion

The results of Experiment 4 are summarised in Table 51.

Table 51. The 'unproductivity' experiment: An overview of the results (top readings only)

Naming unit	*OT*	*PR*	*OPR*
evilen	4	0.178	0.973
sillily	4	0.624	0.998
unill	4	0.212	0.977
two-computered (friend)	4	0.445	0.987
stealer	2	0.760	1.000
sittee	2	0.086	0.887
concentrationhood	4	0.146	0.936
horribility	4	0.194	0.770
nearth	4	0.077	0.852
becomable	2	0.413	0.983

The data obtained indicate that the meaning(s) of 'unproductively' coined naming units can be predictable, in some cases to a very high degree. As a matter of fact, these naming units gained the highest average OPRs from among all the four experiments. Moreover, one of the naming units, *stealer*, in particular, its reading 'a thief', reached the maximum possible value of OPR (1.000). In addition, *sillily* in the reading 'in a silly manner' gained almost the maximum (0.998), and five other potential naming units in their respective top readings approached the maximum OPR value in the range from 0.936 to 0.987. The lowest OPR value from among the ten sample units is 0.770 for *horribility*. This value would be the highest OPR value in Experiment 1, the third in Experiment 2, and the fifth in Experiment 3. These results may be surprising in view of the nature of the naming units tested in Experiment 4: they were coined 'unproductively'; or better, they violate the constraints imposed on productive WF Rules.

The PR values are not so conclusive: only one of the readings is in the high predictability range (above 0.750), in particular, the top reading of *stealer*. One other reading is in the medium predictability range (*sillily*), and two in the low predictability range. All the other top readings fall under the predictability level.

Thus, the results of PR and OPR are in strict contradiction to each other, and raise a number of questions. Let us try to answer them by taking up the individual naming units one by one.

Stealer

This naming unit violates Kiparsky's 'Avoid Synonymy Principle' (1982) which blocks the generation of synonyms by means of the 'Elsewhere Condition'. Monemes like *thief* can block on the grounds that a lexical entry itself is considered to be a rule, i.e., the most specific kind of rule, a 'Lexical Identity Rule'. Since *thief* is an existing lexical entry, representing the most specific rule, it blocks – due to the operation of the Elsewhere Condition – the application of any other, more general rule, for example one which generates *stealer*. A similar view is expressed by Rainer's 'token-blocking' (1988) which concerns the blocking of a morphologically complex word by the existence in the lexicon of a synonymous word. This idea was expressed as early as Aronoff (1976). Aronoff maintains that blocking operates when the following condition is met: the blocking and the blocked units have (as one of three possibilities) the same meaning (1976: 43).

The high meaning-predictability of the blocked *stealer* can be accounted for by resorting to Bauer's assumption that blocking prevents not so much the coining of nonce complex forms as their institutionalisation (1983: 85). By implication, the high PR of *stealer* suggests that the *blocking by synonymy principle is not a serious obstacle to meaning predictability*, especially in those cases where the *structural transparency* and the *strength of a highly productive WF Type* ([Action – Agent] in this particular case) and *Morphological Type* [Verb + -er]$_{N}$ *outweigh the constraint-based 'unproductivity'* of a coinage.

Sillily

Sillily is prohibited by a phonological/morphological constraint on productivity saying that the adverbial *-ly* should not be added to adjectives which already end in *-ly*. This constraint does not seem to impede the meaning-prediction process either. The generation of adverbs from adjectives by the suffix *-ly* belongs to the *most productive WF Types* [Quality → Manner] and *Morphological Types* [Adjective + *-ly*]$_{Adv}$ in English, and its strength is reflected in the high PR of the 'unproductive' coinage. It overcomes the ungrammaticality and boosts the PR of the reading 'in a silly manner, foolishly'.

Two-computered (friend)
This restriction is mentioned by Bauer (2001). It assumes that only those *-ed* compound adjectives are acceptable which express an *inalienable possession*. Hence, *two-computered (friend)* represents a semantic constraint on productivity. The research results indicate however, that *the semantic constraint of inalienable possession does not impose any restriction on meaning-predictability*. In this case, structures with *-ed* participle as a second constituent, such as *blue-eyed, left-handed,* etc., establish a productive pattern. The *productivity of the WF Type* [Stative – State] ('having/characterized by what is expressed by the *-ed*–preceding constituent(s)') and the *transparency of the Morphological Type* [Adjective/Numeral – Noun – *-ed*]$_{A}$ *eliminate the restriction on productivity* as an obstacle to a relatively high predictability.

Becomable
The suffix *-able*, which occurs in the naming unit *sleepable*, does not meet the traditionally adduced restrictions, summarised in Anderson (1992: 186):

(53) WFR: [X]$_{V}$ → [Xəbl]$_{Adj}$
Condition: [X]$_{V}$ is transitive (i.e., [+_NP])
Syntax: 'Object' argument of [X]$_{V}$ corresponds to 'Subject' of [Xəbl]$_{Adj}$
Semantics: '(VERB)' → 'capable of being VERBed'

Becomable is intransitive and there is hardly any acceptable reading for this naming unit that would meet the syntactic condition. From the semantic point of view, it rather features a 'property meaning', to use Plag's (to appear) term. Importantly, as noted by Plag (ibid.), "the forms exhibiting the property meaning are in a clear minority. In fact, this pattern has seized to be productive as early as the 17th century. . ." All these facts imply that *becomable* is a good candidate for my unproductivity test.

As in the previous cases of a relatively *high productivity* the corresponding WF Type [Action – Quality] and the closely related Morphological Type [Verb$_{Tr}$ + *-able*]$_{A}$ and, primarily, the *morphological transparency* of *becomable* and the *semantic transparency* of *-able* derivations in general seem to overcome the unfitting verbal base. These factors enabled the reading 'possible; that can become' to reach the OPR as high as 0.983.

Horribility
The restriction for *horribility* is that the suffix *-ity* is potentiated by *-able* under the condition that the latter is attached to a transitive verb (*readability*, *accountability*, *translatability*). In other cases the competing suffix *-ness* is preferred.

Thus, there is an actual naming unit *horribleness*, and, in fact, *horribleness* was used by several informants to account for the meaning of *horribility*. The *transparency* of *horribility* as a predictability-boosting factor is therefore in opposition to the Avoid Synonymy Principle. We have already encountered this principle in its unsuccessful attempt to reduce the meaning predictability of *stealer*. In the present case, in co-operation with a morphosyntactic restriction this constraint is much more successful (the PR of the top reading 'the degree of horror, awfulness, disgusting condition' is very low – only 0.194; the OPR value (0.770) is the smallest of all the ten sample naming units).

This successful reduction of meaning predictability may be accounted for by unequal circumstances of blocking for *stealer* and *horribility*, respectively. In contrast to *stealer*, blocked by a morphologically unrelated moneme, the *structurally based competition* in the case of *horribility*, including the same WF base, seems to be of principled significance. Moreover, and perhaps even more importantly, while *stealer* fully corresponds to a very productive WF Type [Action – Agent], *horribility* does not comply with the subcategorisation constraint, thus violating the relevant *morphosyntactic constraint*. The morphological transparency of *horribility* and the general (relatively high) semantic transparency of *-ity* derivations were not of much help in this case.

Unill

Unill emerged as an interesting case during the experiment, especially due to 5 occurrences of the reading 'to make null and void, to obliterate'. What makes this reading a surprise is the fact that the informants worked with a written test that does not justify this interpretation. The only kind of motivation that comes to my mind is a 'phantom' motivation by words like *annihilation*. If *unill* is parsed into *un-* + *nill* (with postulating the elimination of one 'n'), i.e., the structure which may have underlain some of the proposals of this reading, we obtain an opposite interpretation: 'to eliminate any null or void'. On top of it, since these informants could not 'see' the *ill* base of the naming unit, they (perhaps subconsciously) disregarded the inadmissible spelling with double *l*.

Yet, the central meaning of *unill*, 'to cure, to become healthy', which violates the semantic constraint requiring a positive stem for negation, has a very high OPR (0.977) in spite of the small PR value (0.212). Once again, this is an outcome of the uniqueness of the top reading as the only reading proposed by a larger number of the informants. One may surmise that it is rather the *low morphological transparency* than the semantic constraint itself that prevented this reading from obtaining a higher PR.

Concentrationhood

While the PR of this naming unit is even lower (0.146) than that of *unill* its OPR is very high (0.936). This naming unit violates two constraints. First, the suffix *-hood* only combines with nouns that refer to people in order to form new nouns. Nouns formed in this way refer to states, conditions, or to periods of time in which something is experienced. Second, it violates the structural-semantic constraint which rejects the co-occurrence of two abstract suffixes. It seems that it is primarily the *structural facet of the structural-semantic constraint* which exerts more negative effects upon the meaning predictability of this naming unit: this follows from the fact that the dominating reading 'the state of concentration' (by a few informants completed with the attributes 'deep' or 'total') roughly corresponds to the meaning of *concentration*. By implication, for the majority of informants the *additional abstract suffix does not mean any obstacle to the interpretation* of *concentrationhood* as such; it, as it were, confirms the semantics introduced by the preceding suffix, but itself does not semantically contribute to the meaning of the naming unit. The informants apparently found *the structure of this naming unit unnatural.*

Evilen

While the majority of the informants voted for the interpretation which is in compliance with the corresponding productive WF Type [Quality (=Result) ← Action] ('Quality resulting from Action') and the related Morphological Type [Adjective – *-en*]$_V$, which is represented in the sample by two closely related readings, 'to do something wicked/evil' and 'to increase the evilness of something', the points assigned are generally very low. One of the reasons for this fact was formulated by one of the informants who stated that *evilen* "just doesn't sound right".

In fact, both structural and phonological constraints are violated in this naming unit. First, it is a disyllabic word, and second, it does not end in an obstruent. In addition, the suffix *-en* is not the only inchoative suffix in English which can, at the onomatological level, express the Actional seme of the onomasiological level. The suffixes which commonly combine with a polysyllabic Adjectival WF base and realise the WF Type [Quality (=Result) ← Action] include *-ize* (*generalise*, *equalize*, *legalize*) and *-ify* (*purify*, *justify*, *simplify*). This list should be completed with the onomasiological recategorisation of the [Quality$\underline{\underline{\text{Result}}}$Action] type, resulting in the change of word-class from Adjective to Verb (*abstract*, *appropriate*, *illegitimate*). By implication, the productivity of the applied WF and Morphological Types is not strong enough

to ensure a more favourable meaning predictability under the circumstances where there are more appropriate and more productive Types at hand.

Moreover, the *morphological structure* is a serious predictability-reducing factor in this case, because it is not transparent enough to contribute to a high Predictability Rate. This was also confirmed by four informants who did not propose any reading.

It may therefore be concluded that the *predictability-reducing effects of the competing Morphological Types* and *the violation of the phonological and the structural constraints* are too serious obstacles to be outweighed by a *limited semantic transparency*, that is to say, the limited 'readability' of the underlying WF and Morphological Types. This is reflected in a very low PR (0.178) which, however – owing to the absence of any competition – does not preclude the top reading from reaching an extremely high OPR (0.973).

Sittee

The selection of *sittee* based on the Morphological Type [V+ *-ee*]$_N$ requires a more detailed explanation. As suggested by Barker (1998:708), the suffix *-ee* can be viewed as a counterpart of *-er*, and "it is possible to entertain the hypothesis that the conditions for use of *-ee* are defined negatively, in contrast to those for *-er*: *-er* picks out subject participants, and *-ee* covers everything else." As he, however, notes this hypothesis faces the problem of the existence of a considerable number of *-ee* nouns referring to subject participants. This is confirmed by Lieber (in press) who points out that while *-er* nouns "most often form personal agent nouns, and *-ee* most often forms patient/theme nouns, not infrequently we find precisely the opposite situation, where *-er* and its cohort form patient nouns and *-ee* agent or at least subject-oriented nouns."

Moreover there are instances of both *-ee* and *-er* attached to the same WF base and have synonymous meaning (*escapee/escaper, absentee/absenter, arrivee/arriver*, etc.) (Barker 1998:709). Based on the analysis of a large corpus, Barker arrives at a conclusion that there are at least three types of *-ee* derivations that are productive: direct object, indirect object, and subject. *Sittee* is the subject type. This possible naming unit safely meets two of three Barker's (1998) semantic conditions imposed on productive *-ee* derivation. First, it meets the condition of sentience of the referent referred to by *sittee*. Second, it meets the semantic constraint of *episodic linking*, according to which "the referent of a noun phrase headed by an *-ee* noun must have participated in an event of the type corresponding to the stem verb" (1998:711). In this particular case, *sittee* participates in a sitting event. Problematic is the third semantic constraint, defined as "a lack of volitional control on the part of

its referent either over the occurrence or the duration of the qualifying event itself or (given a punctual qualifying event) over its immediate direct consequences" (1998:717). Being a subject type *sittee* refers to Agent. However, the category of Agency implies volition. Thus, for a *sittee* to preserve the 'lack-of-volition' constraint it would have to mean the action into which a sitting person is forced somehow – in contrast to *sitter* who, in principle, does his/her activity voluntarily, fully based on his/her will.

It follows from these considerations that *sittee* is a possible naming unit that can be produced by a productive WF and Morphological Type. What made me include this possible naming unit in the 'unproductivity' test is the much more productive competitor, the *-er*-based pattern that underlies the existing Agent noun *sitter*. Thus, while the *-ee* Agent noun is possible it is blocked by a much more productive Moprhological Type that has already produced an established (institutionalised) naming unit *sitter*.

On a fine-grained semantic level, the blocking is eliminated by the 'volition – lack of volition' opposition that can be represented as an opposition between the Agent and the Patient readings. Therefore, one of the questions behind the inclusion of *sittee* in the experimental 'unproductivity' research was whether the informants (native speakers) would perceive this kind of semantic distinction.

The readings proposed by my informants are somewhat surprising. None of the fifty native speaker informants distinguished between volitional and non-volitional action even if they could propose an unlimited number of readings. All proposals are of the Agentive interpretation. Consequently, the proposed Agentive reading faces a tough competition from a much more productive Morphological Type $[V + \textit{-er}]_A$. By implication, the PR of the reading 'someone who sits (a lot) on something' is low (0.086). In fact it is the second lowest of all the naming units in my sample.

On the one hand the situation resembles the above-discussed instances of blocking by synonymy (*stealer* and *horribility*), on the other hand it is substantially different. This difference bears primarily on the existence of the naming unit *sitter* and the related disproportional relation between the strength of the *-er* and the *-ee* rules. The existence of *sitter* may have influenced the informants who proposed the reading 'someone who sits (a lot) on something' (some of the informants stated that their proposal was motivated by analogy with *standee*) to have done it with displeasure, which is illustrated by the low average score (4.0 pts.). The impasse they reached is also illustrated by readings like 'a sofa, a piece of furniture for sitting' (8 informants) and a kind of blend-

reading 'a seat for two'. Some of them explicitly stated that it cannot be a person who sits because this would be a *sitter*.

This conclusion is borne out by research into the sociolinguistic aspects of WF productivity (Štekauer et al. 2004). One of the tasks of native speaker informants was to rate how likely they and other English speakers would be to use the word *swimmee*. 44 of 47 informants interpreted *swimmee* in the sense synonymous to *swimmer*. At the same time, 34 of them identified this reading as 'extremely unlikely', nine of them as 'somewhat unlikely', and one as 'likely' (the remaining classification options were 'very likely', and 'extremely likely'). Three other informants proposed the reading 'swimming trunks'. Thus, while *swimmee* is, in theory, interpretable on the basis of its transparency and the existence of the corresponding WF and Morphological Types, native speakers are reluctant to assign a higher value to its 'swimmer' reading due to the existence of a much more productive WF and Morphological Types and the existence of *swimmer* with a blocking effect.

Sittee thus seems to provide a most compelling example of the direct influence of unproductivity/low productivity/blocking upon unpredictability. In particular, *low productivity of the underlying Morphological Type facing competition from a much more productive Morphological Type*, combined with the *blocking constraint*, reduces the predictability of a novel naming unit.

Nearth

Nearth is the naming unit which gained the lowest support of all the words in the sample. This naming unit illustrates an instance of ambiguous morphological structure admitting more than one parsing, similar to that of *clamprod* (discussed in Section 1.3.4.4) which can be parsed into both *clamp* + *rod* and *clam* + *prod*. Let us recall that Libben, Derwing & Almeida (1999: 385) aptly note that the parsing process, which is vital to the subsequent interpretation, is "highly correlated with semantic plausibility."

In the meaning originally intended by myself *nearth* represents an unproductive Morphological Type of [Adjective + *-th*]$_N$. Despite its synchronic unproductivity, the morphological structure is still transparent, and therefore parsable by language-users. It is this general structural transparency which motivated one of the informants to explain the meaning of *nearth* by *analogy* with *farth*.

On the other hand, *nearth* itself is not transparent enough to avoid ambiguity, because some informants identified the WF base *earth*. This *parsing ambiguity* is reflected in the results obtained: eight informants preferred this blend-based reading ('something near the earth'), and some others proposed

related readings like 'to bring to ground level; on the ground', 'underground habitat of a small Australian marsupial', 'a type of soil'. This fact seems to be responsible for a considerable scattering of the readings and the low predictability of the top reading 'the state of being near to; proximity'. The dominant (even though weak) position of the 'expected' reading acknowledges the postulate concerning the ambiguous parsing instances: the preference of language users for a particular way of parsing is motivated by a higher morphological transparency of one of the parsing options.

It may be concluded that the *structural ambiguity* of *nearth* and the *unproductivity of the Morphological Type* $[A + \textit{-th}]_N$ may be held responsible for the extremely low PR (0.077) of *nearth*. The analogy-based motivation of the interpretation (supported by the existence of *farth*) apparently has not been strong enough to outweigh the unfavourable factors.

4.5.4 Summary 4

The previous analysis enables me to draw the following conclusions, which must be perceived as tendencies rather than rules. Further evidence drawing on more extensive research is required.

1. *None of the productivity constraints can be ascribed consistent and exclusive influence on meaning predictability*. Other relevant factors may promote/reduce the effects. Thus, for the 'blocking by synonymy' constraint the following rough scale can be proposed:
   - *Minimum adverse influence upon meaning predictability*, conditioned by
     a. high structural transparency of a novel naming unit;
     b. high productivity of the WF Type and Morphological Type underlying the novel naming unit;
     c. the monemic nature of the blocking naming unit.
   - *Medium adverse influence upon meaning predictability*, conditioned by
     a. high structural transparency of a novel naming unit;
     b. unproductive WF Type underlying the novel naming unit, due, for example, to the violation of some productivity constraint;
     c. competition of a synonymous productive Morphological Type.
   - *Significant adverse influence upon meaning predictability*, conditioned by
     a. ambiguous morphological structure admitting two different parsings/poor structural transparency;

   b. an unproductive/low productive Morphological Type of a novel naming unit;
   c. a highly productive Morphological Type of a blocking naming unit.

2. *Not all productivity constraints are of equal importance*. The phonological/morphological constraint of avoiding the repetition of phonologically identical morphemes (*sillily*) seems to be of little significance; the same applies to violation of a subcategorisation restriction (*becomable*), provided that the word formation and morphological structures are transparent; in other words, *semantic and morphological transparency tend to outweigh individual constraints*.
3. *In cases of ambiguous parsing, productivity may directly affect meaning predictability*: first, the predictability of meaning is low, and second, those readings are more predictable that are based on more productive WF and Morphological Types.
4. *Semantic constraints* (*two-computered* (*friend*), *unill*, *concentrationhood*) *do not seem to reduce meaning predictability* provided that the morphological structure is well identifiable.
5. The results indicate that while there are many other factors at play (see Conclusion 14 in Section 5.2) and in spite of the fact that the nature of constraints requires a case-by-case analysis, it may be postulated that *WF productivity affects* (to varying degrees) *the predictability of meaning*(*s*) *of novel naming units*.

4.6 Meaning predictability and associative meaning: The experimental results in the light of free association of words

The experimental data discussed above were obtained in two different ways and under two completely different conditions which reflect the different possibilities of encountering a new naming unit by a language user. The central part of the data was obtained without any time limit imposed on the informants, reflecting a situation when a language user has enough time to think over the different possible meanings and to select the most acceptable to him/her. The 'check-group' interpretations were obtained under time pressure, a situation not infrequently experienced by language users. Furthermore, while the former group of informants was asked to propose the greatest possible number of potential readings for each sample naming unit and rate them according to

(the degree of) their acceptability, the latter group was asked to propose just one reading that came immediately to an informant's mind. The idea behind choosing two completely different experimental conditions – in addition to reflecting two different ways of encountering new naming units by language users – was to determine the role of the time factor in identifying the most predictable readings, i.e., to determine whether a sufficient time space for identifying and rating the most acceptable (predictable) meaning(s) is a significant predictability-related factor. The comparison of the main group and the check-group indicates that the time factor does not play a crucial role and has minimal impact on the decision-making of language-users in regard to identifying the predictable readings. (But as indicated in the discussion on *feather-dialler*, demanding cases of meaning predictability may require more time for processing a first-encountered coinage.)

This suggests two things:

1. The dominant position of certain combination(s) of particular prototypical features, captured in the concepts of the objects that are related in the concept of an object to be named, is so striking that its identification is not time-dependent. This applies primarily to the cases characterized by a positive influence of the majority of the 'Predictability-boosting Conditions' (i.e., high productivity of a WF Type, single 'strong' reading, motivation by prototypical seme(s), general knowledge, etc.; see Conclusion 15 in Section 5.2 for a list of Predictability-boosting Conditions and Predictability-reducing Conditions).
2. One also might suspect that there are certain more or less firm connections between the meanings of lexical units stored in the mental lexicon that, in connection with the other Predictability-boosting Conditions, may contribute to a speaker's meaning-prediction capacity.

The view of the interrelatedness of meanings of lexical units stored in the mental lexicon is not new. Both lexical semanticists and morphologists came up with the idea of lexical, semantic, and conceptual fields, based on the paradigmatic relations of hyponymy/hyperonymy, antonymy, and synonymy. Recent psycholinguistic research into associated meanings has confirmed the existence of such connections of different strength. In addition, recent research in the field of pragmatic corpus-based lexicography has emphasised the requirement of well-balanced paradigmatic and syntagmatic relations in determining the meanings of lexical units (Sinclair 1998). With regard to the associative meaning principles, these requirements would, in an ideal case, mean that

1. for complex words with two stems the most predictable meaning(s) of a novel naming unit result(s) from the intersection of the strongest associative meanings (targets) bound to syntagmatically related words (cues), or, at least, from an associative meaning bound to one of the constituents;
2. for complex words with a stem and an affix the most predictable meaning(s) of a novel naming unit results from the intersection of the strongest associative meaning and the general meaning of the affix; and finally,
3. for converted naming units the most predictable meaning(s) of a novel naming unit should be identified by the strongest connection between cue and target.

Therefore the objectives of this section are to check whether there is any relation between the predictability of meaning(s) of new naming units and the associative connections between lexical units stored in our mental lexicon. My research into the relation between meaning predictability and association, was implemented with two predictions in mind:

1. There should be at least some overlap between associates and the experiment-generated meanings.
2. Rankings of associates and meanings should correlate.

For that purpose, I used an associate connectivity database by Nelson, McEvoy and Schreiber available on the Internet. Appendix A of the database includes pairs of 5,019 normed words (cues) and the 72,176 responses (targets). The database provides 31 different types of data which can be variously combined, depending on a user's needs. For the purpose of this chapter I made use of the *FSG parameter* (column 6) indicating the 'forward strength, i.e., the *cue-to-target strength* (i.e., the probability that one word produces another word) (Nelson, Zhang & McKinney 2001: 1153), calculated by dividing #P (the number of participants producing a particular response – column 5) by #G (the number of participants in the group norming the word – column 4).

Importantly, the general free association procedure used to identify a word's associates and the strength of connections is similar to that used for the identification of most predictable meanings by the check-group: "Initially, a word is presented to a large group of participants who are asked to produce the first word to come to mind. The probability of a particular associate is determined by dividing its frequency as a response by the sample size" (Nelson, Zhang & McKinney 2001: 1147).

Since conversion is based on a single word stem (single WF base) it lends itself better than two-stem units for the examination of any links between

meaning predictability and associative meanings. Therefore, the top PR readings of converted naming units of this experiment were confronted with the database associates. I was looking for correlations between the PR and FSG values and between the respective rankings of the two groups of meanings.

Section 4.6.1 presents and comments on the results pertaining to the best predictable readings of the converted naming units of Experiment 2 and the corresponding associates based on Štekauer (in press[1]), and Section 4.6.2 examines possible relations between meaning predictability and associates of two-constituent naming units.

4.6.1 Meaning predictability of conversions and the associative principle

to boy

There is a partial correlation between the results. Thus, the top-rank association 'girl' corresponds with the second-rank PR reading 'to (try to) look/behave like a boy (clothes, haircut, motions, walking) – of girls', and the rank 4 associate 'man' corresponds with the top PR reading 'to act or behave the way boys do (immature)'. A partial and indirect relation between the reading 'to bring up a male child' (3 occurrences in the NNS group) and the 'immature behaviour' (contained in the top reading) may be partly identified in the associate 'child'.

to lion

'King' can be associated with the PR reading 'to have the power and rule/to act as a leader'. It is, however, surprising that (in contrast to my experimental research) there is no associate indicating a lion's physical characteristics, such as 'strength', its look (mane), and behaviour (brave, lordly), i.e., the prototypical features that one would expect to appear in an association database.

to tulip

It seems that one can expect some relation between association meanings and meaning predictability in cases pertaining to *appearance* and *behaviour*. This is clear with animate beings like the experimental naming units *to boy* and *to lion*. Since tulips have no behaviour an important factor of potential association-inspired interpretation is lost. What remains is appearance, which is reflected in the relation between a low PR 'to look beautiful' (4 occurrences in the NNS group) and a high-rank (3) but very low-FSG (0.027) associate 'pretty'. The associates based on the hyperonym-hyponym relation do not seem to have any significant influence on meaning-interpretation. This influence is highly limited and indirect only: any flower (but not any plant) can be picked; any flower

and plant are growable and/or can be used for decoration. In this connection it should be noted that the readings in which a converting noun becomes an Object of Action can hardly be expected to have any counterparts among the associates due to the universal nature of the relevant Action (pick, grow, decorate).

In the check-group 3 out of 25 informants selected the reading 'to be shy, to blush with shame' which may be related to the 'red' associate, but apparently this link is rather loose taking into account the very low FSG (0.020) of the associate and the low frequency of occurrence of the respective reading.

The 'Holland' associate (its FSG approaching zero) has even weaker links to a single-occurrence reading 'to go to Amsterdam' in the NS group and a two-occurrence reading 'to go to Holland' in the NNS.

to planet

The link is indirect and loose in this case. The high PR readings concerning space travel, colonisation, and also discovery of other planets may be related loosely to the associates 'space' (FSG = 0.042) and 'universe' (FSG = 0.018) indicating remoteness, thus establishing a kind of general framework for the identification of more specialised readings. The associate 'orbit' (FSG = 0.013) is related to the corresponding Action by two occurrences with very low rating.

to river

The only strong associate 'stream' (FSG = 0.118) may be related to the low PR 'to flow like a river; the highest PR 'to swim/bathe in a river' relates to the 'swim' associate, the FSG of which approaches zero (0.016), and the rank (19) of which is far from corresponding to the significance of the PR reading. The second-rank PR reading 'to travel down a river (in a canoe, kayak, etc.)' does not get more significant support from the associates 'boat' or 'canoe' either (the FSG of each of them is 0.055). The top-rank associate 'lake' has no PR counterpart.

to conference

The 'meeting' associate (FSG = 0.390) seems to be linked with both of the highest PR readings. The connection between 'to take part in a conference' and 'people' is not strong enough due to a very low FSG (0.027) of the latter. There seems to be a correlation between the third strongest associate 'talk' (FSG = 0.082) and the rank 2 reading 'to meet to talk about/consult/discuss a subject' in terms of their respective rankings. Unfortunately the FSG of the former is not very high.

to triangle

Triangle does not confirm a correlation between the meaning prediction and the association processes either. The 'geometry' (FSG = 0.095) and 'shape' (FSG = 0.027) associates are too general to directly motivate the top readings of *triangle*, 'to draw a triangle' and 'to make a triangle from something/to make triangular things'. The only direct connection seems to exist between the reading 'to be a part of a triangle of lovers' and the associate 'love'. However, the rank (12) and the FSG-value (0.014) of this associate are very low, comparatively much lower than the PR of the corresponding rank 3 reading, to be of any major significance.

The structure of the associates suggests a process leading in a completely different direction compared to the prediction process. While top ranks among the associates are assumed by incompatible geometrical figures, i.e., square, circle, rectangle, they, because of their incompatibility with triangle, cannot and do not play any role in the prediction process of *triangle.*

to morning

While the associate 'early' (FSG = 0.176) points to the topmost PR reading 'to get up early in the morning', a more direct association with this reading is provided by the associate 'wake'. Unfortunately, its rank (18) and FSG (0.014) are too low. Since the verb *to morning* necessarily implies a kind of activity done by an Agent in the morning, the associate 'coffee' (FSG = 0.034) is a good candidate for correlation with the rank 2 reading 'to do one's morning routine; to perform usual morning activities'. Associates like 'sun' and 'sunrise' do not pertain to the Agent's activity, and therefore can hardly motivate the interpretation of the verb *to morning*.

No analysis could be done for the experimental naming units *courtyard* and *cableway* as they are not included in the association database.

4.6.2 Summary 5

The picture obtained from the analysis of associates and predictable readings indicates that the first of the above-mentioned predictions, expecting at least some degree of overlap between associates and the most predictable readings, has been borne out for some of the readings. The other prediction concerning the correlation between rankings of associates and most predictable readings has not been substantiated.

The answer can be sought in the basic principles of the meaning-predictability theory outlined here, in particular, the role of the *prototypical*

semes of naming units (reflecting the prototypical features of the corresponding referents).

Association seems to be based on a variety of relations. A significant portion of them are represented by the *lexical relations* of antonymy and hyponymy/ hyperonymy. The associates based on the former (e.g. 'square', 'circle' for *triangle*), cannot correlate with predictable meanings for the simple reason that the prototypical features of antonymous words are incompatible, mutually excluding, and therefore cannot motivate the interpretation of new naming units. The same is true of associates constituting what Cruse (1986) calls helices as a special type of non-hierarchical lexical configurations ('night – evening – afternoon – day' in relation to the associate 'morning').

The hyponymy/hyperonymy relations must be excluded for a different reason. If the meaning of the word in question is a hyponym of a corresponding hyperonym, the latter necessarily functions as an identification or a classification seme of the former. My basic assumption is, however, that it is the prototypical rather than the too general identification and classification semes that motivate the meaning of novel naming units. It is for this reason that the motivating capacity of hyperonyms ('animal' and 'beast', for example, with regard to *to lion*, and 'man' (in the sense of human being) and 'child' with regard *to boy*) is limited. Co-hyponyms are not good 'motivators' either (for instance, 'tiger' and 'cat' with regard to *to lion*).

A different situation involves those antonyms which, in the interpretation process, are not perceived from a contrastive perspective; instead, the target and the cue enter into the relation of *Similarity*. A case in point is the *boy – girl* 'pseudo-antonymy' where the behaviour of boys serves rather as a Pattern for the behaviour of girls. The other relevant associates, 'man' and 'child' show that the Pattern-based (Similarity) relation may establish a certain correlation of results. Similarly, one of the associates of *to lion* is 'to roar' which is based on the Pattern relation (similar behaviour), in the same way as the 'king' associate (similar power/strength).

Another important observation is that the associative principle does not seem to permit more specific interpretations, such as 'to give birth to a boy', and 'to change sex from girl to boy' where the cue is the Result of Action. The reason is that this kind of reading combines two dimensions, Action and its Result. Contrary to this, association appears to be a single-dimension phenomenon. In principle, since the associative relation between cue and target is single-dimensional it precludes the associates from providing the *level of detail* necessary for the interpretation of novel, context-free naming units.

4.6.3 Meaning predictability of two-constituent naming units and the associative principle

It goes without saying that the assessment of the interrelation of meaning predictability and the associates of a naming unit constituents is more complex compared to conversions because it si necessary to take into consideration two associates, their mutual relation (if any), and the relation of any of the associates to the predictable reading(s). Ideally, the associates of both constituents contribute to the most predictable interpretation of a naming unit. If only one of them contributes to meaning predictability the association with the onomasiological mark may be expected to be more valuable: it is the mark that specifies the general, class-related meaning represented by the onomasiological base.

The following analysis makes use of the experimental two-constituent naming units. Its results suggest that there is only one case (*ball hammer*) in which the two constituents (mark and base) share a common associate, in particular, 'hit'. While the respective FSG values of 'hit' are identical (0.022) they are too low; in addition, the respective rankings of the associates are not high enough (rank 10 and rank 4, respectively) to play a substantial role in the meaning-prediction process. This is also borne out by the fact that the 'hit' associate is not of much use for the most predictable reading 'a hammer (the top of) which has the form of a ball' (PR = 0.420; OPR = 0.681). It may play a limited role at a very general level in predicting the rank 3 reading ('a (special shaped) hammer used for ball-like components') whose PR is, however, very low (0.084).

While there is one relevant associate for each of the constituents of *game wheel* they do not meet in a single interpretation; instead, each of them points to a different predictable reading. The database shows that the most natural associate (as might have been expected) of the onomasiological mark *game* is 'play' (FSG = 0.250). While it specifies the meaning of the base *wheel* in the direction of two most predictable readings ('a wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games': PR = 0.302; OPR = 0.490; and 'a wheel which is a part of a game equipment, a wheel with which a game is played': PR = 0.245), the specification is provided in very general terms. This associate establishes the framework for more precise interpretations as specified in the two top-rank readings of my research. On the other hand, while the FSG of the rank 4 associate 'fortune' of the onomasiological base *wheel* is low (0.084) it directly contributes to at least one, more specific, part of the most predictable reading.

A similar situation is encountered for *baby book*. Its most predictable readings are 'a book for babies (fairy tales, rhymes, pictures, drawings)', 'a book about babies and how to take care of them', 'a book with photos of one's baby, with regard to baby's development . . .', and 'a (very) small book'. Among the associates of *book*, we find 'read' at rank 1 (FSG = 0.130). This associate conditionally correlates with the most predictable reading of *baby book* (PR = 0.727; OPR = 0.574). The associate of the onomasiological base introduces the missing (unexpressed) determined constituent of the onomasiological mark and relates *baby* and *book* by the concept of READING. This interrelation may be viewed as a suitable foundation (first step) for a more penetrating interpretation step. The promising correlation is, however, weakened when one reflects the other part of the most predictable reading, the one referring to pictures and drawings (a more appropriate identification of the purpose of a 'book for babies'). The 'reading' associate does not say much of what actually is done with the 'book' (unless we do not consider looking at pictures and drawings to be the act of 'reading').

The rank 3 associate 'small' of the onomasiological mark *baby* correlates with the rank 4 reading 'a (very) small book' (PR = 0.157) and expresses the [Quality] of the base. Its very low FSG (0.077) reduces the significance of this association.

There are two naming units in the sample in which the mark associate does not contribute to the most predictable reading(s). While the associates of *dog* of the naming unit *dog spade* share no common features with its most predictable readings, a combination of the second-rank associate 'shovel' (0.155) and the sixth-rank associate 'dog' (0.027) of the onomasiological base *spade* might indicate indirectly the use of a spade in the function of a shovel for an ACTION whose object is dog or an ACTION somehow related to dogs. The 'scooping-up of dog's excrement' (PR = 0.137, OPR = 0.548) cannot be inferred from this sort of association. There are no associates pointing to the second-rank reading ('a spade in the shape of a dog') that is based on the semantic components [Shape], [Pattern].

The mark's associate plays no role in the meaning-prediction process of *garden whisky*. That is to say, none of the *garden* associates indicate the most predictable reading 'a whisky served in the garden (on various occasions, for example at parties, barbecues, etc.)' (PR = 0.327; OPR = 0.490) which is based on the [Location] component, or the second-rank reading ('a whisky made from garden products') which is based on the [Ingredient] meaning of the onomasiological mark of *whisky*. Of more use is thus the rank 3 associate of the

onomasiological base *whisky*, i.e., 'drink' suggesting an ACTION taking place in the garden, even if its FSG is low (0.081).

In the next group of naming units it is the associates of the onomasiological mark that are more significant for the meaning-prediction process than those of the onomasiological base.

In the naming unit *hill star*, none of the associates of either *star* or *hill* are directly related to any of the two most predictable readings. An indirect association to the most predictable reading ('a person who is brilliant at hill climbing/running/cycling': PR = 0.152; OPR = 0.521) comes from the onomasiological mark, i.e., *hill's* rank 2 associate 'climb' (FSG = 0.064). However, there is no hint of the 'brilliance' and/or 'excellence' of a human being performing this activity. Furthermore, there is no indication of the relation between the semantic components [Location] and [Visibility] that are crucial to the rank 2 reading ('a star that can be seen beyond/above the hill': PR = 0.119).

With *apple-juice seat*, the rank 3 'drink' associate of the onomasiological mark *juice* indicates the Purpose of the base (*seat*), inherent in the most predictable reading 'a seat for drinking apple juice' (PR = 0.485; OPR = 0.648), although the more specific meaning implying a 'restaurant seat reserved for non-alcoholic drinks' can hardly be inferred from the associate 'drink'.[17] None of the associates of the base *seat* contribute to the interpretation of the compound.

While no indication of the most predictable reading of *shape cloth* comes from the base constituent *cloth* a relatively straightforward reference to the top reading 'elastic cloth shaping woman's figure, very tight clothes shaped by the form of the body' (PR = 0.098; OPR = 0.476) is contained in 'figure' (FSG = 0.161), the second-rank associate of the onomasiological mark *shape*. A similar association is provided by the sixth-rank associate 'body'. However, its FSG value is low (0.039).

The 'old' (FSG = 0.391) associate of the mark constituent *age* is the top-rank associate, unambiguously contained in the most predictable reading 'an old bag; a bag that looks old' (PR = 0.203; OPR = 0.744). None of the associates of the base constituent *bag* help to identify the particular reading of *age bag*. The second predictable reading 'a special bag designed for a specific age category' is indicated by various associates of 'age' referring to a specific age, such as the fifth- and the sixth-rank associates 'eighteen' and 'nineteen', respectively. Unfortunately, both the particular predictable reading and the associates feature very low values (PR = 0.053; FSG = 0.025).

While the clearly dominating, most predictable reading of *feather-dialler*, i.e., 'a feather(-like instrument) used for dialling the phone' (PR = 0.212; OPR = 0.752), is not (and can hardly be) indicated by the associates of either *feather*

or *dialer*, the 'light' (FSG = 0.205) and the 'soft' (FSG = 0.045) associates of the onomasiological mark may suggest the 'light touch dialling' part of the second-rank reading 'a type of telephone that requires a very light touch to dial', even if no direct reference to 'telephone' is contained in the associates. This kind of association is impaired by a low PR value of the relevant reading (0.055).

Finally, we are left with the naming unit *flower hat*. None of its 41 associates for *flower(s)* and 12 associates for *hat* indicate any of the most predictable readings

4.6.4 Summary 6

The analysis of 'two-constituent' naming units has borne out the results obtained for converted naming units. There is some degree of overlap between associates and some of the most predictable readings. There is hardly any correlation between rankings of associates and most predictable readings. If there is any correlation it usually is of a very general nature, and does not make it possible to identify the particular interpretation of a two-constituent naming unit.

The FSG values of those associates that might contribute to the meaning-prediction process are usually very low (mostly under 0.1). An exception to the rule is the rank 1 associate 'old' for *age* that correlates with the most predictable reading of *age bag*. The importance of a relatively high FSG (0.205) of the associate 'light' for the mark *feather* is reduced by the negligible PR of the relevant meaning of *feather dialler*.

The analysis of the micro-sample has also shown that correlation at a very general level may come from both constituents, or only from the onomasiological mark or the onomasiological base. Any conclusions concerning my postulate of a more significant role of the mark associate compared to the base associate are conditioned by implementation of a much more extensive research.

Chapter 5

Conclusions

5.1 General

The primary goal of this research was four-fold: (1) to develop a theory of the meaning-predictability of context-free novel naming units as an integral part of a general onomasiological theory of word formation; (2) to identify the factors that influence the meaning-prediction process either positively or negatively; (3) to propose a method of calculating the Predictability Rate and the Objectified Predictability Rate; and (4) to verify the theoretical considerations and hypotheses in an experimental research, by applying the proposed method to the whole range of potential and non-established naming units falling within five Onomasiological Types.

There are at least two different kinds of meaning predictability; context-free meaning predictability and context-bound meaning predictability. They may also be labelled *system-level predictability* and *speech-level predictability*. Since these two kinds of meaning predictability have a bearing upon two different language levels, they necessarily operate under unequal conditions, and consequently require different methods of research, account, and calculation. The present research focused on the former type. To put it another way, I was interested in the ability of speakers (both native and non-native) of a language to predict the meaning of productively and regularly coined naming units as well as those which violate the principles of productive word formation when they encounter such a naming unit for the first time without any contextual support.

It is important to note that – as with any other research of a limited range and extent – the results should be taken as indicative. Their relevance may be strengthened by a series of analogous research projects. Moreover, as in any other research in which the subjective factor plays a crucial role, the results obtained may have been influenced by subjective evaluation. The relevant factors include:

1. *Grouping of the proposed readings*. In the majority of cases the grouping of the readings proposed did not pose any problem. For illustration, the

paired readings 'comedian' and 'clown' and 'comedian and 'humorist' for *smile-man* clearly belong to the same 'cohort' of interpretations. However, in some other cases, I faced the problem of whether certain readings should be grouped together or classified separately in two different cohorts. For example, while the majority of readings for *boy*$_V$ fall within one of two cohorts, i.e., 'to behave like a boy – of adults' and 'to behave like a boy – of girls', the same group of informants avoided this kind of 'classification' for *lion* and preferred to point out various typical qualities of 'lion', such as beautiful mane, courage, strength, and the manner of eating ('to devour like a lion'). The reason may be that there are no striking and symbolically relevant features for 'boy'. While 'boyhood' is understood rather as a sum of more or less equally significant features, the notion of *lion* is dominated by features like [Predator] in the most predictable reading, [Strong Voice], and especially [Extreme Physical Strength], implying the ruling position in the world of animals, the characteristic Manner of [Behaviour], etc.

Now, the question is whether or not the various readings that are based on the prototypical features of 'lion' should be grouped together (which would, however, mean a kind of overgeneralisation of the individual informants' readings) or evaluated as they are proposed. The latter alternative necessarily reduces the Predictability Rates of these readings.

Similarly, should the readings of *ball hammer* 'a hammer used for doing something to balls' and 'a (special shaped) hammer used for ball-like components' be grouped together or evaluated separately? Should the *game wheel* readings 'a wheel for playing roulette and casino games; a wheel in the Wheel of Fortune type games' and 'a wheel which is a part of a game equipment, a wheel with which a game is played' be evaluated as a single reading or two readings? Should 'to make a triangle from something/to make triangular things' and 'to draw a triangle' be grouped together on the grounds that the Result of Action is an object of a triangular shape, or treated separately because of the three-dimensional vs. two-dimensional nature of the Results of Action? Should drinking coffee, doing one's morning exercises, or applying make-up be included in the reading 'to do one's morning routine/to perform usual morning activities', or should they all be treated separately on the grounds that they are not general enough unlike the hygiene-related morning activities?

I am aware of Coolen et al.'s experiment (1991) in which they employ three independent judges to classify the readings proposed by the informants, according to the classification criteria adopted from Levi (1978). Murphy (1988) also makes use of judgments of two 'raters' to identify the num-

ber of meanings proposed by his informants. While this solution seems to be more objective I rejected it for the reason that Levi's classification is too general and cannot capture subtle semantic differences in the proposed readings. As a result, it does not lend itself to the identification of the individual, usually very specific, readings in a research into meaning predictability within the proposed framework. The grouping task proved to be both demanding and simple. It was demanding because there were no clearcut boundaries between the individual readings (fuzzy edges), and it was simple because the informants themselves solved the grouping problem in the great majority of cases. In principle, in classifying the readings I observed three basic criteria:

a. the classification resulting from the informants' proposals themselves: if two semantically cognate readings were proposed separately as two different readings by one and the same reader, they were classified separately;
b. the semantic identity of the readings (the same interpretation expressed in different words, in different paraphrases) or close semantic cognation;
c. the same degree of generalisation (for this reason, the readings, such as 'to act or behave like boys do' and 'to speak like a boy' were not grouped together).

2. Attention paid to completing the task by the individual informants, i.e. the degree of their *involvement*. The comparison between the Slovak and the native speaker groups indicates that while the native speakers generally proposed a smaller number of readings for the individual primary compounds, the number of proposals of Slovak students was higher. This may indicate two things. First, my students appear to have been more enthusiastic and have taken the research as a kind of creative game. On the other hand, I could not personally influence native speakers' attitude to the research because I did not come into contact with them. Second, the generally smaller number of readings proposed by native speakers may also indicate their higher certainty about the 'appropriateness' of their proposals, which may go hand in hand with their better command of their mother tongue. In any case, these considerations only concern the number of different readings. As indicated in the previous chapter and summarised below, the differences in the PR values between the native and the non-native groups of informants are negligible.
3. *Linguistic competence and awareness* of the individual informants.

4. Their *imagination*, *ingenuity* and '*creativity*'.
5. Their *extra-linguistic knowledge*, including their personal *experiences*. This point appears to comply logically with Aitchison's note (1987: 155) suggesting that there is "some intriguing evidence that highly educated people find it easier to cope with compounds, at least in the absence of strong contextual clues" as well as with the findings of Gleitman & Gleitman (1970) (see Conlusion 6 below). It may be assumed that the level of education affects one's linguistic competence, including the meaning-prediction competence. No attempt, however, has been made within my research to verify this logical postulate.

Other influencing factors include:

6. *Formulation of the task* – the informants were asked to propose as many different meanings as they could think of. No doubt, asking them to propose only those meanings which, in their view, have a chance of being accepted by a speech community (the most predictable readings) would slightly modify the individual PR/OPR values. But it may be assumed that these values would not affect the overall results (ranking of the most predictable readings) or conclusions. These expectations have been, for the major part, borne out by the check-group whose task was to propose only the most acceptable reading for each naming unit. Moreover, asking the informants to propose as many readings as possible made it possible to evaluate the mutual influence of the individual competing readings that come to one's mind when processing a novel, context free naming unit. This competition of readings is captured in the notion of Objectified Predictability Rate.
7. The evaluation relies, *inter alia*, on a *classification of semes* which is only one of several possible classifications and – as any other classification in this field – is open to discussion.

Before proceeding to the evaluation of the research results as a whole one methodological remark is required. J. Grzega (personal communication) proposes some sort of 'onomasiological check' which should answer the following question: for which meanings given by the readers would speakers have ever coined the words? He illustrates the point by *age bag*, and assumes that for readings like 'a period when bags were/are popular' no sensible speaker of English would ever coin a word *age bag*, and adds that the meaning was probably only desperately 'invented' due to the lack of an extra-linguistic reality and in order to have 'just some meaning in the questionnaire'. The same problem, namely the lack of a possible referent/concept, is connected, in his

view, with *apple-juice seat*. Grzega asks: "Why should there ever be a very predictable/probable meaning without an extra-linguistic reality? There must also be an extra-linguistic predictability, an extra-linguistic probability."

This position calls for several comments. I start from the end by claiming that it is necessary to distinguish between the predictability of meaning of a naming unit first encountered by a language user, and the probability of an object. New naming units are always responses to the 'events' or 'phenomena' of extra-linguistic reality. If, for example, a scientist comes up with a new theory introducing new concepts of objects unknown to mankind before (a fairly common case) these objects immediately 'become' an integral part of our extra-linguistic reality. Since extra-linguistic reality is – to us – an open system, there always exists the above-required 'probability' of existence of *any* real and unreal object/phenomenon that can be named. Since language makes it possible to give a name to 'anything' nameable, nameworthiness is a matter of degree.

Second, I agree with Grzega that hardly any speaker would propose the naming unit *age bag* for the reading 'a period when bags were/are popular'. The proposed 'onomasiological check' is, however, provided by the experimental results themselves: this reading gained support from one single informant. This *is* the sort of check required. In addition, this reading – as suggested above – must be rejected due to its unacceptable WF Type. This is another form of check of the admissibility of readings. Due to these circumstances the reading in question is unpredictable, which is the check bringing the results required by Grzega.

Thus, it follows that the method used in this research does have the *mechanisms enabling one to identify inadmissible readings*.

5.2 Conclusions

The following conclusions may be drawn from this research:

Conclusion 1 Meaning prediction capacity
My experiments confirmed hypothesis 4, articulated in Section 3.11, of *an equal meaning-prediction capacity of native and non-native speakers*. This translates into an assumption that a fluent speaker's meaning-prediction process heavily relies on his/her cognitive knowledge and experiences. The hypothesis has been borne out by both the remarkable predictable reading ranking agreement between the native and non-native speakers and by mostly insignificant

differences in the respective Predictability Rate values. Furthermore, this coincidence has been enhanced by comparison of the results of the main and the check groups working under different experimental conditions. At the same time, it was shown that this state of balance may be violated by differences in extra-linguistic reality, cultural traditions, and ways of life.

Conclusion 2 No 'correct' reading

There is no place for the term 'correct' reading in a theory of meaning predictability. This assumption derives from the concept of word formation as creativity within productivity constraints. When coining a new naming unit, a language user can usually select from a range of options available to him at the level of conceptual analysis, at the level of identifying the motivativating semantic components (onomasiological level), and at the level of morphological realisation (onomatological level). As a result, while the coiner produces a new naming unit with one and only one meaning in mind, his/her preferences applied to the onomasiological and onomatological options need not coincide with an interpreter's expectations. A naming unit reading with a high PR need not correspond to the meaning with which the naming unit was coined by a language user. Whether or not the most predictable reading is also the best one from the point of view of the meaning attached to a new naming unit by its coiner is a different issue. Since a coiner may form a naming unit the meaning of which does not correspond to what Murphy (1988: 539) calls the *best fitting slot* in the head noun's schema there need not be any systematic correspondence between the ideal slot filling, the meaning predictability, and the actual meaning of a new coinage. By implication, it is not appropriate to relate meaning predictability with any 'correct' reading. All readings are 'correct' if they are interpretable, that is to say, if they are *acceptable to a speech community as meaningful representatives of a particular novel naming unit*.

Conclusion 3 Prototypical semes

There is an obvious *tendency for the predictable readings to be motivated by prototypical semes* (level 4 semes) or their combinations, reflecting the prototypical features of relevant objects.

This conclusion gives support to hypothesis 3, articulated in Section 3.11. Each of the ten top readings in Experiment 1 is based on at least one prototypical, level 4 seme. Nine of the top readings in Experiment 2 are motivated by a prototypical seme.

Any *figurativeness*, that is to say, any *semantic shift* (metaphor and metonymy) *appears to become a serious obstacle to meaning predictability*. This

assumption only applies to cases in which one or both of the motivating words acquire figurativeness, as it were, within the framework of a new naming unit. This factor is closely related to *context-boundness* as one of the PR-reducing factors.

The research data provide numerous examples, including 'a book with a cover of a baby skin colour', 'a book that smells like a baby', 'one's favourite book' for *baby book*; 'a clumsy person with poor dancing skills', 'an aggressive person' for *ball hammer*; 'a star that sleeps on a hill', 'Noah', 'a famous actor who lives in Beverly Hills' for *hill star*; 'one's life' for *game wheel*, 'somebody very eccentric in dressing' and 'somebody who tries to be perfect in dressing' for *shape-cloth*; 'a group of people of the same age' for *age bag*; 'a spade by which a dog was/is killed' for *dog spade*; 'to turn red from embarrassment' and 'to behave effeminately' for *to tulip*; 'to urinate excessively' and 'to have an orgasm' for *to river*; 'to feel blocked in (like in a courtyard)' for *to courtyard*; 'to be cheerful/ merry' for *to morning*; 'a light beer' for *blondesjoker*; 'someone who talks softly on the phone' for *feather-dialler*; 'human waste' for *refusnik*; 'a paranoid person who thinks all white powder is anthrax' for *anthraxist*; 'a mentally retarded person' for *smile-man*; the contradictory readings of 'to undermine the plans of someone' and 'to support someone', or 'to provide a foundation' for *to befoot*; 'an orphan' for *leave-behind*$_{N}$.

All of these and other similar metaphorical interpretations remained isolated attempts restricted to a single informant. By implication, they have no chances to get above the predictability level.

The meaning predictability of new naming units based on a figurative meaning is a part of a general problem of polysemantic lexemes as well as homonymous words that become constituents of new naming units because the one-many relation between form and meanings hampers the meaning-prediction process. Any necessity to select from optional meanings can be considered as a predictability-reducing factors as it potentially leads to a higher number of interpretations. However, its strength and effects are not directly proportional to the number of potential interpretations. They are influenced by the degree of institutionalisation of the individual meanings, their respective relation to the other motivating constituent(s) of a coinage, and the effects of the other predictability-boosting and -reducing factors.

On the other hand, a well-established, i.e., *institutionalised figurative meaning of a naming unit does not have any negative impact on meaning predictability*. A case in point is the figurative, personified meaning of *star* in *hill star.*

Too general or too specific and idiosyncratic features (semes of levels 1–3 and 5, respectively) *tend to be a serious obstacle to meaning predictability*. The re-

search provided a number of examples of too general semes underlying the proposed readings, which, for this reason, were restricted to a single or two occurrences: 'to blossom' for *tulip*$_V$ (any flower blossoms); 'to exclaim 'boy' for *boy*$_V$ (any Noun → Verb recategorisation can be interpreted this way); 'a spade by which a dog was killed' for *dog-spade* (any [Solid] SUBSTANCE can be used for this purpose); 'a hammer used for doing something to balls' for *ball hammer*; 'to exclaim 'boy" for *boy*$_V$ (any recategorised naming unit could be interpreted this way); 'to blossom', 'to be interested in flowers', and 'to become beautiful in spring' for *tulip*$_V$ (any flower may imply these figurative meanings); 'to buy a courtyard' for *courtyard*$_V$ (any object can be bought); 'a person who constantly cracks stupid jokes' for *blondesjoker* (this can be any 'joker'). Similarly, there were a number of readings unpredictable due to their motvation by a level 5 seme, for example, 'a spade used by dog for playing' in *dog spade*; 'a book in the shape of a baby' in *baby book*; 'a hammer whose one part is spherical' in *ball hammer*; 'a hat with the odour of flowers' in *flower hat*; 'a special seat filled up with apple-juice' in *apple-juice seat*; 'to wear a yellow sweater and green trousers' in *to tulip*; 'to run or move very fast' in *to river*.

A prototypical, level 4 seme and/or the combination of level 4 semes is an important, but not sufficient precondition for a high meaning predictability of novel, context-free naming units. Motivation by prototypical semes does not automatically guarantee a high Predictability Rate. A case in point is the reading 'a book written by a baby' for *baby book* which is based on a 4–4 Seme Level Combination. Here it is an extra-linguistic factor that works against the predictability of this reading: the activated semes of *baby* and *book* are incompatible.

In the Epilogue of *The Indo-European Lexicon: A Full Synchronic Theory* (1981) Beard compares his approach to the range of possible meanings of a derivative with that developed by Clark & Clark (1979). While the Clarks restrict the range of possible meanings of what they call 'contextuals' (conversion or, my 'conceptual recategorisation') by the range of logical possibilities, Beard's answer is based on the system of Indo-European categories. While the Clarks base their approach on performance Beard does justice to competence-bound linguistic regularities, in particular on lexical-syntactic paradigms.[1] He admits that these regularities can be overridden by performance, and therefore the readings of, for example, *to bottle* and *to teapot*, which deviate from those predicted by the respective rules, are treated in Beard's approach as exceptions.

In any case, apart from the differences between the positions of the Clarks and Beard, resulting from their different viewpoints, the two theories are very close to each other. The Clarks' generic knowledge, underlying the interpretation of 'contextuals', is based on the classification of extra-linguistic objects and

relations in terms of *placeables, agents, results, instruments,* etc., a system which is similar to Beard's Indo-European case function system.

Both Clark & Clark's and Beard's classifications specify possible interpretations of novel naming units. While Clark & Clark emphasise extra-linguistic knowledge Beard puts emphasis on linguistic categories. The idea of prototypical semes, employed in this book, bridges these two approaches because it assumes that prototypical semes as semantic units at the level of language (the meaning facet of linguistic signs) are direct reflections of conceptual categories (logical predicates) at the level of the human mind which, in turn, reflect the prototypical features of an object named. From this it follows that meaning predictability, crucially dependent on prototypical semes, is influenced and conditioned by both *linguistic and extra-linguistic factors*. My account in this book has demonstrates that – as noted above – the predictability of novel, context-free naming units is a matter of *competence*, langue, the system of language. In this respect it differs from the meaning predictability of contextualised novel naming units which falls primarily within the scope of performance. This brings us to Conclusion 4.

Conclusion 4 Crucial role of extra-linguistic knowledge and experiences
My research has confirmed the idea that lexical meaning itself is not sufficient for interpreting/predicting the meaning(s) of novel naming units hypothesis 4, Section 3.11). There is abundant evidence that the mere meaning-identification, i.e., the comprehension of objects representing the individual naming unit constituents, is insufficient for the identification of acceptable/predictable readings. Thus, for example, the semantic structures of *baby* and *book* may indicate possible combinabilities: for instance, the semes [+Listening Capacity] [+Perception by Watching] of *baby* can be activated in combination with any of the semes [For Reading/Listening/ Perception by Watching], [Having Some Content], and/or [±With Photos/ Pictures] of *book*, but this combinability cannot identify the subtle semantic distinctions as represented by the three most predictable readings for this possible naming unit, i.e., 'a book about babies – how to take care of them', 'a book for babies', and 'a book with records of one's baby'. These readings result from a conceptual analysis and evaluation of the possible relations between the objects conceptually processed and related within the new naming unit. They reflect language users' knowledge of the class of 'babies' and the class of 'books' and, crucially, also their experiences. Perfect knowledge of the prototypical features of 'book' and 'baby' can hardly be sufficient for proposing a reading referring, for instance, to records about one's baby. What is also needed is a knowledge about/experience

with keeping records of baby development. A member of a primitive tribe living in a rain forest or a six-year old child living in our civilisation can hardly be expected to propose this reading, because neither of them may be expected to have the necessary knowledge/experience of (the habit of) keeping records of babies. This factor prevents them from predicting this reading even if they have the concepts of both 'baby' and 'book' in their minds.

Similarly, the reading 'a spade used for scooping-up a dog's excrement' for *dog spade* cannot be inferred from the lexical meanings themselves. One must have the knowledge of/experience with spades from which one can infer that spades – apart from their basic function – can be and actually are sometimes used instead of shovels. In addition, this reading also requires the knowledge/experience concerning the habit (conditioned by relevant cultural habits/by-laws) of removing a dog's excrement.

The meaning subtleties of the individual predictable readings of *flower hat*, notably, 'a hat with flowers on it', 'a hat made from flowers', and 'a hat with flower patterns' also require a considerable amount of knowledge/ experience. One should know that there are some traditions of decorating hats with flowers, that rustic children have long enjoyed weaving flowers into a decorative cover of the head which, by being worn on the head, might be labelled as a 'hat'.

The reading of *apple-juice seat* as a place in a restaurant, bar, or café, reserved for drinking apple-juice needs a huge amount of additional information: knowledge of the function and equipment of restaurants and similar establishments, knowledge of the existence of temperance bars for young people, and, possibly experience with bars, a part of which is reserved for smokers and the other part for non-smokers, which may suggest that certain seats in bars may also be reserved for '(apple)-juice drinkers'.

The reading of *to boy* 'to (try to) look/behave like a boy (clothes, haircut, motions, walking) – of girls' incorporates the knowledge of the social and cultural development of Western civilisation in the last few decades, including the emancipation tendencies purporting the principle of equal opportunities in women's professional careers and life-style.

The readings 'to travel to other planets/to land on a planet' and 'to populate/colonize another planet' of $planet_{V}$ rely on the knowledge of the possibility to travel to other planets and the desire/plans of mankind to populate them.

The interpretation of *blondesjoker* in the sense of 'a person who tells/makes blonde jokes' is primarily based on the general inclination of the present cultural setting to consider blondes to be less intelligent.

And to take one more example, the interpretation of the naming unit *refusnik* in its established reading of 'a political dissident (from the former

USSR, or from a totalitarian regime)' presupposes highly specialised world knowledge and/or experience. The knowledge of the meaning of its motivating constituents *refuse* and *-nik* does not help much.

Each of these and all the other predictable readings in my research confirm what has been emphasised at various places in this book, and add to a series of examples illustrating *the significance of a language user's extra-linguistic knowledge and experiences* for the meaning-prediction process. For these reasons, I do not share Meyer's (1993) assumption that lexical meaning is sufficient for the interpretation of relational compounds. Consequently, the feature weighting theory (Smith & Osherson 1984; Smith et al. 1988) seems to be untenable.

In fact, every predictable meaning gives support to the assumption that for a language-user to be able to predict the meaning of a naming unit (s)he

a. must know the lexical meanings of the motivating constituents;
b. must be able to conceptually analyse the objects of extra-linguistic reality which are covered by these lexical meanings; and
c. identify their possible relations based on his/her knowledge and/or experience.

Conclusion 5 Preference for stable relationships

The research described here confirms the general assumption of the *preference of stable relationships over fortuitous ones*. This assumption may be reformulated as follows: Those readings of novel, context-free naming units which express stable and habitual relationships and/or are based on prototypical features of the objects named show a higher meaning predictability. Out of 30 most predictable readings in the first three experiments 29 acknowledge this tendency. The only exception is 'a seat with apple-juice spilled on it'. The reason why this compound becomes 'appropriately classificatory' (to use Zimmer's term) may be the serious consequences that spilled juice may have for the functional appropriateness (see the prototypical features [Liquid] [Sticky] [Causing Stains on Cloth] and [For Sitting]) of such a seat. In fact, it makes it unusable (for certain period of time) if one wishes to avoid sticky stains on one's skirt/trousers.

In all other cases it is the general, stable, and/or habitual relation that was proposed by the informants. This follows from the absence in the paraphrases of the present continuous tense, and the use of words like *usually*, *habitually*, *can be*.

Conclusion 6 Acceptable Word-Formation Types

a. My research demonstrated an important role played by the Word-Formation Types, i.e., *Onomasiological Structure Rules* in the interpretation of novel naming units. An analysis of some problematic cases has already been given in Section 3.6.

In this Conclusion let me point out an interesting case that is related to the role of Onomasiological Structure Rules. One would expect that readings based on an unacceptable onomasiological structure may only occur in the group of non-native speakers because the 'knowledge' of the rules like those given in (30) (Section 3.6) should belong to the 'basic outfit' of a native speaker. It should be a part of a native speaker's subconsciously existing competence.

While this expectation is true of, for example, 'a flower in the shape of hat' for *flower hat*, 'a ball destroyed by hammer' for *ball hammer*, and 'vegetable meal with a little whisky', the contradiction between the reading proposed and the relevant onomasiological structure also occurred in the native speaker group. Examples include the reading 'a hill in the shape of star' for *hill star*, and the reading 'a period when bags were/are popular' for *age bag*. In the former case, there are only two occurrences of low scores (2 and 1 points assigned) in the NS group as opposed to four occurrences in the NNS group with higher (not high, though) scores (7, 3, 3, 3 points). In the latter case this type of onomasiological structure only occurred in the NS group! Similarly, a non-native speaker's reading 'a person who fell ill due to anthrax' is accompanied by a native-speaker's 'someone who has caught an anthrax disease'. Both of them ignore the Agentive meaning of the suffix *-ist*, and ascribe it a Patient-based reading which is bound to the suffix *-ee*.

This finding is not new, and is also reported by Gleitman & Gleitman (1970) who, in reference to compounds like *house-bird glass*, relate this kind of misinterpretation to the educational level of language users. Their informants fell within three different educational groups: (1) graduate students and PhD's in various fields; (2) undergraduates and college graduates; and (3) secretaries with high school degrees. In many cases their informants from the group of secretaries proposed various 'unacceptable' readings which, corresponded to the compound *glass house-bird, glass bird-house*, or a paraphrase like *a house-bird made of glass* (in contrast to PhD informants who avoided such errors). The analysis of their research results made Gleitman & Gleitman conclude that there were "very large and consistent differences among these subjects of differing educational background" (1970: 117) and that "[t]he less educated groups make more errors, and to a significant extent make different errors than the

most-educated group" (ibid: 128). Similar cases are reported by Ryder (1994). For example, one of her native informants interpreted *quilt-horse* as 'a quilt made of horse-hair'.

These results suggest that the constraint upon permissible interpretation of new naming units related to onomasiological structure is also a function of language proficiency. By implication, a proficient non-native speaker may be more 'skilful' in this respect than a native speaker of low language proficiency. This seems to be another reason supporting the assumption that if a non-native speaker masters a foreign language at a relatively high proficiency level (no matter how it may be defined) (s)he is not disadvantaged in his/her meaning-prediction capacity.[2]

What seems to be of significant relevance in this connection is the individual language user's intuition, intelligence, and education. Thus, Gleitman & Gleitman assume that giving language judgements – retrieving and making use of one's intuitions – is relatively hard, compared to talking and understanding" and is accompanied by "extensive individual and population differences in performance of harder judgmental tasks..." (1979: 107).

b. *Meaning predictability is influenced by the productivity of Onomasiological/ Word-formation/Morphological Types.* In principle, this influence can be both positive and negative, that is to say, it may have either predictability-boosting or predictability-reducing effects.

The *predictability-boosting effects of productivity* seem to be of a *general* nature in the sense of establishing favourable conditions for good meaning predictability of novel, context-free naming units. These positive conditions, however, are subject to various specific influences (see Conclusion 15 below), bound to a particular novel naming unit. The favourable conditions include

1. The underlying Onomasiological Type 1, 2, or 4, because the potential number of readings for these Onomasiological Types is much lower compared to Onomasiological Types 3 and 5.
2. The higher the productivity of Word-formation/Morphological Type the better chance to identify such a Type as one underlying the naming unit whose meaning is to be predicted.

The *predictability-reducing effects* of productivity concern those cases in which a naming unit can be interpreted on the basis of more than one highly productive, WF/Morphological Types, as illustrated by the Instrumental and Agentive interpretations of the suffix *-er* in the case of *feather-dialler*.

Conclusion 7 Word-formation types and meaning predictability
The general findings articulated in Conclusion 6 are completed with more detailed observations in this Conclusion. It should be noted, however, that the limited sample size does not enable me to draw any definite conclusions concerning the relation between WF Type and meaning predictability. If the top readings are taken into consideration the PR values of the most frequent WF Type, i.e., the [Stative – (State) – Patient], range from 0.727 (*baby book*) to 0.098 (*shape cloth*). The PR values of the Agent-based WF Types range from 0.564 to 0.212. These values indicate that there does not seem to exist any correlation between the internal structure of WF Type (onomasiological structure) and the PR value.

The [Stative – (State) – Patient] WF Type was much more frequent than the WF Types combining an implicit (Action). Actional (dynamic) relations tend to be expressed explicitly (but cf. naming units like *novelist*, *saddler*, etc.). In the [Stative – (State) – Patient] WF Type the Stative constituent may take the form of a Goal, Theme, Pattern, Material, Location, etc.

The research has not confirmed Meyer's assumption that the Purpose-based meaning is the preferred one for primary compounds with Artefact-denoting heads (1993: 111). The dominating meanings of my sample compounds *flower hat*, *ball hammer*, *shape cloth*, and *age bag* contradict Meyer's claim. Moreover, none of the top three readings for the Artefact-headed *flower hat* ('a hat with flowers on it', 'a hat made of flowers', and 'a hat with flower design/pattern/ornaments') express the Purpose relation. The PRs of the Purpose-based readings 'a hat to wear when you work in the garden', 'a hat to protect one from the falling pollen of flowers', 'a hat for protection of flowers', and 'a hat for flowers' fall well below the predictability level.

The only predictable reading of *ball hammer* 'a hammer, a part of which has the form of a ball', is based on the Pattern (Shape) relation rather than on Purpose. The Purpose-based readings 'a special shaped hammer used for ball-like components' and 'a (ball + hammer) toy for children to play with; a game' remained below the predictability level, too. Similar considerations apply to *shape cloth* and *age bag*, with single predictable readings.

The most frequent onomasiological structures for the predictable readings of Onomasiological Type 5 include those of the Result of Action (6 occurrences), Manner/Pattern and Location (4 occurrences each).

Conclusion 8 Onomasiological type and Predictability Rate
The research supports the expectations articulated in Section 3.5.1: the Predictability Rates of the top readings of naming units belonging to Onomasio-

logical Type 3 and Onomasiological Type 5 are, with some exceptions, rather low – the lower range of the low-predictability level (about 0.3). The average PR values of these two Onomasiological Types (top readings) do not differ significantly. The differences between these two Types may be summarised as follows:

1. a much wider range of PR values for Onomasiologial Type 3 compared to a much more compact picture provided by the PRs of Onomasiological Type 5;
2. the values of the non-top readings for Onomasiological Type 3 are higher: this is especially true of rank 3 readings. While four of ten 3rd rank readings in Experiment 1 (Onomasiological Type 3) are above 0.1, there is only one such case with Onomasiologial Type 5. The average PR value of rank 3 readings is much higher for the Type 3 naming units.

The PR values of nine out of ten top readings in Experiment 1 (Onomasiological Type 3) and all ten in Experiment 2 (Onomasiological Type 5) are (most significantly) below the value of 0.5, i.e. below the medium predictability level, which means that *the absence in the onomasiological structure of the Actional constituent* (Onomasiological Type 3) and *the absence of an onomasiological structure* (Onomasiological Type 5) are serious obstacles to good meaning predictability.

It may be hypothesised that the PR values of the top readings of Onomasiologcal Types 3 and 5 are – on average – lower than those of naming units falling within Onomasiological Types 1, 2, and 4.

This postulation follows from the fact that:

1. In Onomasiological Type 1 the polar members of the onomasiological structure (determining constituent of the onomasiological mark and the onomasiological base) are explicitly, and thus, (more or less) unambiguously related via an explicitly expressed determined constituent of mark. Since it is an Actional constituent, it is vital to meaning predictability. Its morphological expression is, therefore, an important predictability boosting factor.[3]
2. In Onomasiological Type 2 the explicit morphological representation of the base and the determined constituent of mark (Actional constituent) is usually sufficient for significant restriction of possible meanings.
3. In Onomasiological Type 4 the close relation between the base and the unstructured mark does not offer many interpretation possibilities.

4. The relation between the two polar members of an Onomasiological Type 3 structure is usually vague because of the absence of the 'relation-establishing' Actional constituent.

The conclusions of the low Predictability Rate of Onomasiological Type 3 naming units do not represent a hard-and-fast rule; rather they should be interpreted as a *strong tendency*. The case of *baby book* (the PR value of 0.727 of the top reading is almost at the high predictability level) suggests that exceptions to this tendency are possible. Meyer (1993) demonstrates that so-called *relational compounds* feature a very high predictability. As an example, the meaning predictability of *cupboard side* may be expected to be very high (for the reading 'a side of a cupboard'). In Meyer's view this is because the modifier satisfies the internal argument of the head noun. From the perspective of my research this fact is accountable by the prototypical, level 4 seme of *side* – roughly, [Surface of a 3-dimensional Object] that may be activated with regard to *cupboard*. In any case, the existence of other possible meanings like 'the side which is near the cupboard', 'the side which is opposite the cupboard', the 'side which is made up of cupboard(s)', activating other semes of *side*, indicates that this type of compound also requires a certain amount of conceptual analysis based on extra-linguistic knowledge and experience in order to give priority to the former reading over all the other readings.

Conclusion 9 Objectified Predictability Rate

The research data unambiguously demonstrate that the objectified parameter of meaning predictability primarily depends on the following mutually interrelated variables:

1. the PR value of the top reading;
2. the Predictability Rate Gap, i.e., the difference in PR between the top reading and the subsequent readings – the higher the PR Gap the higher the OPR;
3. the number of relatively strong non-top readings – the higher the number of such readings the lower the OPR;
4. the R1/R2 and R1/R3 ratios – the bigger the ratio the higher the OPR.

The analysis of the research data has provided me with numerous examples demonstrating that a high Predictability Rate of a top reading does not guarantee a high OPR, and that readings with very low PRs may feature a (relatively) high OPR thanks to the lack of competitive readings. This engenders a paradox: the existence of a relatively large number of semantically fairly well compati-

ble, possible, and predictable seme combinations, implying several acceptable meanings of a naming unit, appears to be an obstacle to the overall predictability due to tough competition between the acceptable (predictable) readings. On the other hand, poor compatibility of the semes of the motivating units need not entail a poor Objectified Predictability Rate owing to the absence of any competition for an (otherwise) low PR reading. For all these reasons the *non-relational* concept of Predictability Rate must be kept strictly apart from the *relational* notion of Objectified Predictability Rate.

While the PRs of the top readings of Onomasiological Type 3 and Onomasiologial Type 5 are roughly the same, the OPR data suggest that there is *a tendency for the Type 5 naming units to be slightly more predictable than those of Onomasiological Type 3*. The differences in OPR are necessarily bound to the unequal strength of non-top readings, reflected in unequal PR Gaps and R1/R2 (R1/R3) ratios.

Conclusion 10 One (two) dominant reading(s), or potential vs. predictable readings

Hypothesis 7, articulated in Section 3.11, has been confirmed. The experiments bear out the dominant position of a single (rarely two) strong reading(s) for almost all sample naming units.

While the rank 3 readings are in general of little significance (the average PR of rank 3 readings is 0.032 for Onomasiological Type 5 and 0.078 for Onomasiological Type 3) in terms of meaning predictability, rank 4 readings are totally insignificant (the average PR value of the ten rank 4 PRs for Onomasiological Type 3 is 0.031; for Onomasiological Type 5 it approaches zero). The rank 3 (0.227) and rank 4 (0.157) readings of *baby book* are the only exceptions to this rule in my sample. This brings us to one of the central conclusions of this research:

While there are many potential readings of novel, context-free naming units, it is usually only one or two that are significant in terms of meaning predictability. This casts doubt on the widely encountered sceptical picture of meaning predictability. The picture emerging from my research is much more optimistic.

Conclusion 11 Single-occurrence readings

The percentage of single-occurrences among the proposed readings is fairly high. This fact can be attributed to the task formulation requiring the informants to give 'as many readings as possible'. In other words, the informants were encouraged to give vent to their imagination. As a result the proportion of single-occurrences exceeded 50% in the vast majority of the sample naming

units in both NS and NNS groups. The same reason can be held responsible for the vague and, in some case, most bizarre readings. It may be assumed that the number of single-occurrence and vague readings would be much smaller for the naming units with clearly dominating readings if the task of the informants were formulated, for example: 'Each primary compound can be interpreted in various ways. Some of these meanings are more likely to occur in a language than others. Propose *only* those readings for the following primary compounds which are, in your view, likely to occur. Use a scale from 10 to 1 with 10 points indicating the highest level of probability of occurrence of a given reading.'

Certainly, single occurrence readings are of little use for meaning predictability. As explained by Nelson, McEvoy and Schreiber (1999: 4) in the introduction to Appendix A of their 'associate' database, the criterion 'two or more' participants "was chosen many years ago on the assumption that idiosyncratic responses given by a single participant would tend to be 'off the wall'."

More relevant to our discussion of meaning predictability is the next factor which may have contributed to a high number of single occurrences and vague readings in this research. It is the *absence of a well-predictable, dominant reading* in some cases. If there is no strong, easily identifiable reading reasoning is replaced by imagination, which results in a large number of bizarre single-occurrence proposals. This observation is in compliance with the assumption of a high diversity of paraphrases for low-interpretable (LI) context-free primary compounds mentioned by Coolen et al. (1991): "[t]he proportion of... idiosyncratic interpretations may be expected to be significantly larger for LI compounds" (1991:349).

This phenomenon can be exemplified by three sample naming units, in particular, *dog spade*, *hill-star*, and *shape cloth*. The very low PRs of their top readings (0.137; 0.152; and 0.098, respectively) bear witness to their 'weakness' engendering a large number of unpredictable single-occurrence proposals.

Conclusion 12 Analogy

In the majority of cases, *templates are insufficient to recognise the subtle shades of individual readings*.

This does not mean that they cannot have predictability-boosting effects. For example, the reading 'a hat to wear when you work in the garden' for *flower* hat was motivated by analogy with *shopping hat* (as explicitly adduced by some informants); 'a seat filled with apple-juice' for *apple-juice seat* may be supposed to have been motivated by 'a bed filled with water', i.e., *water-bed;* the reading 'a hat made of flowers' for *flower hat* may have been motivated

by the existence of *straw hat*; 'a (very) small book' for *baby book* may have been inspired by compounds like *baby car*, *baby ladder*, etc.; 'someone opposed to anthrax' of *anthraxist* was motivated by *sexist* as expressly mentioned by an informant (however, this reading was proposed by only one from among ninety informants!). The same is true of *removage* which, by one informant, was interpreted by analogy with '*Stoneage*' as a period of history.[4]

On the other hand templates may engender negative interference, as in 'whisky made from garden products' (*garden whisky*), analogically to 'wine made from garden products'. In this case the analogy is simply wrong, which is caused by the lack of extra-linguistic knowledge of/experience with the production and ingredients of whisky. Another example of negative interference of analogy concerns the reading 'a book written/drawn by babies/children' (*baby book*) where the abilities and skills of a child are 'smuggled into' the intellectual outfit of a baby.

Conclusion 13 Meaning predictability and associative meaning
While there are some cases of overlap between associates and predictable readings, this relation is far from being systematic. In addition, there is no correlation between the ranks of the most predictable readings and associates. The most important reasons for the *absence of correlation* bear on the *unequal levels of detail* charaterizing these two different phenomena, and the different factors determining the meaning-prediction process and meaning association, respectively (prototypical features vs. lexical paradigmatic relations).

Conclusion 14 Interplay of factors
The analysis of the sample naming units indicates that there is no single factor conditioning the predictability of novel, context-free naming units. On the contrary, it is possible to speak of an *interplay of several factors* conditioning the meaning predictability of novel, context-free naming units, including

a. Onomasiological Type
b. Productivity of the underlying WF Type
c. Semantic compatibility of the motivating words
d. Seme Level Combination
e. Applicability of a 'compatible' combination of semes by the members of a speech community, i.e., the existence in extra-linguistic reality of a corresponding 'object'
f. The nature of the referred 'object' (stable vs. fortuitous, context-bound relation)

g. Morphological Type
h. Availability of other rival Word-formation and Morphological Types
i. Number and strength of predictable readings
j. Analogy-based templates
k. Linguistic competence of an interpreter
l. Extra-linguistic knowledge (knowledge of the world and experiences) of an interpreter

Conclusion 15 Meaning predictability boosting and reducing conditions
Given these factors, it is possible to identify an optimum situation for the meaning predictability of novel, context-free naming units as well as the factors which reduce their meaning predictability. While the Predictability Rate and Objectified Predictability Rate result from a specific interaction of these factors for every single novel, context-free naming unit, the analysis presented in this book has revealed certain *general tendencies*.

I. Predictability-boosting conditions

1. Onomasiological Type 1, 2, or 4 (where neither onomasiological base nor the determining constituent of the onomasiological mark are ambiguous).
2. Productive WF Type underlying the naming unit.
3. Semantically compatible motivating words.
4. 4–4 Seme Level combination, i.e, a combination of prototypical semes reflecting the prototypical features of the motivating objects (level 4 seme in the case of Onomasiological Type 5).
5. Objective justification of a novel naming unit, i.e., the existence in extra-linguistic reality of a corresponding 'object'.
6. Reference to permanent, stable, and constant relations.
7. Productive Morphological Type underlying the naming unit.
8. Unambiguous interpretation of the underlying WF Type and Morphological Type.
9. Absence of competition, i.e., absence of productive rival WF Type and/or Rule.
10. Single strong predictable reading, implying a high PR Gap and R1/R2 (R1/R3) ratio.
11. The possibility to interpret the naming unit through an analogy-based template.

12. Knowledge of the meaning(s) of the motivating word(s) knowledge of the WF Rules, knowledge of the Onomasiological Structure Rules, and any other pertinent competence.
13. Well-established morphemes and affixes.
14. Relevant world knowledge about/experience with the object represented by a naming unit in question.

II. Predictability-reducing conditions

1. Onomasiological Type 3 or 5.
2. Low-productive/unproductive WF Type underlying the naming unit, or a non-rule-governed coining of a naming unit (creativity).
3. Semantically incompatible motivating words.
4. Motivation by non-prototypical features of the motivating objects.
5. Non-existence in extra-linguistic reality of a corresponding 'object'.
6. Reference to context-bound, unstable relations.
7. Unproductive Morphological Type – unusual morphological structure underlying the naming unit.
8. Ambiguous interpretation of the underlying WF Type and/or Morphological Type.
9. Strong competition, i.e., interpretability according to several rival WF Types and/or Rules.
10. Several equally well predictable readings, implying a minimum PR Gap and low R1/R2 (R1/R3) ratio.
11. No possibility to interpret the naming unit by means of an analogy-based template.
12. Poor linguistic competence.
13. Non-established morphemes and affixes, for example, of foreign origin.
14. Absence of the relevant world knowledge of/experience with the object represented by the naming unit in question.

This book aims to contribute to a growing interest amongst psycholinguists and morphologists in the mechanisms of predictability. It presents a model of meaning-prediction process and identifies the factors affecting this process in relation to novel, context-free naming units. Unlike previous studies mostly discussing N+N compounds, its scope is much wider; in particular, my ambition was to develop a unified theory of meaning predictability for all types of naming units, ranging over all of the traditional processes of word formation, represented here by means of five essential *Onomasiological Types*. Through-

out the book, I have tried to draw the reader's attention to the necessity of reflecting the meaning-prediction process in close connection to the word formation process. In other words, it is assumed that word-*interpretation* cannot be treated in separation from word-*formation*, and that a meaning-prediction model should be based on a word formation model. It was proposed that a good point of departure for capturing this inter-relation would be an *onomasiological model* of word formation permitting me to identify the mutual interrelation of three fundamental factors that condition both the formation of words and their interpretation, that is to say, the role of an object of *extra-linguistic reality*, its *conceptual* reflection and processing (by making use of one's knowledge of extra-linguistic reality and experiences with it), and the *linguistic factors*, primarily the interaction between the *onomasiological level* (as the conceptual basis for the act of naming), determining the *Word-Formation Type* of a particular naming unit, and the *onomatological level*, determining its *Morphological Type*. The interrelation between these two linguistic levels, established in the process of word formation by the *Morpheme-to-Seme-Assignment Principle*, appears to play an important role in the meaning-prediction process as a restriction upon the range of acceptable interpretations of a particular naming unit. I demonstrated that the Onomasiological Type underlying a naming unit significantly affects meaning predictability.

I introduced a method of calculation of meaning predictability, including two measures, the *Predictability Rate* and the *Objectified Predictability Rate*. While the former can be advantageously used to get a picture of the competition between various predictable meanings of one and the same naming unit, the latter is proposed to evaluate the most predictable readings of various naming units by employing the PR values of the individual competing readings of each of the compared naming units.

It was demonstrated that the most predictable readings are usually based on (the relation between) *prototypical semes*. The reason is obvious: these semes play a crucial role in the word formation process because they capture the prototypical features of an object to be named. This is another important piece of evidence for the interrelation between word formation and word-interpretation.

Finally, based on my theoretical considerations and an analysis of four experiments, I identified *Predictability-boosting conditions* and *Predictability-reducing conditions*. Their number might be frustrating. But this is not quite so. A good message for language users may be derived from my research: even if the number of potential readings of novel, context-free naming units is rather high, there is a very strong tendency for them to be *dominated by one (rarely two)*

central reading. Another good message concerns non-native speakers. Given a sufficient language proficiency, their meaning-prediction capacity with regard to the type of examined naming units appears to be on a par with that of native speakers.

This book thus discussed some of the intriguing questions of the meaning-prediction process, and offered a new perspective on treating them. As is usually the case with research into unfathomed areas, there still remain many more issues open than solved. In any case, I hope that this work will engender a prolific discussion and instigate further research.

Notes

Introduction

1. In this book, I consistently use the term *naming unit* when referring to units generated in the Word Formation Component by means of productive Word Formation Rules. This term was first suggested by V. Mathesius (1975). In my approach it substitutes for terms like *word*, *lexeme*, and *lexical unit* because of their inconsistent use and varying connotations in the linguistic literature; and because from the perspective of newly formed complex words to refer to newly perceived objects of reality it really is a case of consciously 'naming' the object rather than subconsciously referring to it. From this it follows that an onomasiological theory of word formation deals with the coining of naming units. In principle the naming unit is – like Cruse's term *lexical unit* – a combination of a single form and a single meaning, because each new naming unit is coined with one specific meaning in the 'coiner's' mind. Any additional meaning is acquired in the Lexical Component. On the other hand, lexical unit is a term of much wider comprehension, because it – unlike naming unit – also includes monemes. Cf. Štekauer (1998) for details.

2. J. Grzega (personal communication) proposes to replace the term 'predictability' by 'probability', and restrict the former to forms. In using the selected term I follow M. Dokulil (1978).

3. In Bauer's sense of institutionalisation. Cf. L. Bauer (1983).

Chapter 1

1. Marchand's answer to this case is that it is the emphasised part (topicalised part) of the underlying sentence that becomes the *determinatum*. In *apple-eater* it is the subject (hence Subject Type), in *apple-eating* it is the predicate (Activity Type), in *eating apple* it is the object (Object Type). This topicalisation-based approach was developed into a consistent system of topicalisation rules in Brekle's formalised theory (1970). On this point, see also Kastovsky (1982).

2. For example, *copper coins* or *chocolate bunny* may be accounted for by means of BE or MAKE_2, *party members* by means of HAVE_2 or IN, etc.

3. There have been a number of attempts to identify a final set of relations, some of them drawing on Levi, for example, Shoben (1991) and Gagné and Shoben (1997). Certainly, any such classification within the psycholinguistic framework (see also, for example, Murphy, Hampton, and Wisniewski, discussed in this chapter) serves the main goal of research into conceptual combinations, in particular, to identify whether relational information and other forms of world knowledge are used during the processing of novel conceptual combinations (compounds).

4. Also cf. Günther (1981) who demonstrates the pitfalls of the endeavour to fit specific meanings to a general classification in an experiment focussed on the interpretability of German NN compounds.

5. The other two general observations concerning the interpretation of English compounds, summarised in Lieber (2004: 46) are the following:

– The compound as a whole takes the second stem as its semantic head.
– The first stem in a synthetic compound receives an argument interpretation, often but not always the internal argument interpretation. For example, the stem *truck* in the synthetic compound *truck driver* is interpreted as the internal argument (i.e., the object) of the verb *drive*.

It should be noted that the first of these observations has its predecessor in Marchand who formulated this principle as early as (1967) when defining 'expansion' as follows: "An expansion is a combination AB which is analysable on the basis 'B determined by A', with AB belonging to the same word class and lexical class to which B belongs. All combinations whose determinata are independent morphemes (words) are expansions... Semantically speaking, the *determinatum* represents the element whose range of applicability is limited through the *determinant.* A *steamboat* is basically a *boat.* The applicability of *steamboat* is, however, limited to those boats driven by steam" (1967: 14). These facts were later reformulated as the so-called *IS A Condition* by M. Allen (1978).

6. R. Beard (personal communication) aptly comments on this aspect of Zimmer's approach that categorization may be temporary, contextually induced, e.g. *hamburger dish* may be appropriately classificatory at a specific party even though it serves other purposes before and after the party.

7. The second restriction excludes those compounds, whose first constituent identifies an indispensable part of the head noun – **Klingenmesser* ('blade knife'), **Schalenapfel* ('skin apple'), **Seitenbuch* ('page book').

8. These considerations apply to context-independent interpretations. Certainly, context-bound, and therefore, non-predictable/low predictable, readings are acceptable under specific circumstances with any complex word. Thus, 'a hat like head/in the form of head' is perfectly acceptable in an appropriate context. The same is true, as noted by Beard (personal communication), of *animal horse* in the sense of 'animal like a horse' and *truck vehicle* 'vehicle like a truck'. Another of Beard's illustrations has similar effects: if some chairs are made of balsam and are designed specifically to break apart in a fight and can't be sat in, someone is very likely to say, 'Don't sit on that one; it isn't a furniture chair.' Finally, as he notes in reference to Finin's example, this kind of compound is justified in emphasising the distinction between, for example, regular butter from *soy butter*, *peanut butter*, etc. It should be stressed once more that all these interpretations are context-bound, and therefore their Predictability Rate is low.

9. An interesting case is represented by compounds whose determinant and determinatum are identical nouns. Günther (1981) reports of an experiment, a part of which concerned the interpretation of this type of German compounds labelled by him as 'self-compounds' (Selbstkomposita). He demonstrates that some such compounds are (as confirmed by his informants within an experimental research) interpretable. So, for example, *Bücherbuch*

('booksbook') can be interpreted as 'a book of (some) books', *Frauenfrau* ('women woman') as 'a woman who takes care of women', etc.

10. Finin adds another constraint on the interpretation of compounds by saying that the relation between the compound constituents should not be a negative one: "We would not entertain that *mouse food* might refer to food that a mouse did or will NOT eat" (1980:16). Furthermore, he points out the well-known and frequently discussed fact that the semantic relationships between compound constituents tend to express habitual rather than accidental, temporary, or fortuitious ones. It should be, however, stressed that the latter constraint does not apply absolutely.

11. Inspired by Downing's experimental research, Günther (1981) undertook a similar experiment based on a systematic combination of 14 carefully selected German nouns (which gave 196 compounds) representing the core vocabulary. His experiment, like Downing's, consisted of two parts: (a) a context-free interpretation of NN compounds, and (b) the ranking task based on a three degree scale: common, possible, impossible.

12. Also called a 'dimension-based' approach to conceptual combination (Gagné & Shoben 1997) on the ground that the head noun is construed as a schema with a certain number of dimensions (e.g. colour along with possible values (e.g. red, green, yellow) for each dimension. Gagné & Shoben criticize this type of model because it, in their view, does not account for their finding that thematic relations of the modifier concept significantly influence the interpretation of complex words.

13. Coolen, van Jaarsveld, and Schreuder classified novel compounds prior to the experiments proper into two groups, notably 'high-predictable' and 'low-predictable' compounds, based on their subjective assessment.

14. Regardless of the unequal nature and character of these concepts.

15. See, for example, the analysis of *baby book* in Section 4.2.2.1.

16. In their experiments the 'availability' of a thematic relation for a constituent concept is defined as "the frequency with which combinations containing that concept were interpreted as having the relation in question" (1997:74). The frequencies are determined by using Shoben & Medin's corpus reported on in Shoben (1991).

17. The literature on headedness in word formation is very rich, ranging from Marchand's (1960) discussion of the role of determinatum, through Kastovsky's identification-specification scheme (1982), Williams' Right-hand Head Rule (1981) Selkirk's (1982) revised Right-hand Rule, Lieber's Feature Percolation Conventions (1981), Zwicky's (1985) approach, up to Bauer's (1990) sceptical view of heads. An approach to the notion of head within the onomasiological framework is presented in Štekauer (2001b).

18. A general problem with analogy-based explanation is indicated by Beard's question: "So, where did 'windsock' come from?" (personal communication).

19. In this connection, see the above-mentioned comments on Gagné's (2001) priming-based experiments demonstrating the influence of a recently encountered conceptual combination (compound word) upon the interpretation of a compound word.

20. The frame is conceived of as "a knowledge structure that represents a concept of a stereotypical situation or object" (Wisniewski 1996:435).

21. Rule 1: N1 fills one of N2's roles (*engine repair*, *January flight*, *magnesium wheel*). The application of this rule is based on the identification of the slots of the modified concept that the modifier may fill or specify.
Rule 2: N2 fills one of N1's roles (*drinking water*, *washing machine*, *maintenance crew*). Here, the modified concept fills a slot in the modifier concept.
Rule 3: Thing + Role Nominal (*F4 pilot*, *cat food*, *oil pump*, *dog house*). This rule is applied to the so-called role nominals. The latter are defined as nouns that refer to a role of an underlying concept. For example, *owner* refers to the Agent role of an ownership concept. By the same token, Rule 3 is applied to interpret *F4 pilot* because *pilot* is a role nominal referring to an Agent of the 'flying' event.
Rule 4: Role Nominal + Thing (*newspaper glasses*, *driver education*, *pilot school*). In this case, it is the modifier concept that is a role nominal.
Rule 5: Specific + Generic (*F4 aircraft*). This rule applies to those cases where N1 is a hyponym of N2. The rare occurrence of this type is explained by its redundant nature.
Rule 6: Generic + Specific (*the integer three*, *building NE43*, *President Carter*). Here, the modifier is the hyperonym of the modified concept. As suggested by Finin, this type is more common because it can serve the function of placing a perspective on the modified concept.
Rule 7: N1 be N2 (*woman doctor*). While Finin admits that cases like *woman doctor* might be handled by previous rules (for example, by the occupation role filling), he believes that this rule represents the so-called viewing process in which "an instance of one concept can be transformed into an instance of another process" (1980: 114).
Rule 8: Attribute Transfer (*iron will* 'a will that is strong', *elephant legs* 'legs that are large'). This rule applies to cases in which an attribute or a property of the modifier is predicated of the modified concept.

22. Finin's classification suffers from the same kind of problems as the classification systems discussed above. Thus, L. Bauer (personal communication) asks why, for example, *oil pump* is not assigned to Rule 2, *drive education* to Rule 1, and examples of Rule 6 are traditional appositions rather than compounds.

23. Jan Don (personal communication) confirms Finin's position from the point of view of Distributed Morphology when he says that, in experimental situations, people will give preference to interpretations that are somehow 'known' in the speech community, i.e., those which are listed in the 'encyclopedia'. This might also explain why people (again in experimental situations) prefer 'habitual' rather than 'temporary' interpretations. Only the former will be present in the encyclopedia, while the latter only come into existence when the specific context requires it. J. Don compares this situation to that accompanying the interpretation of a modifier and a head noun in a syntactic construction, like, for example, *red hat*. Usually this will refer to a hat that is red, but it may also, given the right context, refer to a hat in the red corner, or the hat worn by the girl with the red dress, etc.

24. An interesting remark, contradicting Meyer's assumption, was attached to my manuscript by Joseph Feinberg, a proofreader of my manuscript: "When I read the compound, what first occurred to me was 'a fan made of a book' or 'a book used as a fan', i.e. for fanning oneself."

25. Libben proposed the so-called APPLE model (Automatic Progressive Parsing and Lexical Excitation) of morphological parsing which identifies constituent morphemes from left

to right and checks the lexical status of the identified morphemes; hence, *redberry* is parsed into *red* + *berry* rather than *re* + *dberry*.

26. In my experimental research I abstracted away from ambiguous parsing cases.

27. They are treated in the same way as derivations with overt suffixes, for example, *scar*: *to scarify* = *mark*: *to mark*, because the former has the meaning 'cause to have scars', and the meaning of the latter is based on the same formula, 'cause to have marks'. Similarly, to *crystallize* from *crystal* may mean either 'come to have crystals' or 'cause to have crystals'; the same meanings are carried by the denominal *to clot*: 'come to have clots' or 'cause to have clots' (1995:180).

28. A terminological note: 'Functional derivation' changes the basic lexical meaning by means of rules operating over 'grammatical' functions like Subject, Object, Locus, Means, Manner, Possession, Possessive, Origin. 'Transposition', on the other hand, reflects the capability of the lexicon to change the lexical class of a lexeme, without changing the basic lexical meaning.

29. L. Lipka (personal communication) has pointed out that similar extension of the meaning applies to their German equivalents – *Sattler* and *Schreiner*, respectively.

Chapter 2

1. The latest appeal for the incorporation of extra-linguistic factors into linguistic research in general and into the study of words in particular comes from Lipka (2002).

2. In accordance with Štekauer (in press[2]), 'WF Rule' refers to a unique combination of a WF Type and a Morphological Type. For the explanation of these terms see Chapter 3, Note 8.

3. Here I do not intend to discuss at length the problem of whether or not any of the semes are semantic primitives because this issue appears to be irrelevant to the topic of meaning predictability. In any case, however, the hierarchical structure of semes suggests that more abstract semes are decomposable into more specific semes (see below). While Aitchison (1987:65) considers the decomposability of semantic 'primitives' a major problem of the atomic approach to the meaning of words, the hierarchical structure of semes, underlying my approach, appears to be an advantageous tool for the identification of those semes that appear to play a vital role in the meaning-prediction process.

4. Cf. Štekauer (2001b) for an onomasiological theory of headedness justifying the identification of *un-* as the head of *unhappy*, and *re-* as the head of *restart*.

Chapter 3

1. Cf., for example, Lipka (2002).

2. Certainly, there are cases in which the different traditions, culture, habits, etc. are manifested in unequal interpretations of novel naming units – as will also be illustrated by some of my experimental results. These cases can, however, be viewed as exceptions to the general rule.

3. Cf. Štekauer (2001b) for a theory of headedness within the onomasiological framework.

4. For the same view see also Renouf and Bauer (2000:254–255) who mention both metaphor and polysemy as major factors impeding a context-free interpretation of novel complex words.

5. See the discussion on Onomasiological Structure Rules in Section 3.6.

6. Here I refer to instances like *driver*, the Morphological Type of which [V+ *-er*]$_N$ represents two different WF Types: [Action – Agent] (a person who drives a vehicle) and [Action – Instrument] (a mechanism used for driving).

7. Here extra-linguistic knowledge outweighs the principle of salience (Matlin 1989:362) referred to by Ryder (1999:285). The principle of salience refers to "the degree to which something is noticeable in comparison with its surroundings." Ryder also refers to Croft (1991) who maintains that the Agent and the Patient are generally more salient than the Instrument because they represent the natural beginning and end points, respectively, of the event. Finally, according to Langacker (1991:322), the animacy hierarchy of human > animal > physical object > abstract entity suggests that the Agent is, in principle, more salient than the Instrument. In the present example, however, the two thematic roles swap their positions due to extra-linguistic circumstances.

8. *Word-Formation Types* are constituted by semes of the onomasiological level. For illustration, [Object – Action – Agent], [Action – Agent], [Location – Action – Agent], [Factitive – Action – Agent], [Instrument – Action – Agent], [Manner – Action – Agent] are various WF Types unified by the conceptual category of Agent. As such, they represent a single *Word-Formation Type Cluster* (WFTC). Each WF Type may have various morphological representations (*wood-cutter – novelist – writer – cheat – oarsman – transformational grammarian – bodyguard - chimney sweep*). All of these different morphological structures represent various *Morphological Types*. Since they are used to coin new naming units falling within one and the same conceptual category (Agent, in our example), they represent a single *Morphological Type Cluster*.

9. The examples are borrowed from Bauer (1983) and Adams (2001).

10. Cf. Štekauer (1998, 2001a) for a more detailed discussion on blends, clippings, and acronyms.

11. This task is irrelevant for our present purpose. It tested the translation ability of our students.

12. The notion of a Word-Formation Type Cluster is the basic notion of my approach to productivity, and refers to all Word-Formation Types of a particular semantics of the onomasiological base, for example, Agent, Instrument, Process, Negation, etc. A Cluster of WF Types 'guarantees' the coining of a new naming unit of the specific semantics whenever the need arises. Each Cluster is 100% productive. The share of individual options within a particular Word-Formation Type Cluster of the total productivity may be computed internally.

13. 'Stative' is used in the meaning of the 'Complement of State' (cf. Hansen et al. 1982), and can take different forms.

14. This is not to say that other animals, such as hawks, are not used for hunting. These, however, are less typical pets than dogs, or, to paraphrase Aitchison (1987), dog is 'pettier' than hawk.

15. Cf. the basic tenets of their approach in Chapter 1.

16. Here I tacitly disregard the grammatical acceptability of naming units, which means that a naming unit which is coined in defiance of productive WF Rules and/or one which violates any of the restrictions imposed on the formation of new naming units is only evaluated in terms of whether it is interpretable and whether its meaning is predictable. Experiment 4, discussed below, indicates that also 'ungrammatical' coinages may be predictable, depending on the nature of the violated constraint.

17. This assumption reflects several important conditions: (1) Our informants come from countries that do not show principled differences in everyday way of life, culture, traditions, beliefs, value-systems, etc. Therefore, the interpretation of new naming units should not be significantly influenced by these extralinguistic factors. No doubt if the informants came from two completely different cultural backgrounds the differences in interpretation might be expected to be significant on a number of occasions. This is also closely connected with different ways of structuring extralinguistic reality by different speech communities. (2) An obvious condition is that both native and non-native speakers know the relevant meaning(s) of the motivating words and have a standard command of the particular language. (3) Notwithstanding condition (2), the language intereference factor may be at play, closely connected with the unequal structuring of extralinguistic reality by individual speech communities (and thus their languages), unequal language system structures, idiomaticity, connotability, etc. All these factors were observed in the present research – even though their influence was restricted for the above-mentioned reasons.

18. However, I believe that the meaning-prediction process is influenced by certain sociolinguistic factors, such as age, profession, education, because these factors directly affect one's knowledge of the world and experiences.

Chapter 4

1. The naming unit *apple-juice seat* was borrowed from Downing (1977).

2. In this approach, the ± sign does not necessarily indicate the disjunctive, either-or, relation. As is the case of [±Drawing Skill], some features are inherently present in the named objects to a various degree. Thus '+' and '–' represent two poles of the scale.

3. I must admit I did not hear about anything like a *pooper-scooper* before. The information about the existence of such a tool and the corresponding name was provided to me by R. Lieber and L. Bauer.

4. J. Grzega (personal communication) notes that the blocking principle may be outweighed in the case of concepts (and, certainly, words) that arouse 'emotion' in speakers/hearers, and assumes that the prominence of the reading 'a spade used for scooping-up dog's excrement' speaks in favour of this view.

5. Here, [Colour] is level 3 rather than Level 4 because 'flower' as a hyperonym does not have any single characteristic colour. Thus, [Charactersitic Colour] would be Level 4 when talking about any specific flower – rose, violet, etc.

6. L. Lipka (personal communication) assumes that the reason may bear on the singular form of *shape* in this compound.

7. The discussion of conversion-related predictability issues in this book is based on Štekauer (in press[1]).

8. Owing to very low PR values of the rank 4 readings (in the order of hundredths or even thousandths, these readings are not included in the overview, and are disregarded in the calculation of the Objectified Predictability Rate.

9. The contrast based on the seme [Adult] is much more significant for this reading than that based on the seme [Mature] as it is much broader, and as such, it reflects a number of other differences between 'boy' and 'man'.

10. The reason for not classifying the proposals 'aggressive' and 'to roar like a lion' separately is that a large number of informants combined them into a single reading.

11. This naming unit can also be interpreted on the basis of the WF Type [Action – Agent], i.e., as one belonging in Onomasiological Type 2.

12. However, as noted by one of the anonymous reviewers of the manuscript, the assumption of the link between the terrorist attack on the World Trade Centre and the Pentagon on September 11, 2001, may be overstated. The anthrax attacks came after 9/11, and were far less significant than the destruction of the WCT and the loss of life involved there. Therefore, in his view, it is not surprising that some informants did not make the link with 9/11.

13. As noted by Jack Lala, a proofreader of the manuscript, *smile-person* would be politically more correct.

14. From the point of view of meaning predictability, it makes no difference whether we call the phenomenon polysemy, homonymy, or co-functionality.

15. Cf. Marchand (1960: 179–181) for details.

16. As suggested by Salvador Valera, the peculiarity of the structure of *removage* is that there is no other naming unit in English vocabulary (based on the data from the 100-million-word British National Corpus) in which the initial *re-* in the morphological structure [re-Verb-age] is a part of a base in which it does not function as a prefix with the meaning of REPETITION. In *removage*, the base *remove* which goes back to OFr. *remouvoir*, is not synchronically analysable as *re* + *move*. Being a moneme that cannot be analyzed from the word formation point of view one may expect a strong link between *re* and *move*, much stronger than between *move* and *age*. This accounts for the strong preference of the informants for the *remove*+*age* in comparison to the *re* + *movage* interpretation. Interestingly enough, the number of Types of the [*re*+V+*age*] structure in the BNC is very low, with *remarriage* (106) and *reportage* (43) having the highest token frequency. The frequencies of the remaining Types are fairly low. Of these words, there is only one which admits the double way of analysis as appeared with my sample word: *reportage* can also be analyzed as either *report* + *age*, in which case *re* does not mean REPETITION, or *re-* + *portage*. However, it can hardly be surmised that the *remove* + *age* analysis was motivated by *re* + *portage*.

17. Since there is no 'apple juice' in the database the associates of 'juice' are used in my analysis instead of it as its semantics complies with the semantics of the most predictable readings.

Chapter 5

1. Cf. a complete list of these universal functions, including examples, in Beard (1995).

2. Cf. Bauer's (2001) discussion on this issue from a different perspective.

3. See, however, the comments on *blondesjoker*, Experiment 3.

4. Here, however, the informant used an incorrect spelling to 'justify' his reading; *Stone Age* is usually spelled separately and with initial capital letters. Certainly, the respective ways of pronunciation of *Remove Age* and *removage* are different.

References

Adams, Valerie (2001). *Complex Words in English*. Harlow: Longman.

Aitchison, Jean (1987). *Words in the Mind. An Introduction to the Mental Lexicon.* New York: Basil Blackwell.

Allen, Margaret R. (1978). *Morphological Investigations.* Doctoral dissertation, University of Connecticut, Storrs, CT.

Anderson, Richard C., & Pearson, David P. (1988). "A scheme-theoretic view of basic processes in reading comprehension." In P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 37–55). Cambridge: Cambridge University Press.

Anderson, Stephen R. (1992). *A-Morphous Morphology.* Cambridge: Cambridge University Press.

Aronoff, Mark (1976). *Word Formation in Generative Grammar.* Linguistic Inquiry Monograph 1. Cambridge, MA: MIT Press.

Barker, Chris (1998). "Episodic *-ee* in English: A Thematic Role Constraint on New Word Formation." *Language, 4,* 695–727.

Barsalou, L. (1981). *Determinants of Graded Structure in Categories*. Unpublished doctoral dissertation. Stanford University.

Bauer, Laurie (1978). *The Grammar of Nominal Compounding with Special Reference to Danish, English and French*. Odense University Studies in Linguistics 4. Odense: Odense University Press.

Bauer, Laurie (1983). *English Word formation*. Cambridge: Cambridge University Press.

Bauer, Laurie (1990). "Be-Heading the Word." *Journal of Linguistics, 26,* 1–31.

Bauer, Laurie (2001). *Morphological Productivity*. Cambridge Studies in Linguistics 95. Cambridge: Cambridge University Press.

Beard, Robert E. (1981). *The Indo-European Lexicon: A Full Synchronic Theory.* North-Holland Linguistic Series 44. Amsterdam: North-Holland.

Beard, Robert E. (1995). *Lexeme-Morpheme Base Morphology. A General Theory of Inflection and Word Formation*. SUNY Series in Linguistics. State University of New York Press.

Bierwisch, Manfred (1983). "Semantische und konzeptuelle Repräsentation lexikalischer Einheiten." In W. Motsch & R. Ružička (Eds.), *Untersuchungen zur Semantik* (pp. 61–99). Berlin: Akademie-Verlag.

Bierwisch, Manfred (1986). "On the nature of semantic form in natural language. In F. Klix & H. Hangendorf (Eds.), *Human memory and cognitive capabilities* (pp. 765–783). Amsterdam: Elsevier (North-Holland).

Bierwisch, Manfred, & Schreuder, Robert (1992). "From Concepts to lexical items." *Cognition, 42,* 23–60.

Boguraev, Branimir K. (1989). "The Definitional Power of Words." In Y. Wilks (Ed.), *Theoretical Issues in Natural Language Processing* (pp. 7–11). Hillsdale: Lawrence Erlbaum Associates, Publishers.

Brekle, Herbert E. (1970). *Generative Satzsemantik und transformationelle Syntax im System der englischen Nominalkomposition*. München: Fink.

Chomsky, Noam (1970). "Remarks on Nominalization." In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English Transformational Grammar* (pp. 184–221). Waltham, MA: Ginn.

Clark, Eve V., & Berman, Robert A. (1987). "Types of Linguistic Knowledge: Interpreting and Producing Compound Nouns." *Journal of Child Language, 14,* 547–567.

Clark, Eve V., & Clark, Herbert H. (1979). "When Nouns Surface as Verbs." *Language, 4,* 767–811.

Clark, Herbert H. (1983). "Making Sense of Nonce Sense." In G. B. Flores d'Arcais & R. J. Jarvella (Eds.), *The Process of Language Understanding* (pp. 297–331). New York: John Wiley & Sons.

Cohen, Benjamin, & Murphy, Gregory L. (1984). "Models of Concepts." *Cognitive Science, 8,* 27–58.

Coolen, Riet, van Jaarsveld, Henk J., & Schreuder, Robert (1991). "The Interpretation of Isolated Novel Nominal Compounds." *Memory and Cognition, 19,* 341–352.

Coolen, Riet, Jaarsveld, van Henk J., & Schreuder, Robert (1993). "Processing Novel Compounds: Evidence for Interactive Meaning Activation of Ambiguous Nouns." *Memory and Cognition, 21,* 235–246.

Coseriu, Eugenio (1970). "Bedeutung und Bezeichnung im Lichte der strukturellen Semantik." In P. Hartmann & H. Vernay (Eds.), *Sprachwissenschaft und Übersetzen* (pp. 104–121). München: Max Hueber.

Croft, William (1991). *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: The University of Chicago Press.

Cruse, Alan D. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

de Almeida, Roberto G., & Libben, Gary (2002). "Compound Pre-Access Decomposition: Effects of Constituent Disruption." *Folia Linguistica, 1–2,* 97–115.

Derwing, Bruce L., & Skousen, Royal (1989). "Morphology in the mental lexicon: A new look at analogy." In G. Booij & J. Van Marle (Eds.), *Yearbook of Morphology* 2 (pp. 55–71). Dordrecht: Foris.

Dokulil, Miloš (1962). *Tvoření slov v češtině I. Teorie odvozování slov.* Prague: Nakladatelství Československé akademie věd.

Dokulil, Miloš (1964). "Zum wechselseitigen Verhältnis zwischen Wortbildung und Syntax." In *Travaux linguistiques de Prague* 1 (pp. 369–376). Prague: Academia.

Dokulil, Miloš (1968a). "Zur Frage der Konversion und verwandter Wortbildungsvorgänge und -beziehungen." In *Travaux linguistiques de Prague* 3 (pp. 215–239). Prague: Academia.

Dokulil, Miloš (1968b). "Zur Theorie der Wortbildungslehre." *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig.* Gesselschafts- und Sprachwissenschaftliche Reihe 17, 203–211.

Dokulil, Miloš (1968c). "Zur Frage der sog. Nullableitung." In H. E. Brekle & L. Lipka (Eds.), *Wortbildung, Syntax und Morphologie*. Festschrift zum 60. Geburtstag von Hans Marchand am 1. Oktober 1967 (pp. 55–64). The Hague: Mouton de Gruyter.

Dokulil, Miloš (1968d). "Zur Frage der Stelle der Wortbildung im Sprachsystem." *Slovo a slovesnost, 29,* 9–16.

Dokulil, Miloš (1978). "K otázce prediktability lexikálního významu slovotvorně motivovaného slova." *Slovo a slovesnost, 39,* 244–251.

Dokulil, Miloš (1997). "The Prague School's Theoretical and Methodological Contribution to 'Word formation' (Derivology)." In *Obsah – výraz – význam. Miloši Dokulilovi k 85. narozeninám* (pp. 179–210). Praha: FF UK.

Downing, Pamela (1977). "On the Creation and Use of English Compound Nouns." *Language, 4,* 810–842.

Dressler, Wolfgang U. (In press). "Word formation in Natural Morphology." In P. Štekauer & R. Lieber (Eds.), *Handbook of Word formation*. Dordrecht: Springer.

Dressler, Wolfgang U., & Ladányi Mária (2000). "Productivity in Word formation (WF): A Morphological Approach." *Acta Linguistica Hungarica, 47,* 103–144.

Dressler, Wolfgang U., Mayerthaler, Willi, Panagl, Oskar, & Wurzel, Wolfgang U. (Eds.). (1987). *Leitmotifs in Natural Morphology*. Amsterdam: John Benjamins.

Finin, Timothy W. (1980). *The Semantic Interpretation of Compound Nominals*. Urbana: University of Illinois, Coordinated Science Laboratory, Report T-96.

Gagné, Christina L. (2001). "Relation and Lexical Priming During the Interpretation of Noun-Noun Combinations." *Journal of Experimental Psychology*: *Learning, Memory, and Cognition, 1,* 236–254.

Gagné, Christina L. & Shoben, Edward J. (1997). "Influence of Thematic Relations on the Comprehension of Modifier-Noun Combinations." *Journal of Experimental Psychology*: *Learning, Memory, and Cognition, 1,* 71–87.

Gleitman, Lila R., & Gleitman, Henry (1970). *Phrase and Paraphrase: Some Innovative Uses of Language*. New York: W. W. Norton and Co.

Gleitman, Lila R., & Gleitman, Henry (1979). "Language Use and Language Judgement." In Ch. J. Fillmore, D. Kempler, & William S.-Y. Wang (Eds.), *Individual Differences in Language Ability and Language Behaviour* (pp. 103–126). New York: Academic Press.

Günther, Hartmut (1981). "N + N: Untersuchungen zut Produktivität eines deutschen Wortbildungstyps." In L. Lipka & H. Günther (Eds), *Wortbildung* (pp. 258–280). Darmstadt: Wissenschaftliche Gesellschaft.

Haiman John (1980). "Dictionaries and encyclopedias." *Lingua, 50,* 329–357.

Halle, Morris (1973). "Prolegomena to a Theory of Word Formation." *Linguistic Inquiry, 1,* 3–16.

Hammond, Michael, & Noonan, Michael (Eds.). (1988). *Theoretical Morphology. Approaches in Modern Linguistics.* Orlando: Academic Press.

Hampton, James A. (1983). "A composite prototype model of conceptual conjunction." Unpublished manuscript. The City University. London.

Hampton, James A. (1987). "Inheritance of attributes in natural concept conjunctions." *Memory & Cognition, 15,* 55–71.

Hampton, James A. (1988). "Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 12–32.

Hansen, Klaus (1978). "Problems in the Semantic Analysis of Compounds." *Zeitschrift für Anglistik und Amerikanistik, 26,* 247–251.

Hansen, Barbara, Hansen, Klaus, Neubert, Albrecht, & Schentke, Manfred (1982). *Englische Lexikologie. Einführung in Wortbildung und lexikalische Semantik.* Leipzig: VEB Verlag Enzyklopädie.

Heidolph, Karl-Erik (1961). *Beziehungen zwischen Kompositum und attributiven Substantivkonstruktionen in der deutschen Gegenwartssprache*. Berlin.

Horecký, Ján (1983). *Vývin a teória jazyka.* Brastislava: SPN.

Horecký, Ján (1994). *Semantics of Derived Words*. Prešov: Acta Facultatis Philosophicae Universitatis Šafarikanae.

Horecký, Ján, Buzássyová Klára, Bosák Ján et al. (1989). *Dynamika slovnej zásoby súčasnej slovenčiny*. Bratislava: Veda.

Jaarsveld, van Henk J., Coolen, Riet, & Schreuder, Robert (1994). "The Role of Analogy in the Interpretation of Novel Compounds." *Journal of Psycholinguistic Research, 2,* 111–137.

Kastovsky, Dieter (1982). *Wortbildung und Semantik*. Tübingen/Düsseldorf: Francke/Bagel.

Katz, Jerrold, & Fodor, Jerry A. (1963). "The Structure of a Semantic Theory." *Language, 39,* 170–210.

Kiparsky, Paul (1982). "Lexical Morphology and Phonology." In I. S. Yang (Ed.), *Linguistics in the Morning Calm.* Selected Papers from SICOL-1981 (pp. 3–91). Seoul: Hanshin Publishing Company.

Kiparsky, Paul (1997). "Remarks on Denominal Verbs." In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex Predicates* (pp. 473–499). Stanford CA: CSLI.

Labov William (1973). "The Boundaries of Words and their Meaning." In Ch.-J. N. Bailey & R. W. Shuy (Eds), *New Ways of Analyzing Variation in English* (pp. 340–373). Washington D.C.: Georgetown University Press.

Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar. Vol. 1. Theoretical Prerequisites.* Stanford: Stanford University Press.

Langacker, Ronald, W. (1988a). "An Overview of Cognitive Grammar." In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics* (pp. 3–48). Amsterdam: John Benjamins.

Langacker, Ronald W. (1988b). "A View of Linguistic Semantics." In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics* (pp. 49–69). Amsterdam: John Benjamins.

Langacker, Ronald W. (1991). *Concept Image Symbol. The cognitive basis of grammar.* Berlin: Mouton de Gruyter.

Lees, Robert B. (1960). *The Grammar of English Nominalizations*. Bloomington, IN: Indiana University Press.

Lees, Robert B. (1970). "Problems in the Grammatical Analysis of English Nominal Compounds." In M. Bierwisch & K. E. Heidolph (Eds.), *Progress in Linguistics* (pp. 174–186). The Hague – Paris: Mouton de Gruyter.

Levi, Judith N. (1974). "On the alleged idiosyncrasy of non-predicate NP's." *Papers from the 10th Regional Meeting, Chicago Linguistic Society* (pp. 402–415). Chicago: Chicago University Press.

Levi, Judith N. (1975). *The syntax and semantics of non-predicating adjectives in English*. University of California dissertation.
Levi, Judith N. (1978). *The Syntax and Semantics of Complex Nominals.* New York: Academic Press.
Levi, Judith N. (1983). "Complex Nominals: New Discoveries, New Questions." In S. Hattori & K. Inoue (Eds.), *Proceedings of the XIIIth International Congress of Linguists = Actes du XIIIe congrès international des linguists*, August 29–September 4, 1982 (pp. 183–197). Tokyo: Proceedings Publishing Committee.
Li, Charles (1971). *Semantics and structure of compounds in Chinese*. Berkeley: University of California dissertation.
Libben, Gary (1998). "Semantic Transparency in the Processing of Compounds: Consequence for Representation, Processing and Impairment." *Brain and Language, 61,* 30–44.
Libben, Gary, Derwing, Bruce L., & de Almeida, Roberto G. (1999). "Ambiguous Novel Compounds and Models of Morphological Parsing." *Brain and Language, 68,* 378–386.
Lieber, Rochelle (1981). *On the Organization of the Lexicon*. Doctoral dissertation, MIT, Bloomington, IN: Indiana University Linguistics Club.
Lieber, Rochelle (1992). *Deconstructing Morphology. Word Formation in Syntactic Theory.* Chicago: University of Chicago Press.
Lieber, Rochelle (2004). *Morphology and Lexical Semantics*. Oxford: Oxford University Press.
Lieber, Rochelle (In press). "English Word Formation Processes. Observations, Issues, & Thoughts on Future Research." In P. Štekauer & R. Lieber (Eds), *Handbook of Word formation*. Dordrecht: Springer.
Lint, van Trudeke (1982). "The Interpretation of Compound Nouns." In S. Daalder & M. Gerritsen (Eds), *Linguistics in the Netherlands* 1982 (pp. 135–145). Amsterdam/Oxford: North-Holland.
Lipka, Leonhard (2002). *English Lexicology*. Tübingen: Gunter Narr.
Luce, Robert Duncan (1959). *Individual Choice Behaviour*. New York: Wiley.
Marchand, Hans (1960). *The Categories and Types of Present-Day English Word Formation.* Wiesbaden: Otto Harrassowitz.
Marchand, Hans (1965a). "The Analysis of Verbal Nexus Substantives." *Indogermanische Forschungen, 70,* 57–71.
Marchand, Hans (1965b). "On the Analysis of Substantive Compounds and Suffixal Derivatives not Containing a Verbal Element." *Indogermanische Forschungen, 70,* 117–145.
Marchand, Hans (1967). "Expansion, Transposition, and Derivation." *La Linguistique, 1,* 13–26.
Marchand, Hans (1974). *Studies in Syntax and Word Formation.* Edited by D. Kastovsky. Internationale Bibliothek für Allgemeine Linguistik, Bd. 18. Munich: Fink.
Mathesius, Vilém (1975). *A Functional Analysis of Present Day English*. The Hague: Mouton de Gruyter.
Matlin, Margaret W. (1989). *Cognition*. 2nd edition. Chicago: Holt, Rinehart and Winston, Inc.

Meyer, Ralf (1993). *Compound comprehension in isolation and in context. The contribution of conceptual and discourse knowledge to the comprehension of German novel noun-noun compounds.* Tübingen: Niemeyer.

Miller, George A., & Johnson-Laird, Phil N. (1976). *Language and Perception*. Cambridge, MA: Harvard University Press.

Motsch, Wolfgang (1962). "Zur Stellung der 'Wortbildung' in einem formalen Sprachmodell." *Studia Grammatica, 1,* 31–50.

Motsch, Wolfgang (1970). "Analyse von Komposita mit zwei nominalen Elementen." In M. Bierwisch & K. E. Heidolph (Eds), *Progress in Linguistics* (pp. 208–223). The Hague – Paris: Mouton de Gruyter.

Murphy, Gregory L. (1988). "Comprehending Complex Concepts." *Cognitive Science, 12,* 529–562.

Murphy, Gregory L. (1990). "Noun phrase interpretation and conceptual combination." *Journal of Memory and Language, 29,* 259–288.

Nelson, Douglas L., McEvoy, C. L., & Schreiber, T. A. (1999. The University of South Florida word association, rhyme and word fragment norms [On-line]. Available: http://cyber.acomp.usf.edu/ FreeAssociation/AppendixA/index.html.

Nelson, Douglas L., Zhang, Nan, & McKinney, Vanesa (2001). "The Ties That Bind What Is Known to the Recognition of What Is New." *Journal of Experimental Psychology: Learning, Memory and Cognition, 5,* 1147–1159.

Ondrus, Pavel, Horecký, Ján, & Furdík, Juraj (1980). *Súčasný slovenský spisovný jazyk. Lexikológia.* SPN: Bratislava.

Plag, Ingo (1999). *Morphological Productivity. Structural Constraints in English Derivation.* Berlin – New York: Mouton de Gruyter.

Plag, Ingo (To appear). "Syntactic category information and the semantics of derivatinal morphological rules." *Folia Linguistica, 38.*

Raffelsiefen, Renate E. (1992). "A Nonconfigurational Approach to Morphology." In M. Aronoff (Ed.), *Morphology Now* (pp. 133–162). Albany: State University of New York Press.

Rainer, Franz (1988). "Towards a Theory of Blocking." In G. Booij & J. van Marle (Eds), *Yearbook of Morphology* 1988 (pp. 155–185). Dordrecht: Foris.

Renouf, Antoinette, & Bauer, Laurie (2000). "Contextual Clues to Word-Meaning." *International Journal of Corpus Linguistics, 2,* 231–258.

Roeper, Thomas, & Siegel, M. E. A. (1978). "A Lexical Transformation for Verbal Compounds." *Linguistic Inquiry, 9,* 199–260.

Rosch, Eleanor (1978). "Principles of Categorization." In E. Rosch & B. B. Lloyd (Eds), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum.

Rosch, Eleanor, & Mervis, Carolyne B. (1975). "Family Resemblances: Studies in the internal structure of categories." *Cognitive Psychology, 7,* 573–605.

Ryder, Mary Ellen (1994). *Ordered Chaos: The Interpretation of English Noun-Noun Compounds.* Berkeley: University of California Press.

Ryder, Mary Ellen (1999). "Bankers and Blue-Chippers: an Account of *-er* Formation in Present-Day English." *English Language and Linguistics, 2,* 269–297.

Scalise, Sergio (1984). *Generative Morphology*. Dordrecht: Foris.

Selkirk, Elisabeth O. (1982). *The Syntax of Words.* Linguistic Inquiry Monograph 7. Cambridge, MA: MIT Press.

Shoben, Edward J. (1991). "Predicating and Nonpredicating Combinations." In P. J. Schwanenflugal (Ed.), *The Psychology of Word Meanings* (pp. 117–135). Hillsdale, NJ: Erlbaum.

Sinclair, John (1998). "The Lexical Item." In E. Weigand (Ed), *Contrastive Lexical Semantics* (pp. 1–24). Amsterdam/Philadelphia: John Benjamins.

Skousen, Royal (1989). *Analogical modelling of language*. Dordrecht: Kluwer.

Smith, Edward E., & Osherson, Daniel N. (1984). "Conceptual combination with prototype concepts." *Cognitive Science, 8,* 337–361.

Smith, Edward E., Osherson, Daniel N., Rips, Lance J., & Keane, Margaret (1988). "Combining prototypes: A selective modification model." *Cognitive Science, 12,* 485–527.

Štekauer, Pavol (1992). "On Some Issues of Zero Morpheme in English." *Linguistica Pragensia, 2,* 73–87.

Štekauer, Pavol (1996). *A Theory of Conversion in English.* Frankfurt am Main: Peter Lang.

Štekauer, Pavol (1997). "On the Semiotics of Proper Names and their Conversion." *Arbeiten aus Anglistik und Amerikanistik, 22* (1), 27–36.

Štekauer, Pavol (1998). *An Onomasiological Theory of English Word Formation*. Amsterdam/Philadelphia: John Benjamins.

Štekauer, Pavol (1999). "Bracketing Paradoxes: Are There Any?" *Linguistica Pragensia, 1,* 16–33.

Štekauer, Pavol (2000). *English Word Formation. A History of Research (1960–1995)*. Tübingen: Gunter Narr.

Štekauer, Pavol (2001a). "Fundamental Principles of an Onomasiological Theory of English Word Formation." *Onomasiology Online, 2,* 1–42 [www.onomasiology.de].

Štekauer, Pavol (2001b). "Beheading the Word? Please, Stop the Execution." *Folia Linguistica, 3–4,* 333–355.

Štekauer, Pavol (2002). "On the Theory of Neologisms and Nonce-formations." *Australian Journal of Linguistics, 1,* 97–112.

Štekauer, Pavol (In press[1]). "On the Meaning Predictability of Novel Context-free Converted Naming Units." *Linguistics*.

Štekauer, Pavol (In press[2]). "Onomasiological Approach to Word formation." In P. Štekauer & R. Lieber (Eds), *Handbook of Word formation*. Dordrecht: Springer.

Štekauer, Pavol, Franko, Štefan, Horecký, Ján, Stašková, Jaroslava, & Sutherland-Smith, James (1997). *Getting beyond the Rules of Language*. Prešov: Acta Facultatis Philosophicae Universitatis Prešovensis.

Štekauer, Pavol, Chapman, Don, Tomaščíková, Slávka, & Franko, Štefan (2004). *Word Formation as Creativity within Productivity Constraints. Sociolinguistic Evidence*. Manuscript.

Stepanova, Maria D. (1973). *Methoden der synchronen Wortschatzanalyse*. München: Max Hueber.

Szymanek, Bogdan (1988). *Categories and Categorization in Morphology*. Lublin: Katolicki Uniwersytet Lubelski.

Williams, Edwin (1981). "On the Notions 'Lexically Related' and 'Head of a Word'." *Linguistic Inquiry, 2,* 245–274.
Wisniewski, Edward J. (1996). "Construal and similarity in conceptual combination." *Journal of Memory and Language, 35,* 424–453.
Wunderlich, Dieter (1997). "Cause and the Structure of Verbs." *Linguistic Inquiry, 1,* 27–68.
Zadeh, Lofti A. (1965). "Fuzzy sets." *Information and Control, 8* 338–353.
Zimmer, Karl E. (1971). "Some General Observations about Nominal Compounds." *Working Papers on Language Universals.* Stanford University 5, 1–21.
Zimmer, Karl E. (1972). "Appropriateness Conditions for Nominal Compounds." *Working Papers on Language Universals.* Stanford University 8, 3–20.
Zwicky, Arnold M. (1985). "Heads." *Journal of Linguistics, 21,* 1–20.

Dictionary
Webster New World Dictionary (1988). Third College Edition

Author index

A
Adams, V. 270
Aitchison J. 12, 59, 244, 269, 270
Allen, M. R. 4, 13, 14, 40, 48, 266
Anderson S. R. 24, 223
Anderson, R. C. 24, 223
Aronoff, M. 43, 222

B
Barker, M. 226
Barsalou, L. 15
Bauer, L. 3, 4, 11, 44, 56, 196, 208, 216, 222, 223, 265, 267, 268, 270, 271, 273
Beard, R. xvi, 1, 20, 36, 37, 40, 43, 248, 249, 266, 267, 272
Bierwisch, M. 37
Boguraev, B. K. 59
Brekle, H. E. 265

C
Chomsky, N. 5
Clark, E. V. xvi, 1, 34, 35, 40, 63, 64, 248, 249
Clark, H. H. xvi, 1, 34, 35, 40, 63, 64, 248, 249
Cohen, B. 13–16, 40, 47, 63
Coolen, R. 14, 19–21, 26, 56, 80, 242, 258, 267
Coseriu, E. 60
Croft, W. 270
Cruse, A. D. 146, 236, 265

D
de Almeida, R. G. 33, 71
Derwing, B. L. 14, 33, 41, 228
Dokulil, M. 3, 18, 38, 39, 43, 44, 49, 265
Downing, P. 10–12, 31, 41, 63, 131, 132, 267, 271
Dressler, W. U. 32, 33, 94

F
Finin, T. W. 5, 11, 16, 20, 27, 28, 40, 41, 266–268
Fodor, J. A. 69
Furdík, J. 67

G
Günther, H. 266, 267
Gagné, Ch. L. 14, 21, 22, 33, 40, 41, 97, 265, 267
Gleitman, H. 4, 11, 41, 244, 252, 253
Gleitman, L. R. 4, 11, 41, 244, 252, 253
Grzega, J. 244, 245, 265, 271

H
Haiman 59
Halle, M. 48, 216
Hammond, M. 4
Hampton, J. A. 5, 11, 16, 17, 21, 33, 40, 265
Hansen, K. 43, 50, 270
Heidolph, K.-E. 8
Horecký, J. 44, 49, 67

J
Johnson-Laird, P. N 76

K
Kastovsky, D. 18, 44, 73, 265, 267
Katz, J. 69
Keane, M. 13, 17, 40
Kiparsky, P. xvi, 36, 37, 216, 222

L
Labov, W. 94
Ladányi, M. 94
Langacker, R. W. 23, 24, 59, 270
Lees, R. B. 2–4, 7, 10, 40
Levi, J. N. 4–7, 10, 19–21, 37, 40, 242, 243, 265
Li, Ch. 10, 19, 40
Libben, G. 32, 33, 55, 71, 228, 268
Lieber, R. 9, 40, 62, 226, 266, 267, 271
Lipka, L. xx, 9, 62, 131, 269, 271
Luce, R. D. 96

M
Marchand, H. 2, 3, 18, 73, 208, 265–267, 272
Mathesius, V. 50, 265
Matlin, M. W. 270
McEvoy, C. L. 232, 258
McKinney, V. 232
Medin, D. 267
Mervis, C. B. 47
Meyer, R. 11, 21, 28–30, 40, 41, 76, 251, 254, 256, 268
Miller, G. A. 76
Motsch, W. 3, 8
Murphy, G. L. 5–7, 13–16, 21, 31, 33, 40, 41, 47, 55, 63, 79, 111, 242, 246, 265

N
Nelson, D. L. 232, 258
Noonan, M. 4

O
Ondrus, P. 67
Osherson, D. N. 13, 15, 17, 40, 79, 251

P
Person, P. D. 70
Plag, I. 216, 223

R
Raffelsiefen, R. E. 216
Rainer, F. 222
Renouf, A. 56, 270
Rips, L. J 13, 17, 40
Roeper, T. 13
Rosch, E. 47
Ryder, M. E. 14, 23–26, 41, 131, 187, 208, 253, 270

S
Scalise, S. 2–4
Schreiber, T. A. 232, 258
Schreuder, R. 14, 19, 26, 37, 56, 80, 267
Selkirk, E. O. 267
Shoben, E. J. 14, 21, 22, 33, 40, 97, 265, 267
Siegel, M. E. A. 13
Sinclair, J. 23, 231
Skousen, R. 14, 41
Smith, E. E. 13, 15, 17, 18, 40, 79, 179, 251
Štekauer, P. xvi, xvii, 7, 8, 22, 23, 35, 43, 44, 57, 64–67, 78, 82, 84, 96, 228, 233, 265, 267, 269, 270, 272
Stepanova, M. D. 43
Szymanek, B. 20

V
Valera, S. 272
van Jaarsveld, H. J. 14, 19, 26, 56, 80, 267
van Lint, T. 7, 8, 30, 40, 43

W
Williams, E. 267
Wisniewski, E. J. 30, 31, 41, 56, 111, 265, 267
Wunderlich, D. 37

Z
Zadeh, L. A. 17
Zhang, N. 232
Zimmer, K. E. 4, 8–11, 40, 41, 216, 251, 266
Zwicky, A. 267

Subject index

A
acceptability xviii, xix, 8, 22, 25, 30, 60, 68, 92, 94, 95, 97, 103, 108, 150, 181, 195, 231, 271
accommodation 26
actual word 48, 210
AN-compound xvi
analogy xvii, 14, 19, 24, 25, 41, 57, 99, 106, 114, 117, 118, 131, 134, 141, 142, 150, 161, 164, 208, 227, 228, 258, 259
analogy base 24, 25, 131, 141, 142, 150
analogy-based model 14, 26, 41, 99
appropriately classificatory 9, 40, 251, 266
associative meaning 230–240
attribute 11, 16–19, 220, 268
Avoid Synonymy Principle 222, 224

B
blocking 107, 111, 112, 114, 210, 216, 222, 224, 227–230, 271

C
CARIN model 21
classification seme 67, 68, 70, 236
classificatory relevance 9
cognitive grammar 23, 24
combined concept 21, 22, 31, 40, 41
competition xvii, 21, 22, 42, 58, 81, 90, 96, 97, 112, 134, 153, 156, 170, 190, 197, 204, 207, 213, 224, 226–229, 244, 257, 260–262
Competition Principle 96
complex nominal 4–6
componentiality 32
composite concept 18
conceivability 92
concept dominance 17
concept matcher 27
concept specialization model 14, 40
conceptual category 47, 49, 52, 53, 54, 68, 70, 81, 89, 141, 142, 145, 178, 249, 270
conceptual combination 13, 21, 30, 41, 265, 267
conceptual elaboration 5, 7, 40, 111
conceptual level xx, xxi, 21, 32, 40, 44, 46–48, 50, 53, 59–61, 66, 68, 76, 81, 82, 98
conceptualisation 58
conjunction 11, 16, 17
constituency 32
context-free interpretation xv, xvii, 56, 267, 270
contextual 1, 27, 34–35, 56, 63, 64, 241, 244, 248
conversion 34–35, 52–54, 66–68, 81–85, 159–194
cue-reliability 24

D
decision-making operation 71
diagnosticity 18, 19
Discourse Representation Theory 28
distinguisher 69, 83

E
elaboration 5, 7, 15, 31, 32, 37, 40, 111
episodic linking 226
extra-linguistic knowledge 7, 35, 58–62, 65, 66, 72–74, 78, 79, 84,

94, 108, 112, 114, 118, 129, 130, 141, 142, 151, 161, 186, 244, 249, 251, 256, 259, 260, 270

F
feature model 13
feature semantics xx, 62
formal modification 48
frame 27, 267
frame-based representation system 27
fuzzy-set theory 17

G
generic knowledge 35, 248

H
head 5, 7, 10, 11, 13, 15, 18, 21, 22, 29–31, 33, 41, 49, 52, 63, 75, 246, 250, 256, 266–269

I
identification seme 67, 68, 70
idiosyncratic seme 64, 68, 69
independent activation 80
inheritance of roles 16
inheritance rules 16
interactive activation model 20
interpretative semantics xx

K
kernel sentence 2, 6
knowledge representation model 14, 28

L
lexical meaning 28, 37–39, 58–60, 71, 249, 251, 269
linguistic template 24, 25
logical predicate 46, 48, 53, 76, 82, 98, 249
logical spectrum 46, 47, 53, 66, 82, 83

M
matching process 27, 76–78, 81
minimum rule 17
minor speech community 35, 201
modifier 5, 10, 11, 13, 14, 18, 21, 22, 29–31, 40, 41, 256, 267, 268
Morpheme-to-Seme-Assignment Principle 50–52, 71, 72, 262
Morphological Type 42, 55, 60, 71, 72, 88–90, 100, 195, 196, 200, 204, 205, 206, 213, 215, 222, 223, 225–230, 253, 260–262, 269, 270

N
natural function 36, 37, 40
necessary association 8, 40
NN-compound xvi, 11, 28, 29, 60, 266, 267
non-denumerability 63
non-recoverable deletion 4

O
Objectified Predictability Rate xxi, xxii, 58, 74, 90, 95–96, 98, 154–156, 187, 191–193, 244, 256–257, 262
onomasiological base 49–54, 63, 68, 73–79, 81, 86, 103, 197, 198, 237–240, 255, 260
onomasiological category 39
onomasiological level 46, 49, 52, 54, 61, 71, 73, 75, 79, 113, 225, 246, 262, 270
onomasiological mark 49–54, 62, 73–79, 87, 106, 197, 214, 237–240, 255, 260
onomasiological recategorisation 52, 54, 66, 68, 81, 159, 225
onomasiological structure xxi, 44, 49–52, 55, 68, 70–81, 85–88, 122, 142, 204, 214, 252–255
Onomasiological Structure Rule xxi, 60, 70, 85–87, 108, 113, 122, 126, 138, 146, 252, 261, 270
Onomasiological Type 51–54, 73–82, 99, 141, 152–153, 186, 192–193, 195, 200, 204, 205, 212–215, 253–257, 259–262

onomatological level 46, 50, 51, 62, 71, 73–75, 90, 205, 225, 246, 262
onomatology 50

P
paraphrase 3, 145, 205, 252, 270
parsing ambiguity 228
pattern concept 27
phonological level 46, 51, 90
pragmatics xviii, 59, 60, 111, 117, 121, 130, 131
Predictability Rate xvi, xix–xxii, 9, 20, 22, 25, 62, 63, 72–75, 94–95, 151–153, 157, 184, 186–190, 192, 197, 212–214, 248, 254–256, 260, 262
Predictability Rate Gap 90, 96, 97, 178, 213, 256
predominant feature 35, 40
primary compound 2, 8, 10, 12, 13, 14, 19, 20, 23, 29, 80, 99, 103, 145, 152, 187, 204, 243, 254
productivity xvi, xvii, xx, 11, 22, 23, 39, 57, 88–90, 94, 96, 100, 166, 187, 214–215, 216–230, 253
proper name xxi, 65, 66, 82–85, 214
property mapping 30, 31, 41
prototype semantics xx
prototypical feature xix, xxi, 28, 30, 36, 37, 47, 49, 62, 66, 76, 77, 81, 82, 91–94, 98, 157, 231, 233, 236, 242, 246, 249, 251, 259, 260, 262
prototypical seme xix, 17, 31, 35, 49, 62, 68–70, 76, 83, 94, 99, 157, 188, 193, 231, 246, 248, 249, 260, 262
psycholinguistics xviii

R
Recoverably Deletable Predicate 4–7, 20, 21
relation linking 30
relational compound 29, 251, 256
relational distribution 22
relational information 21, 265
relator 7, 8
representation model 1, 14, 19, 28
role fitting 27
root compound xvi, 28, 34, 62

S
salience 14, 18, 19, 35, 270
schema 13, 15, 24, 26, 40, 246, 267
schema model 13, 26
schema 24, 26
selective Modification Model 17, 40, 79
semantic component xix–xxi, 34, 58, 60, 62, 67–70, 73, 76, 77, 80–82, 98, 246
semantic constraint 216, 223, 224, 226, 230
semantic feature xx, 8, 13, 40, 58, 111
semantic feature set 13
semantic formation 47, 48
semantic information schema 26
semantic knowledge 28, 59
semantic marker xx, 48, 58, 69
semantic redundancy 10, 31
semantic representation 19, 36
semantic structure xxi, 7, 42, 49–51, 63, 65, 73, 74, 76, 79, 81, 82, 83, 89, 90, 99, 130, 138, 161, 205, 249
semantic transparency 32, 223, 224, 226
Seme Level Combination xx, 106, 131, 135, 156, 248, 259, 260
slot 5, 13–15, 21, 22, 27, 30, 112, 246, 268
slot-filling 7, 14, 15, 17–20, 40, 41
sortal compound 29
specification seme 67
speech community xv, 35, 43, 44, 46, 48, 59, 60, 64–66, 70, 84, 201, 244, 246, 259, 268
speech-level predictability 241
structural interpretation rule 27
structural transparency 222, 228, 229

subcategorisation 29, 206, 216, 224, 230
system-level predictability 241

T
target concept 27
theta-role assignment 29
transposition 36, 269
typicality 91–93

W
WF Rule 23, 38, 48, 71, 72, 78, 88, 89, 198, 216, 221, 223, 261, 265, 269, 271
WF Type 39, 54, 60, 68, 71, 72, 80, 85–87, 89, 90, 96, 99, 187–190, 194–196, 213, 215, 222–225, 229, 231, 245, 254, 259–262, 269, 270
Word Formation Type Cluster 196, 215, 270

In the series *Studies In Functional And Structural Linguistics* the following volumes have been published thus far or are scheduled for publication:

43 **ANDREWS, Edna and Yishai TOBIN (eds.):** Toward a Calculus of Meaning. Studies in markedness, distinctive features and deixis. 1996. xxviii, 432 pp.

44 **JESSEN, Michael:** Phonetics and Phonology of Tense and Lax Obstruents in German. 1999. xx, 394 pp.

45 **MARTÍN-VIDE, Carlos (ed.):** Mathematical and Computational Analysis of Natural Language. Selected papers from the 2nd International Conference on Mathematical Linguistics (ICML '96), Tarragona, 1996. 1998. xviii, 391 pp.

46 **ŠTEKAUER, Pavol:** An Onomasiological Theory of English Word-Formation. 1998. x, 192 pp.

47 **MARTÍN-VIDE, Carlos (ed.):** Issues in Mathematical Linguistics. Workshop on Mathematical Linguistics, State College, PA, April 1998. 1999. xii, 214 pp.

48 **REID, Wallis, Ricardo OTHEGUY and Nancy STERN (eds.):** Signal, Meaning, and Message. Perspectives on sign-based linguistics. 2002. xxii, 413 pp.

49 **HLADKÝ, Josef (ed.):** Language and Function. To the memory of Jan Firbas. 2003. x, 339 pp.

50 **DUŠKOVÁ, Libuše (ed.):** Dictionary of the Prague School of Linguistics. Translated from the French, German and Czech sources. Author: Josef Vachek †. In collaboration with Josef Dubský. Translated by Aleš Klégr, Pavlína Šaldová, Markéta Malá, Jan Čermák and Libuše Dušková. 2003. x, 216 pp.

51 **CONTINI-MORAVA, Ellen, Robert S. KIRSNER and Betsy RODRÍGUEZ-BACHILLER (eds.):** Cognitive and Communicative Approaches to Linguistic Analysis. 2004. viii, 389 pp.

52 **GORLACH, Marina:** Phrasal Constructions and Resultativeness in English. A sign-oriented analysis. 2004. x, 151 pp.

53 **EDDINGTON, David:** Spanish Phonology and Morphology. Experimental and quantitative perspectives. 2004. xvi, 198 pp.

54 **ŠTEKAUER, Pavol:** Meaning Predictability in Word Formation. Novel, context-free naming units. 2005. xxii, 288 pp.


Sharing is caring
ebook