FOREIGN WORDS AND PHRASES IN ENGLISH

METAPHORIC ASTROPHYSICAL CONCEPTS
IN LEXICOLOGICAL STUDY

Foreign words and phrases in English, 2009

Acta Universitatis Wratislaviensis No 3169

Foreign words and phrases in English, 2009

Marek Kuźniak

FOREIGN WORDS AND PHRASES IN ENGLISH

METAPHORIC ASTROPHYSICAL CONCEPTS IN LEXICOLOGICAL STUDY

Wrocław 2009
Wydawnictwo Uniwersytetu Wrocławskiego

Referee
Barbara Lewandowska-Tomaszczyk

Copy Editor
Martyna Gatner-Siemieniuch

Layout
Urszula Piszczek

Cover design
Hanna Parylak

On the cover:
Kepler's illustration to explain his discovery of the elliptical orbit of Mars, 1609.
Artist: Unknown. Photo Agency MJS

Th e book was published in cooperation with the Faculty of Philology, University of Wrocław
and Wydawnictwo Uniwersytetu Wrocławskiego Sp. z o.o.

ISSN 0239-6661
ISBN 978-83-229-3055-7 (druk)
ISBN 978-83-229-3674-0 (PDF)

Typesetting by
Wydawnictwo Uniwersytetu Wrocławskiego Sp. z o.o.

Printed by TOTEM

To my son, Piotr

Foreign words and phrases in English, 2009

Table of contents

Foreign words and phrases in English, 2009

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to Professor Barbara Lewandowska-Tomaszczyk for the editorial review of this book. Her invaluable remarks and comments helped me to rethink, and, hopefully, more clearly express fundamental theses postulated in this book as well as eliminate numerous inaccuracies contained in the manuscript. I owe a debt of gratitude to Professor Leszek Berezowski and Professor Michał Post for taking the time to proofread the draught version of the present book and for their comments regarding both the contents and the structure of this monograph. Let me also express my gratitude to Dr. Michał Szawerna for his insightful remarks on the ideas discussed in this book as well as his friendly support. My gratitude goes also to Professor William McGregor, Aarhus University, Dr. Anna Ledwina, Opole University, Dr. Michał Garcarz, Wrocław University, and Agata Juraszczyk, Österreich Institut Wrocław, for their assistance connected with the preliminary research on foreign words and phrases. My deep appreciation goes to Weronika Górska from King's College, London, Elaine Horyza and David Bridges, University of Wrocław, Magdalena Sadowska from Cork, Ireland for helping me with the distribution of word-recognition questionnaires across the UK. I owe special thanks to the two physicists, and at the same time my MA seminar students, Ewa and Jacek Woźny, for explaining to me some intriguing complexities of the fascinating world of physical science. I would also like to give my thanks to Mateusz Chmielewski for his technical assistance with the glossary of foreign words and phrases. Last but not least, I would like to cordially thank my family for their endearing encouragement and unfailing support with the progression of my work. Without all these persons, this book would not certainly have taken its present shape.

Foreign words and phrases in English, 2009

TYPOGRAPHIC CONVENTIONS

CATEGORIES	DOUBLE INVERTED COMMAS
LANGUAGE DATA	ITALICS
LONGER QUOTATIONS	FONT SIZE 10, INDENTED AND EXTRACTED FROM THE MAIN BODY OF THE TEXT
MEANINGS AND FIGURATIVE USE	SINGLE INVERTED COMMAS
SHORTER QUOTATIONS	DOUBLE INVERTED COMMAS
TITLES OF ARTICLES	SINGLE INVERTED COMMAS
TITLES OF BOOKS	ITALICS
THESES AND REGULARITIES	FONT SIZE 10, SEMIBOLD

COMMON ABBREVIATIONS

FIGURE	Fig.
FOOTNOTE(S)	Ft(s).
TABLE	Tab.

Foreign words and phrases in English, 2009

PREFACE

This book is about foreign words and phrases in English. The word *foreign* is understood relative to a phonological criterion. Namely, if the word or phrase is uttered with the imitation of the pronunciation of the language of origin, this word or phrase counts as *foreign*. The information about such conceived 'foreignness' is brought by Longman Pronunciation Dictionary. The treatment of this lexicon as the reference base for the elicitation of foreign words and phrases allows us to encompass, besides common words and phrases, a large group of proper foreign words and phrases for the analysis. Such broader analytical perspective not only accords with the conception of 'foreignness' adopted in the book, but is also legitimized on account of the proposed methodology of description (see discussion below).

However, the description of foreign lexis selected along these phonological lines is one part of the story told throughout this book. The second part relates to the outline discussion of the lexicological tradition of research upon foreign lexis. This discussion concentrates only on the facts relevant to the description of the assimilation processes addressed. It thus largely excludes a huge scholarly heritage of research upon borrowings, as these are not the focus in this book.

Last, but certainly not least, the third part of the book is about the search for a viable framework with regard to which the processes of assimilation of non-native words can be described. This may lead to some controversies in the reception of the main goals and assumptions of the book; therefore, explicating the nature of this framework requires additional attention in this place. The framework as such, is argued to be found in astrophysics. It does not of course mean that astrophysics 'engulfs' linguistics as a discipline; making such a postulate would be, to put it mildly, intellectually risky. Instead, the relationship between astrophysical concepts and lexical assimilation processes is argued to be grasped in terms of the overriding explicatory metaphor: LANGUAGE LAWS ARE PHYSICAL LAWS. The metaphor is metadiscursive and metacognitive in its essence. It serves as a convenient conceptual device by means of which modeling foreign lexical assimilation processes can be handled no matter if you are an expert in astrophysics or not. Both naïve and expert views of the physical world, which were epitomized in the respective Ptolemy's and Copernicus' conceptions of the universe, equally count, as the book argues, for the model of lexical assimilation.

But the discussed metaphor is not to be understood in a radical way. It is metonymic in character, i.e. LANGUAGE LAWS actually relate to lexical assimilation processes, whereas PHYSICAL LAWS are reduced to the phenomena, which are sensorily accessible to a human conceptualizer. Therefore, for example, Newton's laws of dynamics, along with the concepts of motion and energy duly apply in the

Foreign words and phrases in English, 2009

analysis as they are grounded in the phenomena experientially akin to a human being. The argumentation advanced in the book is thus ultimately transcognitive in nature. This transcognitive character of the study is particularly visible in the treatment of the overriding metaphor which *is not* posited to reside in the mind of the individual language user, because people normally do not talk about such phenomena in casual conversations. LANGUAGE LAWS ARE PHYSICAL LAWS is thus not to be interpreted in the Lakoffian sense of a conventional metaphor that "we live by."

The metaphor is also not to be interpreted on a purely literal basis. Investigating this literalness of the metaphor would perhaps require neurolinguistic evidence or some other arduous psychological experimentation. This *is not* the ambition of the book. Instead, the book aims to search for a relevant external framework with regard to which the description of foreign lexical assimilation processes can be handled. This external framework perspective entails situating the aforesaid metaphor somewhere in between the literal and non-literal interpretation. Inevitably, this may lead to occasional visits to either more literal or non-literal ends of the interpretative continuum. To put it in a different way, the metaphor, as it is understood in this book, leaves out, as it were, the bounds of conventional human cognition and directs itself to the external environment, but does not escape too far away. Making too remote an escape would mean that the source domain by means of which we describe the linguistic phenomenon of lexical assimilation becomes too abstract, too immaterial and too imperceptible to adequately relate to the target domain. Such would be the case if lexical assimilation were to be described relative to, for example, quantum mechanics. The framework would then lose its transcognitive character which is otherwise guaranteed by positioning the metaphor *outside but close* to human cognition. This location is necessitated by the underlying assumption that the phenomenon of lexical assimilation is a communal (social) property in de Saussure's meaning of the word. This is the property that we normally are not aware of unless our attention is drawn to it.

Therefore, the metaphor LANGUAGE LAWS ARE PHYSICAL LAWS is essentially metacognitive. Conferring such a status onto the metaphor may appear a disappointing solution to the presented study, but the proposal which the book makes is just a first step toward the explication of lexical assimilation mechanisms. The second step would be to potentially admit that the metaphor LANGUAGE LAWS ARE PHYSICAL LAWS is constitutive, i.e. that it is psychologically real. But the book does not decide on this issue. Making thus this first and at the same time final step, the book argues that the presented metaphor along with the following argumentation can at most serve as a *tentative* rather than *definite* description of foreign lexical assimilation processes. This tentative argument is conducted in the book with the aid of indirect evidence, i.e. a frequency-based analysis of almost 3500 words. The evidence, as already said, is indirect, and therefore not entirely

conclusive. The only direct access to the metaphor would be through examination of metadiscursive talks conducted by lexicologists on the subject of lexical assimilation. Samples of such direct evidence are discussed in the book. These appear to favor the metaphor selected for the description of foreign words in English. Certainly, more evidence is needed, but this goes beyond the scope of this book.

All in all, the postulated correspondences between the physical and linguistic world should eventually be read as only *introductory*, yet, hopefully, inspirational for further usage-based and/or neurolinguistically oriented research. The question remains, however, if such neurolinguistic or experimental psychological studies are not perhaps bound to fail, given the presumed temporal and social dimension of the discussed assimilation processes. The temporal dimension requires us to look at the assimilation processes as extending over a longer stretch of time rather than compacted into a temporal bundle and presented as a ready-made package to the everyday linguistic awareness of an individual speaker. On this account, the model of lexical assimilation advanced in this book appears incapable of grasping the psychological reality of the related processes in the minds of average language users. Therefore, usage-based studies which aim at establishing the degree of conventionality of the metaphor are not necessarily much helpful either. Instead, the implicated diachronicity of the model of assimilation processes naturally leads to viewing them as a social rather than an individual speaker's property. Still, the intriguing issue of how the correspondences between the lexical world and physical world entailed in LANGUAGE LAWS ARE PHYSICAL LAWS can be more objectively tested and/or eventually *proven* is not to be resolved at this stage; yet, it presents an open challenge to prospective scholarly research in the field.

INTRODUCTION

> "English is a vacuum-cleaner of a language.
> It sucks words in from any language it makes contact with."
>
> David Crystal (2007b: 59)

0.1. Preliminaries

David Crystal, one of the world's most eminent authorities on language, in the introduction to his book entitled *Words, Words, Words* said what can be viewed as fundamental to any scholarly enterprise in the field of lexical studies:

> No book on words could ever be comprehensive, but it can at least be representative of what is 'out there.' Language is too huge a subject to be discovered by any one person. Everyone has their own linguistic story to tell, and each story is worth the telling.
>
> (Crystal 2007b: VIII)

Whether the story told in this book has eventually been worth the telling is not certainly to be judged by the author of the present research. This monograph, accordingly with Crystal's reservations quoted above, is by no means intended as a comprehensive account of foreign words and phrases in the English language. Postulating such account would be highly unreasonable, if not simply intellectually fallacious. Language is too complex a phenomenon to be reliably encapsulated within the bounds of one volume of the book. The story that will be advanced on the pages of this monograph is thus a perspectivized, non-holistic description of foreign lexical elements in the English language system at the turn of the new millennium. The two words are crucial for our further considerations, i.e. "perspectivization" and "description."

The perspectivization of the study basically entails a degree of limitation as for the number of selected foreign words and phrases under consideration. It is sufficient to say that the average figures concerning the number of lexical items discussed in various dictionaries of foreign words and phrases (see Chapter 4) oscillate between 5000 words (e.g. DFWPCE) and 8000 entries (e.g. ODFWP I). This compared to the present study (ca. 3500 words and phrases) unequivocally indicates that the list of foreign words and phrases displayed at the end of the book represents a selective set. It appears that without such pre-selection the present study might, aside from expanding to unpublishable proportions, offer conceptual chaos and the ensuing 'unreadability.'

The question that naturally arises relates to the criteria of such selection and possible reasons (other than editorial) behind it. The first and foremost criterion

Foreign words and phrases in English, 2009

of selecting foreign words and phrases in the book is phonological. The collection of words and phrases under consideration was made on the basis of two editions of *Longman Pronunciation Dictionary* (LPD), i.e. 1995 and 2004 by John Wells. This dictionary offers a lucid typographical convention of marking those lexical units as foreign which preserve the pronunciation of the language of origin in English. The other criterion is graphemic and involves two aspects of data organization: internal and external.

An internal aspect relates to the organization of data into common words and proper names. This is possible because LPD does not only list the conventionally available pronunciations of common words and phrases, but also presents the reader with a comprehensive list of proper names and phrases. This has led us to adhere, throughout the whole of the present study, to the fundamental distinction between the subset of foreign common words and phrases and the subset of foreign proper names. The categorial demarcation between the two groups of words is sometimes highly debatable (see section 5.4.2); therefore, in this book we generally follow a classificatory rule based on the capitalization of the first letter of an entry. Whenever the lexical unit is presented in LPD with the initial capital letter it has been classified as proper name, and when a given word or phrase begins with a regular, non-capitalized letter, it has been treated as a common word or phrase.

Another external aspect relates to the further processing of data. Namely, each elicited foreign word or phrase has been subject to the investigation in the British National Corpus for its statistics of frequency of occurrence. Naturally, the entry put into the SARA search engine was given as it is spelled rather than as it is pronounced. This produced some basic information connected with the frequency of occurrence of a given lexical unit in the whole of the corpus on the one hand, and the number of texts in which a given form can be found, on the other. This has led us to developing the concept of CRAC that we will discuss in Chapter 6.

The phonological and graphemic criteria are consonant with Crystal's (2007b: 60) view on what makes the word 'foreign'. Bearing in mind that Crystal does not differentiate between loanwords and foreign words and phrases (see sections 3.0, 4.2, 5.2.1), he asserts that "some loans retain their foreign resonance, often because they have a different sound structure to typical English words or they are spelled distinctively."[1] Elsewhere he uses the term 'fresh loanword'. It appears then that one of the main foci in the present study will be the investigation of this 'freshness' of loanwords as these appear to closely correspond to the elicited set of foreign words and phrases discussed above (see especially section 5.2.1).

We can thus summarize the avenues of data selection and categorization mechanisms under the following diagram:

[1] Emphasis mine (M.K.).

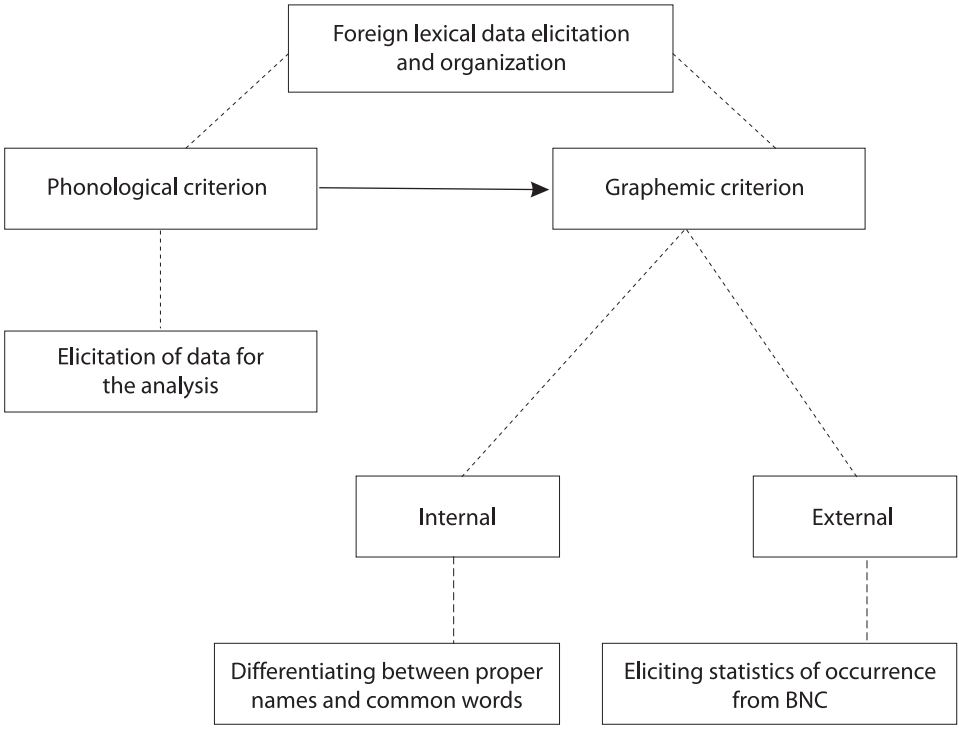


Fig. 1. *Eliciting and organizing foreign lexical data*

0.2. The central claim

In the previous section, we discussed the two pivotal terms, i.e. *perspectivization* and *description* which delineate the general framework of the present research. We already explained how the term *perspectivization* should be understood relative to research objectives adopted in this book. Let us now dwell a bit on the second crucial word, i.e. "description" as it leads us towards the formulation of the central claim constituting the *raison d'être* of the present study. The "description" is understood here as involving two aspects: internal and external. An internal aspect of data description is elaborated on in Chapter 7 and involves a semantic characterization of foreign words and phrases both as common words and proper names in the English language. The criteria of semantic characterization and the details of the adopted procedure are discussed at length in Chapter 7.

An external aspect of the description relates, in turn, to another fundamental goal behind the analysis, i.e. the reconstruction of the model of foreign lexical assimilation processes. This model is postulated to be illustrative of not only foreign words and phrases in English but its ambition is to display the nature of any

foreign lexical assimilation in no matter which target lexical system (see Chapters 8 and 9). These universalist postulates are argued to be substantiated inasmuch as we take the central claim of the book into account. The claim is underlain by the following metadiscursive and explicatory metaphor: LANGUAGE LAWS ARE PHYSICAL LAWS and can be stated as follows:

Language undergoes continual change. Change can be discussed in terms of motion.[2] Language can thus be described as being in constant motion. Motion is determined by forces. Forces that determine language change are analogical to physical forces. There are two fundamental types of forces: centripetal (centre-seeking) and centrifugal (away-from-the-centre) forces. These physical forces are present in circular motion. Therefore language change can analogically be modeled as a circular motion. Languages being in constant motion can be compared to planets. Their mutual interactions are guaranteed by the centripetal force of interplanetary gravity. The atmosphere of a planet (e.g. the Earth [i.e. the English language]) is, however, the residue of opposite centrifugal-like forces. Once an alien body (e.g. a meteor), driven by the force of gravity, enters the Earth's atmosphere, it meets the resistance of densely accumulated air-particles that form the opposite (centrifugal-like) force(s) acting on this alien entity.

Explicatory (metadiscursive) and quintessentially metacognitive, the metaphor LANGUAGE LAWS ARE PHYSICAL LAWS implies that there exist a number of correspondences between the linguistic world and the physical world, especially if these involve modeling lexical assimilation mechanisms. Namely, successful (pre-)adaptation of a foreign lexical unit (a meteor[ite]) into the target language system (in this book English – the planet Earth) occurs if the sum of centripetal force(s) that governs the incorporation of a foreign word or phrase *into* the target lexical system prevails over the sum of centrifugal force(s) that conserves the system. Breaking through the protective barrier formed by centrifugal-like forces thus ensures completion of assimilation process by a foreign word or phrase and its ensuing entrenchment in the target system. Consistently, then, failure to overcome the barrier formed by centrifugal force(s) is analogical to the expulsion of a foreign word or phrase *from* the target system, i.e. its lack of adaptation in the English language.

Naturally, then, the metaphor LANGUAGE LAWS ARE PHYSICAL LAWS along with its set of underlying correspondences will bear a far-reaching consequence for the argumentation advanced on the pages of this monograph. The explication of reasons for the adoption of such status of the metaphor in the present analysis has already been provided in the Preface to this book. The claim about the far-going similarities between how certain aspects of natural language function and how certain aspects of the physical world function is, in turn, elaborated on in detail in Parts III and IV of the research, where the model of foreign lexical assimilation is eventually specified (Chapter 9) and summarized (Chapter 10). Certainly, the target domain LANGUAGE LAWS is not to be equaled with all of the rules

[2] For the discussion of CHANGE is MOTION and CAUSES are FORCES metaphor, see Lakoff (1993: 225).

and principles governing natural language, but should rather be interpreted as a metonymy designating lexical assimilation processes. In addition, the word "law" is not to be interpreted in the classic meaning of the word sanctioned by 19th century historical-comparative linguistics. What is meant by "laws" actually relates to a set of regularities governing the phenomenon of lexical assimilation. These are specified in Chapter 8 of this study. The same 'metonymic' reasoning applies to the source domain PHYSICAL LAWS, which embraces only selected laws connected with force, motion, or energy (see Chapter 6).

Another important claim advanced in the book relates to the manner in which human conceptualizers potentially conceive of foreign lexical assimilation processes. The word "potentially" marks a necessary proviso, here, as the phenomenon of lexical assimilation, with its inherent diachronicity is essentially metacognitive in character. Namely, language users do not focus on this phenomenon in everyday communication unless their attention is specifically drawn to it (see Preface). Postulating such meta-conceptual level of consideration for the description of lexical assimilation processes is not incongruent with the argument that such metadiscursive construal is founded in its source domain on the postulated integration of expert and non-expert knowledge about how the universe is structured. What is even more, the book argues that such conceived domain of PHYSICAL LAWS eventually favors the justification of the adopted model of lexical assimilation – the phenomenon which affects both expert and non-expert language users, although the latter are not normally aware of lexical processes in the same way as they are not normally aware of the fundamental physical forces constituting the reality in which they live. All in all, such integrated approach to the source domain of the above-presented general metaphor appears, then, to be supportive of the model of lexical assimilation of foreign words and phrases as presented in Chapter 9.

The nature of the description of foreign words and phrases can thus be illustrated in the following way:

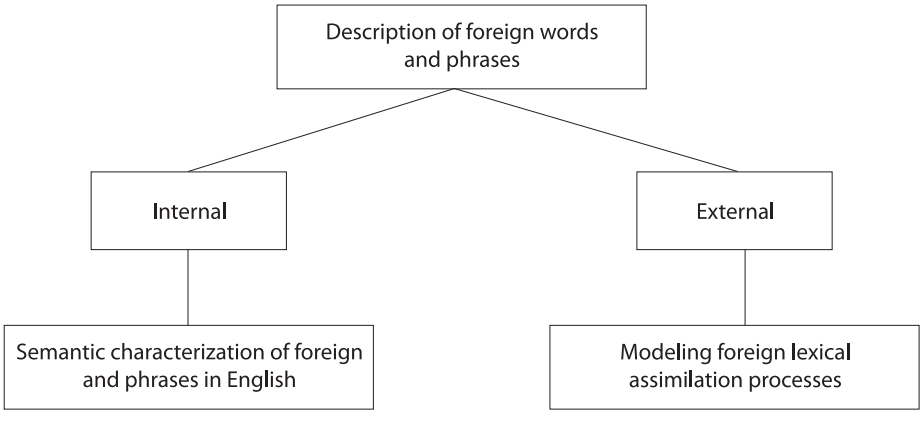


Fig. 2. *Descriptive framework of foreign words and phrases*

0.3. Language contact and foreign lexis

One of the statements constituting the claim presented in section 0.2 is certainly not a new one, i.e. the starting claim about constant language change – one of the fundamental tenets of any historical linguistic research. Quite self-evident, therefore, appears to be another observation that such change must be driven by some language-internal and language-external factors. This book, as already indicated above, focuses on one of such language-external mechanisms and in doing so, treats about the description of foreign lexical assimilation processes as encapsulated within the framework of circular physical motion with all its methodological consequences. Such a description, to the best of our knowledge, has not been attempted so far in the literature on the subject matter.

There are lots of valuable studies conducted on the so-called phenomenon of lexical borrowings (see Görlach 2003, 2007; Fisiak 1985b, 1986, Mańczak-Wohlfeld 1994, 1995, 2006), but these usually concentrate on the final stage of assimilation process, i.e. the situation when a foreign word or phrase is already well adapted to the target language system. It is thus the investigation of all aspects of such integration (phonological, graphemic, semantic, grammatical) that these studies are primarily devoted to. As a result, the processes of lexical assimilation, although well described in the literature of the subject matter,[3] remain quite significantly overlooked when it comes to the description of mechanisms preceding such full integration (i.e. incorporation or pre-adaptation [see Fig. 14 in section 5.3]). If such descriptions are attempted, these are usually confined to some general yet interesting observations.[4]

If borrowings cannot be eradicated from the system of the receiving language,[5] the question that arises is about the reasons for that state of affairs. The answer seems to stem from the *naturalness* of the procedure of foreign lexical as-

[3] See Chapter 5 for the overview of selected works. The key observation is that these works deal primarily with stage 3 of the entire assimilation process (see Fig. 14, section 5.3), which is not the focus in the present book.

[4] Language is the system which undergoes change at all levels. The fastest, the greatest and, as a result, the most noticeable changes occur in lexical sphere. Lexical innovations can happen in a threefold way: via the association of an existing lexeme with a new meaning, the coinage of a new native for, and finally the introduction of a borrowing. The last way appears to be the simplest one and this may explain the popularity of the phenomenon also in the Polish language. It is commonly accepted that there are no natural languages without external borrowings, i.e. foreign lexical elements. It is evident that borrowings cannot be ousted from the language system. As Jakobson said, the faith in self-sufficiency of a language is more illusory than the faith in the self-sufficiency of economy (Buttler *et al.* 1973: 421, after Mańczak-Wohlfeld 2006: 81) – translation mine (M.K.).

[5] Mańczak-Wohlfeld (2006: 11, translation, M.K.) says that the phenomenon of borrowings in language contacts is inevitable. It occurs to a lesser or greater degree in all languages across all epochs, because there are no so-called pure natural languages that would exist without external borrowings, i.e. lexical elements from other languages.

similation process.[6] It is thus in the description of this naturalness that the novelty of the proposal advanced throughout the present book may be ascribed to.

We begin, however, with some preliminary glimpses over some of the relevant issues discussed in the literature on non-native vocabulary, as these foreshadow some of the findings elaborated on further in the book. In the 1998-version of SAE,[7] Haspelmath (1998: 273) distinguishes the following kinds of language groupings (based on typological characteristics):

> a) the nucleus, comprising Dutch, German, French, and northern Italian dialects,
> b) the core (Spanish, Portuguese, Italian, Rumanian) and Germanic languages (English, Icelandic, Norwegian, Swedish, Faroese), West Slavic (Czech, Slovenian) and South Slavic (Bulgarian), and Balkan languages (Albanian, Modern Greek),
> c) the periphery, consisting of East Slavic (Russian, Ukrainian, Belorussian), Baltic (Latvian, Lithuanian), Finnish and Hungarian, Basque, Maltese, Armenian, and Georgian,
> d) the Celtic languages (Irish, Welsh, Breton), which are not part of SAE.
>
> (after Heine & Kuteva 2006: 24–25)

Interesting in the above-illustrated typology is the centre/periphery framework within which the languages are postulated to be encapsulated.[8] Thus, there are languages constituting the nucleus, core and periphery with some languages not making part of SAE. The typological model can, thus, in its essentials be compared to the model of network of languages developed in Chapters 8 and 9 of the present book. Some scholars (e.g. Becker 1948) regard the network to be homogeneous to the extent that they suggest the treatment of European languages as a unit. What is particularly interesting to the present research is that for some other researchers (e.g. Thomas 1975: 47) it is the phenomenon of borrowing

[6] As Crystal (2007c: 89) observes: "You cannot stop language change. You may not like it; you may regret the arrival of new forms and the passing of old ones; but there is not the slightest thing you can do about it. Language change is as natural as breathing. It is one of the linguistic facts of life."

[7] Standard Average European or *Sprachbund*. See also Whorf (1956b: 138).

[8] Similar conception but illustrated on the basis of relations in the lexicon rather than relations among languages is given in introduction to the SOED (1973: x) as well as discussed in Leisi's book (1985: 187 ff), where a synchronic survey of the structure of the English vocabulary is presented (Lipka 2002: 17). This survey is based on the concept of central/peripheral domains of lexis in the structure of language system and can be summarized as follows: "The representation starts out from the assumption that the vocabulary of English contains a large central area, which is common to all media, styles, and social classes. This concept of common English, symbolized by COMMON is equivalent […] to the 'common core' treated in the university Grammar of English" (Quirk, Greenbaum 1973: 1 ff) (e.g. *Father*, *mother*, *drink*, *bed*, *fruit*, *car*), which is present in all varieties of English. […] Above and below we have neigbouring areas of "Literary" and "Colloquial." Connected with these in a radial manner are further areas of the vocabulary, which are also related to each other in content. […] the literary vocabulary contains scientific, foreign and archaic words. […] "Colloquial" English contains dialectal and vulgar elements, as well as words from slang and technical language (after Lipka 2002: 17).

(strictly calquing or loan translation) that plays the most significant role in the rise of such a unit, in that it "produced a convergence of the lexical and semantic systems of the languages of Europe, helping to create the European *Sprachbund*" (after Heine & Kuteva 2006: 5). If the role of borrowings is that crucial, no wonder the interest in the effects of language encounters has been the subject of focal research since the advent of the scientific era of linguistic research, i.e. 19th century (Winford 2005: 6).

As Winford (2005: 6) further claims:

> During the heyday of historical linguistic scholarship in the nineteenth century, research on language contact became an integral part of the field and played a vital role in debate over the nature of language change. Michael Clyne (1987: 453) reminds us, the issue of language encounters was a topic to which such great linguists as Müller (1875), Paul (1886), Johannes Schmidt (1872), and Schuchardt (1884), among others, devoted a great deal of their attention. It continued to be a central topic well into the twentieth century, and was addressed by Sapir (1921), Bloomfield (1933), and early pioneers of structuralism. In the heyday of structuralism during the 1940s to the 1960s, it became rather less central, though not completely marginalized.

As far as the research on aspects of lexical change through language contact, Winford (2005: 42) mentions the 19th century study by Paul (1886), and the subsequent works by Seiler (1907–13) and Kaufman (1939). Winford (2005: 42), however, states that the most noteworthy frameworks (due to their comprehensiveness) appeared as late as 1949 with the work of Betz and his distinction between *Lehnwort* (loanword) and *Lehnprägung* (loancoinage), which is still valid in contemporary lexicological descriptions. The distinction may also be claimed to perfectly mirror the processes of externally induced, but essentially internal language change (*Lehnprägung)*, and simply external language changes (*Lehnwort*). It is of course the description of *the nature* of the latter process that Chapters 8–9 of this book are especially concerned with.

Of quite considerable importance to the research on language contact appear Haugen's (1950a, 1950b, 1953) contributions to the existing typologies of non-native vocabulary based on "the presence or absence of foreignness markers" (Winford 2005: 43). This led Haugen (1950b) to introduce another distinction between "importation and substitution – a dichotomy based on the presence or absence of foreignness markers. Importation refers to the adoption of a foreign form and/or its meaning, and may involve complete or only partial imitation. Substitution refers to the process by which naïve sounds or morphemes are substituted for those in the donor model" (Winford 2005: 43). Haugen's observations on importation and substitution are fundamental to the research conducted in the present book in that the data elicitation has been based, as already indicated above, on the phonological criterion of the retention of foreignness of a given word or phrase.

If the processes of externally induced language change cannot by any means be underestimated, it might be interesting to specify the factors which are often referred to in the literature and which are regarded as contributive to the lexical growth of the English language. It must, however, be remarked that this external impact from other languages is by many linguists (see also section 3.1) viewed as something formidable, the example being the following excerpt: "Our language, for almost century, has by the occurrence of many causes, been gradually departing from its original Teutonick character, and deviating towards a Gallick structure and phraseology, from which it ought to be our endeavour to recal it" (Johnson 1852: IV, after Mańczak-Wohlfeld 2006: 9).

The above-illustrated fears may well be suggested as paradoxically originating from the conception of English as a lingua franca of the contemporary world[9] (see, e.g. sections 6.7.2, 10.2). I say 'paradoxically' because we might intuitively search motivation for such language identity defence mechanisms with speakers whose language has some significant place on the world map of languages, but, nevertheless, it is subject to the threat of domination by some other languages in a given sphere of social discourse like entertainment, sport, media (e.g. the French language, [see also section 3.0]). The position held by English as a dominating language in the universe of media, science and cultural discourse may, however, provoke people to react in a similar, conservative way.[10] An attempt at further explicating the nature of such fears is presented in section 6.9.

Somewhat more optimistic (realistic) view on the character of language contacts is given by Halliday *et al.* (2005: 63–64) who notice that language lexical (semantic) interactions are a natural phenomenon that actually bespeaks a language's own distinct identity from another, because if the opposite were true, there would be nothing to imitate or acquire. In addition, the fact that language differences are real shows, according to the authors, that these variations are not insurmountable, the example being successful foreign acquisitions.

Setting the attitudinal issues to foreign words and phrases aside at the moment, we can observe that the noticeable expansion of English may be argued as stimulated by the following:

[9] See Phillipson's (1996: 8 in Mańczak-Wohlfeld 2006: 10) remark on English as the language on which 'the sun never sets'.

[10] As Dustoor argues (1968: 265): "Linguistic chauvinists might resent the suggestion that English is a superior language. It is not a question of superiority or inferiority. No Indian language is yet fully developed to replace English for the study of scientific and technical subjects. Over 66 per cent of the world's standard books on scientific and technical subjects are written in English. English is considered so important by the Russians and Chinese that they are introducing it as a compulsory subject in their universities. […] with the dawn of freedom there was an urge to remove English from its position, not to the extent to which it is replaced by Hindi, but all at once. It was the result of hypersensitive nationalism, the consequences of which have complicated our problems" [after *Pioneer*, article "The role of English" of the 5th September]).

– the emergence of British Empire,
– industrial revolution,
– economic imperialism as spawned by the development of the United States,
– the growth of science,
– the language of diplomacy,
– the mass media.

(After Mańczak-Wohlfeld 2006: 10–12)

Prior to Mańczak-Wohlfeld's works, Schlauch (1959) attempted to analyze the English language of her day. She concisely summarizes various geographically conditioned factors stimulating the changes of the present day English. She quotes India, that huge country, which "was subjected to a century and a half of English rule, with the inevitable introduction of the conqueror's language for purposes of trade, finance, administration, and government" (Schlauch 1959: 206). She also mentions African territory (Kenya, Ghana, etc): "In these regions too the English language has been introduced for governmental administration, missionary activities, and commercial enterprise" (Schlauch 1959: 207).

Schlauch (1959: 207) further quotes China in the context of both British and American commercial interests. We will see in Chapters 7 and 8 how the factors of the increasing international commerce and trade has led the Chinese language to have grown over the last decade from the minor 'lender' of words into English to one of the major donor languages of the contemporary English lexicon.[11] Schlauch concludes her observations on language contact by asserting the dominating role of English on the world-language map:

> The wide dissemination of English over the surface of the globe by means of peaceful and warlike, commercial and cultural, has left its mark in innumerable other languages encountered on the way. In each situation, the English loan words give an index of the most important contacts: in politics and sport, business, and technology. […] More germane to the history of English is the influx of loan words from the alien languages into it.
>
> (Schlauch 1959: 208)[12]

The last sentence from the excerpt is quite telling in that respect. By this account, studying the foreign words and phrases in English is like discovering little by little its own history. The story elaborated on further will thus be hoped to offer a glimpse at a section of such external history of English as it presents itself at the turn of the 21st century.

The query that needs to be posed in the context of such defined lexicological research is whether description of lexis can be conducted and explained through

[11] See Chapter 7.

[12] As Foster (1981: 72) claims: "Throughout its history the English language has always been hospitable to words from other tongues and while it is doubtless true to say that all forms of human speech have to some extent borrowed from outside models there are grounds for thinking that English is more than usually open to foreign influence as compared with other great languages."

recourse to language structure itself or whether successful explication should obligatorily necessitate intellectual adventures outside the realm of language self-sufficiency (see Schlauch 1959: 218). As the argumentation so far advanced (see Preface and section 0.2) unfolds, we will see that reference to an external world outside language is an essential prerequisite behind the explanatory power of the model of foreign lexical assimilation presented in Chapter 9.

So far, a comprehensive historical discussion of the external influences upon the English vocabulary has been presented by Fisiak (2000) and recently also offered in a concise summary by Harley (2006: 249–266).

Tab. 1. *A modified version of Tab. 9.2 in Harley (2006: 166)*

Period	Vocabulary
Old English 450–1000	Mostly Germanic, a few Latinate borrowings for religious words, Celtic place names, some common Old Norse words.
Norman Conquest 1066–1200	Old Norse pronouns *she, they, their, them* established in English.
Middle English 1200–1500	Approximately 10,000 borrowings from Norman French.
Early Modern English 1500–1750	Rampant borrowing from Greek and Latin for scientific, scholarly and technological terms; coining of new Greek and Latin words using never-before combined roots. Continued cultural borrowing from French and Italian, names for imported goods and ideas from Spanish, Portuguese, Dutch, American Indian languages, Hindi, and other colonial languages.
1750 – Modern Day[13]	Further dynamic development of English vocabulary. English becomes a kind of contemporary lingua franca.

As Rayevska (1979: 214) claims:

> The vocabulary of Modern English is a product of a number of epochs. The existence of the English language began when Germanic tribes had occupied the lowlands of Great Britain. The historical records of English do not go so far back as this because the oldest written texts in the English language (in Anglo-Saxon) date from about 700 and are thus removed by about three centuries from the beginnings of the language […]. The importance of this purely Germanic basis is often overlooked, largely because of the large number of foreign words incorporated in the present-day vocabulary. Many studies of the English language give under prominence to the foreign element, thus leaving an incorrect impression of the foundation of the language.

Manfred Scheler's (1977) in his *Der Englische Wortschatz* appears, in turn, to be mainly preoccupied with such issues as "the etymological heterogeneity of the English vocabulary, the problem of the mixing of languages with its causes and consequences, and finally specific forms of the English vocabulary and its stratification" (after Lipka 2002: 14). Lipka (2002: 14) also quotes Ernst Leisi's (1985)

[13] More general remarks about the contemporary lexical acquisitions into English will be surveyed in sections 4.2 and 4.3.

Das heutige Englisch. Leisi's treatment may be viewed as a synchronic account founded upon on a historical ground. According to Leisi, contemporary English "constitutes a unique mixture of Germanic and Romance elements and this mixing has resulted in the international character of the vocabulary" (after Lipka 2002: 14). As Chapter 8 of the present book reveals, this uniqueness of mixture of the aforementioned languages with respect to English can be deemed as natural in the original meaning of the word. Anticipating a little bit findings presented in Chapter 6, we may argue that this naturalness is anchored in the phenomenon of gravity and its manifestations in different aspects of the physical world.

The description of the process of foreign lexical assimilation (see the postulated model, sections 9.1.2 and 9.2.2) appears thus promising as a significant instrument in accounting for the *nature* of complexity of foreign lexical insertion. Furthermore, the focus laid in the book on the ante-ultimate stages of adaptation (see section 5.3, Fig. 14) is justified inasmuch as it helps to explicate reasons for the resulting constitution of foreign lexical stock in the English language.[14] In talking about "ante-ultimate" I mean concentration on stages preceding the full integration of a non-native word with the receiving language system. One such "ante-ultimate" stage might, for example, involve graphemic aspect of foreign word adaptation (see section 10.1). When Mańczak-Wohlfeld (1994: 150) discusses aspects of graphemic integration of foreign words into the Polish language, she provides the statistics according to which 50% of non-native lexical stock in the Polish language is subject to resistance[15] to spelling adaptation. It may, therefore, be claimed as intriguing to explore and investigate the nature of this 'insubordination' forces as potentially abiding in other languages than Polish. The rationale

[14] Mańczak-Wohlfeld (1994: 151) shows that as many as 94% of borrowings in Polish are classified as nouns, which is in accordance, as the author claims, with the common tendency of borrowing nouns into the system of any receiving language. The second place is held by verbs (3%), then adjectives (1%) and adverbs (1%). A preliminary analysis of 3385 foreign words and phrases discussed in detail in Chapter 7 bears out Mańczak-Wohlfeld's observations. About 95% of all lexical units considered can be classified as nominal. The privileged position of nominalization of lexical concepts cannot only be seen in the process of borrowing. A traditional definition of noun as grammatical category designating a thing, place or person (see, e.g. Burton-Roberts 1986) reflects the average speaker's conceptualizations processes as fundamentally reifying. Forerunning a bit our considerations presented in subsequent chapters, we may claim that this 'nominalization (reification) dominant' in the formation of lexical concepts is in perfect line with one of the central arguments laid down (see Chapter 6) that postulates the analogy between a meteor (a material entity) and foreign lexical unit (a material-like entity that we can hear and see) – the issue which, along with other correspondences between linguistic and physical world, is given special attention in sections 6.9 and 9.1.

[15] See Weinreich (1964: 64–65 in Otwinowska-Kasztelanic 2000: 21) for his discussion of resistance factors in the processes of borrowing. Weinreich attributes these to the relative stability of the system of the recipient language as well as its social value (prestige). Other factors include purist attitudes (see also sections 3.1 and 4.3) of speakers of the recipient language and the resulting native language awareness.

behind this reasoning is that Polish as well as other languages are called *natural*. As such, then, they are argued to be subject to analogous forces affecting their lexicons. The present book attempts to account for this 'universalist' insubordination forces from the perspective of English as a receiver language. The issue appears to be a serious one given rough estimates of total vocabulary stock of the English language according to which ca. 60% of the English lexical repertoire is comprised of borrowings with the remaining 40% pertinent to native origins (see Rayevska 1979: 215).

In some more comprehensive diachronic studies over the contribution of foreign lexical elements into the system of the English language, we observe tendencies that will largely be corroborated in Chapter 8 of the present book. It follows from these observations that English is dominated for the influx of non-native lexical units by a stable grouping of several languages with French incessantly occupying the top position. The reasons for this stability are accounted for by the descriptive framework developed in Chapter 6 and tested on the basis of the presented data in Chapter 8. The phenomenon of this so far undisturbed status quo is also explicated in Chapter 10, where the phenomenon of scale-free networks, as originally investigated within the field of Internet technology, is extrapolated onto the discussion of network of mutually interacting languages. The word "rather" should be emphasised in this place; this reservation is necessary on account of the nature of data studied (see, e.g. sections 0.1, 2.2, 4.1), in which the role of, for instance, Latin and Greek is largely diminished.[16] It is to be reminded that the focus on the present research is on the study of initial and central stages of foreign word adaptation, which may entail quite a high degree of fluctuation in data description, hence the reservations made at the outset to the book, where it is emphasized that the data analyzed are confined to a particular portion of foreign vocabulary[17] recorded within the definite time span, i.e. 1994–2004. All

[16] Cf. Rayevska (1979: 216): "As a matter of fact, three languages have contributed such extensive shares to the English word-stock as to deserve particular attention. These are Greek, Latin and French." However, her observation at that point does not account for the differences in the degree of 'foreignness' (see Ft. 13) of non-native elements. She studies Scandinavian, Latin, Greek, French, Italian, Spanish, Portuguese, German, Dutch, Hungarian, Polish, Chinese, Japanese, Hebrew, and Arabic, Turkish, Russian loan words (Rayevska 1979: 223–238). Most of them are fully acclimatized forms that do not pertain to the criterion of 'foreignness' adopted in the present book and are therefore essentially out of scope of the research.

[17] In a different place, Rayevska does, however, notice the lack of homogeneity in the category of "foreignness" but confines herself to lapidary comments without indicating any larger methodological consequences behind the recognition of the inherent fuzziness: "There are different degrees of 'foreignness'" (Marchand 1960, after Rayevska 1979: 217). As Rayevska (1979: 217) further argues: "Words may appear as complete aliens borrowed from a foreign language without any change of the foreign sound and spelling. These are immediately recognized as foreign words. They retain their sound-form, graphic peculiarities and grammatical characteristics and seem not to have broken their ties with the parent language completely […]. Certain foreign words are not felt to be aliens. They are completely or partially assimilated with already

in all, setting the data-collection reservations and other temporal limitations aside, we have to agree with a general observation by Katamba (2006: 143) who notices that "English has an extremely rich and varied vocabulary because it has enthusiastically borrowed foreign words in very large numbers. Each century the number of words adopted from foreign languages has increased. As we saw earlier, the main source of imports over the centuries has been French" (Katamba 2006: 143). The empirical confirmation of the aforementioned intuitions is given by Bliss (1966: 26):

Tab. 2. *The quantitative contribution of non-native element in the lexical system of the English language over centuries* (after Bliss 1966: 26)

	Medieval	16c	17c	18c	19c	20c
French	19	42	166	316	736	1103
Classical	89	237	371	173	328	250
Italian	–	26	48	100	90	153
German	–	2	2	4	58	240
Spanish	–	13	14	14	47	32
Other European	4	10	13	22	49	53
Non-European	2	12	56	35	97	55
Total	114	342	670	664	1405	1886

Elsewhere, Bliss (1966: 27) presents figures showing the percentage of the total borrowing during each period, which is made up of words from each of these different languages:

Tab. 3. *The percentage of contribution of non-native element in the lexical system of the English language over centuries* (after Bliss 1966: 27)

	Medieval	16c	17c	18c	19c	20c
French	16.7	12.2	24.8	47.7	52.3	58.6
Classical	78	69.4	55.3	25.9	23.4	13.2
Italian	–	7.6	7.2	15.1	6.4	8.1
German	–	0.6	0.3	0.6	4.1	12.7
Spanish	–	3.8	2.1	2.1	3.4	1.7
Other European	3.5	2.9	1.9	3.3	3.5	2.8
Non-European	1.8	3.5	8.4	5.3	6.9	2.9

existing native words and sometimes become indistinguishable from the native element." See also Chapter 5 on the categorization of foreign lexical elements and the resulting implication of the suggested solutions for the oncoming study.

As Rayevska (1979: 216) observes:

> A study of loan-words is not only of etymological interest. Words give us valuable information as to the life of the nations concerned. The linguistic evidence drawn from such an observation is a very important supplement to our knowledge. Loan words have justly been called the milestones of philology.

A detailed investigation of this aspect of cultural import of foreign words and phrases is discussed in Chapter 7 when the semantic-field analysis of lexical units in question is attempted.

0.4. Organization of the book

This book is divided into 4 major parts: Part I (Introduction and Chapters 1–2) places the study of foreign lexical assimilation within the field of lexicology and looks at its major concerns:

- external influences of the donor language upon the receiving one,
- reflecting upon the 'meaning' of a word,
- attempts at modeling the structure of the lexicon, etc.

Part I also provides an outline history of lexical semantics along with its major methodological proposals as well as the latest insights from Cognitive and Corpus Linguistics. We will look at, for example, the notions of frequency or prototypicality, the interplay between which, will turn out to be crucial in the argumentation advanced in the subsequent chapters.

Part II (Chapters 3–5) offers a more detailed, compared to relatively general Part I, survey of issues related to foreign lexical assimilation. Part II is quite heterogeneous but the common denominator of Chapters 3–5 relates to both external and internal aspects of research on foreign lexis. Part II starts with the examination of external aspects and opens up with a short Chapter 3 that discusses what has so far been said, asserted or feared about foreign words and phrases by native speakers of English. Chapter 4 gives an outline history of lexicographic research upon foreign lexis. It looks at most significant dictionaries of foreign words and phrases in search of theoretical digressions concerning the status of foreign words and phrases in language, their usage, and categorization criteria. The aspect of investigation in Chapter 4 is both external (the discussion of status or usage) and internal (the discussion of categorization of foreign lexis). Chapter 5 mainly focuses on the internal, i.e. categorization aspect of research on foreign words and phrases. It attempts at identifying sources of methodological disorder observed in the literature when it comes to the terminology used in the context of foreign lexical assimilation. First, it endeavours to compare and contrast borrowings

with foreign words and phrases and, then, outlines some major trends in the typologization of borrowings, following which the placement of foreign words and phrases within the proposed typology of non-native lexis is suggested. This becomes especially significant for the discussion of the emerging model of foreign lexical assimilation in Chapter 9. Chapter 5 closes with some discussion upon the categorial differences between common words and proper names – the issue already hinted at in section 0.2.

Part III (Chapters 6–7) constitutes a theoretical (Chapter 6) as well as descriptive (Chapter 7) foundation of the book. Chapter 6 elaborates on the claim stated in section 0.2 (see also Preface) and examines in detail a set of multiple correspondences between the physical world and the language world. In consequence, relevant notions and tools of analysis are worked out such as, for example, the concept of force or CRAC (Cumulative Relative Average Count, see also Kuźniak [2008]) – the term which forms, along with the concept of force, the basis for the metaphoric astrophysical model of lexical assimilation processes. The exceptional status of CRAC is that this postulated linguistic term, encapsulating both the information about the number of occurrences of a lemma in the BNC as as well as the number of texts in which it is found, perfectly corresponds to such astrophysical concepts as mass/volume of celestial bodies, especially if these are seen from the perspective of gravity force acting upon these bodies in the universe. This allows us to integrate the findings about linguistic frequency, force, entity mass and/or volume into one coherent model of description (see particularly Chapters 8 and 9). Chapter 7 is a detailed investigation of donor languages (referred to as 'planets' or 'planetoids') and their lexical import into the structure of English lexicon. The investigation involves a semantic characterisation of donor lexical systems, calculating CRAC figures at various planes of their analyses, all of which eventually enables us to draw conclusions concerning the current assimilatory status of a given foreign lexical unit (or at a higher level – a source lexical system) in the English lexicon.

Part IV (Chapters 8–9) offers the integration of theoretical as well as empirical findings from Part III. Chapter 8 gives us a global perspective upon the material presented in Chapter 7. It offers a characterisation of tendencies in the semantic specification of foreign lexical units across the donor languages listed. The characterisation is, for the sake of clarity, conducted in accordance with the criteria of membership formulated both in Chapter 5 (proper names vs. common words), and Chapter 6 where the differentiation of source lexical systems has been carried out by way of analogy to the specificity of the solar system in which planets and planetoids correspond to major and minor language donors, respectively. The chapter closes with the delivery of some compact view of source lexical systems and their lexical input with regard to the English language. The discussion terminates with the elicitation of the prototypical set of foreign words and phrases in use in English at the turn of a new millennium. Chapter 9 is a 'des-

tination point' in the process of investigation whereby the final integrated model of foreign lexical assimilation is proposed. This is based on the findings discussed throughout the book.

Last but not least, in Chapter 10, the concluding part of the research,[18] a brief summary of the most significant observations is presented and some further prospects in the interdisciplinary research upon foreign lexis suggested. These concern the integration of the model proposed in the book with other theories such as the model of atom (Rutherford-Bohr) or recent developments in IT science (the notion of scale-free networks).

[18] The author of this book has deliberately decided to omit the inclusion of separate concluding sections following each part of the book. This is because Parts 1–4 are viewed as interrelated (an elaborate sytem of intratextual references); therefore, providing separate conclusions to each part might produce an undesirable effect of its self-containedness with regard to other components of the book. Still, however, lacking conclusions following particular parts of the book are, hopefully, compensated by the addition of a separate chapter (Chapter 10) which summarizes the discussion presented in parts 1–4 (section 10.1), and also offers general conclusions emerging from the discussion (section 10.1.1). Moreover, the synoptical role of the missing concluding sections is also assigned in each part to special Overview sections.

PART I
THEORETICAL AND RESEARCH PERSPECTIVES

Overview

Part I of the book offers a survey of major theoretical and methodological contributions to the field of lexicology as a discipline of language study. Chapter 1 discusses the status of lexicology in linguistics and investigates *raison d'être* of its research objectives. Firstly, we have a look at the history of lexical semantics, and then examine concisely the most fundamental object of lexicological research, i.e. "word" and discuss some methodological controversies involved in the systematic description of the lexis of a natural language. Chapter 2 offers more specific theoretical and methodological perspectives from which lexical studies can be conducted, i.e. corpus linguistics and cognitive linguistics. These will be taken as the basis for the descriptive and analytical framework adopted in the present research on foreign words and phrases.

Foreign words and phrases in English, 2009

Chapter 1
LEXICOLOGY AS DISCIPLINE OF LINGUISTICS

1.0. Lexicology as a branch of linguistics

If this book were to be classified within the larger research framework, given the research goals and scope outlined in Introduction, this is, without doubt, lexicology. It is interesting to note that although linguistics as an autonomous scientific discipline has begun to develop since the second half of the 19th century (Fisiak 1985a), lexicology started to be perceived as an integral part of language research as late as at the turn of 1950s.[19] That this period was a hard time for lexicology is confirmed by the 'mainstream' philosophies of structuralism and early generativism reigning at the time. It is sufficient to quote Bloomfield's view of the lexicon as "really an appendix of the grammar, a list of basic irregularities" (Bloomfield 1933: 274) or Chomskyean (1965) "syntactocentric" (Jackendoff 2002) approach to language, which resulted in the lack of preoccupation with the nature of words and their meanings. If lexicon was to be mentioned at all, its discussion was confined to the investigation of "rules for inserting lexical items as terminal nodes into syntactic structures ('phrase markers')" (Hanks 2008: 1).

In view of the above observations, Altenberg & Granger (2002: XV–XVI) make the following remark:

> As early as 1951, Stephen Ullmann, in the initial chapter of his *Principles of semantics* (1957: 39) entitled "What is semantics?" proposed a three-dimensional model of structural linguistics. The three dimensions are: 1. Phonology, 2. Lexicology, 3. Syntax, with further subdivisions. All three are said to be fundamental.

The term lexicology here encompassed what might be nowadays subsumed under morphological, semantic, and pragmatic studies. If lexicology was to evolve as a separate section of linguistics rather than denoting a subsystem of language, its goals and scope were to be more specifically defined. But getting to know the details first entails the knowledge of the outline research framework. As Hanks (2008: 1) puts it: "Lexicology is the department of linguistics concerned with lexicon – the study of words and their meanings." Similar is Jackson & Zé Amvela's

[19] As Hanks (2008: 1) argues: "Words are the most obvious manifestation of language as a human artifact, and one might have expected that the study of words would have played a central part in the linguistics throughout its history. It is therefore somewhat surprising to find that lexicology and word meaning were until recently neglected in many, though not all, schools of linguistics – but such was the case. The central concerns for much of the twentieth century of the so-called 'mainstream' in linguistics (a term which generally denotes American linguistics) were phonology, morphology, and above all syntax."

stance (2007: 2), according to which lexicological enterprise is viewed as "the study of lexis, understood as the stock of words in a given language, i.e. its vocabulary or lexicon (from Greek *lexis,* 'word,' *lexikos,* 'of/for words').[20] This working definition shows that notion of 'word' is central to the study of lexicology." When we look at other definitions of lexicology, we immediately observe that there is not any striking difference between them as far as defining the goals and scope of the discipline. The only expected differences lie in the phrasing of the aforementioned research perspectives. A selection of sources from Jackson & Zé Amvela (2007: 3) listed below (except 5) appears to confirm this observation:

1. An area of language study concerned with the nature, meaning, history and use of words and word elements and often also with the critical description of lexicography (McArthur [ed.] 1992).
2. The study of the overall structure and history of the vocabulary of a language (Collins English Dictionary [1998]).
3. A branch of linguistics concerned with the meaning and use of words (Longman Dictionary of the English Language [1991]).
4. The study of the form, meaning, and behavior of words (New Oxford Dictionary of English [1998]).
5. The study of the meaning and uses of words (LDOCE [1995]. In: Altenberg & Granger [2002: XVI]).

In contrast with the above-discussed 'mainstream' pre-programmed abandonment of the study of words underpinning both structuralist and generativist linguistics in the 20th century, we may, however, observe radically different proposals (Mel'čuk 2006) in which lexicological research is viewed as central rather than marginal to linguistic studies:

> Most current theories view a linguistic description of a language as a grammar; a lexicon is taken to be an indispensable but somehow less interesting annex to this grammar, where all the idiosyncracies and irregularities that cannot be successfully covered by the grammar are stored. By contrast, Meaning-Text Theory considers the lexicon as the central, pivotal component of a linguistic description; the grammar is no more than a set of generalizations over the lexicon, secondary to it.
>
> (Mel'čuk 2006: 228, after Hanks 2008: 3)[21]

However, as Altenberg & Granger (2002: 9) continue: "Surprisingly enough, the term *lexicology* was not to be found in 1990 in most medium-sized dictionaries and in various handbooks (e.g. Lyons 1977), nor in English grammars (with

[20] Cf. also the following definition: "Lexicology (from *lexiko-*, in late Greek lexicon) is the part of linguistics, a science which is dealing with the study of words, the relations between words (i.e. the semantic relations), and the whole lexicon" (Internet source 1, see References).

[21] Other eminent representatives of a "pro-lexicon" trend in linguistic research include Wierzbicka, Sinclair, Halliday, Wray or Bolinger (see Hanks 2008: 3).

the exceptions of LDCE, 1985)." One possible reason might be the gradually increasing role of lexicographic research with the advent of the widespread use of computers that started in the early 1990s. This might lead some researchers to use the term lexicography as a cover category to both strictly lexicographic and lexicological research. The other reason might be purely terminological, i.e. the fact that the term *lexicology* was not in widespread use in the English speaking scholarly centers, and the term preferred was *lexical semantics* (see section 1.1).

However, despite discernible similarities between the disciplines of lexicology and lexicography, they can still be distinguished with regard to the end-product of a scholarly enterprise in that lexicography is "the practical business of dictionary making" (Hanks 2008: 1), whereas lexicology is the more directed at modeling the structure of the lexicon with the concurrent study of words, investigating their nature, exploring lexical relations internal and external to the lexicon, addressing the concept of 'meaning,' etc.[22] As Hanks (2008: 3) says:

> For an understanding of how words work – their meanings and implications – it is necessary to turn away from American so-called 'mainstream' linguistic theory of the 20th century to other disciplines – the philosophy of language, cognitive science, anthropology – and to linguistics in other cultures.[23]

In discussing more detailed research perspectives offered by lexicology, Rayevska (1979: 10) notes that there are at least three such areas. One such area relates to the study of word origins, i.e. etymology, and the other two dimensions involve the study within a synchronic and diachronic angle, i.e. semasiology and onomasiology. Semasiology can thus be defined as "the branch of linguistics whose subject-matter is the study of word meaning and the classification of changes in the signification of words or forms, viewed as normal and vital factors of linguistic development" (Rayevska 1979: 10), whereas onomasiology is viewed as "the study of the principles and regularities of the signification of things and notions by lexical and lexico-phraseological means of a given language" (Rayevska 1979: 10). The present research will address both semasiological and onomasiological axes of lexicological investigation. This is the subject-matter of Chapter 7, where criteria of classification for the analysis of domains of foreign words and phrases will be encapsulated within the onomasiological domain, whereas the discussion upon

[22] "Lexicography is the theory and practice of composing dictionaries […]. It is said that lexicography is the practical lexicology, it is practically oriented though it has its own theory, while the pure lexicology is mainly theoretical" (Internet source 1, see References). For some other significant literature on lexicography and the related research problems, see Piotrowski (1986, 1994, 2001).

[23] Hanks (2008: 3) mentions here European structuralism (followers of de Saussure, see also Fisiak 1985a), especially meaning-text theory as developed by Mel'čuk, Maurice Gross's lexicon grammar in French or the followers of the London school of linguistics (in particular John Sinclair). Significant are also studies by such philosophers as Wittgenstein, Quine, Rosch and Brown.

the ambiguity of foreign lexical units[24] in the same chapter will be subsumed under the semasiological aspect of analysis.[25]

According to Hanks (2008: 12), there are at least ten major questions that may be asked of lexicology:

1. How do words relate to the world around us?
2. How do words relate to one another (within the system of a language)?
3. How is the lexicon of a language structured (in so far as it is structured at all)?
4. What is the relationship between words and concepts?
5. How can word meaning be formalized and so be made machine-tractable?
6. How do young children acquire the words and meanings of their native language?
7. How do the words of one language relate to words with "the same meaning in other languages?"
8. How do the meanings of words change over time?
9. What is the difference between the literal meaning of words and figurative meanings such as metaphors?
10. What is the nature of lexical creativity?

It appears that the present book will primarily address questions 1–3 (Chapters 6–9) and the question 10, although the perspective taken will be, as already stated in Introduction, exclusively external, i.e. we will be preoccupied with the phenomenon of foreign lexical assimilation in the English language as only one aspect of the complex process of lexical creativity and development in language system.

In conclusion, we may refer to Dustoor (1968: vii), who views the study of vocabulary of (a) language as necessarily encroaching upon the historical ground both in a narrow linguistic sense of studying how words changed in form/meaning and in broader socio-historical one, in which words are viewed as mirroring history of mankind. Lexicological research may also adopt another perspective in which words are studied from the viewpoint of "how they are introduced – singly and in phrasal combinations – into a language; how they are adapted, consciously and unconsciously, to the different needs of the different peoples who come to use them" (Dustoor 1968: vii). It is intriguing that the problem of external sources of lexical development is treated in a marginal way in Hanks (2008, Vols 1–8). The entire collection of volumes entitled *Lexicology* treats about philosophy and word meaning, theories in the studies of lexis (semantic field theory,[26] structuralist

[24] Researchers that advocate the monosemic bias (see Geeraerts 2006c: 150) in lexical semantic studies include *inter alia* Charles Ruhl (1989), and Anna Wierzbicka (1991), and Claude Vandeloise (1990). As Grice (1989: 47 in Geeraerts 2006c: 150) claims: "The monosemic bias is related to Grice's 'Modified Occam's Razor': the principle that senses are not to be multiplied beyond necessity."

[25] As Rayevska (1979: 11) maintains: "Most perceptive and useful treatments of theory and method in lexicology with much insight on the subject in the light of modern linguistics will be found in linguistic research made in 50–70-ies by O.S. Akhmanova, V.N. Yartseva, A.A. Ufimtzeva, I. Arnold, N.N. Amosova, Y.A. Zhluktenko, R.S. Ginzburg, K.T. Barantsev."

[26] The theory of semantic fields to which Chapter 7 of the present book mostly refers in the context of the analysis of vocabulary structure of particular donor languages, was pioneered by

semantics, componential analysis), prime and universals, extended meanings and polysemy, the relation between cognition and the lexicon, computational lexical studies. There is thus no explicit section in the series of publications devoted to the problem of lexical borrowing or generally the processes of lexical adaptation. I believe that this book, which definitely should be embedded within the tradition of lexicological research, offers some contribution to the field, especially in view of the aforementioned marginal treatment of the problem by Hanks (2008).

Coming back to Dustoor's considerations, we can identify a narrow linguistic profile of research in which we are concerned with the systematic description of foreign lexical assimilation from donor language into a receiving one, or we may look upon the issue from the broader perspective offered by socio-pragmatics. The latter, however, has not been as yet well researched upon for reasons specified in Görlach (2007: 10–11). This book, however, seeks to combine a lexicological descriptive profile of research (Chapter 7) in which a semantic characterization of foreign words and phrases in English is provided along with more speculative socio-cognitive account in which an attempt at modeling lexical assimilation processes and in this way human understanding and involvement in these processes is made (Chapters 6, 8, 9).

1.1. History of lexical semantics

In the previous section, we discussed lexicology as a branch of linguistics with an outline of its major theoretical and research perspectives. We mentioned, *inter alia,* that lexicological research is known better under the term *lexical semantics* in the English-speaking Western scholarly tradition. This section will be devoted to surveying major theoretical contributions underlying this tradition. A survey on the history of lexical semantics can be found in Geeraerts (2006c: 367–397). Dominant theories in the field of lexical semantics can be presented as follows:

J. Trier. His concept of 'linguistic fields' "is known to be based on de Saussure's theory of language as a synchronous system of networks held together by differences, oppositions and distinctive values" (Rayevska 1979: 30). As Lipka (2002: 19) claims, "word-fields" (cf. Coseriu's [1970: 166] "wortfeld") are primary paradigmatic structures consisting of lexical items that share a common zone of meaning and are in direct opposition to each other." More on lexical fields and hierarchies can be found in Lipka (2002: 167–172). On these pages you can encounter synonymous terms for "lexical field": for example, Cruse (1986: 112 ff) speaks of lexical configurations, while for Lehrer (1974) the terms "semantic field" and "lexical field" are synonyms. For Coseriu (in Coseriu & Geckeler 1981: 58 ff), in turn, all lexical fields are necessarily conceptual fields, but the opposite does not need to be true. Another interesting extension of 'lexical field' theory (e.g. the concept of archilexeme) can be found in Faber and Usón's (1997) Functional-Lexematic Model; however, a detailed account of its findings goes beyond the scope of the present publication.

Tab. 4. *Major trends in lexical-semantic research in chronological perspective*[27] (after Geeraerts 2006c: 387)

Main Period	Theoretical framework	Observational Principle	Empirical domain	Explanatory principle	Examples of major works[28]
1870–1930	Historical-philological semantics	Language as the self--expression of an individual or a community	Statements about meaning change	Principles of efficiency and expressivity	Paul (1880); Bréal (1897), Wundt (1900); Sperber (1923); Stern (1931)
1930–1975	Structuralist semantics	Language as an autonomous structure	Statements about paradigmatic and syntagamatic semantic relations	Meaning decomposition, or meaning postulates	Katz and Fodor (1963); Katz and Nagel (1974); Fodor, Fodor and Garrett (1975); Pottier (1964)
1970–	Logical semantics	Language as reference to possible worlds	Statements about the truth conditions of propositions	Algorithmical compositionality	Monatgovian semantics
1975–	Cognitive semantics[29]	Language as a cognitive tool	Statements about the flexibility, internal structure, experiential nature, encyclopedic character of lexical concepts	Natural categorization (prototypicality) and cognitive strategies	Fillmore (1977); Rosch (1977); Lakoff and Johnson (1980); Langacker (1983); Talmy (1985); Sweetser (1984); Geeraerts (1985)

This book in this methodological orientation relies much on cognitive-linguistic tradition of research. The reasons are twofold, both fundamental to the central claim stated in Introduction (section 0.2, see also Preface). Both are also provided by Evans and Green (2006) and are as follows:

[27] A brief outline of lexicological research in India, China, the Islamic world, and Europe as well as evolution of the dictionary in England and thesaurus plus recent developments in lexicology can be found in (Halliday *et al.* 2005: 11–22).

[28] This column has not been originally included in the table and has been added by the author of the present book.

[29] "Among the major conceptual contributions of Cognitive Semantics to the study of lexical meaning, the following should be mentioned:
- Rosch's (1977) prototypical theory of categorial structure
- Lakoff and Johnson's (1980) conceptual metaphors
- Fillmore's (1977) *scenes-and-frames* approach
- Berlin's (1978) *basic level* hypothesis" (Geeraerts 2006c: 405–406).

1) The Cognitive Commitment: "A commitment to providing a characterisation of general principles for language that accords with what is known about the mind and brain from other disciplines" (Evans and Green 2006: 27–28).

2) The Embodied Mind Thesis:

> Cognitive linguistics is not rationalist in this sense, but instead takes its inspiration from traditions in psychology and philosophy that emphasise the importance of human experience, the centrality of the human body, and human-specific cognitive structure and organisation, all of which affect the nature of our experience. According to this empiricist view, the human mind – and therefore language – cannot be investigated in isolation from human embodiment.
>
> (Evans and Green 2006: 44)

The Cognitive Commitment legitimizes a linguist working in a cognitive-semantic tradition to search for the insights from other disciplines if these are to offer tools that enable a linguist to investigate the nature of language. The other commitment, known better under the embodied mind thesis, compels a researcher to account for bodily experiences of language conceptualizers as these may directly contribute to our better understanding of how our mind works. This, in turn, may offer some interesting insights into the complexity of language as language itself is viewed to be an integral part of human cognitive apparatus.

These two fundamental tenets of Cognitive Linguistics constitute an underlying methodological legitimization for tackling the issue of foreign lexical items in the English language – the subject-matter of the present book. The Cognitive Commitment substantiates our research on modeling foreign lexical assimilation processes through a recourse to the science of physics (Chapter 6), whereas the Embodied Mind Thesis appears to be a confirmation of our methodological 'adventure' into the realm of force, mass, volume, motion or energy as these are directly relevant to the every-day experience of a human being. Summing up, the choice of physics as the framework for our linguistic investigations (see especially Chapters 6 and 9) appears as a natural selection that stems from the adoption of the aforementioned principles. It is therefore surprising to find out that cognitive linguists have so far appeared rather reluctant to reach for the insights from physics (except Talmy, Johnson, or Krzeszowski, [see section 6.8]), and instead concentrated on the findings offered by psychology, artificial intelligence, or cultural anthropology (see Geeraerts 2006c below).

Other significant tenets of Cognitive Semantics underlying the descriptive and theoretic part of the book (Chapters 6–10) include the following (after Geeraerts 2006c: 374–378):

> 1) Lexical concepts have vague boundaries, in the sense that they contain peripheral zones round clear conceptual centers (versus:[30] lexical concepts are discrete, well-defined entities).

[30] In brackets you find the opposite view on the nature of semantic enterprise as advocated by structuralist, generativist, and logical strands of research.

2) Lexical concepts are polysemous clusters of overlapping semantic nuances (versus: the various senses of a lexical item can always be strictly separated from each other).
3) Attributes within (or exemplars) of a category may have different degrees of salience (versus: all attributes within the definition, or exemplars in the extension of a concept have an equal degree of salience).
4) Lexical concepts function in a flexible and analogical manner (versus: lexical concepts function in a rigid, algorithmical fashion).
5) Lexical concepts have to be studied as a proper part of human cognition at large (versus: lexical concepts have to be studied as a part of an autonomous linguistic structure).
6) There is no distinction between semantic and encyclopedic knowledge (versus: the semantic definition of a lexical concept is to be distinguished from the encyclopedic data that can be connected with the latter.
7) Semantic studies cannot ignore the experiential and cultural background of the language user (versus: semantic phenomena should be studied apart from user- or culture-specific background data).

The above statements are relevant to the central claim of the research stated in section 0.2 in that statement 1) and 2) takes for granted the vagueness of boundaries of foreign lexical units subject to assimilation into the target system. This non-discreteness of substance out of which a lexical unit is composed is clearly visible in Chapters 7 and 8 of the present book. In Chapter 7 this fuzziness is manifested in the problems of establishing classificatory criteria according to which the meaning (i.e. the substance) of particular foreign words and phrases should be characterized. In Chapter 8 the same problem occurs when an attempt is made at the identification of tendencies in 'mass structuring' (i.e. semantics) of a given planet or planetoid (i.e. a specific donor language). In addition, the implication behind the word "tendency" is in line with statement 3) that assumes the degree of salience of elements structuring the conceptual make-up of a word. Thus, certain elements constituting a lexical concept will be more significant (dominant, [see Chapter 8]) than others. Statement 4) about the analogical manner in which lexical concepts function is in line with the discussion held in Chapter 8 where a number of regularities are discovered with regard to the way foreign lexical items operate in the target language system. The essence of statement 5) relates to the need for interdisciplinary research in linguistics.

This book offers a positive response to this challenge insofar as the science to which the present research refers is elements of astrophysics. The implications obtained from statement 6) have led us to rely on folk-model-oriented encyclopedic projects as a source of information about meanings of foreign words and phrases rather than standard dictionaries (see section 8.4.1). Last but not least, statement 7) along with its message of validating naïve models of the world in semantic studies constitutes a dominant theme in Chapter 9 where both expert and non-expert models of the conception of the universe are seen to be equally significant as illustrating external aspects of foreign lexical assimilation processes (see especially section 9.2.1).

It is important to remark that this cognitive-semantic tradition of lexicological research does not stand in direct opposition to the so-called historical-philological tradition (the foundation for structuralism in the 20th century linguistics, literature and philosophy) in which case the most outstanding similarities can be identified. These are: psychological orientation, significance of experiential factors as well as polysemy (Geeraerts 2006c: 378–385; see also Tabs. 5 and 6 below):

Tab. 5. *A conceptual map of lexical semantics* (after Geeraerts 2006c: 411)

	Semasiology	Onomasiology
'QUALITY': Investigating structure (elements and relations)	Senses and semantic links among senses (metaphor, metonymy, etc.)	Semantic relations among lexical items (fields, taxonomies, networks, etc.)
'QUANTITY': Investigating use and differential weights within structures	Prototypicality as involving salience phenomena, core versus periphery	Differences in cognitive salience between categories, basic level phenomena and entrenchment

Tab. 6. *The contribution of the traditions to the development of the field* (after Geeraerts 2006c: 414)

	Semasiology	Onomasiology
'QUALITY': Investigating structure (elements and relations)	Prestructuralist semantics: mechanisms of semantic change.[31] Neogenerativist semantics: regular polysemy	Structuralist semantics: taxonomies and lexical relations; lexical fields; syntagmatic relations and collocations. Cognitive semantics: conceptual metaphors and metonymies; frames
'QUANTITY': Investigating use and differential weights within structures	Cognitive semantics: prototype theory	Cognitive semantics: basic level and entrenchment

Tables 5 and 6 illustrate what can be postulated as highly relevant to the framework of the research adopted later in the book. As already hinted at in Introduction, semasiological and onomasiological levels of the analysis permeate the discussion upon foreign words and phrases starting from Chapter 5 onwards. Semasiological perspective is present whenever the analysis of polysemy of the discussed words and phrases is undertaken or whenever prototypicality issues are brought to light (see, e.g. sections 8.4.1 and 8.5). Onomasiological perspective is a prevalent one in

[31] It is important to quote in this place Lewandowska-Tomaszczyk (1985, 1991a, 1991b) on modeling semantic change in language. See also Adamska-Sałaciak (1996) on the conception of language change in the 19th century linguistics (the Kazań School) as well as the recent publication on regularity in semantic change by Traugott and Dasher (2002).

Chapter 7 when the identification of semantic domains of the discussed foreign words and phrases is conducted. This, in turn, is inextricably related to the methodological problem whereby certain domains are regarded as more salient than others, which, as it is argued, may lead to the impression of some conceptual disorder in that not all domains identified may be considered at the same level of specificity (see section 7.0 for the discussion and possible solutions to that problem).

Also the "quality" and "quantity" parameters are viable for the description of foreign words and phrases. "Quality" factor underlies all the discovered analogies between a linguistic and physical world spelled out in Chapter 6, as anchored in LANGUAGE LAWS ARE PHYSICAL LAWS overriding metaphor. These are elaborated on in Chapters 8 and 9. The "quantity" factor, on the other hand, is represented at the level of the statistics provided by the British National Corpus in which data about frequency of occurrence of a lexical unit in the entire corpus is displayed. This becomes extremely significant for the development of the notion of CRAC – a technical linguistic correlate of the notion of mass of an entity, the discussion of which is presented in section 6.1.

1.2. The word as the basic object of lexicology

Although linguistics as a discipline aims at a systematic description of natural language, it appears that the notion of 'word' should be one of the most clearly defined concepts as 'words' constitute in the average speaker's mind fundamental 'building blocks' of a communication process.[32] Paradoxically, however, there are endless heated debates over the definition of 'word,'[33] as these are provoked by different theoretical perspectives taken by scholars. The milestones in the conceptions upon 'word' are given by 20th century structuralist linguists, such as Bloomfield (1933), Sapir (1921), Marchand (1960). Thus Bloomfield defines 'word' as "the minimal independent unit of utterance;" Sapir treats it as "one of the smallest completely satisfying bits of isolated 'meanings' into which the sentence resolves itself," whereas Marchand views 'word' as "one of the smallest completely satisfying bits of isolated 'meanings' into which the sentence resolves itself" (after Rayevska 1979: 34–35).

An interesting survey of different positions taken by linguists on the definition of "word" can be found in (Dixon & Aikhenvald 2007). The authors conclude:

[32] As Labov (1973: 340, after Aitchison 1987: 25) argues: "Linguists like any other speakers of a language cannot help focusing their attention on the word, which is the most central element in the social system of communication."

[33] The magic properties of words illustrated on the basis of different cultural communities are discussed in Bauer (1998: 1–5). See also divine characteristics assigned to the word as spelled out in St John: "In the beginning was the Word, and the Word was with God and the Word was God."

> We have found that although many types of definition have been suggested for 'word', there has often been lack of a clear distinction between lexeme and word form, and/or between phonological and grammatical criteria. We suggest that different sorts of criteria should be kept strictly apart – phonological criteria define phonological word, which is a unit in the phonological hierarchy, while grammatical criteria define grammatical word, which is a unit in the grammatical hierarchy. In some languages grammatical and phonological words coincide so that we have a single unit functioning in both hierarchies.
>
> (Dixon & Aikhenvald 2007: 34–35)

In the present book, most of the units listed in Glossary of foreign words and phrases should rather be classified as cases where grammatical criteria and phonological criteria coincide.[34] The so called "foreign phrases" will be understood as units, i.e. phonological words that consist of a number of grammatical words. This is because the phrases discussed are highly conventionalized chunks with fixed order and well-established meaning.[35]

With much help to the conception of word come insights from Polish linguists, especially Szymanek (1989), Mańczak-Wohlfeld (1994), or Grzegorczykowa *et al.* (1984). Szymanek differentiates between word-form (e.g. *look*, *looks*, *looked*, *looking*) and citation form (e.g. *look*) defined as the most representative instantiation of a lexeme, where "lexeme" is defined as abstract meaningful unit represented by one or more concrete units in a text (see also Grzegorczykowa *et al*. 1984: 16). The conception of a citation-form is, in turn, very closely related to the conception of "lemma" developed in corpus linguistics (see Chapter 2). According to Halliday *et al.* (2005: 6):

"The lemma is the basic form under which the word is entered and assigned its place: typically, the 'stem,' or simplest form." The same view on "lemma" can be noted in Biber *et al.* (1998: 29): "When studying a word, it is often useful to consider the different forms of the word collectively […]. The term 'lemma' is used to mean the base form of a word, disregarding grammatical changes such as tense and plurality."

Another recent contribution is offered by Harley (2006) who also extends her discussion of 'words' to cover constructions like idioms. Her findings are to

[34] This is the case of e.g. the English language. The following observation from Darbyshire (1967: 58–59) confirms that, at least as far as the English language is concerned, the linguistic story over 'word' does not necessarily be that complicated: "We said that morphemes were scientific fictions, and some of the difficulties of applying a consistent theory of morphemes to a language like English have led some linguists to think rather in terms of words. But here again difficulties of recognition and definition are likely to arise. Those of us whose native language is English may think that we can easily recognize a word – we are capable, for instance, of dissociating individual words from a sample of connected or continuous speech as we hear it, and when we come to write we have little difficulty in deciding where to put the spaces between words; and indeed, for a language like English the problem is not so difficult as some people have made out."

[35] See Mańczak-Wohlfeld's (1994: 7) extension of the term "lexeme" to cover certain types of phrases, and Kuźniak (2009a, in print) for the discussion upon word and phrase as units in the context of the analysis of English pleonasms.

a large extent compatible with what has been already proposed as terminological solutions in the literature (cf. Szymanek 1988; Mańczak-Wohlfeld 1994, Halliday *et al*. 2005). As Harley (2006: 11) puts it:

> There is an easy way out of this dilemma. One view, the meaning of 'word' has mainly to do with semantics – the part of the definition that refers to the 'minimal meaningful unit', that is an element of the list of sound-meaning correspondences that is one of the two fundamental elements of language. The other, more everyday interpretation of the meaning of 'word' has mainly to do with phonology: the fact that we call whatever we can pronounce in isolation a 'word.' The latter we have simply labeled: *phonological word*. […] The former, the true *minimal meaningful unit,* which includes affixes, like *-s* and *un--,* and idioms like *kick the bucket*, we will call *listemes*.

In a recent publication on lexicology the following simple definition is offered: "Since our main focus is on the definition of lexicology, and in order to avoid a lengthy digression, we use 'word' somewhat loosely in the usual traditional sense of a sequence of letters bounded by spaces" (Jackson & Zé Amvela 2007: 2).[36] As it turns out, the discussion over the concept of 'word' has come a long way to finally reach the point, which intuitively, should rather constitute a place of departure for more systematic scholarly treatises. In an attempt at bringing order to what might appear as a blurred picture of the concept of 'word,' Fig. 3 below presents in an outline the most significant terminological contributions to the definitions of 'word' discussed in linguistics. These can be illustrated in the following way:

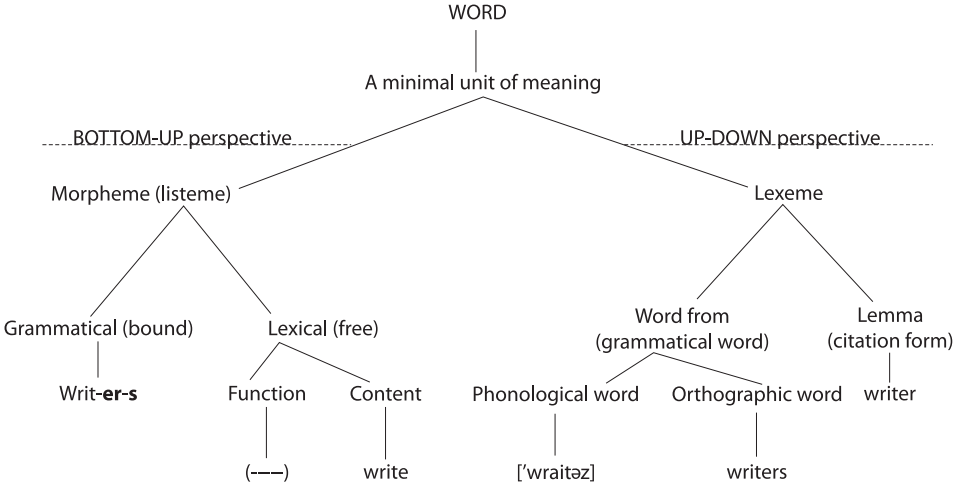


Fig. 3. *Terminological ordering in lexicological considerations upon 'word'*

[36] An interesting survey of the contemporary approaches to the study of units in a mental lexicon is presented in Mikołajczak-Matyja (2008: 13–25).

As we already alluded to above, the research upon foreign words and phrases in English has been conducted upon the basis of SARA software retrieval system, whereas individual 'words' searched for their statistics of occurrence are discussed in the literature as "lemmas."[37] Once a set of lexical units has been collected for the research, the next stage according to the traditional procedure involves "describing the various meanings of lexical items and the structural relations among those meanings" (Geeraerts 2006c: 75). The classificatory criteria, when it comes to the semantic characterization of a given word or phrase, are presented at some length in section 7.0. The overriding principle is the assumed polysemy or mononsemy of a particular lemma or phrase. This is based on the salience of meaning(s) represented for this item in *Wikipedia* (see section 7.0). Such assumption is possible when we accept prototype theory in lexical studies that "does incorporate aspects of use in the structural description of the lexical categories: the fact that some readings are more salient than others in the structure of the category reflects the fact that they are more readily chosen when using that category" (Geeraerts 2006c: 75). It seems that out of the two major types of salience discussed by Geeraerts (2006c: 76), i.e. *perspectival,* and *variational*, the one that we make direct reference to is the perspective based on profile/base or figure/ground oppositional pairs. Thus, if a particular lemma is correlated in *Wikipedia* with more than one meaning, it is referred to as polysemous. If, however, one dominant understanding prevails, it is treated as essentially monosemous, although this must be taken as a methodological oversimplification since perfect monosemy does not exist.[38] Nevertheless, what becomes of interest to our research is the average (a combination of naïve and expert) knowledge about concepts rather than purely

[37] The term "lemma" is compatible with the term "citation form" in that base, uninflected forms made up the ground for the research. "Lemma" can also be viewed as a terminological blend of "lexeme" and "grammatical word." Lexeme/grammatical word distinction is directly in line with Lyons (1968: 197 in Dixon & Aikhenvald 2007: 7) proposal, who defines lexeme as "root or underlying form" and grammatical word as "inflected form of a lexeme." The question that naturally arises is: why is the blended understanding of word adopted in the present book? The answer comes from the nature of the search process in SARA software of the British National Corpus. In search of frequency of occurrence data for particular foreign words, the term "word" is understood in Lyons' sense of "lexeme." What is put into the search bar of the SARA software is uninflected lexical form, which corresponds both to the definition of "lexeme" and "citation form." However, what we obtain as a result of the search is the number of both inflected (grammatical words) and uninflected forms of a given lexeme. As a consequence, the definition of "word" as conceived of in the book assumes a hybrid ontological status. On the one hand, the term "word" fulfils the definition of "lexeme" in the input stage of frequency data search, on the other the term "word" fulfils the definition of a grammatical word in the output stage of the process where frequency data are collected. For methodological convenience then, the term *lemma,* as used in Corpus Linguistics methodology, has been adopted to cover the two stages of word search process. In the light of the above considerations, the notion of "phrase" will thus be understood as a lexical unit composed of grammatical words characterized by conventionalized coherence and meaning.

[38] For the issue of polysemy against the notion of intuitive monosemy (primary/secondary activation), see Langacker (1987) and Geeraerts (2006c: 172).

expert one. The reasons for this are philosophical and derive from the experiential realism as advocated by cognitive linguists (see Lakoff and Johnson [1980, 1999]. See also sections 9.2.1 and 9.2.2 in the present book).

A tentative elicitation of salient meanings[39] of foreign words and phrases in this book can also be considered effective via the reference to semasiological/ onomasiological axes of analysis. As Baldinger (1980: 278) puts it: "Semasiology […] considers the isolated word and the way its meanings are manifested, while Onomasiology looks at the designations of a particular concept, that is, at a multiplicity of expressions which form a whole." In view of this distinction, our discussion upon polysemy of foreign words and phrases with the simultaneous search for prototypical (most salient) meanings of lexical items can be compared to the paradigmatic semasiological type of lexical salience, where, according to Geeraerts (2006c: 90), "the preponderant structural weight of specific senses or members within the semasiological range of application of a lexical category" is accentuated. Onomasiological aspect of the analysis is seen in section 7.0, where classificatory domains for the semantic characterization of foreign words and phrases are discussed. This is based on the so-called global categorical onomasiological salience perspective where "the preference for a specific lexical category as a designation for its range of application, taken as a whole" is taken into account (Geeraerts 2006c: 90).

1.2.1. On the 'meaning' of a word

This section gives an outline of some methodological contributions to the discussion upon the nature of the meaning of a word. The present section is not intended to offer an exhaustive account of the rich philological discussion upon the nature of the meaning of a word. Such comprehensive survey might well constitute the subject-matter of a separate monograph. Instead, in this part, we will focus on the most influential insights into the intriguing issue of word meaning, starting the survey with Aristotle's *Meaning and Essence* and concluding the survey with the recent proposals advanced in cognitive-linguistic research. Last but not least, we will refer to a specific problem of static vs. dynamic perception of word meaning. This will be relevant to the methodology adopted for the creation of the Glossary of foreign words and phrases appended to the present book.

1.2.1.1. Major insights into the study of word meaning. Survey

This section offers only a glimpse into the selected major proposals into the study of word meaning in lexicological perspective. The discussion is based on

[39] The verification of semantic salience could best be conducted through psychological experimental research, but this is beyond the scope of the present book.

the VI volume book entitled *Lexicology* edited by Patrick Hanks which provides the list of authentic texts from various scholars across the history of linguistic and philosophical thought. Directly relevant to the discussion upon the meaning of a word are philosophical reflections by Aristotle, especially those summed up in *Categoriae et Liber De Interpretatione*, in which he discusses his conception of a category explicated in terms of necessary and sufficient attributes which are argued to constitute the quintessence of its ontology. This understanding of the category, and the related conception of meaning as reducible to a well-defined set of elements is particularly discernible in structuralist accounts of word meaning (Pottier 1964; Porzig 2008 [1934]; Lyons 2008 [1969]), as well as the tradition of the so-called componential analyses of meaning (Goodenough 2008 [1956]; Lounsbury 2008 [1964]), where the meaning of a word is analysed in terms of a finite set of semantic features conceived of as binary (+/ –) oppositions.

Also de Saussure's structuralist conception of the levels at which a language operates into syntagmatic and paradigmatic has become the inspiration for the linguists whose analyses of word meaning focused more on discovering intrinsic relations between words rather than deliberating upon meanings of individual words. When it comes to syntagmatic relationships and its emanation in semantic studies, Hanks (2008: 14–16) discusses the Firthian tradition with its stress on collocational as well colligational nature of lexical organization (Halliday 1970; Sinclair 1991; Hoey 2005). On a more paradigmatic basis (see also Cruse 1986, 2000), semantic studies shifted their attention to the investigation of such relations as synonymy, antonymy, hyponymy, hyperonymy, meronymy, and metonymy. These were best epitomized in the monumental work *Structural Semantics* by Lyons published in 1963.

More formalized approach to the study of word (sentence) meaning can be noticed in logic-based approaches. The major foci are the investigation of the relationships between sentences or examination of the truth-conditional value of individual propositions. The former found its quintessence in the enterprise subsumed under the label of propositional logic or propositional/sentential calculus (see Alwood, Andersson & Dähl 1977; Palmer 1988: 177–207). The latter can be best epitomized by Palmer (1988: 195–196), who argues that propositional logic studies focus merely on the truth/falsity of their logical form, whereas the truth-conditional semantics is preoccupied with the examination of the truth/falsity of the meaning of sentences as such. These truth-conditional-based investigations are best illustrated by such works as Tarski (1936); Carnap (1948); Kempson (1977); Quine (1951; 1960).

Highly formalized, but more mentalist and mechanistic perspective upon the study of meaning is offered by the generative tradition (Chomsky 1957, 1965). The major focus of enterprise in this paradigm is directed at modeling the lexicon, but the approach is essentially "syntactocentric" (Jackendoff 2002b). As Hanks (2008: 1) insightfully puts it:

> The great American linguist Leonard regarded the lexicon of a language as no more than "really an appendix of the grammar, a list of basic irregularities" (1933: p. 274). The main concern of the most influential American linguist of recent times, Noam Chomsky, has been the identification of principles of a 'universal grammar' that is presumed to govern all languages, and the principles that govern the generation of well-formed syntactic structures in particular languages – not the nature of words and their meanings. Chapter 4 (1965) is entitled "Some Residual Problems," and section 2 of that chapter – less than thrity pages – is entitled "The Structure of the Lexicon" (pp. 164–192). It is concerned with rules for inserting lexical items as terminal nodes into syntactic structures ('phrase-markers').

The major works in modeling the lexicon in early generative grammar are illustrated by the works of Katz and Fodor (1963) and Bolinger (1965). In modern generative theories of lexicon, the works of *inter alia* Pustejovsky (1991) and Jackendoff (2002b) should be mentioned. A radically different approach is offered by the tradition of Meaning-Text Theory, which in Mel'čuk's (2006: 228 in Hanks 2008: 2–3) words "considers the lexicon as the central, pivotal component of a linguistic description; the grammar is no more than a set of generalizations over the lexicon, secondary to it." The most fundamental works in the tradition of Meaning-Text Theory are works by Mel'čuk (2006).

The universalist trends in the description of a natural language so clearly manifested in generative linguistic analyses are also seen in a concurrently developing cross-cultural semantic studies (Wierzbicka 1985). As Hanks (2008: 3) notices, Wierzbicka's semantics similarly to Mel'čuk's Meaning-Text Theory also highlights the central role to the lexicon. As Hanks (2008: 3) further argues: "syntax is no more than the glue that is used to paste words together. It is in the words themselves, their uses, their cognitive and cultural associations, and their combinations, that meaning – the true meat of linguistic action – lies." The most identifiable characteristic behind the discussed cross-cultural semantic studies is the investigation of semantic primes and universals as best exemplified in the works of Wierzbicka (1985), Bogusławski (2008 [1970]); Apresjan (2008 [2000]), or Goddard (2005), and on more philosophical grounds pioneered by Leibniz in his *Table of Definitions*.

Last, but not least, the cognitive linguistic approach to the study of word meaning is described at some length in the following sections of the present Chapter and in Chapter 2 of this book. For the sake of a complete survey of major approaches to the study of word meaning conducted in this section, we should emphasise that cognitive linguistics is best characterized by the adoption of the non-Aristotelian conception of word-meaning with the emphasis laid on the fuzziness rather than discretness of conceptual boundaries of a word and with the simultaneous recognition of the necessity to encapsulate within the linguistic enquiry encyclopedic aspects of lexical meaning. This methodological perspective has been prompted by, among others, Wittgenstein's (1953) *Philosophical Investigations* and most clearly demonstrated by the works of Putnam (1975),

Labov (1973), Fillmore (2008 [1975]), Rosch (1975) Braisby (2008 [1990]), Lakoff (1987), Langacker (1987), Geeraerts (2006c) and many others.

1.2.2. Static or dynamic?

A crucial issue relevant to the study of word meaning concerns the nature of its perception as a static or dynamic phenomenon. An important remark is given by Taylor, Cuyckens, and Dirven (2003: 5), who note that "users of a language are rarely required to confront the issue of word meanings. If asked to state the meanings even of ordinary words in their language, most speakers find it hard to give satisfactory answers (Johnson-Laird 1987) […]. Language users encounter meaning primarily as a property of utterances, not of decontextualised words." They also suggest that the so-called 'conduit' and 'building-block' metaphors have helped in promoting the idea of word meanings as necessarily "reified objects" independent of the context of situation in which they occur (see Taylor, Cuyckens, Dirven 2003: 21). Challenge to this view is presented by, among others, Alwood (2003). "Alwood introduces the notion of a word's 'meaning potential,' a word provides access to a conceptual complex, only some aspects of which may be highlighted in a certain context" (Taylor, Cuyckens, Dirven 2003: 21). It must be noted that similar ideas have been proposed earlier by *inter alia* Krzeszowski (1997) in his conception of the schematic meaning.

Aside from reflecting upon the ontology of meaning as such, some other scholars pose questions about epistemology of word meaning. For example, Taylor, Cuyckens, Dirven (2003: 35) ask, then, what appears as a fundamental question: "And what, after all, does it mean, to know the meaning of a word? What are word definitions, and what are they for?" Goddard, in his introduction to Wierzbicka's program of definitional semantics, has this to say: "We will assume that we are entitled to expect that an accurate definition will predict the appropriate range of usage of a word. Putting it another way, the reader should be able to trust that the definition is a reliable guide to how to use the word" (see Goddard 1998: 31). The authors continue:

> A definition is a device for generating and evaluating uses […]. Definitions, in short, are linguistic abstractions […]. And mature speakers of a language are able to use words appropriately, not because they have learned the definitions and how to apply them, but because they have learned how to use the words appropriately.
>
> (Taylor, Cuyckens, Dirven 2003: 36)

Therefore, in the word-recognition test (see section 8.6) in which respondents are asked to assess their acquaintance with the selected foreign words and phrases, they are not supposed to provide the definition as a marker of knowledge of the words under consideration, but rather indicate, via introspection, a degree of familiarity with a particular form.

Although as Geeraerts (2006c: 141) concludes, "lexical meanings are not to be thought of as prepackaged chunks of information, but as moving searchlights that may variously highlight subdomains of the range of application of the lexical item in question," we claim that this does not need to stand in opposition to the idea of meaning as reified object. This is because Geeraerts' "subdomains of the range of application of the lexical item" actually denote that the meanings must be conceived of as essentially 'objectified.' If so, the processual (dynamic) conception of word meanings should be at best regarded as complementary to, rather than dismissed as incompatible with, the conception of meaning reification as prompted by 'conduit' metaphor theory Reddy (1979) or 'building-block' metaphor (see also Lewandowska-Tomaszczyk 1986). The reconciliation of the two rivaling positions on the nature of 'meaning' of words can be found in the proposal by Croft (1999: 77) and his 'conceptuality of meaning,' understood as representation of "conceptualizing experience in the process of encoding it and expressing it in language" (Croft 1999: 77, after Divjak 2006: 19).

The conceptualization of experience thus entails some stability in the construal of category in the mind, but does not disregard dynamic processes of multimodal perception in the formation of a conceptual category whatsoever. Salient meanings of foreign words and phrases listed in Glossary of foreign words and phrases appended to the book are thus highlighted subdomains in the range of application of a particular lemma. This should denote some stability in the category structure evoked by a particular lexical unit, especially that postulating 'subdomains' is *per se* an act of a priori introspective *search* by the lexicographer across the labyrinth of meaning evoked by a particular lemma. Postulating thus the prominence of the processual account of meaning with the simultaneous depreciation of the value inherent in the 'pre-packaging' account would in effect diminish the role of dictionaries, thesauruses, lexicons as reliable reference sources of lexical information. This implication is noted by Taylor, Cuyckens, Dirven (2003: 35) who say: "To be sure, this position may seem rather outrageous, because it flouts the well – entrenched role of the dictionary/lexicon in both folk and expert theories of language." The author of the present book agrees that such position is "outrageous" and does not constitute an argument in favour of advancing the processual over pre-packaged account of meaning for the reasons spelled out above.

1.3. Assumptions and theories of the structure of lexicon

This section surveys selected lexicological works featuring significant contributions in modeling the structure of lexicon in (cross-) linguistic studies. In section 1.3.1 we briefly discuss original conceptions underlying the creation of semantic field

theory (Porzig 1934, Trier 2008 [1934]), and look at the contributions to the study of lexicon from the generative grammar perspective (Pustejovsky 1991, Bolinger 1965, Jackendoff 2002a, 2002b). Finally, in section 1.3.2 we conclude the discussion with the illustration of major insights from more usage-based and more dynamic models of lexicon structure (such as Functional-Lexematic Model [Faber and Usón, 1997, Aitchison 1987, Brugman & Lakoff 2006]). The discussion ends up with surveying the findings concerning the historical survey upon the development of field semantics presented in an excellent paper by Wildgen (2000).

1.3.1. Semantic field theory

A very influential paper that discusses the nature of lexical relations comes from Porzig's paper entiltled *Wesenhafte Bedeutungsbeziehungen* published in 1934 whose English translation entitled *Intrinsic semantic relations* is found in Hanks (2008: 3–21). Porzig (2008 [1934]: 4) underlines the significance of de Saussure's observation about the systemic nature of linguistic meanings for the development of scholarly thought over the nature of lexicon in a natural language. His main idea is that meaning relations between words such as, e.g. *drive* a *car, walk* on *foot, ride* a *horse* are organized into intrinsic mutual relations in which the use of one word necessarily implies the other word. As the scholar notices, "*walking* requires *feet*, *grasping* requires a *hand, seeing* requires *eyes, hearing* requires *ears*, *licking* requires a *tongue*, and *kissing* requires *lips*" (Porzig 2008 [1934]: 3). Porzig underscores the significance of attempts made by other scholars in the organization of meaning system as the ones which can be exemplified by the works of Ipsen (1924), who coined the term *semantic field* (Bedeutungsfeld) as well as by Trier's diachronic analysis of changes in meaning relations of the lexis belonging to the realm of reason.

Porzig remains, however, critical of Trier's study in that it fails to predicate on what basis the postulated relations between words can be justified. Trier, in Porzig's words simply takes for granted the existence of certain superordinate semantic fields and, as a result, does not deliberate over their nature much. The key to the discussion upon the nature of a semantic field lies, as Porzig argues, in the recognition of intrinsic meaning relations which hold between two words, which he calls *elementary semantic fields* (elementare Bedeutungsfelder) (Porzig 2008 [1934]: 4). As Porzig eventually concludes: "elementary semantic fields are real both in the sentence and in discourse. They are subject to the dynamics of speech, to the continuous shift and reorganization contained in every speech act. They are the objects of these reorganizations, but they are also operative forces within them" (Porzig 2008 [1934]: 10).

A rivaling position in the formulation of the nature of the linguistic field comes from the work by the aforementioned Jost Trier entitled *Das sprachliche*

Feld. Eine Auseinandersetzung published in 1934. Trier, similarly to Porzig, underlines the importance of the structured organization of lexicon, which comes from Humboldt's observation, adopted by de Saussure, about structure being the most fundamental feature of language (Trier 2008 [1934]: 23). As Trier claims, "we must not think of the lexicon as a treasury, a stock, a thesaurus, but rather – in an analogy with architecture – as constructed and structured space, a construction which through its very constructedness and structuredness establishes and determines the significance of every individual position within the building as a whole" (Trier 2008 [1934]: 23). Trier (2008 [1934]: 24) then defines precisely what he understands by the concept of the field:

> Fields are linguistic entities between individual words and the vocabulary as a whole. With words they have in common that they are structured by the vocabulary as a whole, with the vocabulary they have in common that they can be decomposed into words.

His position thus differs from the one adopted by Porzig as well as Jolles in that his conception of the linguistic field is more broadly defined and involves relations between words based not so much on co-textual restrictions offered by the meanings of predicates that constitute the foundations of intrinsic relations in Porzig's conception of the field, but on the semantic relations of 'higher order.' These semantic relations of 'higher order' are built by the recognition of similarities between particular words, which validates the creation of superordinate categories for the orderly description of these relations. For Porzig these relations can only be substantiated inasmuch as we linguistically prove the existence of such relations through the examinations of the elementary semantic fields or, as Trier (2008 [1934]: 33) calls them, "unidirectional predicative meaning relations" (Prädikative Bedeutungsbeziehungen). Porzig's approach to a semantic field can be called a bottom-up approach, because he starts his considerations from the investigations of what we could now label after the generative-transformational tradition as sub-categorization frames of predicates (see Katz and Fodor 1963) and, as a result, builds up structures of higher-order. Trier argues, instead, that his conception of the semantic field does not need to begin with the investigation of predicate meaning relations in order for them to be classified as linguistically motivated. His linguistic motivation, as he directly admits it, comes from "the entirety of our contemporary shared language and its internal structure." It is on these grounds, that the ontology of the field can be extracted (Trier 2008 [1934]: 35). His perspective thus differs from Porzig's conception in that it offers more up-down methodology and also remains in clear opposition to Jolles' conception of the field conceived of as a collection of semantic fields – opposing pairs like *right/left* remaining, as he argues, "in pairwise isolation" with "their free distribution in the world, their respective unrelatedness." As Trier puts it, Jolles' conception does not "open up a way to a structured whole," whereas Porzig's conception, as it can

be inferred, is unnecessarily restrictive, and as Trier puts it "prevents them from being the starting point from which we can build upwards and get to the structure of the linguistic contents of the whole language" (Trier 2008 [1934]: 42).

A few words should at this moment be spared to the generative tradition of the study of lexicon. These are marginal to our considerations, yet should at least be briefly recalled as these exerted a great impact on the linguistic thought in modern times. One such influential paper by Katz and Fodor entitled *The structure of a semantic theory* published in 1963. Here the role of a semantic theory is to interpret the syntactic structure revealed by the grammatical structure. Semantic (lexical) considerations are regarded as ancilliary to the examination of syntactic structure, the main area of linguistic research (see Chomsky 1965). A detailed exposure of the notion of *semantic marker* coined by Katz and Fodor is presented by Bolinger (1965) in his postulate of the atomization of meaning, i.e. the decomposition of a semantic structure into small and individually idetnifialble elements, which are further divided into *semantic markers*, ie. systematic characterizations of the meaning and *distinguishers*, which as Bolinger (2008 [1963]: 329) puts it, are "the idiosyncratic remainder of a given sense when all the markers have been stripped away." Finally, in modern generative theories, we should note Pustejovsky's (1991) "The Generative Lexicon" and Jackendoff's "What's in the lexicon" published in 2002. As Hanks (2008: 17) notices, Pustejovsky "revived the Aristotelian notion of qualia and argued that qualia are generally sufficient for the expression of polysemy," whereas Jackendoff (2002b) reformulated the division between grammar and lexicon, indicating a "much less rigid divide than usual between lexical items and rules of grammar" (Jackendoff 2008 [2002]: 427). The paper in its postulate of inclusion of regular affixes, stems, and phrasal units into the realm of lexical description loosened the aforementioned boundary between grammatical and lexical study and, in this way, opened up a new way towards, as Jackendoff admits, "a better rapprochement between linguistic theory and psychological studies on language processing than has been possible in more traditional Chomskyan architectures" (2008 [2002]: 427).

1.3.2. Toward some alternative models of lexicon

Summing up the basic domains of investigation in modeling the structure of lexicon, Jackson & Zé Amvela (2007: 14) argue that there are three main areas according to which studies in vocabulary organization have so far been conducted. The first area concerns the relationship between words and their associations; the second involves the examination of lexical fields, and the third one relates to the investigation of relations within word families. This book, in contradistinction to the above-presented areas, does not offer the model of lexicon so much as this would go far beyond the research goals stated in Introduction,

but rather aims at investigating one aspect of its ontology, i.e. the nature of assimilation of foreign words and phrases into the target lexical system as well as the nature of "foreignness" itself (see sections 0.1, 0.2, 6.7.3, 6.9, 9.1.1). Answering this question about the nature of foreign word assimilation certainly entails that we also learn about the nature of the lexicon at least to an extent bounded by the research perspectives outlined in Introduction to this book. Exploring such conceived nature of the lexicon also entails that aspects of its internal functioning are attempted to be sketched, hence the division into internal and external perspectives of the finally emerging model discussed in Chapter 9.

At this moment, however, we should present some non-structuralist and non-generativist proposals given in the field of lexical studies in advancing the model of the structure of lexicon and, if possible, we will relate the findings to the model of foreign lexical assimilation presented in Chapter 9. Another noteworthy trend in lexical studies involves the examination of relational structure holding between words. According to Jackson & Zé Amvela (2007: 14), this approach is on the following assumption:

> Every word is involved in a network of associations which connect it with other terms in the language. Some of these associations are based on similarity of meaning, others are purely formal (i.e. based on forms), while others involve both form and meaning. In de Saussure's graphic formula, a given term is like the centre of constellation, the point where an infinite number of coordinated terms converge […] the use of an arrow and that of *etc.* at the end of each line of associations suggests that the line has no limit and that an infinite number of words can be added to those suggested in the diagram.

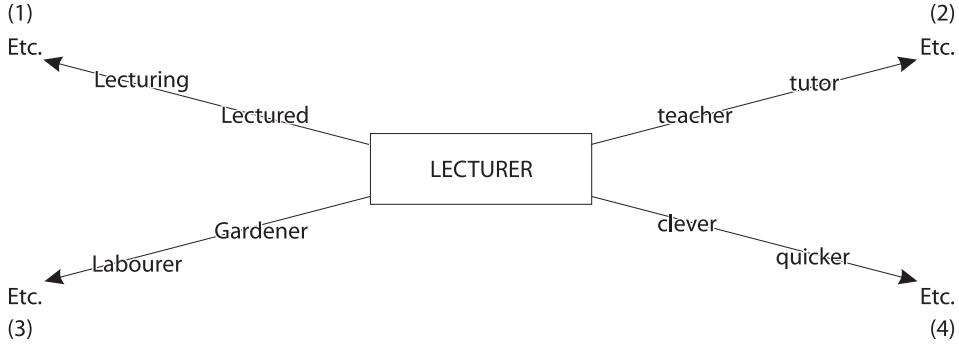


Fig. 4. *The word and its associative field* (Jackson & Zé Amvela 2007: 15)

The study of the associative network of words is an implicitly implemented tool of vocabulary organization in Chapter 7. The foreign words and phrases are grouped into domains on the basis of criteria such as similarity in meaning or, put it in more cognitive terms, on the basis of their correlation with the

same ICM network (Lakoff 1987) at a conceptual level.[40] Once a particular set of lexical units is grouped into semantic domains, the structure of these domains (or fields) may further be explored. This is the second major trend observed in lexical studies.

As Jackson & Zé Amvela (2007: 16) state:

> Some isolated attempts have been made to study the structure of some semantic or lexical fields, such as the hierarchy of military ranks, numerals, colour and kinship terms. Most of these attempts are directly or indirectly connected with what has now come to be known as the theory of 'semantic fields' or 'lexical field theory.' 'Semantic field' or 'semantic domain' is used alternatively for the terms 'lexical field' or 'lexical set.'

Crystal (1995: 157), for example, defines a semantic or lexical field as a "named area of meaning in which lexemes interrelate and define each other in specific ways." Field theory was first advanced by a number of German and Swiss scholars in the 1920s and 1930s. However, according to Lyons (1977: 25) its origin can be traced back at least to the middle of the nineteenth century and more generally to the ideas of Humboldt and Herder (Jackson & Zé Amvela 2007: 17). Some significant remarks are made by Jackson & Zé Amvela (2007: 17) in further characterization of the nature of lexical field that emerges out of studies conducted so far.[41] Their weight relates to the fact that these observations may be viewed as a 'forerunner' of the essentials of the model of lexical assimilation advanced in the present book. The authors note:

> According to lexical field theory, the vocabulary of a language is essentially a dynamic and well-integrated system of lexemes structured by relationships and meanings. The system is changing continuously by the interaction of various forces such as the disappearance of previously existing lexemes, or the broadening or narrowing of the meaning of some lexemes. The system is mainly characterized by the general-particular and part-whole relationships, which hold not only between individual lexemes and the lexical field within which they are best interpreted, but also between specific lexical fields and the vocabulary as a whole.

These insights into the nature of lexical organization within a lexical field, especially a remark about the constantly changing of the system on the basis of

[40] As Brugman & Lakoff (2006: 109) argue: "The category structure utilized here is called a 'radial' structure, with a central member and a network of links to other members."

[41] See also a general observation made by Aitchison (1987: 192): "Words seem to be organized in semantic fields, and within these fields there are strong bonds between co-ordinates which share the same word class. As far as producing speech is concerned, this is a useful arrangement. […] As it is, words which sound similar, particularly at their beginnings and ends, are those which are most closely linked. But from the point of word recognition it is useful to have similar sounding words together. Hearer can then examine several of them together, and find the best fit for what they have heard. The phonological component of the mental lexicon, therefore, appears to be organized primarily in accordance with the needs of recognition (ac. Fay and Cutler 1977)."

action of various forces may be regarded as a very schematic outline of the foundations of the model of lexical assimilation stated in Introduction and developed in Chapters 6, 8, and 9 of this book. It is, however, somewhat bizarre that Jackson and Zé Amvela omit to mention the role of non-native vocabulary in the development (i.e. change) of the lexical system, confining themselves to the discussion of relations between words already existing in the system as if this system was essentially self-sufficient.

Some other problems with lexical field organization pertain to the methodology of assigning particular words to clearly defined domains. According to Crystal (1995: 157), "these difficulties are of three kinds. First, some lexemes tend to belong to fields that are vague or difficult to define […]. Secondly, some may validly be assigned to more than one field […]. The last difficulty concerns the best way to define a lexical field in relation to the other fields on the one hand, and its constituent lexemes on the other" (after Jackson & Zé Amvela 2007: 17). The fact that such difficulties are indicated by lexicologists leads to a conclusion that the ideal of discrete lexical fields where lexemes belong with a particular domain on Aristotelian grounds of yes/no membership is a myth and the actual assignment procedure is sometimes simply discretionary as it is not possible to create one salient homogeneous domain into which lexical items could be placed (see Jackson & Zé Amvela 2007: 19. See also section 7.0). Despite the problems signaled above, some schematic models of vocabulary organization have nevertheless been proposed in the literature and these are presented below as Fig. 5 and Fig. 6.

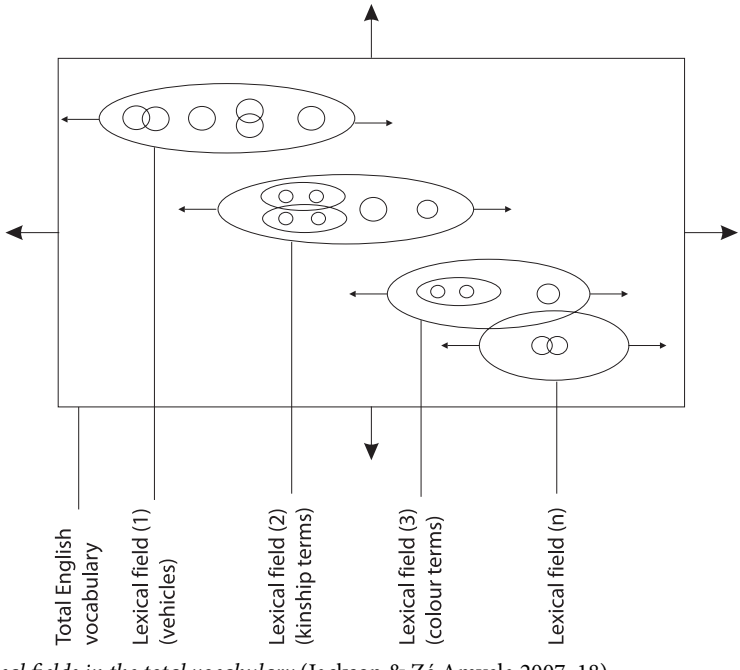


Fig. 5. *Lexical fields in the total vocabulary* (Jackson & Zé Amvela 2007: 18)

Fig. 5 represents a model of the lexicon, whose 'immediate' components are lexical fields with different mode of internal organization. Blank empty arrows suggest the dynamic nature of the lexicon as such, whereas regular black arrows symbolize the changing ontology of various component areas of vocabulary. An example of such internally organized area of vocabulary is illustrated below (Fig. 6).

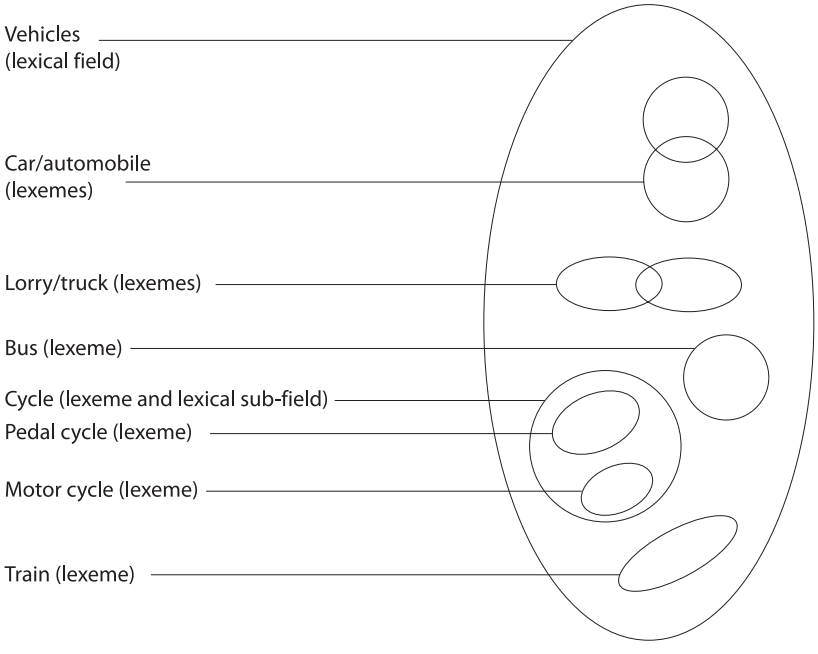


Fig. 6. *Example of a lexical field* (Jackson & Zé Amvela 2007: 18)

As it can be concluded from the above-illustrated figures, the models show some characterization of the relational structure binding lexical units into one organized whole, but what is missing is the reflection upon the nature of the system functioning, especially the nature of forces (tacitly addressed but not elaborated upon by Jackson & Zé Amvela 2007) constituting the organization of vocabulary into the system.

The third major trend in lexical studies discusses lexical organization at the most general level and involves the discussion of words not so much in terms of lexical fields to which they pertain, but rather on the basis of their morphological criteria.[42] As Bauer and Nation (1993) claim this approach is characteristic of French lexicology (e.g. Lehmann and Martin-Berthet 1997). A recent treatment

[42] Another similar trend in lexicological research has come to be known as semantic decomposition. As Altenberg & Granger (2002: 28) put it: "Interconnections within the lexicon have often been analyses in terms of shared primitive components or features. As pointed out by James (1980: 89) and Kittay and Lehrer (1992: 9), one motivation for this has been the economy of description: a small number of components can be used to define a large number of words."

of lexical studies from the viewpoint of 'word family' approach is offered by Jackson & Zé Amvela (2007: 19), who discuss an example of such grouping:

> A family consists of a base form, its possible inflectional forms, and the words derived from it by prefixation and suffixation, e.g.
> (a) State (verb)
> States, stated, stating (inflections)
> Stateable, statement, misstate, restate, understate (derivations)
> (b) Skill (noun)
> Skills, skill's, skills' (inflections)
> Skilful, skillfully, skillfulness, skilless, skilled (derivations)

The morphologically-founded organization of lexicon into families is not the focus of the research in this book. They are relevant to the present discussion inasmuch as they relate to the adaptation stage (Phase 3, see section 5.3, Fig. 13) in the advanced model of assimilation in which case the rules of derivation, inflection are postulated to operate on the already fully entrenched non-native lexical unit in the target system[43] (see section 10.1.1, Fig. 60).[44] This trend of examining relations between words with regard to their morphological composition can also be seen as dominant in the studies upon lexical borrowings (see Görlach 2007), which is consistent with the observation just made that morphological analyses apply to fully adapted units. That borrowings are often taken for granted as fully adapted units in the literature of the subject-matter is discussed in Chapter 5 of the present book. This fact also leads us to make the issue of the categorial differentiation between borrowings as the category subsuming well-adapted non-native lexical units and foreign words and phrases as the category subsuming non-native units that have not yet completed the full adaptation process (see especially sections 5.2 and 5.3).

[43] See Fisiak (1985b, 1986) for some insightful studies of loanwords that relate to the stage 3 of the assimilation process as understood in this book.

[44] The research tradition focusing on the default recognition of a lexical unit as entrenched in the lexicon of a given language permeates the discussion in the field of lexicology. Such taken for granted assumption leads to the following generalization about the lexical structure: "According to a common view of the matter, knowledge of a language can be partitioned into two major components – knowledge of the lexicon and knowledge of the syntax. The lexicon lists the words of the language and states, for each word, its phonological properties, its syntactic category, and its meaning. The syntax comprises the rules whereby elements belonging to certain syntactic categories can be combined into larger configurations. […] The way sentences are pronounced, and the way they are interpreted, thus depend, ultimately, on the properties of words" (Taylor, Cuyckens, Dirven 2003: 4). It must, however, be noted that the research on the second, syntactic component has been so far more powerful as has been observed by Aitchison (1987: 25–26): "In recent years they [theoretical linguists, M.K.] have regarded syntax, which involves combinations of words, as more important than the words themselves. […] This has led many of them to underestimate the complexities of the lexicon and to characterize it as a finite list which concentrates on irregularities and idiosyncrasies." See also Kempson (1977: 102), Bloomfield (1933: 274), Chomsky and Halle (1968: 12) – all quoted by Aitchison (1987: 26).

Aside from the above-presented three major trends in the studies upon the organization of lexicon in the system of language, one may recently observe some sort of shift in the focus on the research of lexical structure. As Faber and Usón (1997: 11) claim:

> In recent years, linguists have become aware of the significance of lexical structure as a means of ascertaining and exploring the organization of concepts in the mind. Indeed, lexical relations at different levels of the lexicon encode a map of conceptual relations and give us a tantalizing glimpse of "mentalese," a language of thought.

In this view lexicon can be seen as a "repository of forms" that contains crucial information about the processes involved in the formulation of thoughts (Faber and Usón 1997: 12; cf. Pinker 1994: Ch. 3). This approach to the study of lexicon has become known in the literature under the notion of Functional-Lexematic Model (or the FLM framework) in which "the lexicon is not a mere storage place for words but rather a dynamic, textually oriented repository of information about words and their contexts" (Faber and Usón 1997: 13). This conception of lexical organization is quite akin to the cognitive-semantic framework that form the background for the model of foreign lexical assimilation discussed in this book. The acceptance of cognitive-semantic framework is an answer to the address formulated by Brugman & Lakoff (2006: 110–111) who expressed their need for more cognitive topological orientation in linguistic studies, which "characterizes structures oriented relative to the human body that apply generally to spatial situations, structures like paths, bounded regions, tops, etc. Structures in a cognitive topology differ from semantic features in a number of ways: they are inherently meaningful (arising from sensory-motor operations), they have an inherent structure, they are analog rather than finitary, and the relationships among them arise naturally via the operation of the human sensory-motor system." The model of foreign lexical assimilation advanced in the present book is then based on the sensor-motor bodily operations by the human being and as such is fundamentally cognitive in nature (see Chapter 6).

Relating the structure of lexicon to the patterning of conceptual mapping in the human mind is quintessentially convergent with the basic assumptions and principles developed by Cognitive Linguistics, and as such FLM framework is worth mentioning at his point. In Brugman & Lakoff's (2006: 109) words: "A network-style mode of storage [of semantic information, M.K.] is cognitively real, and that this allows for a maximum of shared and otherwise related, information between senses." Another related issue arises when we pose more fundamental questions about the validity of postulating the existence of such patterning or network of relations when it comes to the examination of lexicon. In search of more neurolinguisitc evidence in favour of advancing the structural nature of lexicon, we may safely state that "in lexicology the stock of words or lexical items is not simply regarded as a list of isolated elements. Lexicologists try to find out general-

izations and regularities and especially consider relations between elements […]. Lexicology is therefore concerned with structures, not with mere agglomeration of words" (cf. Jackson 1988: 222, after Lipka 2002: 9).

The statement about the structured nature of the lexicon was not shared by Strang (1968: 215) who says that "lexis is the domain of vast lists of formal items about which rather little generalization can be made" (Lipka 2002: 12). The voices about haphazard rather than structured organization of lexicon appear however to be in the minority[45] and most lexicologists agree on a fundamentally binary typology of lexicon structure, i.e. external and internal. As Lipka (2002: 12) says:

> We may distinguish at least two types of a structure in the lexicon: external (to the word) and internal structure. Full words may substitute for each other, i.e. be in opposition, or they may combine with each other. Such paradigmatic and syntagmatic relations constitute external structure. We may also look at the internal structure of lexical items which are either morphologically complex (such as compounds) or simple.

The division into internal and external aspect in the study of lexical structure is upheld in this book (see sections 9.1 and 9.2) in which we attempt to construe the model involving one manifestation of lexical structure, i.e. the aspect of foreign lexical assimilation and its impact upon the target lexical system.

Another issue relates to the differences between book dictionaries and mental dictionaries (lexicon), which actually marks the boundary between lexicographic and lexicological research already indicated in section 1.0. According to Aitchison (1987: 10): "Unlike book dictionaries, human mental dictionaries cannot be organized solely on the basis of sounds or spelling. Meaning must be taken into consideration as well, since humans fairly often confuse words with similar meanings." As Aitchison (1987: 12) continues:

> The biggest difference between a book dictionary and the mental lexicon is that the latter contains far, far more information about each entry. All book dictionaries are inevitably limited in the amount they contain, just because it would be quite impracticable to include all possible data about each word. In any case it is unlikely that anyone has ever assembled the total range of knowledge which could be brought together about one dictionary entry.

[45] Cf. some further critical voices of this haphazard organization of lexicon, e.g. Aitchison (1987: 5–9): "Words are not just merely stacked higgledy-piggledy in our minds, like leaves on an autumn bonfire. Instead, they are organized into an intricate, interlocking system whose underlying principles can be discovered. […] The large number of words known by humans, and the speed with which they can be located, point to the existence of a highly organized mental lexicon." Studies of types of relations between senses are significant in a number of respects. As Brugman & Lakoff (2006: 110) state, "the relations between senses are not arbitrary, but rather principled, systematic, and recurrent throughout the lexicon." The same will be argued for in the present book whereby the systematic relations between different foreign words and phrases will be explored (see especially Chapters 7 and 8).

Some other linguist notes: "There is no known limit to the amount of detailed information which may be associated with a lexical item. Existing dictionaries, even large ones, specify lexical items only incompletely" (Hudson 1984: 74, after Aitchison 1987: 12). In relation to the arguments favoring a strict separation of book from mental dictionaries, the ambition of the present book is to offer an alternative description of foreign words and phrases to the ones provided in book dictionaries, an account which will hopefully illustrate the nature of internal and external structuring of this component of the English lexicon. This, to the best knowledge of the author of the present research, has not as yet been systematically investigated in lexicological studies.

Last but not least, a few words must be spared on the ontology of the model of lexical structure itself and eventually the model of mind. It is again Jean Aitchison who focuses on this aspect of lexicological research. She says that "models of the mind are somewhat like plans of the London Underground system: they are simplified diagrams which encapsulate crucial features of something that is reality considerably more complex. […] Mental maps are unlike real-life maps in that they have to depend on inspired guesswork, since we cannot actually look into the head and see the connections we hypothesize" (Aitchison 1987: 35). This is an important observation for the ontology of the model postulated in Chapter 9 of the present book, which renders it tentative rather than definite proposal.[46]

Another significant issue when it comes to the investigation of lexical structure relates, according to Aitchison, to the rough division of lexicological enterprise into works which favor "atomic globule theories" on the one hand, and other publications advancing "cobweb theories," on the other. Thus,

> atomic globule supporters argue that words are built up from a common pool of 'meaning atoms,' and that related words have atoms in common. Cobweb supporters claim that words are recognized as related because of the links which speakers have built between them. On the one hand, then, words are viewed as an assemblage of bits. On the other, they are regarded as wholes which have various characteristics and enter into relationships with other words.
>
> (Aitchison 1987: 64)[47]

[46] In line with this reservation, see also Chapter 10 for other possible representations of the postulated model of foreign lexical assimilation.

[47] "The network theory (called by Aitchison [1987] the Cobweb Theory) is a theory of the mental lexicon. Behind this theory what we have is a conception of the mental lexicon as a net in which each dot of the net is a word. These words are connected through others to the whole of the system. The relationships established between words have multiple routes. Not all the words are related equally. Distance: some are closer than others to the rest. Words from the same family are closer than those outside this family. There is a problem with this theory. Network theory fails to capture the overlapping of meaning." (Internet source 2)

The present research on foreign words and phrases falls definitely into the "cobweb" type of research and, at the same time is consonant with the criticism voiced over the "atomic globule" theories.[48]

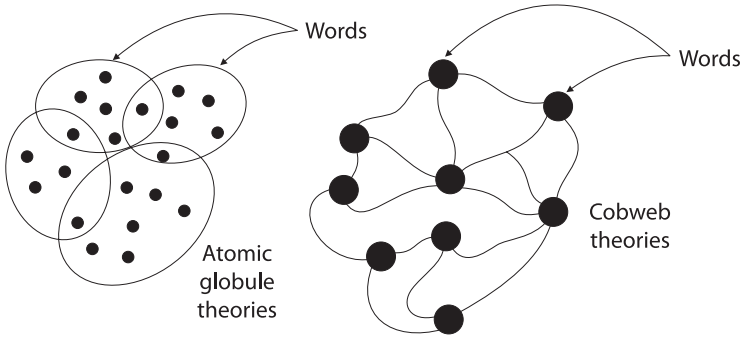


Fig. 7. *Atomic globules versus cobwebs* (Aitchison 1987: 64)

Connected with the "cobweb type" of lexicological study are word association experiments that aim at investigation of semantic networks between words (cf. Jenkins 1970). As Aitchison (1987: 73) points out: "Early work on meaning networks suggested that links between words were formed by habits […]. The close links forged by these habits could be revealed quite easily, it was suggested by means of simple word association experiments." Particularly relevant to the research upon foreign words and phrases are findings according to which "common words are recognized as words faster than uncommon ones, a fact first pointed out over a quarter of a century ago (Solomon and Howes 1951, after Aitchison 1987: 181). This observation is foundational for the model of foreign word assimilation and is elaborated on in Chapters 6, 8, 9. In this book, however, the aforequoted "common words" will stand for entrenched non-native units in the lexical system of the target language (Stage 3 in the assimilation process), whereas 'uncommon' words will constitute the set of items that classifies as belonging within Stages 1 and 2 of the process (see sections 5.3 and 10.1.1).

As Milroy (1987, after Aitchison 1987: 196) observes:

> An alternative way of looking at the connections between these 'lexical' towns therefore, is one which focuses on the inhabitants rather than on fixed architecture. Any town in real life is likely to contain 'social networks', groups of people who know one another and interact fairly often.

[48] As Aitchison (1987: 71) argues: "The theory [the theory of atomic globules, M.K.] into insuperable problems: no one has been able to specify what these atomic globules are, and they leave no trace in the processing of words. The arguments in favour of this viewpoint are based mainly on descriptive convenience and wishful thinking. Our overall conclusion was that they are useful descriptive devices for people such as lexicographers who need to describe things in a succinct and orderly way. But they are unlikely to exist in the mental lexicon."

This alternative perspective is accounted for in Chapters 7 and 8 of the present book. Thus Chapter 7 offers a description of foreign words and phrases and Chapter 8 investigates "social networks" as various regularities holding between lexical elements. As Aitchison (1987: 196–197) further argues:

> This analogy can be transferred to the mental lexicon. Each 'lexical town' will contain numerous clumps of words with strong ties to one another – though each clump will also have bonds, though weaker ones, with other groups. In addition, there will be connections, weaker still, between individuals in different towns. Like social ties, these connections will fluctuate.

Summing up the discussion concerning the more dynamic alternatives to the study of lexicon, we should refer, at this moment, to a very inspiring paper by Wildgen (2008 [2000]), who offers a concise yet contentful summary of the history of field semantics in non-structuralist perspective. He starts the discussion with Raymundus Lullus and his conceptual systems described in *Ars Magna*. As Wildgen puts it, the semantic system – embedded in the tradition of Aristotelian and medieval logic – consists of "three subfields which have a parallel substructure. The first group: 'god,' 'big,' 'eternal,' may be called metaphysical; the second comprising 'mighty,' 'wise,' 'will,' is related to the human mind and body; and the last group of 'virtuous,' 'true,' 'glorious,' may be called ethical" (Wildgen 2008 [2000]: 68). The Lullian system was adopted and elaborated a couple of centuries later by Giordano Bruno, who devised "a new system of conceptual organization based on the analogy between the macrocosm (the universe) and the microcosm (man, his mind)" (Wildgen 2008 [2000]: 69). The types of semantic relations that bind the elements constituting the subfield in Bruno's system correspond to the contemporary lexical–semantic concepts of hyperonymy, hyponymy, synonymy, metonymy, and so on. The system, as such, was very modern and radical in its time (Wildgen 2008 [2000]: 72–73).

When it comes to modern contributions to a theory of semantic fields, Wildgen mentions Peirce's existential graphs. What is particularly novel about Peirce's approach is the concept of valence functioning as elaboration of the relational concept as understood by Lullus. The concept of valence was later reformulated by Fillmore (1977) in his discussion of *frames* (Wildgen 2008 [2000]: 74–75). As Wildgen (2008 [2000]: 75–76) continues:

> Many elements in Peirce's thought (for instance the 'marriage' between natural science and semiotics) and the importance of modern physics and modern experimental psychology for theoretical work in linguistics and semiotics, are also characteristic of Gestalt theory, the next stage in our journey to field-semantics.

One of the fundamental findings of Gestalt theory with respect to field semantics was its observation about non-homogeneity of the field. It was discovered that the field consists of subparts whose status in the composition of the whole is

not equal, i.e. there are more and less prominent (dominant) areas (Wildgen 2008 [2000]: 76). Findings presented by Gestalt psychologists inspired Karl Bühler to postulate two major applications of the concept of field, which he laid out in his *Sprachtheorie* (1965 [1934]). One application involved the semiotic activity and the ensuing use of two fields *Zeigefeld* (deictic field) and *Symbolfeld* (field of symbols). The second application related to the potential usage of the symbolic field in isolation, in which case we can distinguish *Feldmomente* (field moments), which enables "the reconstruction of a whole (a sentence) from the set of its constituents" (Wildgen 2008 [2000]: 77). Field moments further divide into *Stoffhilfen* (material field moments), and *Wortklassen* (word classes). As Bühler (1965: 171) claims "the former open a material domain and a system of interrelations with other elements of the same domain. Thus the term 'salad' opens a field either of garden plants and garden activities or one of kitchens and eating." We can notice far-going correspondences between Bühler's *Stoffhilfen* and the aforediscussed concept of frames as used cognitive-linguisitc studies (see Dirven and Verspoor 2004: 55).

The links between Gestalt psychology and cognitive semantics are particularly strong, which resulted, as Wildgen remarks in the intense development of topological and dynamic semantics between 1976 and 1979. He continues:

> In 1977 Lakoff gave a paper on linguistic Gestalts at the Summer School on Mathematical and Computational Linguistics in Pisa. In the same period Leonard Talmy wrote his articles "Rubber Sheet Cognition in Language" and "Figure and Ground in Complex Sentences," and in 1979 Langacker published the first article entitled "Grammar as Image" – on what would become 'space grammar' and later 'cognitive grammar.' Thus, the new 'wave' of topological and dynamic semantics finally reached California and soon thereafter began to spread through Italy, Germany, and France – cultural areas in which, half a century earlier, the major trends in gestalt theory and corresponding applications to linguistics had been created.
>
> (Wildgen 2008 [2000]: 80–81)

1.3.3. Contributions from neurolinguistics

Out of three most fundamental theories of meaning[49] that of *conceptualism* is most relevant to our considerations as it has given the inspiration to cognitive-semantic studies of language as well as neuroscience. According to this view, concepts exist in our mind as necessary intermediaries between the word and the external reality. This jointly advocated conception of meaning constitutes "a bridging gap" in interdisciplinary searches upon lexis in natural language.

[49] See Ahlsén (2006: 80) for the discussion of *nominalism*, *conceptualism*, and *conceptual realism*.

One manifestation of such investigations relates to the so-called word semantics. Neurolinguists are interested in examining aphasic ways of searching for words, which involves semantic word substitution tests. These researches, once systematically conducted, are applicable in verifying the hypothesis about the nature of organization of mental lexicon, i.e. if it can be disturbed in itself or whether the avenues leading up to the lexicon are susceptible to damage (Ahlsén 2006: 80).

Another aspect that conjoins neurolinguistic and cognitive-semantic research involves the reference to prototype theory by Rosch (1975). The theory that predicts about the prototypical organization of categories with the simultaneous rejection of the notion of category discretness[50] constitutes a set of principles common to both cognitivists and neurolinguists. Furthermore, both cognitivists and neurolinguists agree as to positing different levels of abstraction in the organization of linguistic concepts, i.e. superordinate, basic, and subordinate levels (Ahlsén 2006: 82). These in turn make up the foundation upon the classificatory criteria along which foreign words and phrases are organized in the book (see section 7.0).

The other significant neurolinguistic insight into lexicological studies relates to the aforementioned semantic fields (see section 1.3). Semantic fields, or networks, according to Ahlsén (2006: 83) are "groupings of words according to semantic similarity, or contiguity (co-occurrence), relations. They have been used by several researchers and in a number of different ways. A well-known example is the hierarchical and similarity-based word association network used by Collins and Quillian (1969)."

One more example is Collins and Loftus (1975) model of word meaning which expounds the principle of semantic similarity. The word recognition test (see section 8.6), is here based both on de Saussurian planes of paradigmaticity as well as syntagmaticity at which lexical concepts may be analyzed (Ahlsén 2006: 83). Most weighty 'fruits' of neurolinguistic research upon lexicon comes, however, from the models of lexical comprehension[51] such as Forster's (1976) *active search model,*

[50] As Ahlsén (2006: 82) claims: "Semantic distinctive features have been used in descriptions of linguistic categories since antiquity. The idea that a category can be identified and described with reference to the necessary and sufficient conditions for belonging to that category. This process therefore involves decomposition into more primitive or basic features." As Altenberg & Granger (2002: 29) claim: "Some of the problems inherent in semantic decomposition can be avoided by resorting to the notion of prototypicality (see Rosch 1975, Taylor 1995). Prototypicality indicates degrees (or the 'best fit') of category membership and is a fuzzier notion than other semantic relations. If we accept that meanings can be fuzzy and are better described in cognitive (rather than purely linguistic) terms, certain lexical relations can be characterized more adequately in terms of prototypes. Prototypicality is therefore often used in cognitively oriented taxonomies and semantic field studies based on typological universals (cf. Viberg 1996: 159 f)."

[51] As Ahlsén (2006: 90) argues: "In an *active direct search* lexicon, it is assumed that one searches for matches of incoming stimulus word in 'bins' containing word representations (audi-

Morton's (1969) *logogen model*, Marslen-Wilson and Tyler's (1980) *the cohort model* and McClelland and Elman's (1986) *TRACE model* (see Ahlsén 2006: 90).

The model of foreign lexical assimilation presented in Chapter 9 appears to be consonant with passive, indirect, responsive model, as well as active, direct search model. The compatibility of the models is dictated by the nature of processes governing the inclusion of new words into the target lexicon, whereby the discussed centripetal and centrifugal forces correspond to the processes of accommodating or rejecting the neurological auditory or visual stimuli. Alternatively, the centripetal/centrifugal forces might be represented as cultural forces of social acceptance or rejection of alien lexical items in the target system. More on the representations of physical forces in natural language is discussed in sections 6.2 and 6.8.3.

As the discussion upon foreign words and phrases unfolds, we will see how the findings related to lexical decision experiments permeate the analysis connected with modeling foreign lexical assimilation processes. We will primarily focus of the frequency effect, the word-non-word effect[52] which are be viewed as fully compatible with research results discussed in part IV of the present book.

---

tory or aural) which are ordered by some principle, for example, how frequently they have been encountered. The word representations in the 'bins' are then linked to a 'lexical master file' where phonological and semantic cross-references exist. Similarly, a word can be accessed from a semantic representation in production. […] In a *passive, indirect, responsive model,* like *the logogen* model, auditory and visual incoming word stimuli can activate a specific 'logogen' (recognition) unit if the stimulus is strong enough (i.e., contains enough activation and not too much inhibition in relation to the features of the logogen to get over its activation threshold). The logogens are also connected to a semantic cognitive lexicon, which can increase or decrease the activation and thus affect which logogen is activated. […] A *cohort* model stresses the incremental build-up of activation as, for example, a written word is encountered letter by letter and a particular word is selected at a particular decision point. An early perceptual analysis determines that a set of words are possible candidates and recognition proceeds by eliminating the candidate from left to right."

[52] Ahlsén (2006: 91) discusses the models presented above in that they account for a number of recurrent effects identified in lexical decision tasks. These are the following:

– The frequency effect: the more frequent a word, the faster it is recognized.
– The length effect: the shorter a word is, the faster it is recognized.
– The concreteness effect: the more concrete a word is, the faster it is recognized.
– The word superiority effect: a letter is recognized faster in a word than if it occurs in isolation.
– The word/non-word effect: words are identified as words faster than non-words are identified as non-words.
– The context effect: words are identified faster in context.
– The degradation or stimulus quality effect: a word that is presented clearly is recognized faster than a word that is blurred in some way (e.g. covered by a grid), if it is presented visually.

Chapter 2
LEXICAL STUDIES IN CORPUS AND COGNITIVE LINGUISTICS

2.0. Cognitive Linguistics and Corpus Linguistics

The present study upon foreign words and phrases borrows from theoretical and methodological insights of both Cognitive and Corpus Linguistics. In cognitive and corpus-linguistic studies, the criterion of frequency, for example, serves to objectively describe the degree of usage of a given form in language (cf. Lewandowska-Tomaszczyk 2005: 16). The interest in frequency of occurrence figures leads us to the development of the notion of CRAC in Chapter 6, one of the foundational terms for the model of foreign lexical assimilation presented in Chapter 9. The notion of CRAC is also to be mentioned at this point as it perfectly conjoins the idea of introspective research (see section 6.1) with more objectivised quantitative research (see e.g. sections 8.1 and 8.3).

2.1. Cognitive linguistics

If we look at the matter from cognitive-semantic perspective first, this book with its underlying methodological principles may tentatively aspire to be inscribed within the domain of "empirical linguistics"[53] as it is governed by the Generalization Commitment, i.e. the "commitment to characterize the general principles governing all aspects of human language" (Lakoff 1990: 53). This induces interdisciplinary character of a linguistic enterprise especially that we see language functioning as one of possible windows into how we think about the world around us (see Gries 2006b: 1). So predictably, in Parts III and IV of the book, we heavily rely on the insights from the science of physics in accounting for how language assimilation processes work, on the one hand, and how we, human conceptualisers, are able to make sense of it, on the other. Lewandowska-Tomaszczyk (2006: 251–254) notices that there are a couple of fundamental assumptions held by cognitive linguistics

[53] As Gries (2006b: 3) maintains: "Given the overall cognitive orientation in general […], it comes as no surprise that this empirical perspective also included various experimental paradigms such as sorting (e.g., Jorgensen 1990, Sandra and Rice 1995, Bencini and Goldberg 2000), elicitation tasks (e.g. Rice 1996, Raukko 1999), priming and reaction time studies (cf., e.g. Rosch and Mervis 1975)."

that have constituted its *raison d'être* ever since the discipline came into being.[54] These are the following:

1. Priority of semantics over other aspects of meaning,
2. Lack of distinction between semantics and pragmatics,
3. The existence of iconicity relations,
4. Pervasiveness of metaphor and metonymy in language,
5. Existence of blending processes in language (Fauconnier and Turner 1996),
6. Prototypicality (Rosch 1973), along with profiling, and the conception of active zones (Langacker 1987, 1991),
7. Priority of polysemy,
8. Experiential realism as a philosophical doctrine. The embodied mind thesis (Lakoff 1987).

Out of the eight principles enumerated above, the present book directly relates to points 1, 4, 7, and 8. The study upon foreign words and phrases is a cognitive-semantic study, which means, we primarily focus on the semantic composition of forms in question (point 1). This semantic priority is given prominent place in Chapter 7 in which we provide a systematic characterization of all foreign words and phrases selected for the analysis. This semantic priority is also inextricably connected with point 4, where the existence of conceptual metaphors and metonymy is regarded as a natural mechanism via which we construe concepts in language. Thus in investigating meanings of words, we inevitably fall back on deeper conceptualisation mechanisms, as semantic representation equals conceptual representation (see Evans and Green 2006: Ch. 5). No wonder, then, the foundation for the semantic characterization conducted in Chapter 7 is conceptual and represented by the overriding metaphor LANGUAGE LAWS ARE PHYSICAL LAWS already presented in Introduction and discussed at a length in Chapter 6 (particularly section 6.9). Finally, points 7 and 8 are directly relevant in that the issue of polysemy and some attempts at resolving the ensuing methodological problems are discussed in Chapter 7 (section 7.0) and Chapter 8 (section 8.4.1). Last, but not least, the doctrine of experiential realism (point 9) is directly evoked in this book in that the model of description of foreign words and phrases finally proposed (Chapter 9) along with its attributes emerges as a result of the nature of everyday human interaction with physical environment. The description of the nature of this interaction has already been stated in the central claim of the book (section 0.2, see also Preface), and is given further detailed account in Chapters 6, 8, and 9.

Another significant factor that cognitive linguistics takes into account in the examination of the conceptual structure of lexicon relates to the so-called image-schematic patterns which are viewed as foundational for lexical concepts in language (cf. Lewandowska-Tomaszczyk 2006: 259; Gibbs *et al.* 1994; Krzeszowski

[54] See also Geeraerts (2006a: 1–28).

1997). As it is claimed, a cognitive approach to lexis and metaphor induces network kinds of relations among lexical units, which exhibit high correlations with lexical networks manifested in such big corpora as *WordNet* (Miller and Fellbaum 1991), which describe the lexicalised knowledge of reality (Lewandowska-Tomaszczyk: 2006: 277). The present study – cognitive-semantic in spirit – remains consistent with the above observations, and therefore in Chapters 6 and 9 a discussion of image schemas and their role in constituting the foundation for the network among foreign lexical units is presented. As it is discussed in Chapter 9, the model of foreign lexical assimilation will be argued as essentially grounded in the activation of two fundamental image-schematic patterns, i.e. UP–DOWN and CENTER–PERIPHERY (see particularly section 9.1.2). It is on the basis of these two image-schemas that the network model of foreign words and phrases is construed.

Let us right now look at other theoretical insights into the cognitive-semantic lexicology that appear relevant to the study in the present book. When we look at fundamental theoretical prerequisites of Cognitive Linguistics to lexicological research, we inevitably recall Theory of Markedness (Prague School). This theory basically states that things are conceived of by humans in a twofold manner, i.e. as marked or unmarked. This division is based on two factors: salience and frequency. The rationale behind the candidature for being assigned the status of marked or unmarked entity is clear and relates to the notion of salience.[55] Similarly, if we take the second frequency factor, we notice that the more frequently something occurs in the external reality the more salient it becomes. Salient and frequent things are filtered out more easily than the rest (Internet source 2, see References). This observation is particularly applicable to the conception of foreignness as undertaken in the present book, wherein salience is the measure of the degree of familiarity of a foreign word by a native speaker, whereas frequency (measured in terms of CRAC, see section 6.1) is another factor that is linked with salience. This perceptual regularity is tested in Chapter 8 where the results of survey upon word-recognition tests are presented and discussed (see section 8.6). As Geeraerts (2006c: 346) argues:

> The intractability of polysemy involves the absence of a coherent set of criteria for establishing polysemy; a more charitable way of wording things would be to say that distinctiveness between senses of a lexical item is to some extent a flexible and context-based phenomenon.

This gives an argument in favour of taking *Wikipedia* as a relevant source of lexical information in cognitive-semantic perspective.[56]

The notion of structural weight as discussed by Geeraerts (2006c: 348) is also of much help to the present study in that the most salient meanings are marked as

[55] See Geeraerts (2006c: 74) for the discussion of salience phenomena in the context of lexicological research.

[56] On the problems of polysemy itself, see sections 7.0 and 8.4.1.

meanings of a particular entry that are displayed upon 'clicking' the search button on the *Wikipedia* website. Thus, "differences in salience involve the fact that not all the elements at one level of analysis have the same structural weight […]. The basic reading, in other words, is the center of semantic cohesion in the category; it holds the category together by making the other readings accessible" (*ibid.*).

Other significant issues discussed in cognitive-linguistic studies that are particularly relevant to the investigation of lexis are found in Lipka (2002: 199–203) and involve:

– natural taxonomies and hierarchies (Cruse 1986: 145),

– Brent Berlin's studies upon Tzeltal classifications of plants as folk-oriented); the problem that became particularly relevant for the description of the psychologically 'basic level' in categorization (what Berlin called "the folk-generic level"),

– global knowledge structures, i.e. frames, scripts and scenarios, as well as inferencing processes (Brown/Yule 1983),

– cognitive and cultural models (Lakoff 1987).[57]

All of the above underlie the methodological framework adopted for the investigation of foreign words and phrases laid out in Parts III and IV of the book. Thus, insights from natural categorization are applied in the discussion of typology of foreignisms (section 6.7.3) as well as in section 7.0, where criteria of classification for semantic domains underlying foreign words and phrases are presented. As far as theories of knowledge representation are concerned (the aforementioned theory of scripts, scenarios, global knowledge structures) and the models of inferencing, these underlie the descriptive framework of foreign lexical assimilation emerging through Chapters 6–9. Without the insights from this field, we would not be able to account for the integration of naïve and expert knowledge of universe (see sections 9.2.1 and 9.2.2) structuring the model representation of foreign lexical assimilation processes in the target language.

2.2. Corpus linguistics

This study cannot by any means be labelled as corpus-based in the strict meaning of the term. As already indicated in Introduction to the book, we do not attempt at investigating the usage of foreignisms in the contemporary English language, but rather seek to explore the mechanisms governing the externally induced changes in the lexicon of a natural language. Although systematic investigation in the usage of foreignisms would make up a valuable research contribution to the field of

[57] Further lexicological research in integrated cognitive-linguistic and corpus perspective can be found in Geeraerts, Grondelaers, and Bakema (1994) and continued in works such as Geeraerts, Grondelaers, and Speelman (1999) or Speelman, Grondelaers, and Geeraerts (2003); all after Geeraerts (2006b: 30).

lexicology, it would definitely require a separate monograph to be devoted to the problem. In consequence, this book can be at most representative of a corpus-assisted approach (see Lewandowska-Tomaszczyk 2005) in which we combine statistics of frequency retrieved on the basis of the British National Corpus with the traditional introspective tools.[58]

Let us then have a closer look at corpus linguistics[59] along with its methodological notions and principles in search of direct relevance to cognitive-semantic approach adopted for the study of foreign words and phrases. For that purpose, we need to first glimpse into the source and nature of corpus linguistics.[60] As McEnery & Wilson (2004: 103) notice:

> The importance of corpora in language study is closely allied to the importance more generally of empirical data. Empirical data enable the linguist to make statements which are objective and based on language as it really is rather than statements which are subjective and based upon the individual's own internalized cognitive perception of the language.[61]

---

[58] As Geeraerts (2006b: 32) argues: "Corpus data too were used early on in the history of Cognitive Linguistics. The methodology of European studies in Cognitive Linguistics in particular has tended to be more corpus-based than the early American studies, which were predominantly introspective. The use of corpus materials […] was already part of early European studies like Dirven, Goosens, Putseys and Vorlat (1982), Dirven and Taylor (1988), Rudzka-Ostyn (1988), Schulze (1988), Geeraerts, Grondelaers, and Bakema (1994)."

[59] The significance of Corpus Linguistics as the discipline concerned with the examination of real life language usage is manifested in the following: "Linguistics is a tradition that sees language as a corpse on a slab, under the dissecting scalpel of the professional linguist. Luckily for Lancaster however, the past 20 years have seen new life breathed into this corpse. The clinical skills of the pathologist have given way to skills in understanding how real, living language is used in real life situations" (Aitchison 1994: 9 after Svartvik 1996: 3). See also Biber *et al.* (1998: 1).

The very term "corpus" is defined as "the body of written or spoken material upon which a linguistic analysis is based" (Aston & Burnard 1998: 4) or "a collection of pieces of language, selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language" (Sinclair 1996, after Aston & Burnard 1998: 4). Another eminent linguist, Lewandowska-Tomaszczyk (2005: 40) does not mention the category of text as definitional for the notion of corpus. She concludes that corpus constitutes a systematic collection of language data to be investigated, which represents an attempt at maximally objective and comprehensive reflection of linguistic reality. This meaning of "corpus" is closest to the understanding adopted in this book where the list of foreign words and phrases as based on the phonological criterion has systematically been compiled (see Introduction). Upon investigation of types of corpora listed by Aston & Burnard (1998: 11–12), it appears that a collection of foreignisms gathered for the analysis in this book resembles a genre-specific corpus.

[60] As McEnery & Wilson (2004: 2–3) claim: "While not identifying themselves with the terms corpus linguistics, field linguists, e.g. Boas (1940), and later linguists of the structuralist tradition all used a basic methodology that we can undoubtedly call corpus-based. […] The corpus was subjected to a clear, stepwise, bottom-up strategy of analysis. The corpus underpinned the methodological approach of pre-Chomskyan linguistics in the twentieth century." On the theoretical problems related to corpus linguistic study, see also Piotrowski (2003: 143–154).

[61] The importance of empirical data in Cognitive Linguistics has also been discussed in the literature. For more discussion, see Geeraerts (2006b: 23–24) and Itkonen (2003). Suffice it say

This foundational statement of corpus linguistic enterprise is fully compatible with the principles and assumptions held by cognitive linguists who are also distanced from subjective statements and instead focus on the nature of conventionality (intersubjectivity) of a linguistic construal of the external reality. Likewise, this book does not so much aim at ultimate presentation of intellectually coherent model of foreign lexical assimilation, which could be argued as purely speculative, but rather attempts to minimize the degree of such speculation by reliance on statistical data retrieved on the basis of the British National Corpus and some facts established through direct questionnaires (see section 8.6).

Another relevant terminological differentiation concerns the qualitative and quantitative corpus studies. As McEnery & Wilson (2004: 76) put it:

> The difference between qualitative and quantitative corpus analysis, as the terms themselves imply, is that in qualitative research no attempt is made to assign frequencies to the linguistic features which are identified in the data […]. Qualitative forms of analysis offer a rich and detailed perspective on the data. In qualitative analyses, rare phenomena receive, or at least ought to receive, the same attention as more frequent phenomena […], because the aim is complete detailed description rather than quantification. […] Quantitative analysis thus enables one to separate the wheat from the chaff: it enables one to discover which phenomena are likely to be genuine reflections of the behavior of a language or variety and which are merely chance occurrences.

Relating this statement to the methodology undertaken in the book, we may safely conclude that the aforementioned terminological separation does not necessarily need to apply to what constitutes the nature of the research upon foreignisms in this book. Thus, the research advanced in Parts III and IV is qualitative in that it aims at offering a rich and detailed account of the data selected for the analysis (see particularly Chapter 7), which is later integrated into the model presented in Chapter 9. However, the research may also be called quantitative in that the frequency factor is one of the crucial arguments in favour of claiming the ontological make-up of the model of foreign lexical assimilation proposed in the book. In this sense, the present research on foreignisms integrates both qualitative and quantitative approach into one. Of course, the analysis undertaken in the book does not employ the hard and fast type of statistical method in the sense of using sophisticated mathematical formulae in the investigation, discussion and presentation of the relevant data. No such ambition has ever been phrased.

that the empirical trend is not regarded as dominant in the cognitive-linguistic paradigm: "First, systematic theory comparison focusing on specific empirical phenomena is not dominant activity in linguistics. Even with the realm of Cognitive Linguistics, for instance, an analytical and critical comparison of different approaches as in Langacker 2005, Nuyts 2005, or Croft and Wood 2000 is exceptional" (Geeraerts 2006b: 26). As Geeraerts (2006b: 42–43) concludes: "There seems to exist a tension, in other words, between a broad methodological tendency in Cognitive Linguistics that considers introspection the most or perhaps the only appropriate method for studying meaning, and a marginal but increasing tendency to apply empirical methods that are customary in the other cognitive sciences."

Rather, the book tries to show that by reference to very simple tools like averaging between the two values, the final picture although essentially idealized, does not provide the reader with necessarily inconclusive (or even misleading) results. Idealization of data is after all inscribed in any quantitative description (see McEnery & Wilson 2004: 77).[62]

As Leech (1966: 73, after Svartvik 1996: 4) notices:

> Counting occurrences, in a large number of cases, is merely a laborious way of coming to conclusions one has already arrived at subjectively. More detailed quantitative analyses (requiring large corpuses and the aid of computers) can be expected to produce results beyond the insight of the native speaker.[63]

The results of a quantitative analysis undertaken in the book appear to substantially corroborate intuitions of a native speaker when it comes to their judgment of the degree of impact of particular donor languages upon the English language. Still, however, this statistics helps, via the idealization of data presented, to form different regularities (see e.g. section 8.2.4) and to formulate predictions (see section 10.2) about the possible routes of evolution of the existing patterns of influence in the future.[64]

Other common assumptions and methods adopted by Corpus Linguistics include the following:

> – The analysis is based on a corpus or corpora of naturally-occurring language which are machine-readable so that the retrieval of the search patterns is computerized.
> – The analysis is or at least attempts to be systematic and exhaustive, meaning that the corpus does not simply serve as a database of examples from which some can be chosen *ad libitum* and other neglected, but the whole (sample) of the corpus is taken into consideration so that even less frequent patterns must somehow be integrated or at least addressed.
> – The analysis proceeds on the basis of frequency lists (of words, morphemes, grammatical patterns, etc.), concordance lines in which the word of interest is shown in its natural context, and collocations, i.e. lists of tables in which for the word of interest the most frequent neighboring words are given.
>
> (Gries 2006b: 4)

[62] As Gries (2006b: 5) says: "On the one hand, corpus-based work often makes it necessary to operationalise subjective qualitative phenomena on the basis of quantification, i.e., by using frequency data from corpora. On the other hand, and this is the more important interpretation here, corpus-based studies differ as to the role quantitativity plays in the evaluation of the results. For example some corpus-linguistic studies are rather qualitative in the sense that their contribution is mainly based on which categories are observed and which are not what this implies."

[63] Emphasis mine (M.K.).

[64] Cf. Labov (1973) conception of grammar as a system of social communication which contains elements preferred by language users and those that are on the margin of frequency of usage (after Lewandowska-Tomaszczyk 2005: 10). Thus criteria of frequency do not apply only to lexicon, but are also manifested at other levels of linguistic representation, which testifies to the pervasiveness of this phenomenon in natural language.

Another significant aspect in Corpus Linguistics relevant to this study concerns the so-called level of granularity at which the study of language data is conducted (see Gries 2006b: 5). For example, "some studies take lemmas as their central focus in order, for instance, to be able to collapse individual word forms and, thus, make more general statements" (cf. Atkins, Kegl, and Levin 1988, Atkins and Levin 1995, Hanks 1996, Stefanowitsch and Gries 2003; all after Gries 2006b: 5). Therefore, the list of foreign words and phrases discussed in the present book is actually the enumeration of foreign lemmas, as this is the level at which information about frequency of occurrence is retrieved (see section 1.2). All in all, the present study does not adopt "a statistically – informed perspective." It is, as Gries (2006b: 5) notices, rather like the majority of studies which "restrict themselves to reporting frequency data and usually attribute some importance to the different frequencies with which particular categories are attested in the data and the consequences this has for the phenomenon under investigation […]. Well-known examples include Meyer (1991), Berglund (1997), Stubbs (2002)."

Lewandowska-Tomaszczyk (2006: 278) summarizes the main points that the researcher needs to observe in a reliable corpus linguistic study. The crucial issue appears to be the overriding assumption in which introspective analysis is combined with empirical investigation. At least, these are not viewed as mutually exclusive (cf. McEnery & Wilson 2004: 19). Thus the parameters employed in a corpus-linguistic approach include:

(I) the size of a language sample,[65]
(II) the number of users,
(III) objectivity of assessment expressed in the type and format of investigation,
(IV) the construal of a model,
(V) interpretation of research results on the basis of one's own introspection and intuition.

Most of the above-presented parameters appear to be observed in the present study on foreign words and phrases. We say "most" because the nature of the analysis undertaken minimizes the significance of parameter (II), as sociolinguistic implications do not constitute the core of interest in the research. Otherwise, the most outstandingly used parameters are (IV) and (V). This is because the finally emerging model of foreign lexical assimilation (see Chapter 9) turns out to be a reconciliation of both introspective aspect of research (see physical foundations for the model discussed in Chapter 6) and empirical one (see the discussion of statistically-based research results in Chapter 8).[66]

[65] As Knowles (1996: 36) says: "An important consequence of handling large amounts of data is that it enforces rigour and discipline in data organization."

[66] As Biber *et al.* (1998: 9) argue: "A crucial part of the corpus-based approach is going beyond the quantitative patterns to propose functional interpretations explaining why the patterns exist. As a result, a large amount of effort in corpus-based studies is devoted to explaining and

In discussion upon some of the caveats inherent in corpus-linguistic methodology, Lewandowska-Tomaszczyk (2005: 15) notices that a researcher may only find positive evidence of the existence of a particular form in language corpora. The negative proof cannot be directly accessed, but may be deemed as highly probable once the existence of a particular form is not witnessed in the corpus. In the present study, a number of foreign forms, although represented in LPD and consequently labeled as "foreign," are not actually represented in the British National Corpus. The conclusion is that these forms are classified as residing on the peripheries of "foreignness" and assigned the status of "unknown" forms (see, e.g., section 5.3). In other words, the lack of existence of a given foreign form in the corpus is not *sensu stricto* the proof of its absence in the lexicon, but rather serves as the basis for formulating the statement about the perception of these forms as *likely* to be unknown by the language users.

The other significant issue discussed in corpus linguistic studies concerns the problems of disambiguation of linguistic data (Lewandowska-Tomaszczyk 2005: 84–90). This is a weighty issue in the case of so-called homographs. As Aston & Burnard (1998: 9) put it: "In order to disambiguate homographs or to identify particular uses of words or structures, it may be necessary to inspect the lines in the output, classifying them individually." Unfortunately, the phonological criterion adopted for the investigation of foreign words and phrases appears to make the problem of disambiguation of forms virtually impossible to resolve. This is due to the fact that the researcher does not have much of an access to the pronunciation of the discussed forms. In this respect we rely on LPD as an authoritative source. The criterion that we base on in the examination of frequency of occurrence is, in effect, graphemic (the suggested spelling form of a foreign word prompted by LPD). This entails that the problem of disambiguation is of not primary importance to the research as the actual investigation relates to the presence or absence of foreign written forms in the British National Corpus, regardless of their semantic import in particular contexts evidenced in the corpus.

This mixed phonological-graphemic criterion of the status of foreignness of particular forms is thus auditory-perceptual in nature. The question that arises relates to the verification of the validity of the adopted criterion of foreignness. The auditory (phonological) part of the criterion has been taken for granted as fundamental for the description of foreignisms via the aforementioned reliance upon LPD. The perceptual (spelling) part of the criterion of foreignness has been subject to a two-step verification. First, the elicitation of the frequency of occurrence associated with a particular form has been conducted, and, second, this statistical information has further been confirmed through the word-recognition

exemplifying quantitative patterns." On the necessity of combining top-bottom (introspective) and bottom-up (empirical) methodology in linguistic research, see Lewandowska-Tomaszczyk (2005: 12–13). On the role of introspection in corpus linguistic research, see also Aston & Burnard (1998: 5–6, 13).

test, the format of which along with the consequent results have been presented and discussed in section 8.6.1.

It must, however, be remembered that accordingly with the reservation made by McEnery and Wilson (2004) (see Ft. 28 in section 10.1.1), the presence or absence of diacritics in the graphemic representation of a given word or phrase is not viewed as directly relevant to the research, as this aspect is usually ignored because it does not affect our judgment of a given form as foreign. On the other hand, it is argued that the essence of foreignness is manifested in the subtle intricacies regarding phonological-morphological make-up of a word. The rule applied throughout this book is that the spelling form of a given foreign word without diacritic marks is the basis for the BNC search. This is because of two reasons. One is practical and relates to the fact that not all diacritic marks are present as default symbols in the SARA search engine. The other reason is theoretical and should be viewed as compatible with assumptions about the model of lexical assimilation adopted in the book. As already indicated in section 0.2, the process of assimilation naturally encounters forces of resistance that affect formal (phonological--graphemic) composition of a given lexeme. This causes lots of foreign words and phrases to lose original diacritics present in their original spelling forms in the process of assimilation (see section 10.1.1 on formal aspects of adaptation). This fact is manifested in LPD, which lists as primary the 'non-diacritic version' of the spelling form of the foreign word. Exceptions to the rule are cases of words whose original spelling has become preserved in the process of assimilation (e.g. *fiancé*, *café*). Often the reason for this preservation stems from the potential ambiguity which may result if diacritic symbols are lost. Namely, the risk involves homographic pairs with already well established native lexical units. Such is the case with, e.g., *blasé*, *curé*, *engagé*, *exposé*, *pavé*, *protégé* where we have preserved the original spelling as the basis for further explorations (see also section 2.3 below).

2.3. British National Corpus

As the British National Corpus constitutes a fundamental framework of reference in the empirical part of the research on foreign words and phrases, some relevant information about this source must be provided. The British National Corpus (henceforth BNC)[67] is a 100 million word collection of samples of written and

[67] BNC is also referred to as a reference corpus. As Halliday *et al.* (2005: 118) notice: "Reference corpora are being used for a multitude of purposes. Reference corpora contain the standard vocabulary of a language […], we need reference corpora, the larger the better, for investigating lexical semantics. A typical reference corpus will represent what the discourse community agrees to be what a fairly educated member of the middle class would read outside of work, mostly in printed form, but also handwritten or typed; and, in principle at least, it should also contain a sample of what they would hear, in conversation, at more formal social events, or on the radio. It

spoken language from a wide range of sources, designed to represent a wide cross-section of British English from the later part of the 20th century, both spoken and written. The latest edition is the BNC XML Edition, released in 2007. The written part of the BNC (90%) includes, for example, extracts from regional and national newspapers, specialist periodicals and journals for all ages and interests, academic books and popular fiction, published and unpublished letters and memoranda, school and university essays, among many other kinds of text. The spoken part (10%) includes a large amount of unscripted informal conversation, recorded by volunteers selected from different age, region and social classes in a demographically balanced way, together with spoken language collected in all kinds of different contexts, ranging from formal business or government meetings to radio shows and phone-ins.[68]

The BNC is classified as a referential corpus. As Lewandowska-Tomaszczyk (2005: 30) points out, this fact entails that the BNC must be maximally representative of language data characterizing the state of a given language in particular place and time. Such corpus must be proportionally large to ultimately represent a general, objectivized picture of vocabulary, phraseology, etc. occurring in a given language.[69]

However successful the BNC project may be, it is important to remark some significant caveats that cast an important light upon the results of research presented in the present book. The first reservation concerns the indiscriminate reliance upon the statistics about frequency provided by the BNC (see Aston & Burnard 1998: 36–37). However, at no point throughout this study have we suggested that the statistics serve as unfailing evidence for drawing far-reaching conclusions. Rather the opposite has been implied. Thus, although the information about the frequency of occurrence of particular foreign words has led us to formulate some conclusive generalizations (Chapters 8 and 9). These have, however, been made with the necessary proviso in view, i.e. the necessary idealization (simplification) of the presented results. This "idealization" prerequisite emerges from the awareness of various caveats hidden in the BNC.

The very "all-inclusiveness of the BNC" does not preclude other errors, especially the ones connected with the aforementioned problem of disambiguation. As Aston & Burnard (1998: 38) point out:

> Quotations in languages other than English are also occasionally to be found, which may lead to confusion where they include forms which are identical to English words – for example, a fragment in German may contain many occurrences of the word *die* but have nothing to do with mortality.

---

is carefully construed, with a deliberate composition. The British National Corpus of 100 million words, compiled in the early 1990s, is a good example."

[68] Source: www.natcorp.ox.ac.uk/corpus/index.xml (ED: 01/09 2007). See also Svartvik (1996: 8) and Aston & Burnard (1998: Preface).

[69] More characterization of the BNC is given in *inter alia* Aston & Burnard (1998: 29).

Foreign words are thus regarded as some special category of unexpected homographs in the corpus: others instantiate names, abbreviations, and acronyms, as well as misprints (Aston & Burnard 1998: 38). Although, as already indicated above, the issue of disambiguation is not particularly vital in the research due to the adopted criteria of foreignness (see Introduction), the problem has appeared as significant in the case of foreign forms whose graphemic representation was isomorphic with the common English word or phrase. In such cases, the data about the frequency of such homograph could not be regarded as negligible,[70] as the final picture of the foreign word representation in the corpus would be largely falsified, which would in effect lead to formulating counter-intuitive judgments.

Other caveats are related to the composition of the BNC. As Aston & Burnard (1998: 40) put it:

> Some results may be biased by the fact that the corpus was collected at a particular time, with the result that certain 'buzzwords' occur more frequently than might otherwise have been the case. Others may be influenced by typically frequent recurrences in one or a few particular texts.

This is another 'cost' behind the corpus-linguistic study, which must be treated as inscribed in the risk of the enterprise. It cannot, however, function as an argument against the validity of the BNC project as such. It appears that the ultimate benchmark for a linguist is a commonsensical judgment as well as introspection that may serve as a 'safety valve' against making too definite statements based exclusively on the BNC statistics. Hence the aforesaid need to combine empirical and introspective tools in linguistic analyses.[71]

2.3.1. Frequency

The issue of frequency has already been alluded to in the preceding sections, nevertheless, its significance as a 'gap bridge' between Cognitive and Corpus

[70] In less conspicuous cases, the deliberate negligence of a potential distortion of statistical facts entailed in the problem of homographic pairs has been the rule and ultimately relegated to the inevitable "cost" of data idealization procedure undertaken in the research.

[71] One has to mention at this point the role of computers in corpus-linguistic studies as their growing significance introduces a new 'quality' into the discussion upon the dichotomy between empirical and introspective studies. Thus: "Without the advent of computers the approach to lexis propounded by Halliday would never have had the tremendous impact it has already had and continues to have on the field of linguistics. […] For the first time ever, linguists have been able to rely on non-impressionistic large-scale frequency data. Although the reliability of frequency studies was questioned from a relatively early stage, this did not put an end to them but, instead, merely prompted corpus linguists to gather bigger and more tightly controlled corpora" (Altenberg & Granger 2002: 4). On a comprehensive discussion of corpus-based approaches to lexicography, see Francis (1992), Sinclair (1985, 1987, 1991), Stubbs (1995), and Zernik (1991); all after (Biber *et al.* 1998: 54).

Linguistics has not yet been precisely formulated. McEnery & Wilson (2004: 82) put it as follows:

> The most straightforward approach to working with quantitative data is simply to classify items according to a particular scheme and to perform an arithmetical count of the number of items (or tokens) within the text which belongs to each classification (or type) within the scheme. […] Sometimes the classification scheme used in frequency counts may be a simple one-to-one mapping of form on to classification. This can be the case with word frequency analysis, where each graphical word form is equivalent to one type within the classification scheme.

This is exactly the approach adopted in the present book, where *a priori* selected set of foreign words or phrases is further subject to quantitative investigation in that the graphical form (type) is checked for its frequency of occurrence in the corpus. The validity of such frequency based study is that the kind of quantitative data thus retrieved is this type of information that cannot be directly accessed through introspection. Although, McEnery & Wilson (2004: 11) claim that "at times intuition can save us time in searching a corpus," humans appear to possess "only the vaguest notion of a frequency of a construct or word." This motivates a researcher to conduct a systematic investigation in the field to empirically test our often inaccurate introspective judgments regarding the status of a particular unit in the intersubjective consciousness shared by a speech community.[72]

Related issue is the size of the corpus as an adequate representation of particular forms under investigation. McEnery & Wilson (2004: 80) recall a pilot study by Biber and notice that "frequent items are stable in their distributions and hence small samples are adequate for these. Rarer features on the other hand show more variation in their distributions and consequently require larger samples if they are to be fully represented in the corpus, as de Haan (1992) has also observed." This observation appears to justify the general corpus frequency annotation for foreign words and phrases applied in the present book rather than restrict the search to a specific set of subcorpora. This justification is twofold: first, foreign words and phrases by definition do not belong within the target lexical system as "fully-fledged members," so their status as uncommon lexical units prompts the search for their occurrence in the whole of the corpus. Second, this global corpus search is also motivated by fundamental assumptions advanced in the book that view foreign words and phrases as 'visitors' to the English language conceived of the lingua franca – the planet Earth (see Chapter 6).

Frequency-based studies have already resulted in the growing number of dictionaries to be published that include this type of information into the descrip-

[72] The idea of lexical dispersion studies was to systematically investigate the dispersion of a particular unit in different texts of a given corpus, whether it is spread evenly throughout a text or occurs in small clusters (McEnery & Wilson 2004: 21).

tion of particular entries.[73] Such is LDOCE (1995) with its indication of whether a particular lexical unit belongs with the most frequent words in spoken and written English (see Summers 1996). But we also notice that there are attempts at producing dictionaries of word sense frequencies rather than word forms (West 1953). This type of research still, however, waits for some more updated challenge in corpus linguistic studies conducted at present (McEnery & Wilson 2004: 108).

A very interesting observation is made by Quirk & Stein (1996: 29) when it comes to the practice of frequency-based lexicography in the context of non-native lexical units. As the authors put it:

> R.E. Allen's splendid 1990 edition of the *Concise Oxford Dictionary,* probably – it is modestly confessed – 'the best known household dictionary' (p. vii), has the very proper aim of embracing expressions 'that are well attested in current English,' and it marks as *lit.* or *obs.* or *offens.* or *math.* or *hort.* or *NZ* or *Ind.* (etc.) items that are deemed by contrast to be restricted in currency. But no such restricted marking appears with many entries that surely are current in only a minority of 'household': for example, *burka, dojo, hajj, hajii, nacho, tzatzik*: not forgetting, of course, *glasnost* and *perestroika*. Some of these are indeed proudly mentioned in Mr Allens's preface as illustrations of his thoroughness in updating the book, but no explicit measures (such as corpus occurrence) are offered to show how the criterion 'well attested in current English' is established.

The problems discussed in the aforequoted passage bring forth what has been deemed as unresolved in lexicological studies conducted so far, namely, an attempt at systematic investigation of the relation between the category of foreignness and the category of frequency. It is hoped that Parts III and IV of this book will offer some insight into the description of the nature of this relation.

And now we reached the point where we can bring up the notion of frequency as conjoining the realm of corpus and cognitive linguistic study. This issue is undertaken by Bybee (2006: 9) who concludes:

> The topology of the cognitive organization of language is neither flat, regular, nor permanent. High-frequency words and phrases grow strong with repetition and loom large, forming looser connections with other items, while low-frequency words and expressions are less prominent but gain stability by conforming to patterns used by other items.

This phenomenon whereby the notion of prominence is correlated with the notion of frequency has more systematically been investigated and eventually labeled under the Conserving Effect mechanism, which has to do with fact that

[73] As Bybee (2006: 5) remarks: "Most of the twentieth century facts about the frequency of use of particular words, phrases, or constructions were considered irrelevant to the study of linguistic structure." This is attributed to "a major theoretical factor working against an interest in frequency of use in language" i.e. the distinction traced back to Ferdinand de Saussure (1916), "between the knowledge that speakers have of the signs and structures of their language and the way language is used by actual speakers communicating with one another" (Bybee 2006: 6).

repetition of forms strengthens memory representations for linguistic forms and ensures their better accessibility than linguistic forms whose frequency of occurrence is relatively lower (Bybee 2006: 10).[74]

In some other research (Nosofsky 1988), the insights into the relation between frequency and cognition have been discussed. It has been suggested that the so-called token frequency exerts a considerable impact on the perception of the center of a given category as well as its boundaries (Bybee 2006: 15). Similarly, as a result of some semantic categorization studies (Bybee and Eddington 2006), it was found that "the high-frequency pairs served as the center of some of the most productive categories" (Bybee 2006: 15). Thus, the links between frequency-based data (the domain of corpus linguistics) and prototypicality (the domain of cognitive linguistics, see Geeraerts 1988: 207) can be viewed as particularly strong. This is very significant for the present book where the procedure of prototypical foreign word elicitation is fundamentally grounded in a frequency-based study (see particularly sections 8.6 and 8.6.1).[75]

Frequency of occurrence itself becomes a "heuristic tool in the pinpointing of prototypes" (Geeraerts 1988: 222) or it may be regarded as instantiation of entrenchment in the cognitive system (cf. Radden 1992, Schmid 2000, after Gilquin 2006: 168). However, we must be cautious to indiscriminately link up the high frequency of occurrence with prototypicality. Examples where such correlation does not necessarily need to hold true, are provided by e.g. Sinclair (1991: 36) who "notes about common words that as a rule, the most frequent meaning is not the one that first comes to mind" (Gilquin 2006: 169).[76] As Gilquin (2006: 178) further states:

> A number of hypotheses can be put forward to explain the lack of correspondence between the literature and the corpus data. The first one is that cognitive salience (i.e. "prototypicality") is simply different from frequency in language. What comes first to people's minds may rely on principles such as the primacy of the concrete over the abstract, which are not at work (or, at least, not to the same extent) in language usage. It thus appears that what scholars may intuitively classify as prototypical model of some language phenomenon may not have a confirmation upon closer empirical scrutiny.

[74] Similar observations have been made by Mańczak (1980) who has written that "the most frequent forms of a paradigm are the most likely to resist change and to serve as the basis for change in other forms" (Bybee 2006: 10); cf. also Zipf (1935).

[75] As Aitchison (1996: 55) notices: "Obviously, frequency of usage is likely to have some effect: in California nectarines and boysenberries are commoner than mangoes and kumquats, so it is not surprising that the former were regarded as 'better' examples of fruit than the latter. However, the results could not be explained away solely on the basis of word frequency. On the furniture list, rare items of furniture such as *love seat*, *davenport*, *ottoman* and *cedar chest* came out much higher than *refrigerator*, which is a standard part of every American household […]. So people genuinely feel that some things are better exemplars of a category than others, a feeling which is not simply due to how often one comes across the word or object in question."

[76] See other studies that support this lack of correlation, e.g. Roland and Jurafsky (2002); Nordquist (2004), after Gilquin (2006: 169). For some opposite views on the matter, see Kemmer (2001) and Stefanowitsch (2001).

Gilquin (2006: 180) finally concludes:

> In view of the results of this and other studies, it looks as if prototypicality is perhaps best described as a multi-faceted concept, bringing together (1) theoretical constructs found in the cognitive literature and relying on deeply-rooted neurological principles such as the primacy of the concrete over the abstract, (2) frequently occurring patterns of (authentic) linguistic usage, as evidenced in the corpus data, (3) first-come-to-mind manifestations of abstract thought, as revealed through elicitation tests and (4) possibly other aspects that contribute to the cognitive salience of a prototype.

Certainly, the notion of prototype[77] has become some sort of "catch-all notion" (Wierzbicka 1985: 343) in that its understanding has in many cases substantially departed from the original conception held by Rosch (1977). This departure has led to the situation in which diverse kinds of phenomena have been subsumed under the label of "prototype" (Geeraerts 1989: 606). This has some vital consequence for both cognitive linguistic and corpus linguistic studies. As Gilquin (2006: 159) claims:

> In particular, cognitivists tend to consider the prototype as the cognitively most salient exemplar, while corpus linguists often equate it with the most frequently-attested item (cf. Stubbs' [2004] equation of "prototypical" and "high frequency" exemplars). Most of the time, the often implicit assumption is that the two concepts coincide with one another. Yet some voices have been raised to claim that corpus linguists and cognitivists examine different things when they study frequency and salience, respectively.

The conclusion that emerges from the discussion is that prototypicality (salience) of a given linguistic feature does not necessarily be correlated with its high frequency in language. This is precisely the claim that is postulated in the context of the research on foreign words and phrases in the book. As the results of frequency-based examination in Chapter 7 show, the category of foreign words and phrases can be pretty well defined with respect to its center and periphery (see sections 8.6 and 8.6.1, Chapter 8). What emerges as crucial from the discussion is that prototypicality of a foreign word or phrase is not bound up with its high perceptual salience in the corpus, but rather is associated with the frequency value that constitutes an averaged count between the highest and lowest extremes of the continuum. Thus, the novelty proposed in the book is that the formulation of the statement of prototypicality of foreignness can be approximated via relatively simple mathematic set of calculations, where a particular value obtained from such calculation (CRACn3 value, see section 6.1) serves a 'yardstick' for indicating the degree of foreignness of particular forms. This is, in turn, largely corroborated by the word-recognition test (section 8.6.1), where quantitatively collected information is tested for its perceptual validity. Such conceived prototypicality makes up a blend of both quantitative (frequency) and qualitative (perceptual salience) factors determining its nature.

[77] For the notion of prototypicality, see Rosch (1975), Heider (1971, 1972), Fillmore (1977: 68–69), Taylor (1989), Givon (1986).

PART II
STUDIES IN FOREIGN LEXIS

Overview

Part II of this book opens up a discussion on foreign words and phrases as a research problem in lexicological studies. The Chapter 3 discusses various facts and myths connected with the reception of foreign words and phrases by language users. The perspective taken is thus sociolinguistic. Chapter 4 gives a close-up on the studies conducted so far which bring up the issue of foreign lexicon as an integral part of their research interest. This chapter, rather unsurprisingly, will thus briefly discuss major related lexicological (lexicographic) publications that came out both in Great Britain and the United States. Finally, Chapter 5 attempts to find a solution to the problem of categorization of non-native lexis from the perspective of the recipient language. The ambition of this chapter is to postulate criteria that might bring some order into the typological framework of foreign vocabulary in lexicological research.

Chapter 3
SOME MYTHS AND FACTS ABOUT FOREIGN LEXIS

3.0. Foreign lexical studies. Preliminaries

The issue of the place and role of foreign[1] lexis in the English language has been the subject of many debates (see, e.g. Kuźniak 2007). The reason why this issue raises so many controversies stems from the position of English as a lingua franca of the contemporary global communication (see Dirven and Verspoor 2004: 236, Mańczak-Wohlfeld 2006).[2] A natural corollary of this fact is that a wealth of monographs and articles have been published on the role of English as a donor language from which many smaller languages borrow (see, e.g., Görlach 2003, 2007).

However, as the discussion in Chapter 4 shows, not so many publications (besides dictionaries) have as yet been devoted to a systematic description of foreign lexical element in English as a recipient language.[3] This strange observation was made by Crystal (2007b: 59) who said that English is a regular borrower that records over 350 languages as contributors to the expansion of its lexical system.[4] Crystal, in a characteristically witty manner, indicates the problem in that "when one language 'borrows' from another, it does not give them back. 'Steals' would be more appropriate. But whatever we call the behavior, the consequences are evident" (Crystal 2007: 59). I believe that 'stealing' is no more fortunate a proposal as

[1] The understanding of the term "foreign" in the present chapter is more general that the one that constitutes the subject matter of the study in the present book. The term "foreign" is to be interpreted as equivalent to "non-native" thus subsuming lexical categories at various levels of adaptation to the receptor language (see section 5.3).

[2] As Crystal (2007b: 59) observes: "In the twentieth century, as English became an increasingly global language, there was a huge reverse movement in the direction of borrowing. All over the world, languages found themselves inundated with English words." On the other hand, Winford (2005: 29) notices that "most English speakers would be surprised to learn that 75 percent of the words in their language were 'borrowed' from other languages during the course of its history," and continues: "A great deal, perhaps the majority of lexical borrowing results from only marginal contact with other languages. Such contact may be due to travel, exploration, or conquest or it may be due to exposure to the donor language in the mass media, foreign language instruction, and the like. Loveday (1996) refers to these as settings involving 'distant' contact with the external language" (Winford 2005: 31).

[3] "Of course, borrowing in the opposite direction, from subordinate to dominant language, also occurs, though not usually to the same degree" (Winford 2005: 34).

[4] For Mańczak-Wohlfeld (1995: 13) borrowings belong to one of the three fundamental processes of lexical innovation besides coinage and meaning extension – the mechanisms that relate to the already well entrenched lexical unit.

Foreign words and phrases in English, 2009

it implicates that the donor language has been deprived of its own original items in favour of the donor language. This does not happen either. As the analysis in Parts III and IV will show, we need to reformulate our terminology connected with the processes of borrowing[5] and search for some framework of reference in other disciplines of science, like astrophysics (see particularly Chapter 6). Setting this issue aside for the time being, we notice what is fundamental to this book that Crystal does not discriminate between loanwords (borrowings) and foreign words. Foreignness is reduced to a quality possessed by loanwords. It does not have an existence of its own (see Kuźniak 2009b on the opposite view).

What is more, Crystal states, which on the other hand is in compliance with the basic tenets advanced in this book, namely that most of loanwords (borrowings) have already lost traces of foreignness due to extensive contact between people (Crystal 2007b: 60). This loss of the trace of foreignness is discussed in Chapter 6 (section 6.7.3). The key word in Crystal statement is "most," which implies that some minor portion of non-native vocabulary actually displays this feature. This is particularly weighty for this book's research on foreignisms in that it undertakes the problem of this minor section of lexis. This, as the analysis will hopefully show, have some lasting methodological and model theoretic consequences. First, some significant criteria of typologization of this section of vocabulary are proposed, and second, some more general characteristics constituting foreign linguistic assimilation in the system of the receiving language can be further indicated.

The common reason for the implementation of non-native element into the target lexicon includes "the asymmetry in power and prestige of the languages involved" as well as the intensity of contact between the donor and receiving language (Winford 2005: 34).[6] On more pragmatic grounds, this implementation is effective when "a word is needed to give a name to an unfamiliar animal, thing, or cultural phenomenon" (Katamba 2006: 138).[7] The same author also quotes another factor behind importing foreign words, i.e. identity. He says:

> Language is much more than simply a means of communication. It is also a badge that we wear to assert our identity. By using a particular language, bilingual speakers may be saying something about how they perceive themselves and how they wish to relate to their interlocutor.
>
> (Katamba 2006: 139)

Bauer (1998: 16) gives another 'mundane' kind of motivation behind the use of foreignisms, i.e. personal prestige of the speaker:

[5] See Otwinowska-Kasztelanic (2000: 21) for some structuralist accounts of the processes of borrowing, e.g. Bloomfield (1933), Weinreich (1964) or Haugen (1968).

[6] A similar observation is made by Weinreich (1953: 56) and Mańczak-Wohlfeld (1995: 18).

[7] See also Mańczak-Wohlfeld (1995: 18).

> Sometimes people borrow words for reasons of prestige. In these cases it might fairly be claimed that there is no real need for the borrowed word or phrase. In such cases the use of the borrowed word may tell us more about the speaker or writer than about the state of the language. Words and phrases in this category are sometimes (though not always) used consciously to impress or to display learning. […] For people who speak the donor languages, some of these expressions undoubtedly express cultural values as well as their purely linguistic content, and are thus felt to be more effective than their English translations.

Also foreignisms are said to function as euphemistic devices for rendering meanings that would otherwise cause embarrassment or annoyance if we use words of one's own language (see Katamba 2006: 142). However, as Bauer (1998: 15) remarks: "Some language communities, such as Iceland and France, make positive attempts not to borrow words from outside, this sometimes being seen as demeaning to the borrowing language. Other communities are much more relaxed about this." We relate to this issue in section 3.1.

The above-mentioned reasons can be classified as language-external. However, Mańczak-Wohlfeld (1995: 18) lists also a number of factors that promote the introduction of non-native vocabulary, which she classifies as language-internal. These are the following:

> 1) A low frequency of use of some native words. This causes them to become obsolete and hence the need to replace them with borrowings.
> 2) The occurrence of homonymous pairs, which leads to the substitution of one homonym by a foreign lexical unit.
> 3) The loss of expressiveness of some native words, which contributes to the influx of foreign words with a clearly encoded expressive load.
> 4) Insufficient variation of semantic fields in the receiving language in relation to a donor language.
> 5) Negative associations evoked by some native lexemes.

Quite controversial appears, however, another division of non-native words quoted by Mańczak-Wohlfeld (1995: 19), where she identifies two types of foreignisms according to the criterion of their usefulness. She divides the vocabulary into "necessary" and "redundant."

"Necessary" type includes:

> 1) Exoticisms, i.e. expressions connected with the culture of the donor language,
> 2) Names of referents or concepts non-existent in the target language,
> 3) Internationalisms, which include words that already function in many languages.

The other type is referred to as "redundant" and includes the introduction of non-native words into the target system as motivated by fashion or snobbery caused by the supremacy of culture or technology of the donor language.[8] I argue

[8] Morawski (1992: 80) uses the term *nadwyżka leksykalno-semantyczna* (Eng. lexical-semantic surplus) to refer to this type of vocabulary (after Mańczak-Wohlfeld 1995: 19). Other negative views about the use (abuse) of some non-native words (e.g. English) in Polish are quoted by Mańczak-Wohlfeld

that the division is controversial because it seems to be based on some non-explicitly defined normative criteria of what counts as useful or redundant in language. The scepticism behind the validation of normative principles underlying the judgement of words on an axiological basis stems from the principles adopted by Cognitive Linguistics favouring a usage-based model of the description of language. This practically entails that a language user – a rational being – is driven in his/her linguistic choices by various needs that must be viewed as essentially justified (see Kuźniak 2009a). This is another dimension of the philosophy of humanism inherent in principles and approaches constituting Cognitive Linguistics. Therefore, we agree with Hope (1963: 38, quoted in Mańczak-Wohlfeld 1995: 19) who claims that every reason for borrowing should be regarded as substantiated.

It is true that the protection of a language identity is an important issue that must be given due care and attention. However, I believe that this protection should implicate positive rather than negative measures to be taken. These should involve promotion of one's own language rather than frowning upon, or what is even more absurd, prohibiting the use of foreignisms by means of laws and regulations. The speech community system is sufficiently conservative to hamper the unjustified spread of foreign terms in its own language towards their conventionalisation. I would not be, therefore, so much concerned with the use of *sale* for the Polish word *wyprzedaż* or *shop* for the Polish word *sklep*, because these forms will only be accepted in language if the speakers find sufficient ground for their validation as well-entrenched units of their lexicon. Once such validation is found, there is nothing for the linguist to argue against, but concede to the overwhelming power of convention (cf. Mańczak-Wohlfeld 2006: 80).

Extensive information on the motivation of use with the emphasis on borrowings from English in major European languages can be found in Görlach (2007). In the introduction to the book, Görlach (2007: 11) notices that the special status of a loanword as 'foreign' determines its usage in the target language, which often stands in contrast with everyday usage of that word in the donor language. The areas of vocabulary in which loanwords can frequently be found relate, according to Görlach, to at least three sections:

a) Technical terms from fields like economics, medicine, and especially computing.
b) The jargon of pop music, drugs, etc.
c) Slang.

What follows Görlach's insightful introduction is the survey of major theoretical and empirical input from linguists dealing with the influence of English on major European languages. For this moment, the survey will be restricted only to the insights directly relevant to the issues of research upon usage of

(2006: 76). These relate to the abuse of interjections like *wow, yep, yeah* by young Polish people. The matter appears particularly significant as these words relate to the subtle emotional sphere of language, the sphere that is infected by foreignisms (see Kowalik 1999: 4). See also Grybosiowa (2000) who claims that foreignisms serve only the group identification function (after Mańczak-Wohlfeld 2006: 76).

English borrowings in the receiving languages. As regards German, for example, Busse and Görlach (2007: 28) indicate that detailed sociolinguistic and stylistic research in the field is to a large extent missing due to the dynamically changing data whether lexical, stylistic, social or attitudinal. Nevertheless, some general observations regarding the usage of English borrowings in German relate us to the following areas:

> 1. 'Technical' (restricted to the terminologies of sciences, technologies, and other jargons); these tend to be infrequent, incompletely integrated, written, and attitudinally neutral.
> 2. 'Colloquial/slang.' This tends to be frequent only in youth language, journalism, and advertising and more typical of spoken use; the degree of integration is somewhat unpredictable, with lack of competence in English in some speakers being partly compensated for by the fashionable prestige of near-English pronunciation.
>
> (Busse and Görlach 2007: 28)

As far as the Dutch language is concerned, Berteloot and van der Sijs (2007: 52) observe that the lexical domains mostly affected by English borrowing predominantly relate to technology and music, but can also be found in advertising, films and books. The authors notice that the attitude of Dutch people towards English is rather positive, but Dutch does not seem to be as affected by English as we might think one of the reasons being that English "is often used, but not always understood" (Bertellot and van der Sijs 2007: 52).

Usage of English borrowings in other languages will now, for convenience, be compacted into Tab. 7 with the following categories addressed: a) name of leading researcher; b) language; c) areas of usage and d) attitudes of native speakers of the receiving languages towards borrowings. The data presented in the aforementioned table are adapted on the basis of Görlach (2007).

Tab. 7. *Usage of English borrowings in selected European languages* (after Görlach 2007)

Source	Language	Usage	(Predominant) attitudes
Graedler (2007: 76)	Norwegian	Culture; entertainment; sport, commerce (oil industry), slang	Positive (especially from young people)
Kvaran and Svavarsdóttir (2007: 102)	Icelandic	Aviation, computer science, food, clothing, pop music, slang	No explicit information
Humbley (2007: 120)	French	Technical sectors, colloquial, informal and journalistic contexts	Rather negative (among intellectual elites)
González (2007: 145–46)	Spanish	Mostly in the written language. Many borrowings have a learned character. If used in the spoken language, the usage is mainly restricted to music, sport and drugs (young people)	No explicit information

Tab. 7 – cont.

Pulcini (2007: 162)	Italian	Scientific and technical, youth language	No explicit information
Constantinescu, Popovici and Ştefânescu (2007: 185–87)	Romanian	Journalese, youth language, computer science	Unequivocally positive
Maximova (2007: 207)	Russian	Science, technology, and professional jargons; colloquial or slang words, journalism and advertising	No explicit information
Mańczak-Wohlfeld (2007: 226)	Polish	Technical or colloquial fields	No explicit information
Filipović (2007: 236–37)	Croatian	Technical terms (mostly written) or youth language and journalese	No explicit information
Alexieva (2007: 255–56)	Bulgarian	Technical terms (aviation and maritime terminology), youth language, journalese, lexical gap fillers	Both positive and negative attitudes can be found
Battarbee (2007: 272–73)	Finnish	Markedly lesser impact of Anglicisms in the area of science and technology compared to other languages. Otherwise the same general observations can be made as determined by transnational trends	No explicit information
Farkas and Kniezsa (2007: 287)	Hungarian	Sports, Medical language, fashion, cinematography, advertisements, computation techniques, banking and trade, brand names, youth language (music and rock)	No explicit information
Ködderitzsch and Görlach (2007: 299)	Albanian	Marginal usage	No explicit information
Stathi (2007: 323–24)	Modern Greek	Technical or colloquial or slang registers	As far as technical register, Anglicisms are perceived as attitudinally neutral

3.1. Attitudes to foreign words and phrases

Attitudinal factors pertinent to the perception of foreign words and phrases in English can be conveniently discussed in terms of image schemas as understood by Johnson (1987). The study of attitudes, however, also implicates a valuative (axiological, see Krzeszowski 1997) component. Combining Johnson's findings about image schemas and Krzeszowski's findings about the axiological dimension of image-schematic patters, we argue that foreign words and phrases can

be viewed as alien units in the target system, which are correlated with the FAR pole of NEAR-FAR schema (see Krzeszowski 1997: 117–118). This is due to their inherent property of foreignness. As such, then, they exhibit absolutely negative valuation in that an axiological (+)/(–) parameter built into the pre-conceptual schema profiles NEAR pole of the schema in a positive way, whereas FAR pole is profiled in a negative way. We can conclude that, prototypically speaking, foreign words should actually arouse negative than positive attitudes among speakers of the receiver language community. This rather negative reception of foreignisms is also accounted for by reference to the notion of frames, i.e. "that summon rich knowledge structures, which serve to call up and fill in background knowledge" (Evans and Green 2006: 11, see original works on 'frame' by Fillmore 1982).

Taking this into consideration, it is easier to understand why the use of foreign word and phrases often results in so much resistance from the so-called language purists, who believe that the influx of foreignisms into their own language brings forth the images of the plague or other kind of virulent infection. Besides the institutionalized forms of protectionism from the influence of foreign words, there is also the pragmatics of everyday talk in which the mention of a foreignism frequently causes the interlocutor to have a feeling of disrespect for the speaker. This is because the use of a foreign word or phrase is to a large extent associated with snobbish behaviour as a principal emotive aspect governing the act of communication[9] (see section 3.0).

As Tab. 7 in section 3.0 indicates, however, attitudes to foreign words and phrases systematically vary with the age of speakers, where younger generation is positive or at least neutral to non-native lexis and older generation being rather negative about foreignisms. This sociolinguistic observation however does not appear to cancel what has already been suggested, that is the absolutely negative reception of foreignism, viewed globally. All seems to crucially depend on the definition of a foreign unit. It is tentatively suggested that, given a phonologically-based definition of a foreign word adopted in the book, we may safely stipulate that an act of imitation of the pronunciation of the language of origin leads to rather a negative reception from the hearer. Young people may exhibit positive attitudes to foreignisms, when these are conceived of as already entrenched units in the lexicon at least as far as the process of their phonological adaptation is concerned (see section 10.1.1). Thus, although, the presence of the pronunciation of the language of origin constitutes a prerequisite to classifying a particular form as foreign in the book, the pragmatic motifs that underlie its usage are claimed to be substantially different, depending on the degree of assimilation of a unit in the

[9] As Gooden (2005: vi) remarks: "There are less practical reasons for employing foreign terms. One might almost say less creditable ones. We use words to impress as well as to communicate and choosing an appropriate and exotic term may be a short cut to impressing others. If you have chosen the right words and the right audience, that is, you also run the risk of being misunderstood or thought pretentious."

target lexicon. In effect, the regularity observed is that the degree of assimilation of a particular foreign word is directly proportional to negative reception of a native pronunciation variant of that word (see the aforementioned section 10.1.1). This may be illustrated (Fig. 8) by the intersecting horizontal axis of assimilation and vertical axiological axis of acceptance of the native pronunciation of a given form. Symbol (+) refers here to a positive reception; (0) to a neutral one, and finally (–) mark designates a predominantly negative attitude towards the use of the native pronunciation variant of a given foreign unit. It must be emphasized at this point that whether these intuitive observations are true must certainly be subject to a more systematic investigation – a good research problem for a separate monograph.

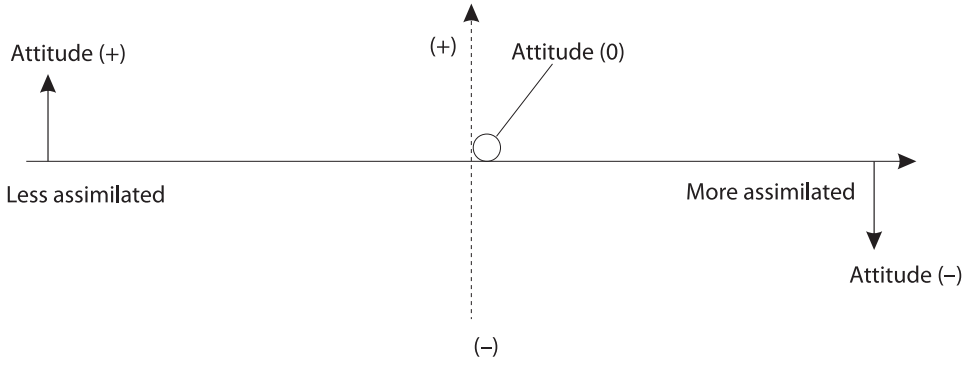


Fig. 8. *Attitudinal and assimilatory aspects of a foreign word (a phonological perspective)*

Another perspective (somewhat more optimistic) from which the study of attitudinal factors can be approached relates to the usage of foreignisms based on the criterion of a speaker's individual competence of the receiving language rather than the above-presented notion of conventionalization:

> People unacquainted with foreign expressions popular in English are often inclined to think that those who scatter such phrases in their writing or conversation are pedantic. There is no doubt that this practice can be overdone, and that at times it may be in bad taste. However, to those who are acquainted with subject being discussed, the use of such terms is often a short cut to making a meaning clear. […] Similarly, to people with literary interests certain expressions may brighten a conversation by their aptness or sum up a concept that might otherwise require several sentences for a clear explanation.
>
> (DFPA: XV)

Interesting is also the observation made by Foster (1981: 72), who attributed to the English people the quality of liberalism associated with a free trade policy when it comes to the reception of foreign words and phrases. He says: "English speakers, who seem to believe in a species of linguistic free trade and argue that

if a term of foreign origin is useful it should be put to work forthwith regardless of its parentage" (Foster 1981: 72). Similar remarks are made by Gooden (2005: v) who claims that "in the basic and simplest sense of the phrase, language is a free market. […] Among world languages, English has some claim to provide the freest market of them all, not only because it is compounded from a variety of sources but also because it has made itself open to linguistic influences from around the globe." This practical rather than ideological approach to foreignisms makes the English language highly adoptive with the resulting consequence that "the most nondescript expressions from foreign languages and set them to work with a minimum of fuss" (Foster 1981: 72).[10]

Foster (1981) also provides a concise survey of the most interesting loan words from various foreign languages in the English language. In his analysis, he focuses on the usage of the borrowed term in the target language as well as its meaning in the language of origin. He discusses the influences from French, German (to which he devotes most of the pages in the chapter in question), Danish (Dutch in connection with Afrikaans), Italian, Spanish, Russian, Czech, West African languages; Chinese, and Japanese. His analysis, which is conducted on pages (72–116), is mostly confined to the investigation of maximum of 10 lexical items from the languages that most powerfully shape the English lexicon of today to one or two items in the case of less powerful languages. His analysis cannot thus be claimed as highly comprehensive.

All in all, stereotypically speaking, the common feeling about foreignisms is rather negative, which is connected with the implantation of foreign cultural practices and patterns of thought into the receiving language. In effect, borrowing foreign words must be done with caution because it may pose a threat to the identity of the recipient language (Crystal 2007b: 60). As Crystal (2007b: 62) continues, loan words become even more problematic when they are said to replace items already existent in the recipient language. We come back here to the implicitly indicated problem of "redundant" borrowings, some critical comments on which have already been presented above. Yet as Gooden (2005: v) rightly points out, language is in constant change so the process of borrowing words from other languages is a daily occurrence. The main barrier the foreign word needs to overcome to be fully assimilated in the target lexicon is, however, that the speakers of the receiving language consider that foreign unit useful enough to be validated as part of their everyday communication[11] (cf. Gooden 2005: v).

[10] "Some influential figures like Sir John Cheke (1514–57) robustly resisted the influx of what came to be known pejoratively as INKHORN TERMS, i.e. words imported from Latin and Greek and anglicized in order to give gravitas to one's discourse" (Katamba 2006: 140–141).

[11] Gooden (2005: vi) poses the following questions: "Is there something lacking in our home-grown English? Why do we want to use foreign expressions? The simplest answer is that we need them."

Chapter 4
SURVEY OF SELECTED LEXICOGRAPHIC RESEARCH IN FOREIGN LEXIS

4.0. Selected literature of foreign words and phrases in English

The following sections discuss a portion of significant literature on foreign words and phrases in English. The word "portion" appears to be the most appropriate in this place as reaching all relevant publications has proven virtually impossible. It is believed, however, that the following survey lists the most significant contributions to the research in this area of the English language study.

4.1. Major dictionaries

Kuźniak (2007) surveys only a part of lexicographic research on foreign[12] words and phrases that has so far been published. These, however, constitute the foundation for the approach taken to foreign words and phrases adopted in the present book. Of particular significance is *Oxford Dictionary of Foreign Words and Phrases* (first edition [1997], henceforth ODFWP I, and second edition [2008], henceforth ODFWP II). This significance of these dictionaries lies in the fact that they were published in the UK by Oxford academic circles, the source of whose scholarly inspiration is also the BNC – the reference corpus from which statistical information is drawn in the analysis presented in Parts III and IV. The other dictionary, i.e. *A Dictionary of Foreign Words and Phrases in Current English* (henceforth DFWPCE) offers one of the most significant contributions to the description of the foreign lexis in English. This contribution is manifested in the so far most detailed lexicographic introduction to the presented material (see section 4.2). In this sense, then, the aforementioned publications are classified as "major." For a start, the ontological status of "foreign" words and phrases as it is described in these dictionaries is presented. As it appears, a part and parcel of the discussion upon "foreign" words and phrases hinges on establishing the criteria of what classifies as a "foreign" word.

[12] The word "foreign" in this chapter is viewed as equivalent to "non-native." See the typological distinctions discussed in section 5.3.

Foreign words and phrases in English, 2009

In Preface to ODFWP I, the criteria in question primarily relate to spelling, pronunciation and grammatical aspect of the marker of plurality as formal exponents of the "foreign." Another criterion relates to the typographical convention of italicization, which as ODFWP I concludes "is a helpful but not infallible clue to the extent of assimilation, the rule of thumb being that the more 'foreign' a word is felt to be the more likely it is to be distinguished in this way." Last but not least, the phonological criterion is evoked where the retention of accents emerges as another indicator of awareness of the foreign origin of a word.

ODFWP II (ed. A. Delahunty) constitutes, as mentioned above, a second edition of ODFWP I (ed. J. Speake) and can be briefly described as its natural continuation. The best indication of this continuity is publication of Preface to the first edition along with a very brief Preface to the second edition in this dictionary. At the same time, all of the theoretical remarks about the nature of foreignness of non-native vocabulary presented in the first edition are consequently upheld in the second edition. Delahunty thus briefly in a few general sentences repeats the significance of the English language as one of the biggest 'borrowers' in the world. Moreover, a few words are spared to discussing French as the most significant source of lexical imports in English. What differentiates ODFWP I from ODFWP II is that the latter work presents an updated account of the foreign lexis compared to the former publication, given ten years of time gap between the first and second edition. Thus forms marked as foreign in ODFWP I are no longer recorded in ODFWP II due to naturalization processes, and conversely, many words and phrases absent in the first edition of the dictionary have been added, as these gained some popularity in English during the time. Delahunty gives the following examples: *issei, reiki, doosra, jalfrezi, bruschetta, macchiato, jilbab, niquab, catenaccio*, or *galactico*. It appears, however, that both ODFWP I and ODFWP II argue in favour of postulating the ph+onological criterion as one of the most significant benchmarks against which the judgment of the degree of foreignness can be made in language.

In view of the above-quoted remarks, it seems quite natural to refer to some specialist pronunciation dictionary as an authoritative source for further lexical research, given the prominence of phonological criterion as the exponency of foreignness of non-native lexis. LPD (p. xxii), thus, identifies a lexical category as "foreign" in that, first and foremost, its pronunciation in the language of origin along with the potentially existent anglicized variant is provided. Without much theorizing upon the status of "foreign" words and phrases, LPD (*ibid*.) gives practical reasons for this:

> The obvious one of giving the curious user information about their foreign-language pronunciation, and the less obvious one that those speakers of English who also have some knowledge of the foreign language in question may well pronounce such words or names in a way that imitates the phonetics of the foreign language.

It seems that the underlying principle of categorization in LPD relates to the criterion of native speaker's competence and their resulting judgment upon what constitutes the "foreign" in the vast repertoire of lexical items at hand.

The aforementioned DFWPCE seems, however, to be most 'preoccupied' with finding adequate tools for the objective description of "foreign" words and phrases. This search for the objectivity of description that enables the readers/speakers distinguish foreign from non-foreign lexical items is much in keeping with the traditional Aristotelian view of categorization that admits of no gradation, but instead favours the equal status of category members, categories themselves being exhaustive and discrete. This is particularly noticeable in the following quotation (p. 3): "In this dictionary I have tried to supply the needs of the general reader in search of a single work of reference which will explain at least the majority of the foreign words and phrases likely to be encountered in current English."

DFWPCE (p. 8–11), thus, presents criteria by means of which "foreign" words can be distinguished from those which are "not foreign." These are divided into positive and negative ones.

> Positive criteria:
> – The use of italics in printing,
> – The use of accent-marks or other diacritic signs,
> – The presence in the accepted pronunciation of sounds never heard in native English words, e.g. use of French nasal vowels or German front rounded vowels like ö or ü,
> – A different correspondence of sound and symbol from the normal correspondence in native English words, e.g. rouge, whereby the use of 'ou' to represent the sound usually spelt 'oo,' and the use of 'g' to represent the sound normally heard only in such words as "pleasure" and "invasion," mark the word out as "foreign,"
> – The use in English of foreign feminine or plural forms.
>
> Negative criteria:
> – The existence of derivatives formed in accordance with English rules,
> – The use of a different spelling from that in the language of origin,
> – The existence of a different meaning from that in the language of origin.

Helpful and convenient as the above-presented criteria appear to be for the description of a not insignificant portion of words and phrases, on the same pages the remark can be found that "the application of these criteria is difficult, particularly since the results derived from the negative criteria often seem to conflict with those derived from the positive criteria" (DFWPCE p. 13). The dilemma that the search of the objective tools of categorizing lexical items as "foreign" may appear practically infeasible is more stressed in the following passage: "Nevertheless, there will remain a substantial proportion of words to which none of the criteria is applicable. For such words there is no alternative but to apply a wholly subjective criterion: a word is 'foreign' if it feels foreign to the speaker or writer who uses it" (DFWPCE, p. 14).

This so much theoretically unwanted but so much psychologically valid criterion of the subjective judgment of a native speaker upon what counts as "foreign" is indicated earlier in DFWPCE (p. 8) where "in point of fact, words of foreign origin form a spectrum graduating imperceptibly from words like *faith* at one end, the foreign origin of which would be obvious to the professional student of language to words like *éclat* which no one would consider anything but 'foreign,' at the other." It appears, then, that a native speaker's subjective judgment as to whether the word or phrase counts as "foreign" depends largely on the kind of linguistic competence that the language user represents. Thus the more comprehensive the native speaker's competence is, the larger the spectrum of "foreign" words and phrases he or she has. In consequence, the more reliable judgment of the status of a word or phrase as "foreign" such a native speaker provides. The issue of the subjectivity as a relevant criterion for the evaluation of the lexical item as „foreign" is tackled in the following section.

4.2. Other dictionaries of foreign words and phrases

The following is a more detailed account of various dictionaries of foreign words and phrases that appeared both in the USA and the UK. Below there is the list of available books discussed according to the year of their publication. The survey is divided thematically into 4 parts: Motivation, Source and Scope, Status, and Criterion. "Motivation" component of the survey lists quotations from a given dictionary that treat about reasons why foreign words and phrases are used. "Source and scope" component discusses any indication of the role of donor languages in contributing to the lexis presented in the dictionary as well as the thematic range to which the selected foreign words and phrases are restricted. "Status" part of the survey elicits any attempts made by lexicographers at pondering over the nature of a foreign word as a lexical category. Finally, "Criterion" section of the following survey shows what particular criteria have been chosen by lexicographers in excerpting foreign words and phrases discussed in a given dictionary. If there is no information given relevant to any of the above-presented parts, "no mention of this issue in the dictionary" phrase is provided below.

Dictionary of Foreign Terms (DFT 1934 [1963])

I Motivation
No mention of this issue in the dictionary.

II Source and scope
"Foreign words and phrases from more than fifty languages, including not only foreignisms in familiar use but those which occur in literature, law, logic, phi-

losophy, science, music, fine arts, diplomacy, literary criticism, etc., with special reference to those foreign expressions that have gained currency since the World War or that have acquired new meanings and connotations" (p. v). As far as the issue of scope itself is concerned, DFT notices that "the foreign terms made use of by English and American writers are as numerous and diverse as the subjects on which they write and the countries they seek to portray" (p. vi). DFT also makes, what we might call, quite an obvious remark, i.e. it correlates the number of the foreign words and phrases discussed in the dictionary to its size, following the simple rule: the larger the dictionary, the more words are discussed.

III Status

The status of a foreign word as part of the landscape of the receiving language system is confined in DFT to the process referred to as naturalization. However, instead of obtaining conclusive answer, we are left with a query: "When does a foreign word cease to be foreign; in other words, when may it be said to have become truly naturalized?" (p. vii). As the argument further goes, "the naturalization of words is an undocumented process unaccompanied by the forswearing of old allegiance or the avowal of new loyalty." Finally, DFT attempts at relating naturalization of a given unit to its global presence in the English-speaking communication system:

> Now when a foreign word becomes known and used throughout the English-speaking world, it is unquestionably naturalized. On the other hand, when its use is restricted to certain geographical areas and it is not generally understood outside those localities, its claim to rank as an English word is not apparent.
>
> (DFT: viii)

IV Criterion

As for the criterion of selection, DFT emphasizes an arbitrary decision about the inclusion or omission of a word or phrase:

> This work presents in one volume a varied assortment of foreign expressions such as occur more or less frequently in literature, periodicals, and the daily press. […] In preparing this book, I have introduced many such terms into the vocabulary, preferring to treat them as foreign, even though all dictionaries may not agree with me.
>
> (DFT: v)

Dictionary of Foreign Words and Phrases (DFWP 1950, 1986)

I Motivation

DFWP (p. 3) in searching for the growing use of foreign words and phrases in everyday talk refers to the fast-advancing modern communications technology that has caused an unprecedented intensity of contact between people speaking different languages.

II Source and scope
DFWP says the following in the matter:

> They (foreign words, MK) turn up in newspaper articles written by foreign correspondents, in novels and plays with foreign settings, in advertisements of perfumes, lingerie, or wearing-apparel, on sheets of music or program notes, in lectures and in sermons. There is no escaping them.
>
> (DFWP: 3)

Worth underlining is the last sentence that implicitly suggests negative attitude taken towards the foreign lexis as an all-pervasive phenomenon.

III Status
The comment on the nature of foreign words and phrases can be said to be characteristic of many publications that stress the often unresolved status of foreignisms on the map of the lexicon of a given language: "The very definition of what constitutes a foreign term is disputed. Neither is there any widely-accepted convention among writers of English regarding the spelling or diacritical marking of foreign terms" (p. 4). Interesting is the observation about the frequently transient, elusive character of non-native words as in many cases their appearance in the receiving language system is rather incidental (cf. Chapter 6 of the present book where meteoric status of a foreign word or phrase is elaborated on): "The absorption of these foreign phrases into our language is always much too rapid for assimilation, and as a result, many of them never reach the ordinary dictionaries" (p. 3).

IV Criterion
Again, the selection of foreignisms under consideration is rather subjective, although some vague mention of statistical grounds for elicitations of words in question is also provided: "The choice of terms has also been based on frequency of occurrence. In the absence of statistical criteria on this point, the choice has necessarily been subjective. The test of its validity must rest with time and the consultants of this dictionary" (p. 5).

Dictionary of Foreign Phrases and Abbreviations (DFPA 1965)

I Motivation
DFPA offers somewhat positive explication of motivation that underlies the verbal choices of speakers who make use of foreignisms in their communication. One such reason maybe the speaker's pedantry, the other, more significant, the speaker's wish to avoid wordiness with the simultaneous preservation of communicative intentions:

> However to those who are acquainted with a subject being discussed, the use of such terms (i.e. foreign words and phrases, MK) is often a short cut to making a meaning clearer. In numerous areas, such as law, business, philosophy, medicine, and music,

> certain foreign phrases have a fixed meaning that is immediately understood by those who are familiar with the subject. Similarly, to people with literary interests certain expressions may brighten a conversation by their aptness or sum up a concept that might otherwise require several sentences for a clear explanation.
>
> (DFPA: 5)

II Source and scope

The major source of reference from which DFPA borrows in its description of foreign words and phrases is Latin. As it is further stated in DFPA, "[…] of the modern languages French predominates, with Italian a distant second" (p. 6). The editor of this dictionary relies on other dictionaries of quotations and other similar publications and excludes the following lexical units:

> Items longer than a couplet have been omitted since they can hardly be classed as phrases; at the other end of the scale, single words, with a few exceptions, are not included since these are proper material for foreign language dictionaries. Latin and Greek nomenclature used in the biological sciences is not included since this is technical information available in manuals on these subjects.
>
> (DFPA: 7)

III Status

There is not the slightest digression concerning any problems in delineating the ontology of a foreign word. The category of a foreign word is taken for granted as well defined and uncontroversial. The word "phrase," on the other hand, is spared some attention and is conceived of "in the recognized and not uncommon meaning of a pithy, quotable expression, such as a proverb, motto, and maxim. It is not necessarily taken in its narrow grammatical use, as, for instance, a preposition followed by an object" (p. 5).

IV Criterion

Similarly to the aforelisted dictionaries, DFPA does not offer any conclusive criterion that served as a guideline for the selection of foreignisms under consideration:

> Such compilation as this, is bound to reflect in some measure the tastes and interests of the compiler. Much material, however, must be included because of frequent usage. The problem, then, is one of selection from among thousands of phrases that might be included.
>
> (DFPA: 7–8)

The Dictionary of Foreign Terms in the English Language (DFTEL 1973)

I Motivation

No mention of this issue in the dictionary.

II Source and scope
DFTEL, in contradistinction to the above-presented dictionaries remains rather 'immodest' in that it assumes rather an all-inclusive than self-exclusive perspective in terms of repertoire of units and the resulting range of themes discussed. DFTEL sets its goal in the following way:

> To arm the reader against such interruptions […], DFTEL (abbreviation, mine) contains a selection of words and phrases that represent a wide range of subjects and which are taken from more than thirty different languages. Included are foreign terms from medicine, philosophy, law, the culinary arts, science, history, politics, religion; included are also ancient proverbs and modern quotations, familiar mottos and literary aphorisms, famous exclamations and infamous oaths, titles of nobility, entrees from a menu, musical terminology; words used in the office, in the classroom, in reference libraries, all are here.
>
> (DFTEL: vii)

III Status
Interesting is the consideration upon the ontology of a foreign word in this dictionary, namely a foreign word is viewed as an alien body in the organism:

> This book is designed for the scholar, the teacher, the linguist; it is also meant for the casual reader who has occasional difficulty remembering the difference between *a priori* and *a posteriori* or who has often puzzled over the untranslatable foreign quotations which so many authors sprinkle throughout their prose. It is, in fact, a book for anyone who has at one time or another been stopped in the middle of a passage by a foreign word.
>
> (DFTEL: vii)

IV Criterion
DFTEL offers a graphemic criterion for the identification of what counts as foreign phrase, but no further justification in favour of this criterion is provided. This same goes with the phenomenon of 'polysemy' which is alluded to only superficially:

> In using this dictionary the following facts should be noted. First, certain words included have been assimilated into English and are now part of the language proper. […] When naturalized words are listed, they are set in roman type; consequently, all italicized listings are foreign words and should be kept in italics during textual usage. Also where one word has more than one meaning, these meanings are numbered in sequence, beginning with the definition most commonly intended; when the meanings of a word differ only slightly, these meanings are divided by the word 'also' or by a semicolon.
>
> (DFTEL: vii–viii)

Dictionary of Foreign Terms (DOFT 1974)

I Motivation
No mention of this issue in the dictionary.

II Source and scope

DOFT is rather 'economical' in its presentation of source and scope of the presented material. It confines itself to a couple of lapidary statements in the matter:

> The average educated English speaker with little or no knowledge of foreign languages often is baffled by foreign terms and phrases that occur, without translation, in the books, magazines and newspapers that he reads, or that he may hear by way of television, radio, films and theater.
>
> (DOFT: 7)

III Status

DOFT does not make any serious claims as regards the categorial status of foreignisms as lexical categories. Instead, there is a noticeable terminological confusion in which foreign words and phrases are treated on a par with 'unknown' words and phrases.[13] This appears highly unsubstantiated, given the lack of reference to possible gradation that various units boast in terms of the level of naturalization (assimilation) to the receiving language system. All foreign words presented in this dictionary are treated as if they had the same status in the lexicon of the target recipient language:

> Unless he has a large reference library at his immediate disposal, he will have no idea of their meaning or of how to pronounce them (in the spoken media, more often than not, he will hear them pronounced!). This dictionary aims to assist him by telling him from what languages these terms and phrases come, often with some further interesting information about their history.
>
> (DOFT: 7)

IV Criterion

The criterion of selection the material for the analysis is the intuitively felt continuum spectrum of familiarity and unfamiliarity of a unit to the average language user:

> In selecting the entries to be included in this dictionary the authors have had to consider and reject far more than could be included. The aim has been at all times to include only useful, interesting and timely material. Obscure classical quotations, highly technical terms, personal and geographical names, titles or opening lines of songs and arias, and terms that have become wholly naturalized in English have been omitted. […] The authors have included terms and phrases not only from the more familiar languages such as French, Spanish, Italian, German, and Latin, but also from Greek, Russian, Sanskrit, Hebrew, Arabic, and from little-known languages such as Malay, Tagalog, Afrikaans and scores of others.
>
> (DOFT: 8)

---

[13] In the present book the distinction is made between:
a) borrowings/loanwords,
b) foreign words and phrases,
c) 'unknown' words and phrases.
Detailed discussion is presented in sections 5.2, 5.3 and 6.7.3.

A Dictionary of Classical & Foreign Words & Phrases (DCFWP 1982, 1981)

I Motivation
Motivating factors discussed in DCFWP can be divided into two parts: global and personal. Global motivation of the use of foreign words and phrases is linked in the dictionary with the cultural dominance of the donor language: "The presence within a language of a foreign word or phrase is evidence of the cultural, political or technological dominance of that foreign country" (p. 9). The personal motif, on the other hand, has more positive resonance and is related to the use of a foreign word as an exponent of higher educational upbringing: "Latin was for so long the language of general education that those who wished to give the impression of being educated decorated their English with the Latin 'tag'" (p. 10).

II Source and scope
DCFWP provides what appears as mutually exclusive explication of the source of the presented material as well as its scope: "This book is intended to help […] those who do not know the meaning or correct use of a term from a foreign language which is commonly used in English, but not always understood, even by 'sophisticated' people" (p. 7). The paradox involves the juxtaposition of the common usage of discussed words with the fact that these are actually 'unknown' by language users. This sounds like a clearly counterintuitive judgment. As far as the major languages upon which this dictionary relies, the leading position is inarguably taken by Latin.

III Status
No mention of this issue in the dictionary.

IV Criterion
The criterion of selection is based on a vaguely formulated criterion of common usage of the words and phrases in question: "This dictionary defines the foreign words and expressions commonly used in English" (p. 7). The issue remains open whether 'foreignness' of a unit and its frequent usage are not mutually exclusive criteria. As the analysis of foreign words and phrases presented in this book unfolds, the argument will be advanced in favour of rejection of such risky correlation. In other words, as Chapters 6 and 8 reveal, a frequently used term does no longer bear the feature of foreignness but rather is a likely candidate for a perfectly naturalized unit in the recipient language, which, accordingly with the nomenclature adopted and elaborated on in Chapter 5, classifies this unit as at least a candidate for a borrowing.

A Concise Dictionary of Foreign Expressions (CDFE 1984, 1982)

I Motivation
CDFE restricts itself to presenting only some general external factors governing the frequency of occurrence of foreign words and phrases in the contemporary

English: "The origins of modern English, its subsequent international nature and the history of Britain, have all contributed to the frequency with which foreign words and expressions are found in English" (p. vi).

II Source and scope
Surprisingly, CDFE does not provide its reader with any information regarding source and scope of the presented material.

III Status
This dictionary offers some interesting observations concerning the typology of foreign words and phrases. One may safely risk the claim that the following phonologically-grounded discrimination between different types of non-native vocabulary is in perfect conjunction with the criterion of selection adopted in the present book:

> With the passage of centuries, many such importations have lost their foreign identity and feel along with, in many cases, their original spelling or pronunciation, and are now recorded in dictionaries as fully fledged English words […]. Other importations are still in transit, so to speak: some of them have more than one spelling or pronunciation – their original ones and anglicized ones – and dictionary makers have differing views about their status. A third group is, however, undeniably foreign, even though words and expressions from this group are sometimes recorded in English dictionaries with a clear (often typographical) indication that they are not yet assimilated into English.
>
> (CDFE: vi)

An inarguable value of this division is the tertiary classification of units according to their degree of assimilation which is bound up with the phonological factor. The present book elaborates on this perspective and additionally seeks to search for some ways of resolving the long-lasting controversies regarding various pitfalls in classifying foreign lexis (see particularly sections 5.2.1 and 8.2).

IV Criterion
Unfortunately, the dictionary appears to focus on what we refers to in the present book as "unknown words and phrases" or words at the borderline between those literally unknown to the vast majority of speakers of the recipient language and the words that are described as "in transit" towards assimilation, which implies some degree of their entrenchment in the target system: "The present collection is drawn mainly from this third group, together with a few 'transit' words that still have enough of their original flavor for the general reader to expect to find them here" (p. vi).

Loanwords Dictionary (LD 1988)

I Motivation
Similarly to DCFWP, the motivational factors behind the use of foreignisms are of two kinds: external to the language user and internal to him/her. External argument is the following:

> Evidence of borrowing can be found throughout the historical record of English, as speakers and writers sought to capture the essence of a phenomenon or some facet of experience. Hence the lexicon of the language bears witness to the diverse cultural contacts and influences that have shaped it.
>
> (LD: vii)

Internal motifs of usage are formulated in a straightforward fashion and assume negative interpretation: "Another, perhaps more subjective reason for the wide use of loanwords in English has to do with simple vanity." However this negative motivation is 'weakened' by other more positive aspects of personal motivation: "Familiarity with a foreign culture and language, whether directly or through education, is taken by some as a mark of sophistication, refinement, even erudition" (p. ix).

II Source and scope
LD presents us with source and scope of the presented material reference to areas well known from the aforementioned dictionaries:

> We have tried to focus on those terms that are likely to be encountered either in speech or in writing, by a person who is attuned to literature, culture, and society. We have included loanwords from certain technical areas, such as cookery and music, on the grounds that these and other such areas are familiar to a broader audience than terms from, say, medieval philosophy and pharmacology.
>
> (LD: viii)

What is quite novel is further specification of the scope of material presented which is conducted by way of peculiar exclusion of potentially relevant material. Thus what remains excluded are "highly technical or specialized terms of occasional interest mainly to experts or aficionados in certain fields. Also excluded were many words from widely studied languages – French, German, Italian, Spanish – that are of high frequency in those particular languages (and hence well known to anyone who has studied them), but not common in English" (p. ix).

III Status
Status of foreign words as viewed in LD can be classified as concurrent with the one (i.e. second group of foreign words and phrases) discussed in CDFE. LD,

however, uses the term "loanword" to refer to these units which are reserved to designate well-adapted units (see section 3.0). Still, loanwords in contrast with the regular words found in English are said to retain an ephemeral flavor of foreignness: "What distinguishes the loanwords, however, and allows treatment of them as a subset of the English lexicon, is that despite their use in English contexts, they still, in some regard, *seem* or *feel* foreign"[14] (pp. vii–viii). Their status as a psychologically real category is, however, viewed in this dictionary as clearly indefinite: "We could mention hundreds of other examples of words that are felt by some to be more or less assimilated, but by others to be totally foreign" (p. x).

IV Criterion

LD offers quite an elaborate discussion upon criteria of typologization of non-native vocabulary. This is formulated in terms of the following propositions:

> 1. The term maintains some measure of its foreign orthography, pronunciation, or flavor. Put differently, the term has a widely recognized association with a specific non-English language or culture.
> 2. The term is freely and commonly used in English contexts; that it is employed by users of English as if it were part of their language.
> 3. The term, if originally specialized or technical, has become generalized in application or is from a field that attracts broad general interest (such as cookery or literature).
>
> (LD: viii)

LD, however, does not conclude its consideration upon giving the list of circumstances which, once fulfilled, classify a given unit as relevant for inclusion in the dictionary. Other criteria relate to the notion of naturalness and frequency of usage of the discussed terms in English:

> The constant intention of the editors in compiling *Loanwords Dictionary* was to create a dictionary that is essentially English, one containing words that are used in English contexts with some degree of naturalness or frequency, so that a wide-reading and curious English speaker likely to encounter them in English sources may use them.
>
> (LD: viii)

Quite debatable is also the orthographic criterion as infallible in the identification of loanwords: "One of the most obvious distinguishing features of many a loanword – often a strong suggestion that the term has not been fully assimilated – is the presence of diacritical marks" (p. xii). As it appears from the analysis of foreign words and phrases in the present book, which roughly correspond to "loanwords" suggested in LD, many of these units are observed to have, in fact, lost diacritic marks originally existent in their native languages, which is *per se* a marker of this transitional stage to which they pertain.

[14] Cf. Crystal (2007b: 60).

Dictionary of Foreign Phrases and Abbreviations (DFPA 1990, [1965])

I Motivation

In discussing motivational factors underlying the use of non-native vocabulary, DFPA presents a quite clearly articulated positive stance on foreign words and phrases in respect of the recipient language system. Other dictionaries presented so far remain rather indifferent if not ostentatiously negative about this issue. In DFPA, then, motivation for the use of foreignisms is rather global and viewed as social instrument of enriching a recipient language an expressive potential:

> The language of England is rich in foreign infusions. Thanks to the many invasions that have enlivened the history of that 'right little, tight little' island, what we today know as English is far from the language of the Angles. A much changed, and still changing, potpourri, it echoes the world of the Romans and of the Franks, of fur-clad Gaels and raging Norsemen and archly opportunistic French barons. […] It echoes, too, colonial times, when the English were on the giving end of invasions. From Shakespeare's day down to this century, the language has always profited from foreign contacts, as explorers and sea captains and merchant princes continued to swell the ancient word hoard with contributions from the Hindu Kush and the forests of the Wampanoags and New Spain. […] Nor has the process stopped today. In the English-speaking world generally, and in the United States particularly, the mix of immigration and geopolitics, of electronic journalism and ethnicity, has continued to enrich the so-called King's English with snippets of Russian and Yiddish and even Sanskrit. As has been the case ever since the Romans' entry into Celtic Britain in the first century B.C., the Mother Tongue continues to prosper from foreign borrowings.[15]
>
> (DFPA: ix)

II Source and scope

DFPA relies in about 40% of its material on Latin and French as major donor languages. This is explained by the following: "Latin was the language of the educated England throughout the long Middle Ages, and French picked up its fallen mantle from the Renaissance into our own time. Hence over 40% of my entries come from one of these two tongues" (p. x). These prerequisites stand in sharp contrast with the assumptions laid down in the present book where Latin is almost excluded from the discussion as its words and phrases do not fulfill the phonological criterion established for the identification of a foreign word or phrase. This is because the vast majority of Latin words functioning in English have a perfectly developed anglicized pronunciation variants, which is certainly aided by the status of Latin as a dead language.

III Status

No explicit mention of the issue of a foreign word as a lexical category in the dictionary.

[15] See also Crystal (2005: Ch. 7).

IV Criterion
The question of the criterion employed for the selection of lexical units for the consideration in the dictionary is not much deliberated upon. DFPA confines itself to rather lapidary a statement in the matter: "I have tried to present here a selection of those foreign terms now used in English but that are neither on the one hand thoroughly Anglicized nor on the other merely pedantic" (p. ix).

A Dictionary of Foreign Expressions (DFE 1990)

I Motivation
Similarly to DFPA, DFE highlights rather positive aspects of the usage of foreign words and phrases: "English grows not only through coinage of new words and employment of existing English words in novel ways, but also through adoption of foreign words and phrases that give opportunities for colourful and precise expression" (p. x). The issue of the presence of non-native lexical element in the English language is not to be regretted but rather attributed to the utilitarian *raison d'être*" manifested by English speakers; all embedded in the perspective on intercultural communication, which itself is judged as a positive phenomenon:

> Aside form, this dictionary provides other benefits. Finding out how people of other cultures express their thoughts affords insight into attitudes, prejudices, and practices – even perfunctory browsing will show that people everywhere have thoughts that are marvelously alike and astonishingly different.
>
> (DFPA: x)

II Source and scope
DFPA provides an extensive list of the sources it relied on in compiling the list of foreign words and phrases. This list includes words and phrases "from more than fifty languages frequently used in conversation or likely to be encountered in the fields of literature, law, science, politics, music, art, diplomacy, fashion, travel, food, and dining" (p. xi).

III Status
Considering the issue of the status of foreign words and phrases in DFE, the editors appear not to contribute much to ordering the terminological chaos present in the literature of the subject matter. This is instantiated by their equating borrowings with foreign words and phrases: "Some of the terms are so well established in English that we sometimes fail to recognize them as borrowings, but others are used in full knowledge that they are of foreign origin" (p. ix). The valuable observation is the implicitly indicated fuzziness of the levels of foreignness represented by particular items, which is in accordance with the assumptions presented in Chapter 5 of the present book.

IV Criterion
Again, quite analogically to the aforementioned dictionaries, the issue of the criterion of selection is not clearly laid down and, instead, is limited to somewhat imprecise formulation of selection procedure: "From the time the *Dictionary of Foreign Terms* was first complied in 1934, the invention of its editors has remained constant: to create a single volume source that explains foreign phrases and words likely to be encountered in American and English literature" (p. ix).

Dictionary of Borrowed Words (DBW 1991)

The editors of DBW are the same as editors of LD (1988). This may explain why the issues related to motivation of use of foreignisms, their status, etc. are subject to the same argumentation. The only dubious qualitative difference involves switching the labels in that instead of "loanwords" (LD 1988) the editors choose the term "borrowing."

The Wordsworth Dictionary of Foreign Words in English (WDFWE 1995, 1991)

This dictionary conception of foreign words and phrases seems to be the closest to the definition accepted in the present book.

I Motivation
Motivation of use of foreignisms as presented in WDFWE is definitely external to a language user and is related to the period of commercial and imperial expansion: "English is the way it is because it is such an inveterate borrower from other languages. […] In the period of commercial and imperial expansion [English borrows, M.K.] from virtually all the major languages of the world, and many minor ones" (p. ix). Interesting and quite innovative, in this respect, is the analogy to somewhat negative phenomenon of kleptomania: "this linguistic kleptomania has given the language an extraordinary richness of vocabulary, enabling English in many cases to make subtle distinctions lexically, where other languages must have recourse to circumlocutions" (p. ix).

II Source and scope
No explicit mention of this issue in the dictionary.

III Status
It is in this dictionary that the target group of foreign words and phrases is identified according to the phonological criterion with so much of the introspective input. Thus, although borrowings constitute a superordinate category of non-native vocabulary, which stands in contrast to the proposal advanced in the present book, the realization of the degree of foreignness of particular sub-categories of

borrowed terms as based on the aforementioned phonological criterion is, however, indicated:

> But this borrowing has not stopped. It is an ongoing process. And many of the words and other expressions acquired during the past couple of hundred years have not yet become thoroughly anglicized. […] They retain an aura of 'foreignness.' When they appear in printed texts, they are often italicized, as if they are not quite fully paid-up members of the language. When we say them, we often attempt to reproduce their pronunciation in their original language. If they have accents or other diacritic marks in the language they come from, these are generally retained.
>
> (WDFWE: ix)

IV Criterion
The criterion of selection is correlated with frequency in that "a thousand of the most widely used of such words in present-day English" (WDFWE p. ix) are presented. There is, however, no mention of any more specified criteria that could objectivize the notion of "wide usage."

Chambers Dictionary of Foreign Words and Phrases (CDFWP 1995)

I Motivation
The issue of the pragmatics of usage of foreign words and phrases is not deliberated upon in this dictionary.

II Source and scope
The dictionary bases its repertoire of foreignisms on corpora of English texts in which they occur: "It should be emphasized that words and phrases are only included if they are used *in English* contexts" (p. viii). Unfortunately, no further specification of the nature of these corpora is provided. Instead, the source of the discussed foreign words and phrases is attributed to quite general socio-political trends that shape the influx of non-native items into the lexicon of English:

> In the twentieth century, new influences are discernible. For example, political conflict in the Middle East has given rise to a whole glossary of culture-resonant terms of which the most notable are *fatwa, intifada,* and *jihad* […]. The growth of international communications and the holiday industry have made us more aware of the cuisines of other countries, with remarkable consequences for the language, as will be seen from many terms included here from languages as diverse as Russian, Japanese, Spanish, and Arabic, in addition to the rich gastronomic language of French.
>
> (CDFWP: vii)

III Status
This dictionary devotes some attention to the categorial status of foreign words and phrases. Interestingly, the category of foreign words and phrases functions as a superordinate one (see, e.g., section 5.3.1), subsuming borrowings as conceived

of as naturalized units adapted to the English language in terms of orthography, pronunciation, grammar (inflectional system), and so-called "loans," i.e. non-naturalized forms which constitute the subject matter of the dictionary. The weakness of this reasoning is the division of non-native vocabulary into two groups, i.e. naturalized and non-naturalized as if more intermediate, transitional stages in the process of assimilation were non-existent. The strong point is that the dictionary makes such subcategory differentiations at all, and that it shifts its attention to forms whose status as stable elements of the English lexicon is dubious.

IV Criterion
The criterion of inclusion of foreign words and phrases into the dictionary is based on a clear, yet psychologically controversial basis, i.e. typographic:

> Some of the noteworthy naturalized words of more recent origin are included here. Most of the items included, however, are non-naturalized words and phrases that conventionally appear in italics in print and are thought of as loans rather than belonging to the vocabulary of English.
>
> (CDFWP: vii)

This typographic convention, however, ignores the frequency of occurrence factor, which appears to be the crucial in measuring the degree of naturalization of a given unit in the recipient language. As a result of reliance on the aforementioned typographic convention, the dictionary classifies items such as *blasé, catchet* or *quid pro quo* as belonging within one category, which seems to be rather counter-intuitive especially if we confront the frequency of use of *blasé* with and the frequency of use of the two other words.[16]

World Dictionary of Foreign Expressions (WDFE 1999)

I Motivation
The motivation of use of foreignisms as presented in WDFE can be described as highly instrumental and viewed as restricted only to a group of professionals using a particular jargon at work or at most, to a group of individuals who have not much leisure time and simply intend to kill time. This gives an impression as if foreign words and phrases were otherwise not present as significant tools in everyday communicative encounters: "Busy professionals who need accurate information quickly will get exactly what they require from *World Dictionary of Foreign Expressions.* Leisurely word-lovers will be instructed and entertained" (Foreword). Subsequently, however, the editors present the list of personal factors guiding the use of foreignisms in English. These include:

[16] According to BNC statistics of occurrence, the discussed forms have the following occurrences: the first value stands for global occurrences in the corpus, whereas the second value stands for the number of texts in which the items are recorded: *blasé* (18/17); *catchet* (0/0); *quid pro quo* (44/40).

> a) Enhancing the profundity, richness, and quality of scripts.
> b) Wishing to avoid the use of an English expression which seems obscene or indecent.
> c) The need to portray and capture the local associations and color of foreign institutions.
> d) The usage of foreign words in specific contexts where their replacement with English translations would detract from their charm, forcefulness, or succinctness.
> e) Some are technical terms used in certain professions.
> f) Finally, such expressions are the hall mark of scholars of diverse and solid educational background.
>
> (WDFE: xvi)

II Source and scope

In Foreword to this dictionary, editors, continuing their argument about the restrictive usage of foreignisms in the present-day English, make the following remark: "This is a very useful book. The English of lawyers, physicians, academics, journalists, and others is replete with foreign phrases, many of them traps for the unwary." Not surprisingly, then, the sources of the lexical material presented are limited to professional literature:

> The emphasis is on legal, philosophical, historical, and literary material. While some scientific words are included […] readers are advised to consult specialized dictionaries […] for vocabulary in the health care professions. Nor will the reader find the vocabulary of scientific taxonomy here. […] Also excluded is the foreign vocabulary which has enriched English references to flora and fauna in words like *tomato* (Nahuatl), *rhinoceros* (Greek), *aardvark* (Afrikaans), and *dachshund* (German). Foreign words for food and drink like *saki* (Japanese) and clothing like *sari* (Hindi) are also not included here unless they have other significance, especially religious, like *yarmulke* (Yiddish).
>
> (WDFE: xvii)

It appears that the weakly justified overspecification of various circumstances that include or exclude a given set of words for the consideration in the dictionary gives the impression of some methodological inconsistency in collecting the lexical material for the analysis.

III Status

Again in WDFE there is no differentiation introduced between foreign words and phrases and loan words: "Users of this dictionary may, at times note the absence of an occasional foreign word or phrase. Such lacunae are inevitable in an opus of this scope […]. Linguists would call the entries in this dictionary loan words or phrases" (p. xviii).

IV Criterion

The criteria of differentiation remaining vague, the editors confine themselves to a few lapidary comments whereby the criterion of anglicization of a foreign word constitutes a benchmark for the selection of words under consideration.

Unfortunately, the editors do not bother to elaborate on this crucial criterion. Instead they emphasize their focal interest in the written rather than spoken material, reverting to the phonological criterion in the end:

> Other loan words, like major, minor, restaurant, and trombone, have been so thoroughly Anglicized that they are included here for the sake of illustration […]. Since the emphasis in this dictionary is on written rather than spoken material, no pronunciation guides for foreign words and phrases are provided […]. Another factor to consider is that the more a word or phrase is used in English, the more Anglicized the pronunciation becomes.
>
> (WDFE: xviii)

Cassel's Foreign Words & Phrases (CFWP 2000)

I Motivation

CFWP takes the phenomenon of the presence of a non-native element in the English language for granted and attributes that fact to commonly known factors like geographical/historical contacts of English with other languages, especially French:

> In modern times there is hardly a language in the world that has not contributed elements of its vocabulary to the English language. England's geographical and historical links with France have not surprisingly resulted in the adoption of a large number of French words and phrases.
>
> (CFWP: vii)

What is quite novel in this dictionary (Preface: vii–viii) is an attempt to link a set of stereotypes to particular donor languages in such a way that French is correlated with diplomacy, dinner, dressing table, Italian with music, Spanish with sea, German with technical, political, military, literary fields, Hebrew with religion, etc. To what extent, if at all, these stereotypical associations can be confirmed is illustrated by the survey of particular donor languages presented in Chapter 7.

II Source and scope

CFWP boasts the selection of ca. 5000 words for the analysis, basing its criterion of elicitation on the category of perception, which unfortunately is taken for granted and not subject to further deliberation: "This new dictionary aims to provide a full background to some 5000 words and phrases that, although used by English speakers and writers, are perceived to be foreign and not fully assimilated into the language" (p. vii). A valuable contribution made by the dictionary relates to the introduction of foreign proper names for the investigation – an aspect of studies on foreign lexis that has so far been rather neglected than given due attention:

> A special feature of the dictionary is the inclusion of a number of familiar phrases or even sentences that rarely appear in other dictionaries of this type. Proper names will also be found, such as the names of the ancient Hebrew months (such as *Abib)* and the months in the Revolutionary calendar (such as *Fructidor*), together with personal names that have gained a figurative sense, such as *Mata Hari*.
>
> (CFWP: ix)

III Status

The discussion of the status of foreign words as a lexical category is unfortunately quite 'economical' and limits itself to making adequate but general remarks about the often encountered spelling and pronunciation variations with foreignisms:

> An observation should finally be made regarding spelling and pronunciation. Many foreign words and phrases are notoriously difficult to spell and pronounce correctly. For some there are alternative spellings. For others there are alternative plurals. Others again have variant pronunciations.
>
> (CFWP: ix)

IV Criterion

The problem establishing the criterion that CFWP employed for labeling non-native vocabulary as foreign remains unresolved due to lack of some more investigation into intuitively felt fuzziness in the degree of assimilation displayed by different non-native lexical units. As a result, the final decision about the inclusion or exclusion of a particular unit for the analysis becomes rather arbitrary. On the other hand, the editors include items which, although rare in the English of the day, have been recorded as part of its system. Thus, we have mutually exclusive criteria adopted by the dictionary: the criterion of the degree of assimilation which is bound up with frequency, on the one hand, and the criterion of presence which does not allow for frequency effects, on the other:

> Conspicuous by their absence in the dictionary are names of plants and animals […] this is because almost all names of non-indigenous flora and fauna remain foreign, and it is often hard to tell to what degree they have become fully assimilated into English.
>
> (CFWP: viii)

As it is further argued: "The degree of assimilation of many learned Latin and especially Greek terms are also hard to establish […]. On the other hand, a number of relatively esoteric or even rare words and phrases are unashamedly included since they are or have been current in English" (p. ix). Aside from these inconsistencies, a handful of insightful remarks are given in the dictionary concerning the phonological component of a foreign word's make-up, the observations of which are in compliance with the basic tenets about foreignisms presented in section 0.2 of the book:

> The usual English method of dealing with the spoken form of a foreign word or phrase is the typical British compromise. Anything close to the original language can sound affected, while a complete Anglicization can seem coarse […] because foreign words and phrases are by definition not assimilated into the language, therefore their 'correct' or educated pronunciation remains variable, and the dictionary here can only indicate what is normally regarded as the best or at any rate standard.
>
> (CFWP: x)

The Facts on File Dictionary of Foreign Words and Phrases (FFDFWP 2002)

I Motivation

The last (but not least) in our survey is FFDFWP published in 2002, which in quite peculiar, or put it mildly, poetic way describes the global motivation behind the use of foreign words and phrases in English: "The American poet and essayist Ralph Waldo Emerson described the English language as 'the sea which receives tributaries from every region under heaven'" (Introduction).

II Source and scope

The discussion of the source and scope of the presented material is confined to the enumeration of appropriate fields:

> This dictionary has been compiled as an accessible guide to expressions that are derived from foreign languages. […] Words and phrases that have become part of the English language cover a wide range of fields: entertainment […], food and drink […], language and literature […], law, music […], politics and economics […], and religion […].
>
> (Introduction to FFDFWP)

III Status

There are no clearly formulated ideas about the status of foreign words or phrases as lexical categories.

IV Criterion

Considerations over the criterion of selection of foreignisms in question are reduced only to a lapidary indication of different degrees of assimilation manifested by different non-native words in English. In effect, the dictionary includes in its analysis not only foreign words proper, but also borrowings, which should naturally fall out of the scope of research interest in such a book, especially if such an observation is made by the dictionary: "some such words and phrases [those that have been included, comment mine] have been fully assimilated into the language (*igloo, bonanza*); others are still thought of as foreign *(de rigueur, magnum opus*). The dictionary does not however proceed to further explicate why, given the title of the dictionary, it also bothers to investigate well-adapted units that have lost their quality of foreignness altogether.

The Browser's Dictionary of Foreign Words and Phrases (BDFWP 2006 [2001])

I Motivation

BDFWP does not, as most of the aforelisted dictionaries, discuss internal or external aspects underlying a language user's motivation to employ foreignisms, but, instead, assumes more pedagogical perspective in which the everyday character of foreign lexis is stressed (this is again quite innovative) – the lexis that requires semantic clarification to the reader. This goal appears to be successfully implemented throughout the dictionary: "Veteran reference book compilers and editors Mary Varchaver and Frank L. Moore, who are known especially for their recently published *Dictionary of the Performing Arts*, have turned their talents to enlightening us on meanings of foreign words and phrases we meet in our daily lives" (p. v).

II Source and scope

A consequence of adopting a practically-minded policy in the dictionary is also discernible in the type of source material relied on: "It is rather, a dictionary to help you elucidate what you come across every day in newspapers or hear on television. Its choice of terms and directness of style reflect the immediacy of everyday discourse. Thus, it is a unique and exceptionally useful addition to the genre of special dictionaries" (p. v).

III Status

There is no explicitly signaled passage in the dictionary on the categorial status of foreignisms as elements of lexicon.

IV Criterion

The criterion of elicitation is based not so much on a frequency factor, but rather on a diachronic basis whereby the most recent adoptions into the English are classified as foreign in contrast with the older ones that got completely absorbed into the English language. What appears risky in such reasoning is the automatic correlation of the time of incorporation of a non-native unit into the recipient system with the recognition of its status as foreign or non-foreign unit in that language:

> The entries focus on those words that an American reader will recognize as truly foreign. Thousands of words in American English have foreign sources, such as *dollar* or *ketchup*, but their meanings have become so thoroughly absorbed into our everyday language that they are no longer classified as foreign. […] In this *Browser's Dictionary of Foreign Words and Phrases,* the authors have stressed foreignness by choosing words that are relatively recent additions to the language. They also included some older adopted words that have different or expanded meanings in current usage.
>
> (BDFWP: vi)

4.3. Other publications

Dictionaries of foreign words and phrases constitute one major contribution that addresses the issue of non-native lexical units in the English language. Aside from these, however, we may find some other significant monographs related to the subject matter, the survey of which is presented below.

One such publication is by Boese (1988). In the book, the author resorts to a paradoxical statement in which, on the one hand, foreignisms are viewed as part of the exclusive lexical competence of an educated language user, whereas, on the other, the same words are presented as appearing not only in specialist literature accessible to the educated, but also manifested in ordinary written English:

> Have you ever needed a book that hadn't yet been written? […]. I had this feeling when I taught college English to adults at an Air Force base. […] They needed a book explaining in simple language the common allusions and foreign terms that educated readers are expected to know – even when they are reading an ordinary news magazine […]. To my frustration I found that there was no such book.
>
> (Boese 1988: i)

The criterion of selection of foreignisms is quite clearly formulated. This is the analysis of an extensive corpus material (2000 words and phrases) collected from 750 magazines, however, the notion of foreignness is presented as taken for granted, which practically leads to the discussion of a wide spectrum of qualitatively different units ranging from those quite well adapted to the English language to others whose presence in the English lexicon is rather incidental. The disadvantage behind the methodological premises underlying the book is that it puts on equal footing all of the aforementioned items, which are treated as homogeneous in respect of their degree of foreignness (Boese 1988: i).

Another publication related to foreignisms is that by Whitcut (1996). The author discusses the place of English among the languages of the world, pointing out to the lingua franca status of that language[17] – the language spoken by 300 million people as their mother tongue scattered all over the world (1996: ix). Other markers of the leading position of English on the map of communication include its significance as the working language of the UN, the main tool of mass media communication as well as science, technology, transport, sport, travel, and pop music (*ibid.*).

Interesting is the set of remarks presented by Whitcut concerning the English and its relation to foreignisms. According to the author (1996: xi), English has always been a 'free admitter' of non-native elements to its lexicon:

[17] See Chapter 9 which reverts to the problem of English as a lingua franca from astrophysical perspective.

> What it does have is an unusually rich vocabulary; it has been said that English is the only language that has, or needs, dictionaries of synonyms such as *Roget's Thesaurus*. It has always freely admitted new words[18] whenever it needed them. For instance, we learned a lot about music from Italy, and our musical words – soprano, diminuendo – are still Italian.
>
> (Whitcut 1996: xi)

What should be emphasized is that Whitcut at this stage of the book indicates a certain rationale behind the phenomenon of borrowing, where the words introduced do not simply replace the native ones but add new meanings that enhance expressive potential of the recipient language: "As Dr Johnson wrote in the next century [18th century, M.K.]: Words are seldom synonymous; a new term was not introduced, but because the former was thought inadequate" (Whitcut 1996: xi) – this clearly brings to mind the concept of *ecological niche* and the ecological concept of the lexicon (see Radden and Dirven 2007). Johnson's optimism about foreign words and their role as contributors to the semantic enrichment of the target lexicon was not, however, unconditional and he felt it a duty to warn against uncritical importation of foreignisms into English. This is formulated as follows:

> Johnson was fairly happy about this [borrowings, M.K.], accepting that language has been 'suffered' to spread, under the direction of chance, into 'wild exuberance'; although he did feel it his duty to 'warn others against the folly of naturalizing useless foreigners to the injury of the natives.'
>
> (Whitcut 1996: xi)

Whitcut (1996: xii) also mentions the role of academies and various institutions as safeguards against an uncontrollable influx of foreignisms into the English language (see section 6.8.3). Characteristic is the use of language filled with the images of violence and opposition[19]:

> Academies have been instituted to guard the avenues of their languages, to retain fugitives, and *repulse intruders*, but their vigilance and activity have hitherto been vain; sounds *are too volatile* and subtle for legal restraints; to enrich syllables, and to *lash the wind*, *are equally the undertakings of pride*, *unwilling to measure its desires by its strength*.
>
> (Whitcut 1996: xii)

---

[18] For convenience, the elements, which constitute metaphors about how foreignisms are perceived as potential candidates for regular words in the lexicon, have been italicized.

[19] For ease of reference, the crucial lexical correlates of this warlike and dynamic conceptualization are italicized. The other similar words and expressions found in the context include (after Whitcut 1996: xii–xiii):

- The *influx* of religious words,
- A *flood* of French words,
- English as *assailed* by new words,
- *Thrust* new vocabulary into our language,
- Words that have *passed* into English.

Metclaff (1999) in contrast to Whitcut (1996) appears a little more realistic (optimistic) about the role non-native lexical stock has played in the formation of the present-day lexicon of the English language:

> If you speak English, you know at least a bit of a hundred languages. Or more. It's true. You are a savant in French, a genius in Latin, a philosopher in Greek. If you made it through kindergarten, you have mastered a bit of German. If you have a yen to be a *tycoon*, whether or not you become one, you are speaking Chinese and Japanese. If you *trek* to *paradise*, you are going through Afrikaans and Persian.
>
> (Metclaff 1999: x)

It can, however, be clearly seen that author does not bother to discriminate between different sets of non-native vocabulary, putting them into the same large homogeneous category. Metclaff (1999: x) seems not to bear any illusions about the inevitability of the process of borrowing, seeing it as a natural phenomenon observed throughout centuries in the development of the English language:

> Of the half a million words that have accumulated in the vocabulary of present-day English, only a small minority are native to the language as far back as we can trace it, that is to the Anglo-Saxon invasions of England some 1500 years ago. The rest have immigrates in subsequent years. The immigrants have come in such numbers that they seem as much at home as words of native stock.

Unavoidable as the process of borrowing presents itself, the process of adaptation is not viewed as automatic and subject to various circumstantial restrictions:

> Though new words and meanings *knock at the door every day, few are admitted*. A word can't bully its way in; it has to have special value to get its linguistic green card. Successful foreign words may name a plant, animal, or food that has no counterpart in the borrowing language […], or the word may name an exotic cultural trait worthy of comment and sometimes emulation. Often words change their meaning as they become truly *naturalized and adapted to* English-speaking circumstances.
>
> (Metclaff 1999: xiii)

As visitors to the English lexical system, non-native words, according to Metclaff, are not regarded as detrimental parasitic units, but rather as useful contributors to the lexical system of the receiving language: "While words enrich the receiving language, they have the advantage over some other kinds of imports in taking nothing from the *giver and doing no damage to the receiver or the environment*" (1999: xiii).

Another monograph by King (2001) also touches upon the issue of English and its contact with foreign words and phrases. The author remains rather skeptical about the necessity of using foreign words and phrases in English, quoting as a significant voice in the discussion a grammar authority, H.W. Fowler:

> The great grammarian H.W. Fowler did not mince his words on the practice of using foreign terms and expressions designed to go over the heads of the average English reader. To use French words, he wrote, 'that your reader or hearer does not know or does not fully understand […] is inconsiderate and rude. Display of superior knowledge is a great vulgarity as display of superior wealth.
>
> (King 2001: vii)

What parallels this resentful attitude to foreignisms is realistic acceptance of them as stable elements of the lexical landscape of a given language, which testifies to the pervasiveness:

> Nevertheless, this *infusion of foreign expressions* into our reading, in newspapers and magazines and books, is a fact of life whether we like it or not. We have the choice, when encountering yet another burst of italics, of groaning and fetching the dictionary, with an excellent chance of searching in vain for the word or phrase, or blithely leaping over the *intrusion* and reading on, saddled with guilt and ignorance.
>
> (King 2001: vii)[20]

Finally, the book provides the criteria of selection that King follows in his analysis of foreign words. These are at first sight appealing because they relate to the commonsensical usage of the terms in every-day English, but the weakness accompanying this choice involves a high degree of subjectivity from the author. As King (2001: 3) puts it:

> But here, at least, is a concise selection that should prove useful to the contemporary reader and writer. It excludes many foreign words that now reside in any good comprehensive dictionary. It also excludes specialist terms including those connected with music and opera; these will be found in most English dictionaries. On the other hand, it includes a good many colloquial and even vulgar expressions not to be found in dictionaries but which are common enough in speech and print.

Jacot de Boinod's book (2006) – another recent contribution to foreignisms in English – focuses on the discussion of extraordinary words from around the world. It appears to deal with what is discussed in the present book as "unknown" words and phrases (see section 6.7.3). The mode of analysis takes the shape of a lexicographic presentation. Examples of items taken into consideration include (Jacot de Boinod 2006: viii–ix):

- *Nakhur* (Persian) 'a camel that won't give milk until her nostrils have been tickled,'
- *Aerodjarekput* (the Inuit) 'to exchange wives for a few days only,'

[20] Cf. Chapter 6 and the discussion of foreignisms as analogous entities to meteors which are also said to attack our Planet Earth in vast quantities against which the Earth forms the protective shield. Thus the pervasiveness of the phenomenon of borrowings as alluded to by King (2001) is preserved under the set of various epistemological correspondences between the world of linguistic borrowing and the world of meteors with the related astrophysical laws (see especially section 6.9).

– *Zechpreller* (German) 'someone who leaves without paying the bill,'
– *Neko-neko* (Indonesian) 'one who has creative idea which only makes things worse,'
– *Serein* (French) 'the rain that falls from a cloudless sky.'

The author's analysis of foreign lexical assimilation processes is, as in the case of the aforementioned publications, full of metaphorical imagery[21]: "The English language has a *long-established and voracious tendency to naturalize* the best foreign words: *ad hoc, feng shui, croissant, kindergarten.* We have been *pinching words from other cultures* for centuries. Here are some we missed" (2006: ix). Sadly, however, the book does not introduce any objectivized tools by means of which it would be possible to describe the phenomenon of naturalization, if this is presented as a criterion of demarcation between different non-native words. Instead the author relies on subjective, impressionistic ideas of what constitutes foreignness. Although the criteria of selection based on introspection cannot be ignored in cognitive-linguistic research, however, they appear to be largely insufficient if more reliable results are to be achieved. Thus, although, similarly to Jacot de Boinod (2006), the present book relies on the selection of words for the analysis as motivated by introspection (see section 5.0), it further investigates the preselected set of words using more quantifiable methods, like those based on the information about frequency of occurrence of particular words from BNC (see sections 2.3.1 and 6.1).

The last monograph to be presented in this section is Ayto (2007). This book offers the conception of the lexicon as rather a dynamically organized entity, which differentiates the proposal from the static representation of lexicon as it is modeled in structuralist linguistic methodology. Ayto formulates this idea as follows:

> Do words shape an age or are they shaped by it? Are our actions and modes of thought channeled by the vocabulary available to us for giving expression to them, or do we continually create our lexicon anew to reflect a changing world? The two propositions are far from mutually exclusive, but no doubt the rationalist would wish to emphasize the second of them. Every year that passes throws up new ideas, experiences, and inventions for which no name has hitherto existed, and since names are indispensable cogs in the machinery of communication, our natural human propensity for coining them soon plugs most gaps.
>
> (Ayto 2007: 1)

In this book, the lexicon of the English will also be viewed as quintessentially dynamic in which the processes of assimilation constantly operate (see section 0.2), ensuring the incessant language change.

As far as the discussion on non-native lexical imports is concerned, again, we notice no clear methodological proposal as to the typological organization of this stock of vocabulary. As a result, there is no differentiation postulated between

[21] Again, for ease of reference, the crucial evaluative statements are italicized.

borrowings, foreign terms and internationalisms, all of which being classified as belonging within the same level category. Interesting is the metaphorical description of the process of introduction of non-native items in the target recipient system in which these items are conceptualized as 'thieves' passing into the target system unnoticed:

> It is not uncommon for words to escape immediate notice if they *slip unobtrusively into* the language rather than being announced with a fanfare and slang and other colloquial items routinely take some time to *find their way into* the written record. Sometimes English is slow to *adopt a term* for a phenomenon, movement, etc. which to modern eyes seems the obvious one, and sometimes, no doubt, the record of an earlier adoption has yet to *come to light*.
>
> (Ayto 2007: 2)[22]

[22] Other books worth mentioning in this context, which the author of the present book has not, however, accessed at the moment of writing this monograph include:

a) Finkenstaedt, T. & Wolff D. 1973. *Ordered Profusion: Studies in Dictionaries and the English Lexicon*. Carl Winter.

b) Lyovin, A.V. 1997. *An Introduction to the Languages of the World*. Oxford: Oxford University Press.

Chapter 5
CATEGORIZATION IN FOREIGN LEXICAL STUDIES

5.0. Establishing the criteria of foreignness. Preliminaries

This chapter presents issues relevant to the placement of foreign words and phrases in the categorial structure of the English lexicon with the simultaneous differentiation of that category from other manifestations of non-native lexis. The discussion begins with establishing the criterion of foreignness as understood in the book. Kuźniak (2007: 110–111) offers some glimpses into the possible avenues of search for the criterion by quoting an excerpt of the e-mail correspondence with John Wells (the editor of LPD), whose subject matter centers on the criteria of classification of words and phrases as "foreign" in the dictionary. It is to be reminded at this place that LPD constitutes the fundamental source of reference on the basis of which a preselection stage of foreign words and phrases is conducted (see section 0.2). The following are the most relevant parts of the aforesaid correspondence:

> MK[23]: In LPD, you quote a substantial number of foreign words and expressions. Alongside many of the entries you provide anglicized pronunciation and the pronunciation in the language of origin. This procedure works, as I said, for most cases, e.g. all proper names and places and many regular foreign words, however not with all of them. For example, *fandango* or *falsetto* (to quote just two entries), which are definitely foreign words (recorded as such in ODFWP), are not equipped with the pronunciation in the language of origin in your dictionary. Therefore I would like to ask you to comment on the criterion that you followed in providing the double pronunciation pattern for some foreign lexemes, and just single pronunciation for others, as this is the thing that keeps puzzling me for some time.
>
> JW: It is a matter of judgment how far words and names from foreign languages are still perceived as foreign. Personally, I have known the words *fandango* and *falsetto* since boyhood – the first as the name of an English folk dance, the second from my training as a choirboy – i.e. from before I studied foreign languages and became aware of the foreign origin of these words. I should think the same is true for many English people. Accordingly, it seems to me to be an irrelevance to give their Spanish and Italian pronunciations. In English no one pronounces them in an imitated Spanish or Italian way.
>
> The LPD entries that DO have a foreign-language transcription are those where people are – in my judgment – generally aware that the words are foreign and may therefore attempt to imitate the foreign pronunciation, or as dictionary users may seriously wish to know what the pronunciation is in the donor language.

[23] PC: MK – Marek Kuźniak; JW – John Wells.

> It is the difference between Warsaw and Gdansk. The English all call Warszawa /"wO: sO:/, and its Polish form [var"Sava] is irrelevant. The traditional English name of Gdansk is /"dæntsIg/, but even those who know this also know that it now has a different, Polish, name, which they are not sure how to pronounce properly – so it is my job to supply the authentic information.

As Kuźniak (2007: 112) further argues, the above-cited excerpt of e-mail correspondence appears to support the argument already alluded to in section 4.1, namely that the most vital criterion of classification of words and phrases as "foreign" at least in the preselection stage of their analysis inevitably relates to a native speaker's intuition. More reliable these judgments, however, appear to be if speakers' intuitions are confirmed in the specialized lexicographic source such as LPD edited by such the worldwide recognized authority[24] on language as John Wells.

Kuźniak (2007: 111), then, concludes that other criteria than phonological may provide us with some air of objectivity regarding the systematic description of the lexical categories in question, but as he says, "they finally appear rather inadequate as infallible tools in the analysis, which, as it has been presented above, is actually asserted by the authors of these classifications themselves" (see also section 2.1). In the light of all this, the set of foreign words and phrases analyzed in the book is based on LPD, which means adopting its criterion of foreignness as grounded in the phonological domain of a lexical unit (see also Preface).

Some justification for the above-mentioned phonological criterion can also be found in Görlach (2003: 84, after Mańczak-Wohlfeld 2006: 55), who argues:

> Although generalizations are risky, it can clearly be stated (but is not excitingly informative) that all languages here sampled, and more and more speakers of these, have become increasingly guided by native English pronunciation as regards phonological structure, allophones, and even articulatory features as a natural consequence of a vastly increased competence in spoken English.

It follows, then, that the primary level, at which foreign lexemes are "imported" (see Haugen 1950b) into the target language system, is via phonological structure.[25]

5.1. Definition and nature of borrowing

If the notion of borrowing appears to be pivotal for drawing possible demarcation line differentiating types of non-native vocabulary, particularly the category of foreign words and phrases, we have to examine in some detail the most recurrent

[24] For the discussion on the role of authority in language study, see Crystal (2007c: 81–87).

[25] See section 5.3 below, where it is argued that the validity of the phonological criterion of foreignness is associated with different foci within foreign lexical assimilation processes from the ones addressed to in the most significant works on borrowings.

stances on the definition of the term "borrowing" in the literature. Winford (2005: 42) poses the question directly:

> What exactly is a lexical borrowing? […] some are close imitations of foreign items [emphasis mine, M.K.]. […] Others are items that have been thoroughly transformed in shape […], while still others are inventions that employ only recipient language materials in imitation of some foreign pattern.

Winford, similarly to other eminent scholars (e.g. Görlach 2003, 2007; Mańczak-Wohlfeld 2006) in the field, conceives of the borrowing as a superordinate category[26] encompassing other types of more or less adapted instantiations of foreignisms. That this approach is not without problems is discussed in Kuźniak (2009b).[27] The major query relates to the inconsistency in the treatment of foreignness, which, on the one hand, is perceived as an attribute of borrowings,[28] while, on the other, regarded as exponency of the categorial status of a set of lexical items with intuitively distinct ontological status as non-native forms (see discussion below).

Certainly, the phenomenon of borrowing is called into existence via the contact[29] between languages. This common knowledge is clearly expressed by Winford (2005: 2):

> Most, if not all, languages have been influenced at one time or another by contact with others. In some cases, externally induced changes do not even require speakers of the

[26] The superordinate status of borrowings is also discernible in Haugen (1950b: 212), who argues that "the heart of our definition of borrowing is then attempted reproduction in one language of patterns previously found in another" (after Mańczak-Wohlfeld 1995: 13). Cf. also *Encyklopedia językoznawstwa ogólnego* (1993: 6161), where the category of borrowing is defined as the element (most often lexical, more seldom prefix or suffix) taken over from a foreign language (after Mańczak-Wohlfeld 1995: 13).

[27] On the definition of the term borrowing and some doubts connected with the semantics of the term "borrowing" itself, see Bauer (1998: 14).

[28] "The attributive nature of foreignness as pertinent to borrowings is clearly manifested in the following: Cases of maintenance may involve varying degrees of influence on the lexicon and structure of a group's native language from the external language with which it is in contact. This kind of influence is referred to as 'borrowing.' Since this term has been used in a variety of senses, it is necessary to emphasize that it is employed here, following Thomason and Kaufman (1988: 37), to refer to 'the incorporation of foreign features into a group's native language by speakers of that language'" (Winford 2005: 12).

[29] The nature of the contact between languages and the phenomenon of language change are widely discussed in the literature of the subject matter. The survey of the literature goes, however, beyond the scope of the present book. At this place we just quote Katamba (2006: 140) who attributes this contact to the pragmatic rationale of verbal behavior evinced by a speech community: "The concentration of borrowed words in certain semantic fields reflects the nature of the contact between speech communities. It reflects the areas where new words had to be acquired in order to fill a perceived gap." Cf. Buttler *et al*. (1973: 41) who notice that there is no correlation between the number of borrowed terms in the receiving language and communicative efficiency in that language (Mańczak-Wohlfeld 1995: 15).

> different languages to have actual social contact. For instance, lexical borrowing can be accomplished through book learning by teachers, writers, lexicographers, and the like who pass on the new vocabulary to others via literature, religious texts, dictionaries, and so on.

What is crucial for the argumentation advanced in the present book is the notion of contact between the two entities as this, according to laws of physics, implicates the action of force. This fundamental observation is included in section 0.2 (see also Preface) and elaborated in detail in Parts III and IV of this study.

If we conceive force to be determinative of the nature of the process of borrowing from the donor language, it is easier to understand the division of borrowings quoted by Katamba (2006: 135) into "direct" and "indirect" subtypes: "It is useful to distinguish direct borrowing and indirect borrowing. If a language takes a word directly from another, as English got *omlette* from French, we call what happens DIRECT BORROWING. But in other cases a word may be passed indirectly like a relay baton from one language to another, e.g. *kahveh* (Turkish) […] – coffee (English). This is called INDIRECT BORROWING. If a word is directly borrowed, the chances of its undergoing drastic phonological modification are considerably less than those of a word that is indirectly borrowed." This observation is in compliance with the methodological framework of description presented in Chapter 6, in which the lesser degree of phonological modification in the case of "direct borrowings" is attributed to the Law of Conservation of Energy (see section 6.5). This regularity also appears to be at work in the case of semantic shift which may be witnessed in the process of borrowing. Again this works accordingly with the principle: the more direct the borrowing, the less of the semantic alteration it undergoes in the receiving language (Katamba 2006: 136).

Naturally involved in the contact between the two languages is the cultural character of the phenomenon. As a consequence, the term "cultural borrowing" was coined which shows us "what one nation has taught another" (Bloomfield 1933: 458).[30] Earlier, Sapir (1921: 207–20) as one of the first researchers, emphasizes the connection between linguistic borrowings and the development of culture.[31] That this process is natural in the strict meaning of the word is also noted by Jakobson who observes that borrowings cannot be eradicated from a given language and links this illusionary hope about the self-sufficiency and the resulting 'purity' of a language to the hope about economic self-sufficiency (Mańczak-Wohlfeld 1995: 15).

Quite neutral assessment of the process of borrowing is provided by Deroy (1956), who argues that borrowings enrich, on the one hand, the lexical system of the recipient language, and on the other, supersede the already existent elements, which contributes to its impoverishment (Mańczak-Wohlfeld 1995: 15). This ob-

[30] See also Mańczak-Wohlfeld (1995: 14)

[31] *Ibidem*. The scholar quotes Cienkowski (1964), who similarly to Sapir argues that the analysis of borrowings has a significant place in the research on the evolution of culture, and cultural relations (Mańczak-Wohlfeld 1995: 14).

servation about the balanced nature of the phenomenon of non-native lexical takeovers stands in clear opposition to the aforequoted Katamba (2006: 140, see Ft. 29 in section 5.1). All in all, we agree that borrowings are primarily understood in the literature as items that are permanently assimilated[32] to the system of the recipient language. If this conception of borrowing is to be upheld, its impact on the system of a recipient language should be multilevel. Katamba (2006: 148–149) discusses these multidimensional effects of borrowings.[33] He distinguishes the following aspects:

> a) Phonological – introducing new phonemes into the English phonological system,
> b) Grammatical – introducing new allomorphs of the plural morpheme,
> c) Semantic effects: "Adding a new word may disturb the equilibrium of the words already in the language, causing semantic narrowing, for example. The new borrowed word may take over part or most – but not all – of the meaning of the original word, and the original word may survive with a restricted meaning."

In conclusion, major outcomes of language contact, which itself can vary from casual through moderate to intense are the following (after Winford 2005: 23):

- contiguous geographical location,
- intra-community multilingualism,
- intense pressure on a minority group,
- intense inter-community contact (trade, exogamy).

It appears that geographical proximity as well as "intense inter-community contact" will constitute the most relevant factors determining the process of lexical assimilation, the discussion of which is presented in more detail in section 6.8.3.

5.2. Foreign words and phrases in typological perspective

5.2.1. Typology of borrowings

Consistently with the argument presented in the preceding sections, borrowings can be claimed as lexical units viewed as non-native by nature, i.e. the ones that have either been fully assimilated into the recipient system or those that function

[32] Cf. Polish equivalent of the term borrowing, i.e. *zapożyczenie*, where the derivational prefix *za-* designates the end-result (completion) of the assimilation process. See section 5.3 (Fig. 10) for the argument in which a borrowing is viewed as the subcategory of non-native lexis that has undergone a complete process of adaptation.

[33] Degrees of interference of a borrowing in the recipient languages are also discussed by Thomason and Kaufman (1988: 40), who differentiate 4 levels: very strong, strong, moderate, weak. They support their arguments discussing substratum influence of Yiddish upon English.

as a cover term designating other subcategories of non-native lexical or grammatical stock. Thus, we receive two mutually conflicting representations of borrowings at a conceptual level, i.e., once as basic level category, the other time as superordinate. In discussing the typology of terms undergoing lexical contact with the items belonging within another system, Haugen (1953) proposes quite original typology with a highly comprehensive category of "lexical contact phenomena," which he posits at the highest level of conceptual schematicity. According to Haugen:

> Lexical contact phenomena can be divided into two broad categories – *lexical borrowings,* which involve imitation of some aspect of the donor model, and *creations*, which are entirely native and have no counterpart in the donor language. Lexical borrowings can further be subdivided into two categories. First, there are *loanwords*, in which all or part of the morphemic composition of the loan derives from the external donor language. Second, there are *loanshifts*, in which the morphemic composition of the item is entirely native, though its meaning derives at least in part from the donor language. […] Loanwords may be divided into two categories; "pure loanwords" and "loanblends." Pure loanwords may consist either of single words like *rendezvous* or compounds like *chincibiri*.
>
> (After Winford 2005: 43)

This morphologically based typology, although aiming at ordering the categorial organization of the non-native vocabulary from the view point of the recipient language, suffers from some drawbacks in that the term "pure loanwords" is not, for example, further explicated. The use of the adjective 'pure' conventionally implicates the existence of "impure" or "not so pure" categories of non-native lexis. Haugen, unfortunately, discusses the typology of linguistic borrowings as if all of the categorial subdivisions were discrete in character. In other words, Haugen does not admit of any gradation of membership of a given non-native form to a particular subcategory. Another problem involves the use of the aforesaid morphological criterion in differentiating between different categories of non-native units. This criterion is just one of the aspects of a potential imitation that occurs in the process of the lexical transfer from one language into another. The other aspects implicated in the definition of "lexical borrowings," aside from the aforementioned morphological one, are however, not dealt with in the typology. Setting these methodological problems aside now, we may conclude that the category of foreign words and phrases as we understand it in the present book approximates the concept of "pure loanwords" (see also sections 6.7.3 and 10.1.1).

According to Katamba (2006: 137) there are just two kinds of borrowing: LOANWORDS and LOANSHIFTS (see Haugen 1950b). Katamba (2006: 137) continues:

> A loanword is a word belonging to one language which is IMPORTED or ADOPTED by another, e.g. *catamaran* was imported into English from Tamil and *shopping* was imported into French *le shopping*. By contrast, a loanshift involves taking on board the meaning represented by a word in a foreign language, but not the word-form itself. Loanshifts are also called LOAN TRANSLATIONS or CALQUES.

The advantage of this typology is that the non-native vocabulary is divided into two conceptually salient and distinct groups. The weakness is that the category of "loanwords" is presented as no further analyzable homogeneous entity. That such homogeneity is not justified has already been tentatively addressed above and is elaborated on in section 5.3 where the argument is advanced about the different degree of adaptation of non-native units and hence the resulting different psychological status possessed by these units in the lexical competence of the users of a recipient language.[34]

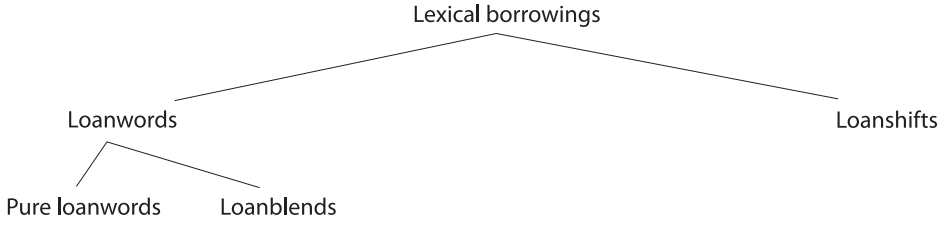


Fig. 9. *Typology of borrowings*[35] (After Haugen 1953)

The typologization issues are also discussed in Mańczak-Wohlfeld (1995). The classification of lexical borrowings can be instantiated by the following:

1) Loanword: *budget/budżet*,
2) Loanblend: *drewland*,
3) Loanshift: *teenager/nastolatek*.

(Mańczak-Wohlfeld 1995: 17)

Mańczak-Wohlfeld (2006: 49) continues to deal at a length with borrowings.[36] She offers a heterogeneous explication of "borrowing" which encapsulates:

[34] See, e.g. Katamba (2006: 144–145) on the explicitly formulated gradient nature of the category of foreign words and phrases: "Many foreign words that are borrowed become fully NATIVISED. In the case of English, nativisation means ANGLICISATION. The words become assimilated and undistinguishable from indigenous English words."

[35] As Winford (2005: 51–52) argues: "Hierarchies of borrowing were proposed as early as the nineteenth century by Whitney (1881), and later by Haugen (1950b) and Muysken (1981). The most comprehensive of these is the following from Muysken: nouns> adjectives> verbs> prepositions> co-coordinating conjunctions> quantifiers> determiners> free pronouns> clitic pronouns> subordinating conjunctions […]. Syntagmatic constrains relating to the morphological and syntactic properties of lexical classes may also operate to favour or inhibit borrowing. This may explain why categories like verbs or prepositions, which govern other categories and assign case, tend not to be as heavily borrowed as nouns and adjectives."

[36] See also Mańczak-Wohlfeld (1994, 1995). It is to be emphasized that Mańczak-Wohlfeld excludes from her considerations over English borrowings in Polish those items that constitute the core of the analysis in Parts III and IV of the present book. Thus, Mańczak-Wohlfeld (1995: 8) does not on the whole take into consideration items, which she labels after Doroszewski (1938: 47) as *cytaty* (Eng. citations). These are characterized by the spelling and pronunciation of the language of origin.

– Assimilated lexical items in the target lexicon (Category 1),
– Non-assimilated but frequently occurring borrowings (Category 2),
– Citations, which are unknown to many readers, yet introduced by journalists to attract readers' attention (Category 3).

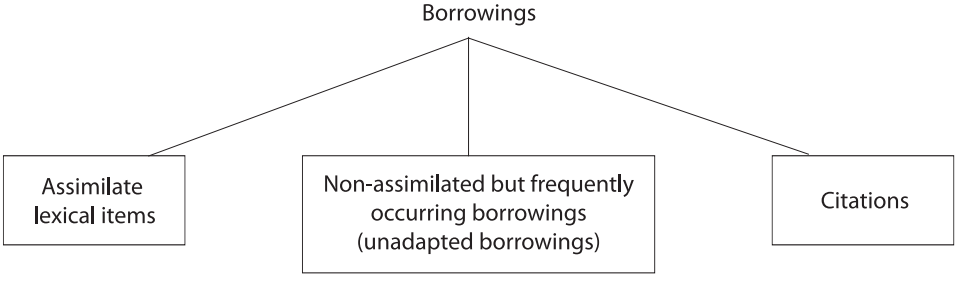


Fig. 10. *Typology of borrowings according to Mańczak-Wohlfeld (2006: 49)*

A triadic model of classification of borrowings is also upheld by major European scholars[37] who usually distinguish the following types of borrowings (after Busse and Görlach 2007: 29), which:

a) relates to totally unadapted and not felt to be part of the receiving language,
b) refers to words still looking foreign in form or entirely unadapted,
c) subsumes fully integrated items.

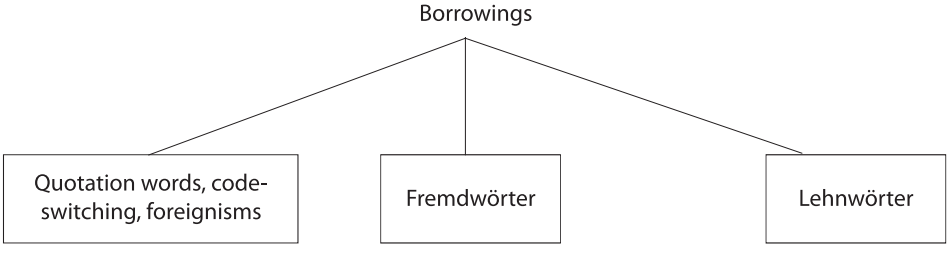


Fig. 11. *Busse and Görlach's typology (2007)*

An interesting typological discussion is conducted in Introduction to DEA. DEA offers a highly insightful pioneer investigation into the status of English words in 16 selected European languages. What is particularly relevant to the

[37] See, e.g. Rusiecki (1980) for his triadic typology of borrowings (loans) into *fully assimilated*, *unassimilated* and *part-assimilated* units. The criterion of assimilation adopted by the scholar is, however, formal (i.e. grammatical). Thus, part-assimilated loans are categorized on account of their potential to be inflected or combined with derivational suffixes. Part-assimilated loans should, therefore, be viewed as distinct from fully assimilated units in that the latter undergo adaptation at graphical, phonological, morphological, and semantic level. Part-assimilated loans should, in turn, be carefully distinguished from unassimilated loans in that their original spelling is preserved (Otwinowska-Kasztelanic 2000: 19).

ongoing discussion is the description of the category of "anglicism." According to DEA, an anglicism is a superordinate category encompassing non-native items with varying degree of integration with the system of the receptor language at least at one of the following levels: phonological, morphological, graphemic, semantic. As it is argued in DEA (p. xxiv), an "anglicism" represents "a cline of increasing integration (and also of frequency and acceptability)," which is realized in the following stages:

- The word is not known but a calque or another equivalent is provided;
- The word is known but it is a foreignism – that is, it is used only with reference to British or American contexts;
- The word is in restricted use in the language;
- The word is fully accepted and found in many styles and registers, but is still marked as English in its spelling, pronunciation, or morphology;
- The word is not (or is no longer) recognized as English.

Although DEA rightly points out to the gradual character of the integration processes, it appears to unnecessarily go into detailed specifications as the ones shown above, especially in view of discussion upon the nature of foreignness in the dictionary. All in all, as DEA concludes, the ultimate judgment about the "foreign" status of the analyzed word or phrase appears to be a part of the subjective judgment of the lexicographer, whatever the criteria for exclusion or inclusion of English words under consideration: "Needless, to say, all these categories provide many doubtful cases, and in consequence, subjective decisions – a problem magnified by the large number of collaborators" (DEA, p. xxvi).[38]

5.3. Towards the target typology of foreign words and phrases

In the present book we propose a triadic typology of foreign words and phrases but the difference between the aforementioned proposals (see Figs. 10 and 11) and the one presented in the current research lies in the perspective taken, i.e. elevation of foreign words and phrases to the status of basic level category with the category of "non-native" lexis as elevated to the status of the superordinate category subsuming all instances of non-native words and phrases with varying degrees of adaptation to the target language system. An advantage of such approach is that it allows some terminological ordering in the relevant lexicological research. The terminological confusion observed in many lexicological or lexicographic works involves either the frequently interchangeable usage of the term "borrowing" and that of "foreign

[38] An interesting account of the definitional status of anglicisms in Polish can be found in Piotrowski (1998: 271–273; 2005: 503–510). See also an interesting discussion on the issue of the place and function of selected anglicisms in the Polish language in Miodek (1971, 1980).

word" (cf. section 5.2), or understanding borrowings as hypernymous with regard to foreign words and phrases (see Görlach 2007 and Fig. 11). Thus, we are confronted with the problem of the correct identification of the target category of foreign word or borrowing at a conceptual level of representation.

A possible disadvantage behind the proposal advanced in this book is that presenting such typology may lead to the criticism about redundant proliferation of typologies in view of the already established conventions in lexicological studies (cf. Görlach 2007). That this new typological framework is not claimed to be a pure terminological shift but offers a new cognitive value to it, is argued below (see Fig. 12 and the ensuing discussion) and is subject to further investigation under the descriptive framework developed in Chapter 6 and summarized in Chapter 10. The typology of foreign words and phrases proposed in this book is, thus, as follows:

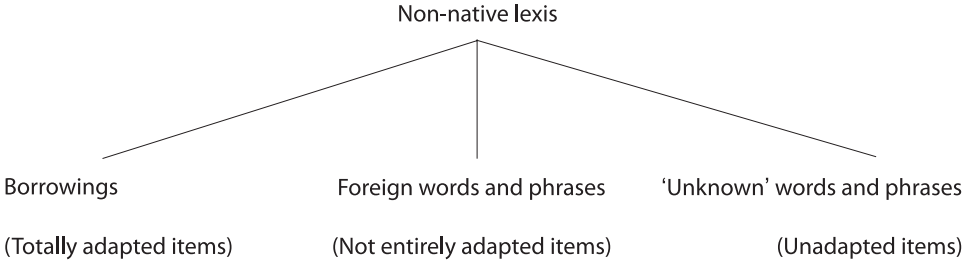


Fig. 12. *Typology of foreign words and phrases*

A significant argument in favour of postulating the above-illustrated typology comes not only from the already indicated considerations in Chapter 6, but also from the model of assimilation processes elaborated on in Chapter 9, whereby three identifiable stages in the process towards complete adoption in the target language system can be pinpointed. These three stages or phases in assimilation are naturally in accordance with the triadic typological distinctions discussed above (Figs. 10 and 11), whereby the criteria of typologization range generally along the continuum from the full integration with the target language system to the complete lack of adaptation.[39] This mechanism is characteristic of all typologies

[39] Problems in typologization of non-native words and phrases are also noticed by Katamba (2006: 148) who states the following: "As foreign words become fully integrated in a language, the pressure to make them conform to the standard rules is often irresistible. But, until they are fully assimilated, it is to be expected that speakers will treat them differently." Earlier, he says: "Evidently, there is no real problem in cases where the foreign import has been fully adopted and integrated into the English lexicon for so long that anyone who is not especially knowledgeable about etymology would be unable to sniff out its foreignness. […] Words may resist nativisation to a greater or lesser degree. Even after a long period of use in English some words fail to become fully adopted. Instead, they remain on the fringes, as tolerated aliens with one foot in and the other foot out of the English lexicon" (Katamba 2006: 145, cf. Mańczak-Wohlfeld 1995: 14).

Finally, Katamba (*ibid.*) states as follows: "There is no set of exact, scientific principles that can enable us to infallibly separate foreign words used in code-switching from nativised borrowed

discussed in the field no matter whether the postulated superordinate category is "borrowing" or "non-native lexis." As we already suggested in Introduction to this book, we selected a phonological criterion as primary for the collection of foreign words and phrases conceived of as a basic-level category (see section 5.3.1). It is, thus, the adaptation at the phonological level that constitutes an underlying mechanism of general assimilation processes discussed in detail in Chapter 9, the outline of which is delineated below:

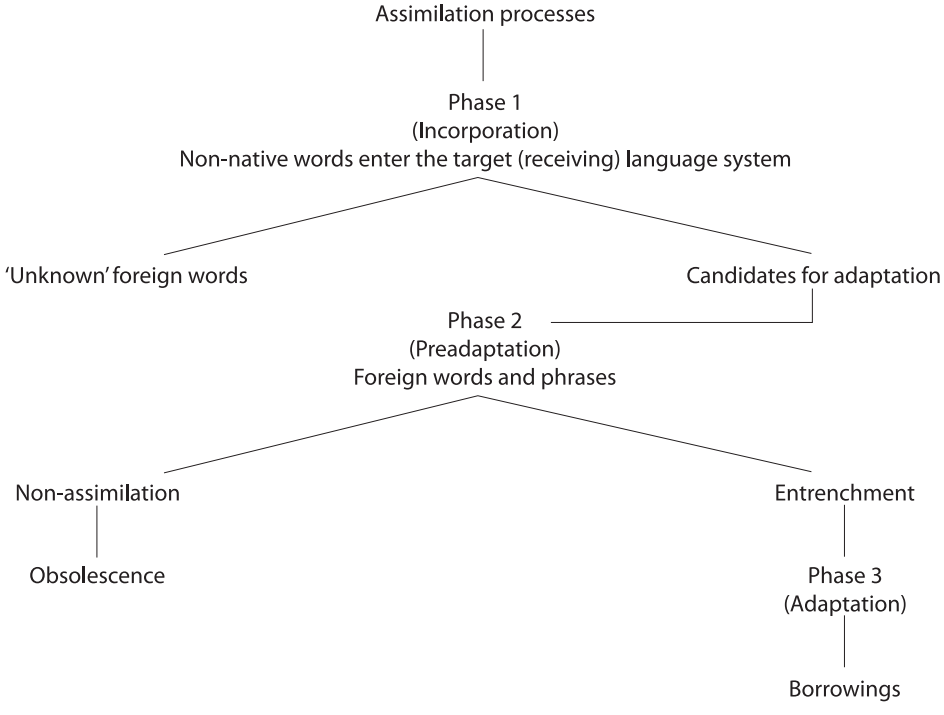


Fig. 13. *An outline of the model of foreign lexical assimilation processes*

Italicization of the three levels in the above-presented figure is a typographical aid that should help the reader place the discussion on categorization of

words." If the clear-cut demarcation line cannot be approximated in a systematic way though scientifically rigorous methodology, the criterion of differentiation is thus reconstructed via introspection and displays the following tendencies:

1. Foreign grammatical properties may be ignored when a borrowed word is assimilated into the grammatical system of English […].

2. If a word is not perceived as foreign any more, writers stop giving it special treatment. Any foreign marks and diacritics used in its spelling disappear. They stop italicizing it or putting it in inverted commas, or offering a gloss, or doing anything to draw attention to it any more than they would an indigenous word […].

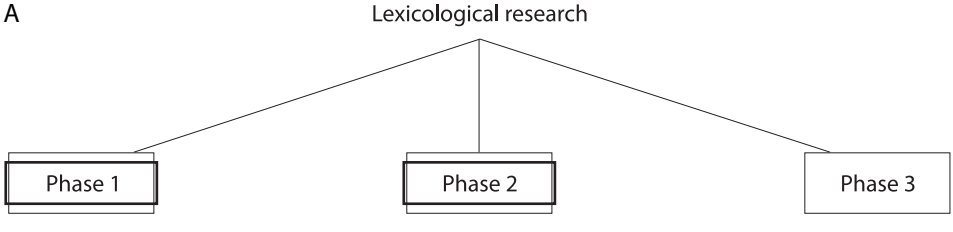
3. In the spoken language, the more nativised a borrowed word is, the more it is made to fit in with the standard rules that govern the pronunciation of words in the host language (after Katamba 2006: 146–147).

foreign words and phrases in the book along with the proposed distinction introduced in Fig. 12 (see also Fig. 18). We thus observe that the elevation of foreign words and phrases to the category of basic-level appears to be consonant with the impressionistic top-bottom ICM of assimilation processes, whereby "foreign words" along with "unknown words" and "borrowings" constitute identifiable salient stages in the process of adaptation, which posits these categories at the same level of conceptual organization. What is highly relevant to the argument is that the category of foreign words and phrases represents the central stage in the assimilation process. This leads us further to the investigation of the core of this central category in Chapter 8 (see section 8.5).

It may therefore be argued that the elevation of borrowings into the status superordinate term is justified inasmuch as the lexicological research in the field (cf. Görlach 2007; Mańczak-Wohlfeld 1994, 1995, 2006, 2007) concentrates mainly on Phase 3, i.e. the description of phonetic, graphemic, morphological and semantic integration of a foreign unit with the target system. No wonder, then, a major preoccupation in such research is the systematic analysis of the degree of adaptation of non-native lexis to the target language with regard to the three levels of integration mentioned above. It is thus natural that the most important part in such discussion is spared on lexical units that fulfill the definitional criteria of borrowings, i.e. fully adapted non-native lexical units.

In view of arguments just formulated, it appears more than whimsical to see the necessity of conceptual reordering proposed in this book in that foreign words and phrases are suggested to occupy the basic-level position on a category hierarchy. In the same way as we accounted for the elevation of borrowings into the superordinate level in the discussed lexicological works by pinpointing the authors' foci on the description of Phase 3 with the concurrent relegation of foreign words to a subordinate status, we, consequently, postulate to elevate foreign words and phrases to the basic level status in the current research. This leads to viewing foreign words and phrases as a salient conceptual category associated predominantly with Phase 2. As the category is prototype-based, its boundaries are fuzzy and thus merging with the boundaries of other conceptual categories, i.e. "unknown words" and "borrowings" associated with Phases 1 and 3, respectively (see section 5.3.1).

Our preoccupation lies, then, in the description of Phases 1 and 2 (see Parts III and IV) as Phase 3 – characteristic of borrowings – has already been widely discussed in the literature. This difference in the focus on research is displayed in Fig. 14:



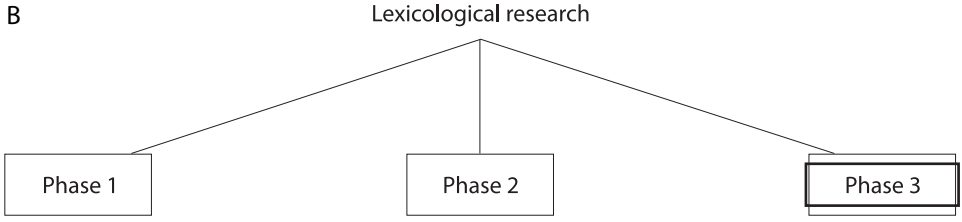


Fig. 14. *Different foci in lexicological research upon assimilation of non-native lexis*
A – The present book; B – Major lexicological research on non-native lexis (Görlach, Mańczak-Wohlfeld[40], Fisiak 1970 and others)

One more illustration (Fig. 15) shows what has been signaled above, i.e. prominence of the phonological level in the assimilation process, which in contrast to other levels (graphemic, morphological, semantic), underlies all three identified stages, not only Phase 3.

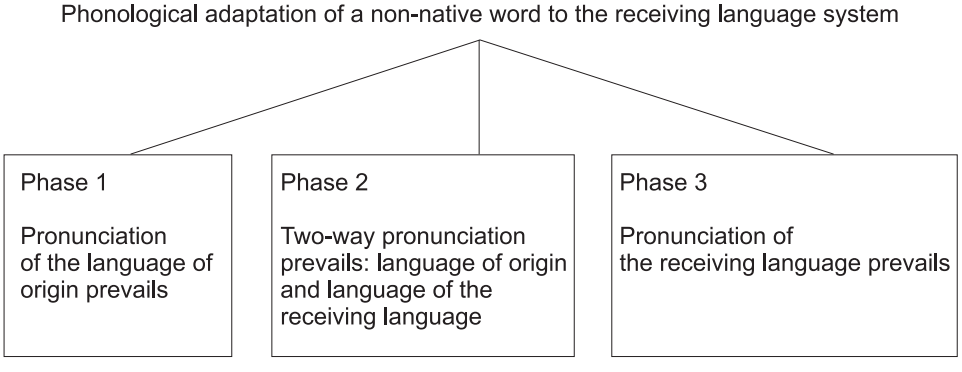


Fig. 15. *Phonological adaptation in respect of the phases of assimilation*

5.3.1. Foreign lexis in conceptual categorization framework. Consolidation

Once the criterion for the description of the ontological status of a foreign word or phrase has been provided (see sections 5.0–5.2), it may appear relevant, as a conclusion, to situate the category in some larger categorization perspective. In this respect, it might be helpful to quote Scheler's (1977: 139) typology. Scheler illustrates English vocabulary by means of the concentric circles representing four levels ranging from the centre to the periphery. The core of the English lexicon is occupied by *common* vocabulary (the first level). The second level comprises *formal*

[40] See the definition of borrowing quoted by Mańczak-Wohlfeld (1995: 14), who argues that borrowings constitute the elements of a foreign language that are assimilated in the recipient language and entrenched in it. This clearly corresponds to Phase 3 illustrated in Fig. 11.

(literary)/colloquial vocabulary. The third level is the most numerous in terms of categories and includes *slang*, *dialect*, *technical*, *scientific*, and *foreign* words. The fourth – the outermost level – is occupied by *poetical*, *archaic*, *vulgar*, *cant* varieties of vocabulary.

If we extrapolate Scheler's taxonomy onto the findings presented in section 5.3, we obtain the following tentative taxonomy[41]:

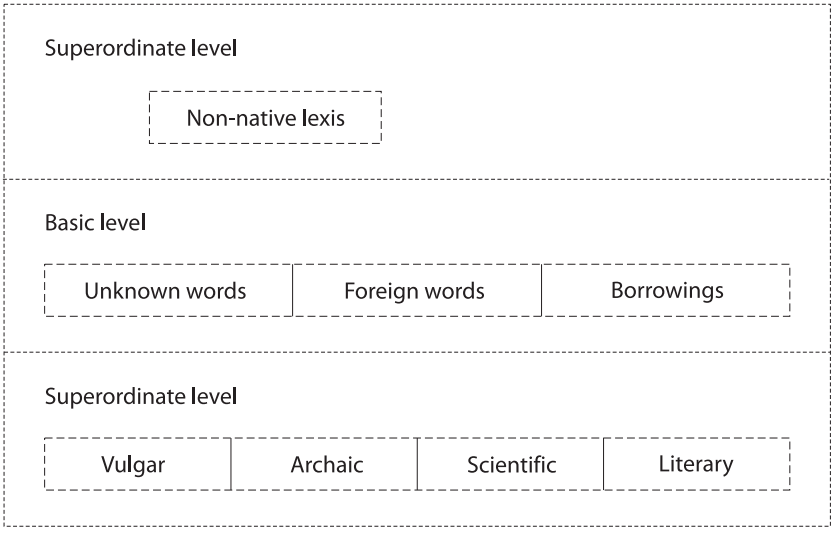


Fig. 16. *Conceptual organization of foreign lexis* (See Rosch's [1977] levels of categorization)

Superordinate level represents the most general division of the English non-native lexicon. The postulated term "non-native" appears as the most adequate at this level as it comprises all different instantiations of foreign vocabulary regardless of their level of adaptation to the receptor language. The non-native lexis is sub-divided into "everyday" or conversational vocabulary characteristic of spoken English on the one hand, and formal repertoire of lexical items reserved for the primarily written institutionalized English, on the other. It is along the lines of the colloquial/formal dichotomy that the lexicon at the basic level is organized. The categories at this level exhibit a higher degree of conceptual salience in that they are maximally distinct from each other.

As Fig. 16 indicates, foreign words and phrases on which the present book focuses occupy the basic level. This is in line with argumentation advanced in section 5.3. The boundaries between respective levels are of course fuzzy, but placing foreign words and phrases at the basic level carries important cognitive implications, one of the most significant being the aforementioned conceptual salience

[41] Cf. Kuźniak (2007). Also notice the close connection between the conceptual organization of foreign lexis presented here to the typological model advanced in sections 6.7.3 and 9.1–9.2.

(see Rosch 1977, Taylor 1995). Bearing this in mind, the criterion of category assignment based on the introspective judgment of the native speaker (human conceptualiser) assumes an additional cognition-supported validity (see section 5.0). The subordinate level comprises different styles that provide the characterization for the categories at the basic level. Thus, foreign words and phrases may be characterized by the predominance of scientifically-oriented vocabulary, but other styles are also identified.

This should be rather seen as a tentative observation, which practically means that a far lot more comprehensive research should be conducted to systematically capture all the intrinsic relations between the presented levels. The important is the realization of the fuzzy boundaries existent between particular conceptual levels (vertical dimension of the taxonomy) as well as between the categories representing particular levels (horizontal dimension). These non-discrete boundaries are typographically marked by dotted lines.

At this point we are able to precisely formulate the research objectives and designate its scope. We will henceforth focus on foreign words and phrases understood as a prototype-based category posited at the basic level of conceptual organization (bold dotted lines in Fig. 17), the periphery of which determines the scope of the lexical material analyzed and discussed in Parts III and IV of the book (regular dotted lines, see Fig. 18).

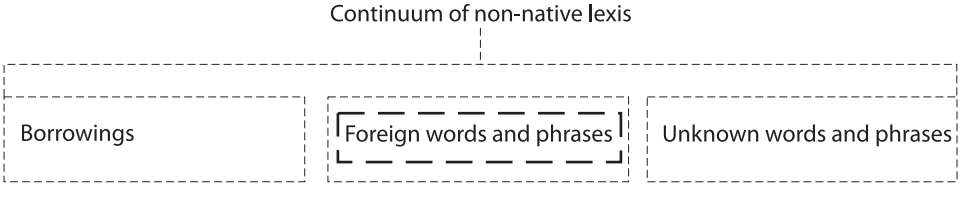


Fig. 17. *The category of foreign words and phrases in view of research scope and objectives in the present book*

Once we consolidate our findings from Figs. 16 and 17 into one typological representation as based on the aforementioned phonological criterion (see sections 0.1 and 5.0), we obtain the following hierarchically organized conceptual framework of foreign words and phrases (see Fig. 18).

Fig. 18 displays three levels of conceptual organization of foreign words and phrases. The top-most level is occupied by non-native lexis which represents all types of foreign language lexical import in the recipient language. At this level, phonological criterion is not relevant. The middle or basic level is occupied by three distinct sets of non-native lexis. These are differentiated on the basis of phonological criterion, yet the category boundaries at this level are fuzzy as shown by dotted lines (see also Fig. 17). The bold-typed marking of the category of foreign words and phrases indicates the area of focus in the present study (see sections 0.1 and 5.0). Finally, the lowest level represents different instantiations of foreign words and

phrases. Dotted lines conjoining particular subtypes of foreign words and phrases illustrate the continuum nature of the category as such. Again, prototypical foreign words and phrases, defined later as "core meteorites" (see section 8.5) are marked in bold-typed lines for emphasis. The differentiations manifested at bottom-most level anticipate the results of the research presented in sections 8.6–8.6.1 of the book.

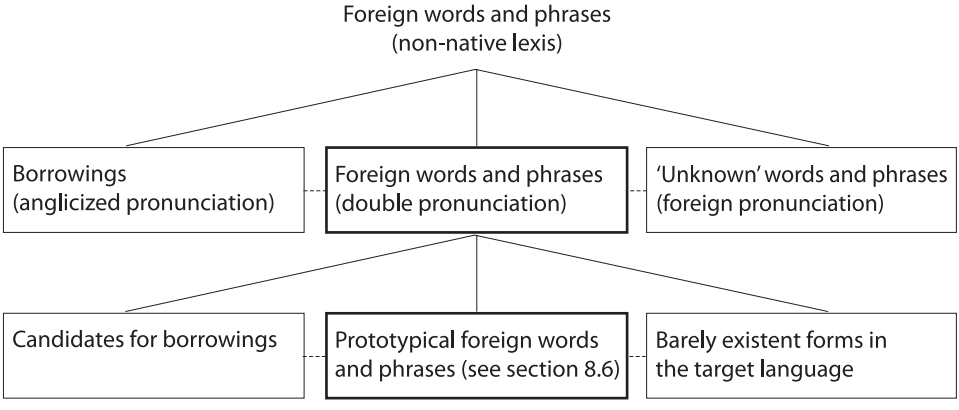


Fig. 18. *Target typology of foreign words and phrases*

5.4. Proper names and common nouns

5.4.1. Proper names

This chapter concludes with the discussion of proper names[42] as a subcategory, along with common words and phrases, of foreign words and phrases. As regards linguistic contributions to the study of proper names, these are rather seen as incidental, and if any systematic analyses are undertaken, these are usually confined to various typological classifications.[43] What is noteworthy, however, is that

[42] For classic publications on proper names, see: Mill (1843), Frege (1892), Russell (1918), Strawson (1950), Wittgenstein (1953), Searle (1958), Geach (1962), Donnellan (1972), Kripke (1972), Dummett (1973); all after Berezowski (2001: 9).

[43] Proper names are usually divided into two fundamental groups: personal names (Pol. *antroponimy*) and geographical names (Pol. *toponimy*). Other categories identified for proper names include: animal names (Pol. *zoonimy*), plants (Pol. *fitonimy*), firms and institutions (Pol. *ideonimy*) and other artifacts like cosmetics, food, appliances, cars (Pol. *chrematonimy*) – see Grzenia (1998: 19–20). Fraurud (1996: 81, in Berezowski 2001: 25) includes the following list of typological distinctions (cf. Allerton 1987): "Persons, animals, some classes of artifacts, such as certain vehicles (trains) and vessels (boats, ships), works of art (books, paintings, sculptures), periodicals (newspapers, magazines), social organizations (institutions, political parties, companies),

proper names as such have never posed a significant research problem to linguists (Berezowski 2001: 10).[44]

Traditional linguistic approach has seen proper names as essentially nominal categories (Priscian and Donatus), but as Berezowski (2001: 45) continues: "Grammarians following this line of reasoning view thus proper names as a subdivision of nouns (hence their preference to talk about proper nouns rather than proper names) and try to separate them from common words." In the present book, this traditional 'nominalist' account of proper names is to a large degree upheld and, as the argument shows (see section 6.7), this treatment receives some cognitive justification, given the descriptive framework developed in Chapter 6.

In the literature on proper names we find different criteria by means of which an attempt is made to differentiate between the types of nouns. Thus, morphosyntactically, proper names differ from common nouns in that they exhibit:

a) Lack of inflection for number (Bloomfield 1933: 205),
b) Lack of postmodification by restrictive modifiers (Seppänen 1974: 272–274),
c) Definiteness in spite of absence of any overt determiners (Sloat 1969: 26–30).

(after Berezowski 2001: 45)

The above-presented features display of course just tendencies and many counterexamples can be found. Berezowski (2001: 52) thus concludes:

> The traditional approach fails because of its insistence on finding some morphological/syntactic properties which would set common words off from proper nouns and does not appreciate the key insight of the logical approaches to the issue. Logicians and philosophers realized long time ago (e.g. Strawson 1950) that proper names primarily serve to refer, i.e. they differ from other expressions not in their structures or grammatical properties but in the use they are most frequently put to. The notion of the proper name is thus inherently pragmatic and any syntactic or morphological attempts at defining it are doomed to fail since all they can do is focus on some marginal properties of proper names and not on their most fundamental characteristic. However, in order to capture these key features it is necessary to peer beyond the syntax and morphology of proper names and recognize the fact that their closest neighbour is not the common noun but the definite description.[45]

---

and geographical locations. While persons, works of art, periodicals, organizations, and some varieties of locations obligatorily have names, the naming of animals and artifacts such as vehicles and vessels is optional."

[44] What may partly explain this lack of interest in proper names by linguists is that some of them (e.g. Algeo 1973: 63; Napoli 1997: 185–188) appear to relegate them to the phenomenon that lies outside the bounds of language (Berezowski 2001: 62).

[45] See also Jespersen's (1924) and Kuryłowicz's (1956) study of proper names, in which they are viewed "as the word class with a maximum of content but a minimum of referents, i.e. the intension of a noun considered inversely proportionate to its extension" (after van Langendonck 2007: 39).

The above-cited observation appears only to bear out the argument that proper names and common nouns belong ontologically within the two different categories (this is actually what traditionalists also claimed), but what is crucial in these considerations is that the boundary between the two categories is essentially fuzzy, i.e. the judgment of whether a particular form pertains to proper or common words is in some cases the matter of degree. This brings forth a cognitive-linguistic approach with its non-Aristotelian view of categorization in which the subtleties and intricacies involving category membership assignments are well accounted for.[46]

Worth quoting at this point is the so-called pragmatic view of proper names (Coates 2006). According to van Langendonck (2007: 66):

> For Coates, proper names are meaningless[47] NP expressions that refer at discourse level. Coates does not seem to recognize a linguistic level of denotation (extension), at least not for proper names, since properhood is defined at the pragmatic level of discourse, i.e. of language use. […] For a linguist, the everyday fact that the same name may apply to more than one individual must deal a fatal blow to the notion that to be proper is ESSENTIALLY to denote uniquely.

If the prerequisite of unique reference is not absolute for the identification of nature of proper names as opposed to common words, we need to look at the issue in a little more detail if we are to advance the argument about the psychologically real foundation for the categorial difference between the two sets of lexical units (see sections 8.4.2.1 and 8.4.2.2).

As it turns out, according to Lyons (1999: 21), the category of proper names is highly heterogeneous instantiating expressions which, on the one hand, possess internal grammar or/and exhibit descriptive content (e.g., South Farm Road) and those that, on the other hand, do not display such characteristics (e.g. John). Similar observation is made by Strawson (1950: 338, in Berezowski 2001: 36) who argues that "besides 'pure' names there are also the 'impure' ones, i.e. those which do have descriptive content but 'have grown capital letters' and occupy a midpoint position between proper names and definite expressions, e.g. the Round Table."

---

[46] In addition, van Langendonck (2007: 39) makes the following observation: "Cognitive linguistic approaches to proper names (e.g. Marmaridou 1989 and Langacker 1991) seem to initiate a revival of the maximum meaningfulness thesis." As van Langendonck (2007: 51) continues, Marmaridou (1989), for instance, suggests that proper names are to be rather seen from a communicative angle as "a more efficient and economical means of communication," whereas Langacker (1991), analyses proper names from the perspective of common words, which according to van Langendonck (2007: 51) is a manifestation of a lack of concern for some essential attributes of proper names. As he concludes: "As a consequence of the analysis of proper names which I set out in this chapter is that this nominal subclass is considered to represent the prototypical noun since its primary function is to refer to an object or person. By contrast, common words are less prototypical nouns since they contain a predication, which is in the first place a verbal feature" (van Langendonck 2007: 51).

[47] Certainly J. Coates was not the first to claim that proper names do not mean, but refer (see Ft. 42 in section 5.4.1 for the classic literature in this respect).

To sum up, proper names can thus be viewed as a prototype-based category of nominal concepts (see also van Langendonck 2007: 79), the salient attributes of which are as follows:

> – They are bestowed in the speech act of naming[48] by duly authorized individuals,
> – They refer to predefined referents,
> – They are not sensitive to changes in the referent/context,
> – They are used successfully if the hearer has been inducted into the chain of people familiar with the reference.
>
> (after Berezowski 2001: 60)

5.4.1.1. Foreign proper names

Since proper names, aside from common words and phrases, constitute the fundamental typological distinction of foreign words and phrases analyzed in the present book, it appears necessary to examine in some detail how foreign proper names have so far been dealt with in linguistic studies. One such aspect of investigation relates to spelling convention. As Grzenia (1998: 21) puts it: "In the majority of cases foreign proper names are spelt accordingly with the original convention." This is, however, not so straightforward in some cases in which proper names are not spelled accordingly with the conventions of the language of origin due to a high degree of entrenchment of a lexical unit in the recipient language. The use of such original spelling would then appear as highly pretentious. Sometimes, the reason for spelling simplification is other than the one related to the frequency of the unit in language. For example, such enforced simplifications occur when texts are spread via electronic channels (Grzenia 1998: 22).

An interesting part of Grzenia's considerations over foreign proper names concerns the postulated process of accommodation of foreign proper names into the target language system. This is very significant to the present book as it accords, at least as far as the outline framework is concerned, with the processual account of foreign lexical assimilation model developed in Chapter 9. Strangely enough, Grzenia's remarks are limited to proper names, although, it may be argued that the discussed stages can also be found applicable to foreign non-names (i.e. common words). Thus, Grzenia (1998: 28) identifies the following stages in the process of the aforementioned accommodation:

[48] See Searle (1969) for the pioneer analysis of name-giving acts. As van Langendonck (2007: 92) claims: "Such name-giving acts take place where proprial or non-proprial lemmas are assigned *ad hoc* to unique entities. In the case of proprial lemmas, these speech acts can generate prototypical names such as first and family names like *John* and *Mary,* or place names when applied to new places, like the many towns in the USA called *Oxford* or *London*, and so on. In the case of non-proprial lemmas, name-giving acts are assigned *ad hoc* to unique entities, such as film *Gladiator*" (after van Langendonck 2007: 92).

> 1st stage: *Przejęcie* 'capturing' of a word or phrase from a foreign language. In the beginning it functions like a quotation as it preserves foreign phonetic and graphemic make-up.
>
> 2nd stage: Gradual phonetic adaptation of a quotation to the target language phonological system. The pronunciation is close to the original but with distinctive traces of target language influence.
>
> 3rd stage: Morphological adaptation of a foreign word or phrase. The words with untypical suffixes will remain uninflected (refers primarily to synthetic languages).
>
> 4th stage: Graphemic adaptation of a foreign word or phrase to the target language.

Most of foreign proper names remain as quotations. This is because of the three reasons:

> a) Introducing alternative target language spelling would unnecessarily multiply already quite a numerous group of proper names.
> b) Most of the proper names are ephemeral while process of accommodation usually is long lasting.
> c) Identificational function of proper names prevents the emergence of variant spelling forms due to the nature of the function itself which is the more effective the lesser spelling variation exists for a given proper name.
>
> (after Grzenia 1998: 28)

The above-quoted process of adaptation of foreign proper names is interesting because it points out, as already indicated above, to its processual character. Some doubts can be raised when it comes to what appears pivotal for the description, i.e. stage 2. Grzenia argues that at this stage foreign proper names have adapted, to some extent, their pronunciation to the phonological system of the target language. Sadly, the scholar does not further explain what near-original pronunciation with "distinctive traces of target influence" actually means. It seems that foreign proper names may alternatively be viewed to display at this stage a twofold pronunciation (the one that constitutes more or less successful imitation of the original pronunciation), and the one that constitutes an attempt at adaptation of the original pronunciation to the target language phonological system (see section 5.0).

It, therefore, appears that if there are no phonotactic constrains upon the pronunciation of a given foreign word or phrase, these should be pronounced accordingly with the phonological rules of the target language (cf. Grzenia 1998: 35 – the example of *Bergman).* The conclusion that Grzenia arrives at is, however, less restrictive in terms of the ultimate criteria determining pronunciation choices made by speakers of the receptor language. As he says, the pronunciation of foreign words and phrases is largely dependent upon the linguistic convention. If there is no clearly established pattern, we should fall back on the original pronunciation. In the Polish linguistic context, all foreign word sounds which have no Polish equivalents may be replaced by similar sounding phonemes. Hence we often receive two phonetic variants of the pronunciation of a foreign word or phrase (Grzenia 1998: 37).

5.4.2. Comparing proper names and common words

According to a traditional view, common words (*nomina appelativa*) differ from proper names (*nomina propria*) referentially.[49] That is, proper names have no meaning. They only refer to entities in the external world, whereas common words relate us to concepts that mediate the formal representation with the representation in the external world (see Grzenia 1998: 16). Proper names also differ from common words in that proper names are primarily characterized by identificational function (apart from ideational). Without the possibility of individual variation, no society could function properly. For example, Henryk Sienkiewicz provides a better characterization than 'The Polish author of the novel *Krzyżacy* and *Potop,*' which can be best evidenced in any encyclopedia in which Henryk Sienkiewicz is one, but many Polish authors of the novels (Grzenia 1998: 17).

Another criterion of differentiation is formal and relates to the capitalization of the first letter of a lexeme in the case of proper names (Grzenia 1998: 18). This visual criterion is an interesting one because it allows the scholar in the process of data elicitation to make rather definite groupings into two identifiable types of lexical sets, given, of course, the lack of clear semantic and functional discriminative criteria. This visual criterion as based on an authoritative source such as LPD has aided in dividing the collected material into "proper" and "common" sub-categories in the present book (see section 0.2).

Otherwise, we notice that finding the differences between the two categories poses much of intellectual challenge. This challenge has to some degree been successfully taken by van Langendonck (2007) who attempts to enumerate different circumstances that possibly differentiate the functional, semantic, syntactic[50] behavior of the two groups of words, however, the perennial, insurmountable obstacle that appears to come in way during such investigation is the prevailing fuzziness of the area upon which both proper and common words appear to inevitably encroach. Thus, van Langendonck (2007: 169) says: "linguists and onomasticians often have difficulties in telling proper names and common words apart (see Harvalik 2005)." Van Langendonck (2007: 169–171)

[49] "Jonasson (1994: 19) proposes to dissociate the linguistic category of the proper name from its referential function and to define it in cognitive terms as a class of linguistic expressions associated in long-term memory with an individual, i.e. with some specific and not some general knowledge (as in the case of common words) (after van Langendonck 2007: 57). See also Hansack (2004: 56, in van Langendonck 2007: 58): "The essential difference with common words, then, is that a common noun is a form which indicates a more-than-one-element class meaning in the brain whilst a proper name indicates a one-element class meaning in the brain. Proper names and common words have in common that they refer to a class with meaning."

[50] As van Langendonck (2007: 62) says, "the proper name provides the denotation, while the common noun characterizes the denotatum, for instance, Burns is a poet (and not the poet is a Burns)."

presents a list of circumstances that may serve as differentiating proper names from common words. These are:

> – Like common words, but unlike personal pronouns, proper names constitute an open class of words, and, hence, are lexical rather than functional (grammatical).
> – Like common words, but unlike personal pronouns, proper names can take adjectival and other modifiers, at least in Germanic languages.
> – Unlike common words and proper names, personal pronouns can be defined exhaustively by means of grammatical features such as person, number, gender. Personal pronouns lack the categorical meaning attributed to proper names and (most) common words.
> – Unlike common words, proper names and personal pronouns display only presuppositional meanings, of whatever nature these are.
> – Proper names and personal pronouns appear in identificational sentences, not in predicational ones. Common words display both possibilities.
> – Unlike common words, proper names and personal pronouns are inherently definite.

The overall conclusion bespeaks, however, the inherently fuzzy conceptual boundaries between the two categories in question. Van Langendonck (2007: 171) admits that himself: "We can hold that proper names can be situated between common words and personal pronouns although they share more characteristics with common words than with personal pronouns."

As the discussion presented in the current and preceding sections on proper names shows, it is impossible to find satisfactory criteria by means of which it would be possible to clearly discriminate between proper and common types of lexical units. This observation is in accordance with Crystal (2007b: 74), who notes with some tinge of scholarly optimism (realism), that "there is no sharp dividing line between common words and proper names. They feed off each other." Later he says:

> The conclusion is clear. When we study words, we have to study names too, for everything influences everything. Names become words. Words become names. We look for meaning everywhere we go. And if there is no immediately obvious everyday meaning in a name, there is someone waiting in the wings to invent one.
>
> (Crystal 2007: 79)

What is particularly significant to the ultimately adopted typology of proper names accepted in the present book (see section 7.0) is, however, the following remark made by Crystal (2007b: 79):

> Person names and place names are the two major domains in the field of name-study. […] But these two categories by no means exhaust the human penchant for naming. In a 1990 edition of the Radio 4 series *English Now,* over a thousand listeners sent in information to me about the things they named at home. The list included cars, computers, washing machines […].

Crystal's observation, in consequence, paves a path towards the further subdivision of foreign proper names into the three categories: *personal*, *place*, and the third encompassing other foreign proper names, i.e. *non-personal* (see the aforementioned section 7.0).[51]

[51] The inherent fuzziness of the boundary between proper names and common nouns has led the author of the present book to make the following arbitrary classification decisions, which are allowed for in the analysis presented in PART III. Thus, definite descriptions, e.g. *Markov process* have been classified in the analysis as common rather than proper phrases, given the argumentation provided by Berezowski (2001: 52). Another controversy relates to place-related lexemes, e.g. *Provencal* and *Quebecois*, which have been classified in the analysis as proper names vs. *nicoise*, *madrilène* which have been classified in the analysis as common words. The criterion adopted here has been formal and based on orthographic convention followed by LPD.

PART III
FORCE, FOREIGNNESS AND OTHER THINGS

Overview

Part III of the book is composed of two chapters (Chapters 6 and 7) and along with PART IV constitutes the analytical contribution to the book. Chapter 6 offers an insight into the methodological framework behind the study, which involves uncovering a series of far-going correspondences between the phenomena occurring in the physical world and the phenomena occurring in the linguistic world. We will first look at some significant statements relevant to the science of astrophysics and their postulated relevance to our linguistic research. We will show that the assumption about the primary basis of the physical world for the discussion of other nonphysical phenomena is far from iconoclastic. It is rather an attempt at the fulfillment of one of fundamental assumptions of Cognitive Linguistics where human bodily experience along with the implicated interaction of a human being with the direct physical environment is regarded as source of linguistic conceptualization. For this reason we will look at some significant contributions from cognitive linguists to the study of the relationship between physical phenomena and language structure. More specifically, we will examine two of such fundamental contributions, i.e. the concept of FORCE image-schemas (Johnson 1987) and the concept of force dynamics (Talmy 2002). The insights offered by the aforementioned publications will serve as point of departure for further investigation of the place and nature of force in language (section 6.8). Additionally, a set of other conceptual correspondences relevant to a linguistic (strictly lexicological) research will be presented. These will relate to the discussion of the concept of Cumulative Relative Average Count – the technical notion introduced into the analysis, in which the relationship between the corpus statistics of lexical frequency and our physical experience of weight and height will be correlated (section 6.1). We will also look at other notions such as force, mass, volume, energy, motion, gravity, to name just a few; all seen as highly contributive to the discussion upon the description of the influence of the vocabulary of other languages on English.

Chapter 7 is primarily devoted to the detailed examination of the semantic architecture of languages affecting the lexical system of the English language at the turn of the 21st century. By looking into semantic architecture is meant investigation of semantic domains that can be identified in the structure of these donor languages in the context of their contribution to shaping the structure of foreign lexical 'mass' of the English language. The uncovered correspondences from Chapter 6 will serve here as source of metaphorical language used where the English language is consistently referred to as 'the Earth' and other languages as 'planets' or 'planetoids' depending on the classificatory criteria discussed in sections 6.7.1 and 6.7.2.

Prior to the study proper of the foreign lexis of each individual donor language (also referred to as "planets/planetoids" [section 7.1]), however, criteria of classifi-

cation of semantic domains are discussed (section 7.0) as these already highlight some of the problematic issues in the oncoming analysis. All in all, it is believed that through the examination of a set of multiple physical-linguistic analogies we will be able to first of all account for a more systematic analysis of the lexical input offered by donor languages to the English language and then, which is actually the objective of Part IV, endeavour to sketch the model within which the explication of the nature of foreign lexical assimilation processes can be made.

Chapter 6
PHYSICS AND LANGUAGE

6.0. Astrophysics and its relevance to research

As already signalled in Overview to Part III, the objective of Chapter 6 is the examination of some crucial concepts discussed in astrophysics as they correspond to processes observed in lexical assimilation. That such correspondences may be justified has already been stated in Preface as well as Introduction (section 0.2) to this book. For the sake of clarity, let us recapitulate central claims underlying this research. First and foremost, the argumentation presented in the research is hinged on the metadiscursive metaphor LANGUAGE LAWS ARE PHYSICAL LAWS which, metonymically, informs us that lexical assimilation processes are to a large extent correspondent with some physical processes shaping our external reality. These processes include centripetal and centrifugal forces present in circular motion, the ubiquitously palpable force of gravity, the concept of mass entity and volume, to name the most significant ones. The aforementioned physical notions are discussed relative to astrophysical knowledge about the structure of solar system (e.g. Kepler's law of interplanetary gravitation) or what we know, both on an expert and non-expert level, about meteors, meteorites, or the moon. Such conceived PHYSICAL LAWS serve then as the source domain for explicating foreign lexical assimilation processes in language (here English). Let us quote a relevant passage from Preface to this book:

> **The framework [of the analysis, M.K.] as such, is argued to be found in astrophysics. It does not of course mean that astrophysics 'engulfs' linguistics as a discipline; making such a postulate would be, to put it mildly, intellectually risky. Instead, the relationship between astrophysical concepts and lexical assimilation processes is argued to be grasped in terms of the overriding explicatory metaphor: LANGUAGE LAWS ARE PHYSICAL LAWS. The metaphor is metadiscursive and metacognitive in its essence. It serves as a convenient conceptual device by means of which modeling foreign lexical assimilation processes can be handled no matter if you are an expert in astrophysics or not. Both naïve and expert views of the physical world, which were epitomized in the respective Ptolemy's and Copernicus' conceptions of the universe, equally count, as the book argues, for the model of lexical assimilation. But the discussed metaphor is not to be understood in a radical way. It is metonymic in character, i.e. LANGUAGE LAWS actually relate to lexical assimilation processes, whereas PHYSICAL LAWS are reduced to the phenomena, which are sensorily accessible to a human conceptualizer. Therefore, for example, Newton's laws of dynamics, along with the concepts of motion and energy duly apply in the analysis as they are grounded in the phenomena experientially akin to a human being. The argumentation advanced in the book is thus ultimately transcognitive in nature. This transcognitive character of the study is particularly visible in the treatment of the overriding metaphor which is not posited to reside in the mind of the individual language user because people normally do not talk about such phe-**

nomena in casual conversations. LANGUAGE LAWS ARE PHYSICAL LAWS is thus not to be interpreted in the Lakoffian sense of a conventional metaphor that "we live by."

In section 0.2, we offer a more detailed characterization of this general metaphor, this time restricting the discussion to the outline of the most fundamental correspondences between lexical assimilation processes and astrophysical phenomena:

Successful (pre-)adaptation of a foreign lexical unit (a meteor[ite]) into the target language system (in this book English – the planet Earth) occurs if the sum of centripetal force(s) that governs the incorporation of a foreign word or phrase *into* the target lexical system prevails over the sum of centrifugal force(s) that conserves the system. Breaking through the protective barrier formed by centrifugal-like forces thus ensures completion of assimilation process by a foreign word or phrase and its ensuing entrenchment in the target system. Consistently, then, failure to overcome the barrier formed by centrifugal force(s) is analogical to the expulsion of a foreign word or phrase *from* the target system, i.e. its lack of adaptation in the English language.

These observations will constitute the foundation for the description of donor languages in Chapter 7 and the ensuing argumentation presented in Chapters 8 and 9 of this book. As the history of science shows, attempts at finding laws that would regulate phenomena apparently far apart one from the other have found their most prominent reflection in physics. As Kowalski-Glikman (2008: 41) says, all the greatest physical laws turned out to be the laws of unification. One of them was the theory created in the 17th century by Newton and known under the three laws of dynamics. Newton discovered that laws which cause an apple to fall onto the ground are the same as laws that regulate sea tides, or the motion of planets in the solar system.

Newton's revolutionary observation upon the discovery of laws governing phenomena which apparently have nothing in common leads us to make the following claim: if the theory of unification was possible within the domain of physics, why not search metaphorically conceived laws of unification across disciplines such as physics and linguistics? Why not, then, try to explore the processes occurring in the physical world and see them as subject to the correspondent forces as other phenomena observed in the linguistic world? One of the central postulates in the book is that a positive answer to these somewhat brave assumptions can be found. First and foremost, the present study is held within cognitive-linguistic spirit, which entails interdisciplinary nature of linguistic studies, therefore, there is no methodological counterpremise not to undertake this sort of search. A good ground for the kind of analysis where 'unifying laws' governing physical and linguistic world can be identified turns out to be lexis, particularly when it is examined in the context of external influences that a given (target) vocabulary system receives from other lexical systems (i.e. donor languages).

In order to elucidate the nature of these correspondences, let us first examine the notions of "height" and "weight" as these appear to be fundamental to our bodily interaction with physical environment and see how these dimensions cor-

relate with phenomena that may be applied in the current research. Then, the most fundamental notions concerning Newton's laws of dynamics will be discussed and, again, an attempt will be to consistently relate them to a set of phenomena existent in language.

6.1. Weight, height and CRAC

The category of "height" and "weight" are the two most fundamental dimensions according to which the body of a human being is measured. When we are born, our height/weight ratio is noted in official records. The information about a further increase in height/weight values of an infant is of vital importance in diagnosing potential health problems that a newborn baby may face. Later as we grow up, we learn that getting taller with a proportionate increase in weight value is a desirable course of events, not only from purely physiological point of view but also social standpoint. We, thus, learn that having adequate proportions of height to weight is perceived as more attractive than in the case of individuals whose height/weight proportions are socially viewed as distorted. Of course, there is not any universal key according to which we may judge the aesthetics of a human body, based only on the criteria of 'proper' height/weight ratio as this 'appropriateness' appears to be culture-specific. Nevertheless, it seems that the categories of "height" and "weight," whatever the cultural-aesthetic norm is established in measuring 'ideal' proportions of a human body, are intricately related one with the other. In further examination of this relationship, we may thus arrive at some correspondences with the linguistic world.

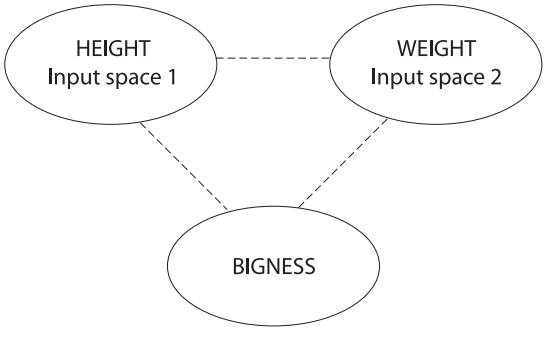


Fig. 19. *The concept of BIGNESS as a blended category*

One such relationship that holds between the category of "height" and "weight" relates to the concept of BIGNESS. We tend to judge a human being as big on the basis of criteria of "height" and "weight." These should not be considered separately but rather viewed as properties that invariably combine in making up the "gestalt" perception of BIGNESS of a human body. The results of the cross-cultural research I conducted appear to be supportive of the tentative claim

where BIGNESS is viewed as a blended category with the category of "height" and "weight" regarded as input categories.[1]

In the research, a sample group of 94 English-speaking students from the University of Wrocław, Poland and the University of Aarhus in Denmark were instructed to indicate which individual in the considered pair is bigger. The format of the survey was the following:

Tab. 8. *Survey format*

Sex	Height	Weight	Tick an appropriate member of the pair	Average[2]
Male	200	100		150
Male	190	110		150
Male	190	85		137.5
Male	180	105		142.5
Male	165	70		117.5
Male	180	75		127.5
Male	160	70		115
Male	190	80		135
Male	190	100		145
Female	205	85		145
Male	185	85		135
Female	195	75		135
Male	195	105		150
Female	165	65		115
Male	170	65		117.5
Female	205	105		155

The hypothesis goes like this: the higher the average value discrepancy between the individuals discussed in a pair, the higher the likelihood that a respondent will select the one with higher value regardless of the sex of a human being. Consistently, in the case of pairs where the average value estimated is equal, a great deal of variation in answers has been expected. The research is intended to verify the hypothesis whether the category of BIGNESS is cognitively relevant

[1] For the discussion of blending in linguistic categories see Fauconnier and Turner (1996).

[2] This column was not originally inserted into the survey in order not to suggest answers in any way.

as a cover term for the two aforementioned categories of "weight" and "height" or not. The results of the research are presented in Tab. 9.

Tab. 9. *The results of the research*

Sex	Height	Weight	Average	Preference in percentage
Male	200	100	150	0.47
Male	190	110	150	0.53
Male	190	85	137.5	0.11
Male	180	105	142.5	0.89
Male	165	70	117.5	0.22
Male	180	75	127.5	0.78
Male	160	70	115	0.20
Male	190	80	135	0.80
Male	190	100	145	0.59
Female	205	85	145	0.41
Male	195	105	150	0.96
Female	165	65	115	0.04
Male	170	65	117.5	0.05
Female	205	105	155	0.95

As one can easily notice the tendencies manifested appear to corroborate the aforementioned hypothesis in a positive way. The more of the discrepancy there is between the two individuals in terms of the average value, the higher the agreement among the respondents as to which individual is bigger. An interesting phenomenon can be observed with respect to the value 100. It follows from the research that once the average in a considered pair is the same, say 145, respondents tend to select, albeit not in so much clearly determinate way, an individual whose weight is equal or exceeds 100. It appears, then, that the integer 100 may be a part of symbolic ICM of BIGNESS, at least in Denmark and Poland where the research was conducted. However, because the research was made in an English speaking multicultural environment, the claim about some universalist perception of BIGNESS may tentatively be postulated with regard to a large part of the speech community in the whole world. Certainly, much more comprehensive research would need to be done to more objectively verify this observation.

All in all, the goal of the above-presented survey is to test the hypothesis whether human conceptualizers construe, or not, the categories whose constituent sub-

domains (in this research, weight and height) are not reducible to a common denominator, at least in a logical-mathematical sense. Since the positive answer to the hypothesis has been found, a tentative proposal is to correlate the results of the research onto the field of lexicological studies, whereby the value of CRAC may serve as a cognitively substantiated category for the discussion of salience of a given lemma in the corpus. CRAC is an abbreviation from (Cumulative Relative Average Count) and is the value calculated on the basis of eliciting the average out of two parameters:

- occurrence of a lemma in the corpus – (*x*),
- number of texts in which the lemma is encountered – (*y*).

For example, the French word *femme* occurs 81 times in the whole of the BNC, but the number of texts in which it is recorded amounts to 61. Given these two figures, CRAC value for this particular lemma is 71. The notion of CRAC can also be written down in terms of the well-established mathematical formula[3]:

$$AM = \frac{1}{n} \sum_{i=1}^{n} a_i .$$

In this book, however, simpler notation is used in that three parameters are taken into account, i.e. the aforementioned values of (*x*), (*y*) and the resulting CRAC, i.e. the arithmetical average that emerges from the calculation of (*x*) and (*y*). CRAC is, thus, "cumulative" because its value is always viewed as average of the sum of two parameters; it is "relative" because its value is always calculated relative to some other dimension; it is "average" because the mathematical calculation of "averaging" appears to be the closest conceptually to the cognitive process of blending as presented in Fig. 19. Finally, it is referred to as "count" because its value is the result of a mathematical calculation of "averaging."

The relative nature of CRAC is manifested in the fact that CRAC value maybe instantiated as CRACn1, CRACn2, or CRACn3 depending on the reference domain according to which it is analyzed. Thus, CRACn1 is the value obtained relative to an individual lemma recorded in a given set (see the discussion of *femme* above). CRACn2 corresponds, in turn, to the value obtained relative to a set of lemmas identified for a particular donor language subcategory, whether "proper names" or "common words." Eliciting CRACn2 value may be illustrated schematically as follows:

Tab. 10. *Eliciting CRACn2 value*

Lemma (e.g. proper names)	Language	CRACn1
lemma1	e.g. Polish	124
lemma2	e.g. Polish	75
lemma3	e.g. Polish	48
lemma4	e.g. Polish	1200
CRACn2 = 361.75		

[3] See www.en.wikipedia.org/wiki/Average (ED: 12/08).

Finally, CRACn3 value is the value obtained for a set of lemmas relative to the entire subgroup of "proper names" or "common words" constituting the foreign lexical input regardless of individual assignment of a lemma to a particular language. CRACn3 is then the value calculated as the average of all CRACn1-s identified for a particular group whether "proper names" or "common words." Eliciting CRACn3 value may be illustrated schematically as follows:

Tab. 11. *Eliciting CRACn3 value*

Lemma (e.g. proper names)	Language	CRACn1
lemma1	e.g. French	1240
lemma2	e.g. German	750
lemma3	e.g. Polish	125
lemma4	e.g. Welsh	64
CRACn3 = 544.75		

The facts presented above lead us to argue that CRAC value is a cognitively salient category although *prima facie* it emerges as a result of putting together two incompatible variables at least in the strict mathematical sense. The postulate that I am making is that the two parameters, i.e. number of occurrences and number of texts can be brought together under one common category, that is CRAC, in the same way as categories of "height" and "weight" can be discussed in terms of the cumulative category of BIGNESS.

Thus the lemma with CRACn1 value, say 150 ($x = 200$ and $y = 100$) is postulated to have a greater salience than the lemma whose CRACn1 value is 125 where ($x = 130$ and $y = 120$). Of course since CRACn1 values in the discussed lemmas are not that discrepant, the hypothesis will consistently be that language users may vary as to their judgment of which of the two lemmas is more recognizable (familiar) to them. Therefore, consistently with the result of the research presented above, the assumption taken in the book is that the greater the difference between the CRACn1 values pertinent to particular lemmas, the more likelihood that a language user will rate the one with considerably greater CRAC value as more familiar than the one with relatively lower CRAC value. This has some vital implications for the forthcoming study (see particularly section 8.6).

The aforementioned significance of the category CRAC is attributed to the fact that CRAC serves in the book as a relevant criterion to our considerations over the category of foreign words and phrases. The postulated importance of CRAC is particularly discernible with regard to a broader phenomenon of lexical assimilation, where CRACn1 value is correlated with a potential assimilatory force of a given lexical unit (see section 8.1). The concept of CRAC in its three instantiations n1, n2, and n3 will also be highly significant in modeling the processes of foreign lexical assimilation as shown in Chapter 9.

6.2. Force

Having discussed the significance of CRAC for the current analysis of foreign words and phrases, we may take a closer look at other significant notions relevant to the research. One such concept is "force." "Force" is defined by Hewitt (1998: 61) as follows: "In the simplest sense, it is a push or a pull. Its source may be gravitational, electrical, magnetic or simply muscular effort." For the time being, the two manifestations of "force" will be of interest to us, i.e. centripetal force[4] and centrifugal force. According to Hewitt (1998: 128):

> Any force that causes an object to follow a circular path is called a *centripetal force*. *Centripetal* means "centre-seeking" or "toward the centre" […] Gravitational and electrical forces can be transmitted across empty space to produce centripetal forces. The moon, for example, is held in an almost circular orbit by gravitational force directed toward the center of the earth. The orbiting electrons in atoms experience an electrical force toward the central nuclei. Centripetal force is not a new kind of force but is simply the name given to any force, whether string tension, gravitational, electrical, or whatever that is directed towards a fixed centre.

Centrifugal force, on the other hand, is the opposite kind of force. Again as Hewitt (1998: 129) claims: "Sometimes in circular motion we seem to experience an outward force. This apparent outward force is called *centrifugal force*. *Centrifugal* means 'centre-fleeing' or 'away from the centre.'" The idea of centrifugal force is rather an illusion in the strict meaning of the word; however, it has a viable psychological reality status because it is perceived as existent by humans.[5] It is due to the psychological realty status of that force that we treat centrifugal force as equally valid for the considerations over lexical assimilation processes as centripetal force, which is, in turn, fundamentally significant for the conceptual model of lexical assimilation discussed in Chapter 9 (sections 9.1 and 9.2).

Continuing the discussion upon the nature of physical force, one may show how centripetal and centrifugal forces, being prototypical instantiations of force, function as parameters defining other identified sources of force. This somehow bespeaks 'universalist' nature of these two fundamental types of forces.[6]

---

[4] The term centripetal force was first used by Issac Newton in *Philosophiae Naturalis Principia Mathematica*, a three-volume work published in 1687, whereas the conception of centrifugal force appears to have its origins in Christiaan Huygens' paper *De Vi Centrifuga*, written in 1659 (see www.en.wikipedia.org/wiki/Centrifugal_pseudo-force#Rotating_frame [ED: 06/09]).

[5] As Hewitt (1998: 130) argues:"It appears thus that depending on our frame of reference (e.g. being inside the rotating system) we feel the centrifugal force as real as the pull of gravity. […] However its status is only interpreted as force, it is not a real force as gravitational, electromagnetic or nuclear one."

[6] This is in line with how Newton developed his conception of "centre-seeking" force as gravity force regulating phenomena which have not been regarded as necessarily existent in straight correlation (e.g. sea tides, planetary motion, etc.). It is also to be reminded here that the notion of "prototypicality" is understood in cognitive-semantic terms, i.e. it relates to the primary (most

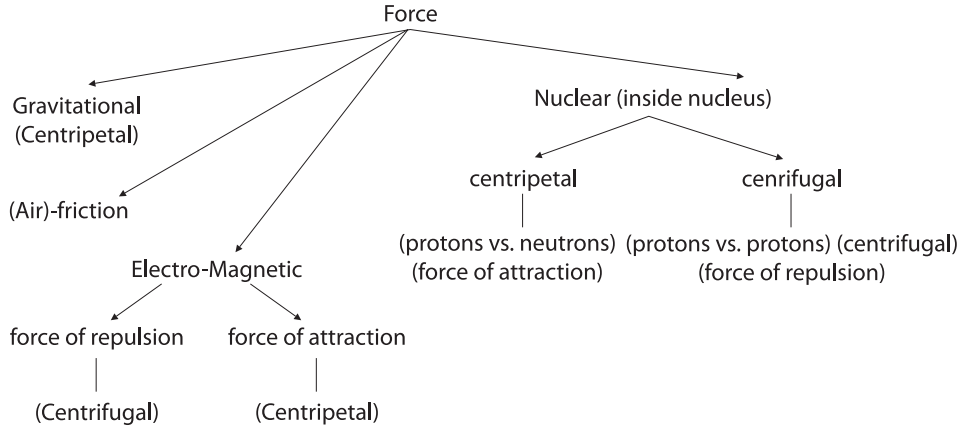


Fig. 20. *Types of force as different manifestations of centripetal and centrifugal force*

The linguistic correlation of the discussed centripetal and centrifugal types of force is that they will henceforth be regarded as analogous to the processes of lexical assimilation of a foreign word or phrase. Thus, the working of centripetal force may be viewed as analogous to the incorporation of a given lexical unit within the 'atmosphere' of a target lexical system, whereas centrifugal force may be regarded as analogous to the forces of resistance, whether of social or grammatical (morpho-phonological constraints) provenance, regulating the eventual entrenchment or disappearance of a lexical unit in that system (see especially Figs. 52 and 53, section 9.1). More about the significance of force in linguistic research in general as well as the current study on foreign words and phrases can be found in section 6.8.

6.3. Motion

Inextricably connected with the concept of "force" is the concept of "motion." This is valid to the present discussion for two reasons. One is that lexical assimilation process is viewed as *dynamic*, which is in accordance with how linguists view the processes of language change. In order to evolve, the language needs to be 'in constant motion.' This motion is determined, in the case of foreign lexis, by the creative forces of lexical incorporation on the one hand, and by the conservative forces of resistance to that change, on the other. Legitimate is, thus, further comparison of languages to planets in constant circular motion, which are clasped in the "shackles" of interplanetary gravitational forces in the solar system in the

basic) experience that a human conceptualizer makes sense of in his/her dealing with the phenomenon of force.

same way as languages are intertwined in mutual interactions with one another within the intricate network. This is another illustration of the validity of centripetal force of gravity as fundamental to our research, as it is coherent with the idea of circular motion which is, in turn, indicative of the motion of languages in the aforementioned processes of change.

Our considerations over the nature of motion are also partially compatible with Aristotle's conception of natural motion and his famous observation about a feather and a lump of clay: "A feather properly fell to the ground but not as rapidly as a lump of clay. Heavier objects were expected to strive harder. Hence, objects were thought to fall at speeds proportional to their weights: the heavier the object, the faster it was thought to fall" (Hewitt 1998: 18). An important reservation is made in the word "partially" because, although it is true that heavier objects fell to the ground faster than lighter ones, they are subject to varying gravitational forces ($F = mg$ [sometimes referred to as IV Newton's law of dynamics]) in that they do not fall proportionally to their weights. The scientific explication of this mechanism is again provided by Hewitt (1998: 66):

> A heavy object is attracted to the earth with more force than a light object. The double brick, for example, is attracted with twice as much gravitational force as the single brick. Why then, as Aristotle supposed, doesn't the double brick fall twice as fast? The answer is that the acceleration of an object depends not only on the force – in this case, the weight – but on the mass as well. Whereas force tends to accelerate things, mass tends to resist the acceleration. So twice the force exerted on twice the inertia produces the same acceleration as half the force exerted on half the inertia.

The conclusion that emerges from the current discussion is that in modeling lexical assimilation processes we may safely take into consideration the two notions of "mass" and "weight" and the concept "acceleration" as compatible with our everyday conception of greater objects accelerating with greater magnitude than lighter objects. This observation is also in accordance with the scientific view, i.e. Newton's 2nd Law of Dynamics (see section 8.4.1). The significance of the above remarks to our research on foreign words and phrases becomes evident already earlier (section 6.4).

Aristotle's concept of motion was valid for hundreds of years and contributed to the geocentric model of the world (see section 9.2) in which the Earth is in the centre of the universe and obviously there is not any force acting upon it as it has its proper place in the universe. The conclusion was that because there is no need for any force to act on the Earth, it does not move (Hewitt 1998: 19). Aristotle's conception of motion was based on our everyday observation of how things move and the belief that vacuum is impossible. For Aristotle it was rather self-evident that the motion of an object is conditioned by the existence of some force acting on that entity, whether of 'push' or 'pull' kind. This idea was later argued against by Galileo who claimed that in the absence of forces acting on a given object, it will keep moving in a straight line. This phenomenon was subsumed under the term of *inertia* (Hewitt 1998: 21–22).

The concept of Galileo's inertia and the ensuing new perspective on motion may serve as explication in the realm of language contact, namely, how languages remain in a permanent state of change. Regrettably, however, it does not account for the concept of language death,[7] the issue that we also briefly endeavour to tackle in section 8.2.4.

6.4. Newton's laws

Of extreme relevance to the research are Newton's Laws of Motions. They combine the notions of force and motion already discussed above and constitute one of the greatest theories of unification in the history of physics. The following is a concise presentation of the aforementioned laws along with their applications in physics, after which their correlation to the world of language is illustrated. In this way, the theory of unification is further extended onto the social ground (cf. de Saussure's concept of language as a social phenomenon [Fisiak 1985a: 27]).

Newton's First Law of Motion[8]

> LAW 1: Every object continues in its state of rest, or of uniform motion in a straight line, unless it is compelled to change that state by forces impressed on it.

Linguistic correlation: unless subject to forces of assimilation, foreign lexical items remain in a state of rest (i.e. they are irrelevant) in respect of the target lexical system. In compliance then with the facts about interplanetary gravitation, foreign lexical items, which have not been incorporated into the "sphere" of the English lexical system, revolve round the globe in on their circular orbits without much affecting the target lexical system. They can be said to be literally unknown to the vast majority of English speakers. These regularities can be presented as follows[9]:

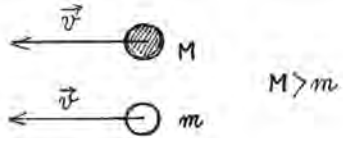


Fig. 21. *1st Law of Motion and two alien bodies "M" and "m"*

In Fig. 21 we have two bodies *m* and *M*. Body *M* has greater mass than *m*. They move through the space with constant velocities due to the absence of force or its

[7] For a comprehensive multi-perspective discussion of the issue of language death, see Crystal (2007a). Among many complex factors leading to the death of a language Crystal (2007a: 70–90) mentions the two main categories: a) physical endangerment of people speaking that language and b) diversified culture changing mechanisms.

[8] Newton's Laws of Motions have been quoted here after Hewitt (1998: 56–73).

[9] My special gratitude is owed in this place to the two physicists Ewa i Jacek Woźny for their hand-made drawings illustrating the working of Newton's Laws of Motion (see Figs. 21, 23, 25).

negligible character. Their velocity will not change until some unbalanced force appears. This unbalanced force may appear when the two bodies approach the Earth. In the same way in language, the two foreign words or phrases are equally significant to the target system unless they fall within the target's language sphere of attraction. In this case the word with higher CRACn1 will be attracted with more force to that system than the word with lower CRACn1.

Newton's Second Law of Motion

> LAW 2: The acceleration of an object is directly proportional to the net force acting on the object, is in the direction of the net force, and is inversely proportional to the mass of the object.

Linguistic correlation: the analogy in the linguistic world relates to the concept of CRAC (i.e. mass) and the force of incorporation acting on the unit. This action can be formulated in the following way: The greater the mass, the greater the force of incorporation is. On that account, the successful assimilation of a foreign lexical unit will be viewed as a resultant force of prevailing forces of incorporation, on the one hand, and of negligible forces of resistance, on the other (see also section 9.1.2).

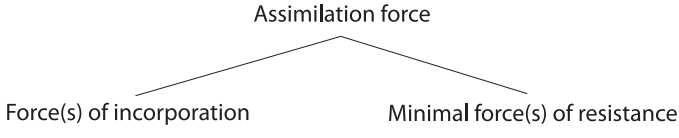


Fig. 22. *Successful assimilation as a resultant force*

These observations can be more accurately presented by way of a drawing based on the findings from the 2nd Law of Motion:

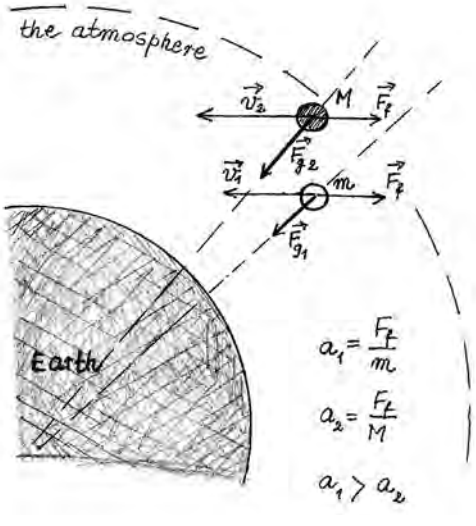


Fig. 23. *2nd Law of Motion and two alien bodies entering the atmosphere of the Earth*

Thus *M* corresponds to a foreign word with a higher CRACn1 value, whereas *m* correlates to a foreign word with a lower CRACn1 value, whereas a_1 and m_2 stand for the negative acceleration. *M* moves faster, i.e. is incorporated much quicker into the target system (the Earth). Chances of assimilation are thus greater for *M* than for *m* due to *M*'s greater mass coupled with greater velocity. In consequence, a word or phrase with a higher relative CRACn1 may in effect, mark a more durable place in the target lexical system, i.e. become entrenched as a stable element of it.

Newton's Third Law of Motion

> LAW 3: Whenever one object exerts a force on a second object, the second object exerts an equal and opposite force on the first. For every action there is an equal and opposite reaction.

Linguistic correlation: the force of lexical incorporation into the target system meets the force of lexical resistance which comes from that system. It appears then, that the 3rd Law of Motion is not only about forces acting on entities in the physical world but is also a fundamental statement about the process of lexical assimilation whereby a foreign lexical unit exerts a force of incorporation (action) onto the target lexical system, and the target lexical system exerts an equal opposite force of resistance (reaction) on that unit.[10] Certainly, our naïve identification of force as pertinent to "action" or "reaction" depends on cognitive salience of the two entities in contact (Fig. 24). A schematic illustration of the 3rd Law of Motion relevant to the analysis of foreign words and phrases is displayed below. More detailed description is given in Fig. 24:

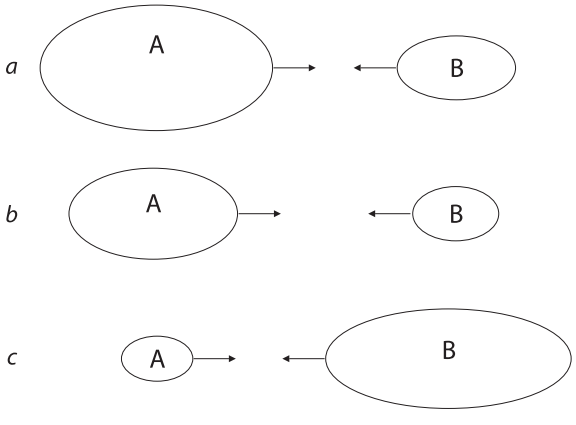


Fig. 24. *3rd Law of Motion: Which falls toward the other, A or B?* (Hewitt 1998: 73)

Hewitt (1998: 73) addresses the above-formulated query as follows: according to the 3rd Law of Motion the forces between the bodies indicated as A and B

[10] On the nature of resistant forces, see section 6.8.3.

are equal. One thing is this equality in magnitude of force; the other thing is the noticeability of acceleration of one body toward the other. Thus, in *a* the acceleration of planet A is hardly discernible with respect to planet B due to its relatively greater mass, whereas in *c* the acceleration of planet A is much more noticeable due to is relatively smaller mass with regard to planet B.

It becomes, thus, evident that 3rd Law of Motion is in line with the central observation made in section 0.2 in that it corroborates the validity of the claim in favour of the existence of far-going correlations between the physical world and the linguistic world. These correlations are confirmed here by the analogy where a lexical item corresponds to a celestial body, strictly a meteor. In this light, the process of lexical incorporation will be best represented by the illustration *c* in Fig. 24, in which A represents a particular foreign lexical item, and B is illustrative of a recipient language (in this book the English language). Thus A is perceived as moving towards B. Naturally, then B (English) may either accept or reject A (a foreign word) within the sphere of its lexical system. If we proceed with investigation of other epistemic correspondences entailed in this analogy, we may bravely postulate the (air)-friction that the meteor (foreign word) meets on its way towards planet B as a manifestation of centrifugal force of reaction (see Fig. 19, section 6.2). These regularities can also be illustrated by a more detailed drawing:

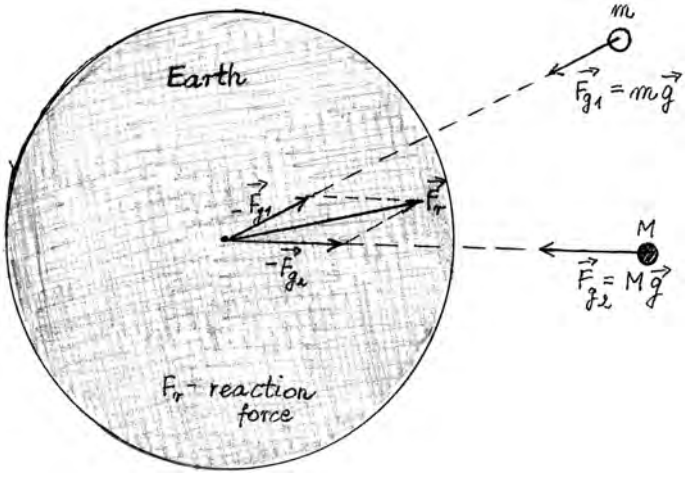


Fig. 25. *3rd Law of Motion. Two alien bodies "M" and "m" entering the atmosphere of the Earth*

Strictly correlated with the above-presented 3 Law of Motion is the notion of mass and weight. According to Hewitt (1998: 57), mass relates to "the quantity of matter in an object. It is also the measure of the inertia or sluggishness that an object exhibits in response to any effort made to start it, stop it, or change its state of motion in any way," whereas weight is viewed as corresponding to "the force

upon an object due to gravity."[11] Again, the above-quoted dichotomous pair of "mass" and "weight" can be easily extrapolated on to the linguistic ground, given the concept of CRAC.

Accordingly with the analysis presented in section 6.1, CRACn1 relates in a straightforward way to the concept of mass which designates a potential assimilatory power of a lexical unit, whereas the concept of weight is connected with the actual assimilatory force of a unit as weakened by the opposing force of resistance originating from the target system. If the mass of an entity can be easily accessed through the BNC data and presented in terms of CRACn1 value, the actual weight is much harder to investigate due to the absence of any stable parameters regulating the action of resistant forces in either the entrenchment of the item in the target system (actual lacking or minimal resistant forces, see Fig. 22), or its disappearance from the system (prevailing resistance forces). The aforementioned lack of stability connected with these resistance forces is attributed to the complex emotive factors (such as, e.g. positive or negative attitudes to foreign words by speakers of a recipient language) structuring the intensity of these forces. A detailed set of correspondences between the physical and linguistic world is discussed in section 6.9.

6.5. Energy

The significance of the concept of energy to the present discussion is illustrated by the following excerpt from Hewitt (1998: 100):

> Perhaps the most central to all of science is energy. The combination of energy and matter makes up the universe. Matter is substance, and energy is the mover of substance. The idea of matter is easy to grasp. Matter is stuff that we can see, smell, and feel. It has mass and occupies space. Energy on the other hand, is abstract. We cannot see, smell, or feel most forms of energy […].

Even more interesting is the following remark on the concept of potential energy as it directly relates to the present study on foreign words and phrases. Hewitt (1998: 103) states that "an object may store energy because of its position relative to some other object. This energy is called potential energy (PE), because in the stored state it has the potential to do work." Again the concept of PE is correlated in the linguistic world with the concept of CRAC. Thus, CRACn1 relates to the mass of an entity (and thus the potential assimilatory energy of the item), whereas the actual assimilatory energy associated with the weight of a unit is subject to

[11] See the discussion upon the concept of mass discussed by Harwit (1973: 81–83), whereby the distinction is made between *inertial mass* and *gravitational mass*. This distinction appears to be correspondent with the distinction between "mass" and "weight" as conceived of in the present book, respectively.

non-systematic variation.[12] The only thing that can, in effect, be systematically measured is CRACn1, and thus the following hypothesis may be stated: the higher the CRACn1 value of a lexical item the higher the likelihood that a given lexical item will become adapted into the target lexical system (see section 8.1).

The above-quoted correlation between the mass of an entity and CRACn1 value of a lexical item is further confirmed in the world of physics by the notion of gravitational potential energy. As Hewitt (1998: 103) claims: "Work is required to elevate objects against Earth's gravity. The potential energy of a body due to elevated positions is called *gravitational potential energy*. Water in an elevated reservoir and the ram of a pile driver have gravitational potential energy." It thus appears that gravitational potential energy may be viewed, at least at the level of non-expert worldview, as a ratio of two categories, i.e. height and weight. That is the higher the object is placed from the ground, the more potential gravitational energy is accumulated in the object, which has direct impact onto the weight of an entity. The linguistic correlation is again quite straightforward and relates to the aforementioned concept of CRAC as psychologically real blended category, which grounded in height/weight distinction, was corroborated for its ontological status by the results of the survey discussed in section 6.1.

Last but not least, a very significant place in physics is occupied by the so-called Law of Conservation of Energy. In Hewitt's (1998: 107) words: "Energy cannot be created or destroyed; it may be transformed from one form into another, but the total amount of energy never changes." Again the law is fully compatible with our emerging model of lexical assimilation in that assimilation of a lexical unit in the target lexical system tends to be achieved via overcoming the opposing force of lexical resistance (see section 6.4). Energy expended on assimilation is not lost, but transformed into the energy expended on weakening the forces of repulsion.

6.6. Gravity

As we already argued in section 6.2, the centripetal force and centrifugal force have been considered the most fundamental forces upon which the process of lexical assimilation is hypothesized to be based. We also said in the same section that the prototypical source of the centripetal force is connected with gravity. This brings us back to one of Kepler's laws of planetary motion that reads: "Each planet moves in an elliptical orbit with the sun at one focus of the ellipse" (Hewitt 1998: 144). As Hewitt (1998: 144) further states:

[12] Again as in the discussion provided in the section above, the actual weight of an entity and thus the correspondent energy of a foreign lexical item are virtually impossible to evaluate. To achieve this goal, we would need to systematically investigate attitudes of speakers towards these 'guest' items, the issue which has so far posed some methodological problems in lexicological studies (see the discussion on the category of usage of non-native lexis in Görlach 2007).

> From the time of Aristotle, the arc-like motions of heavenly bodies were regarded as natural. The ancients believed that the stars, planets, and moon moved in divine circles, free from any impressed forces. As far as the ancients were concerned, this circular motion required no explication. Newton recognized that a force of some kind must be acting on the planets; otherwise, their paths would be straight lines.

These observations marked a few steps away from formulating one of the greatest laws in physics, i.e. the Law of Universal Gravitation. It is described by Hewitt (1998: 145) as follows:

> Everything pulls on everything else in a beautifully simple way that involves only mass and distance. According to Newton, every body attracts every other body with a force that for two bodies is directly proportional to the product of the masses of the bodies involved and inversely proportional to the square of the distance separating them […]. Thus the greater the masses m_1 and m_2, the greater the force of attraction between them in direct proportion to their masses. The greater the distance of separation, the weaker the force of attraction – in inverse proportion to the square of the distance between their centres of mass.

Harwit (1973: 73), in turn, adds what appears quite significant in the selection of centripetal force as prototypical in the characterization of lexical assimilation processes: "Whatever the nature of the force acting on the planet may be, it is clear that this force acts along the radius vector: Such a force is called a *central force*. A planet is pulled toward the sun at all times; and the components of a binary star are always mutually attracted." As Harwit (1973: 74) further argues: "We know that in planetary motion we are dealing with a central attractive force and that the force should decrease more rapidly than the inverse first power of the distance between attracting bodies. We postulate that attractive force is an inverse square law force."

The linguistic correlation with the above presented facts is quite telling and is also evoked in section 8.4. Firstly, the nature of interaction between the lexical systems can be adequately depicted in that the greater and closer the two lexical systems are in respect of one another, the greater the force of attraction between them. Secondly, the force of attraction may be conditioned geographically (the centripetal force of geographical proximity) or culturally (the centripetal force of cultural proximity).[13] Both of these forces are 'equal' in magnitude. The two of them may be equally prominent and, therefore, contribute with double intensity in attracting the donor system and the target system. The chances for the number of lexical items of being anchored in the target system are then the greatest. The intermediate stage is when only one of the aforementioned types of forces becomes prominent. The degree of attraction between the two systems is then relatively smaller. Finally, when centripetal forces of geographical as well as cultural proximity are weak (the source system is both culturally and geographically distant from the target one), the

[13] More on non-physical manifestations of centripetal force is found in section 6.8.

chances of lexical attraction between the two systems are the smallest. All in all, the degree of force acting between the two lexical systems is calculated analogically to Newton's formula (see Fig. 26).

$$F \sim \frac{m_1 m_2}{d_2}$$

Fig. 26. *Symbolic statement of the Law of Universal Gravitation*

6.7. The solar system

What has been stated so far in the present chapter is, hopefully, a sound legitimization of the central claim laid out at the Introduction to the book (section 0.2, see also Preface), namely that linguistic phenomena are analogous to a large extent to the events occurring in the physical world. I would like to show that some more general analogies can be drawn on the basis of correlations discussed above. When we think about the comparison of lexical assimilation processes, relating them to the action of two fundamental forces, i.e. centripetal force of attraction and centrifugal force of repulsion, and when we combine this discussion with Newton's Law of Universal Gravity as well as Kepler's Law of Planetary Motion, it becomes a matter of cause and effect reasoning that another correlation between the linguistic and physical world can be drawn. This correlation concerns the juxtaposition of the solar system from the world of astrophysics with a network of languages affecting the English speech.

This correspondence is strengthened by what we know as basic facts about the existence of the solar system. There is the Sun, the centre of gravitational force, and there is 9 planets[14] revolving round the Sun. The central role of the Sun in the solar system may be in some conflict with what has already been alluded to above, where it was rather suggested that it is the Earth that functions as the centre of the universe. That these two conceptions of the universe can be successfully reconciled and subsumed under the one coherent model of lexical assimilation is argued in section 9.2. We may postulate, however, that the boundaries of the solar system will be the boundaries of the universe as investigation of other systems than solar, and the resulting knowledge we possess is from both the viewpoint of science as well as folk more than scarce.

In the present book we give some priority to Ptolemy's geocentric model of the universe (see Fig. 54 in Chapter 9) as it appears to be more consonant with how lexical assimilation processes can potentially be modeled from the viewpoint of the 'average' human conceptualizer. At the same time, however, we do not, by

[14] Although Pluto has been recently deprived of the status of a planet by scientists, this book, cognitive-linguistic in spirit, makes allowance for the naïve human conceptualization of the universe, according to which, Pluto is still viewed as the farthest planet of the solar system.

any means, ignore what post-Aristotle physicists (Galileo, Copernicus, Newton, Brahe, and Kepler) stated about the functioning of the universe. This leads us to the view that the two models, i.e. Ptolemy's and Copernicus' are not necessarily exclusive but, as already indicated above, rather complementary (see section 9.2). This complementation view is possible within the framework of cognitive-linguistic study, which accepts two modes of knowledge, i.e. naïve view and expert view of the world as equally relevant for the study of language and cognition (see Evans and Green 2006). Ultimately, then, these two modes of knowledge contribute to modeling a metacognitive framework within which the assimilation processes can be grasped (see the aforementioned section 9.2).

Thus in relation to Ptolemy's conception of the universe, we may, by way of analogy, compare the English language to the Earth – the centre of the universe and lexical systems of other languages to planets or planetoids, whose elements, i.e. lexical items (or meteors), get through the atmosphere of the Earth and then either fall onto the surface of the Earth (become preadapted into the system) or simply disappear (i.e. get 'expelled' from the system as non-adapted elements) due to (air)-friction. We may see that the just presented set of analogies constitutes a little bit of simplification in respect of what will be said in the forthcoming sections, but the general idea about the nature of the process of lexical assimilation remains fixed throughout. We are right now at the point when we can examine the aforementioned elements one at a time, looking first at planets, then planetoids, and meteors.

6.7.1. Planets

In his book, *Astrophysical Concepts,* Harwit (1973: 500) defines planets as "a variety of different objects orbit the sun. Together they make up the *solar system*. The earth is representative of planetary objects. *Planets* are large bodies orbiting the sun. They are seen primarily by reflected sunlight." This short excerpt is at the same time a very concise, yet quintessential illustration about how planets behave in the universe with respect to the centre, that is, the Sun. The linguistic correlation appears to be the following: planets correspond to the greatest lexical systems revolving round the Earth (Ptolemy's conception of the universe). The mass of the planet can be measured with some approximation (see section 6.7.1). The same can be done with respect to the planets, i.e. lexical systems of donor languages affecting the English language (see section 8.1). Also similar is the systemic nature of both the solar system and the system of languages, hence a handful of far-reaching analogies can be detected.

Furthermore, it may be claimed that the most powerful lexical system affecting the English language cannot only be granted the status of the biggest planet orbiting in the system (under the expert view of the universe) but also a very pres-

tigious status of the Moon (under a naïve view of the universe), as it is the Moon that people associate as having the greatest impact on the Earth.[15] Moreover, it is the Moon that can be easily seen every night, and it is this celestial body along with the Sun, which are first learned and thus identified by children as existent in the universe apart from the planet Earth. Additionally, there are a lot of myths and legends created about the Moon in almost every culture all over the world. This all, entitles us to the claim that the biggest discovered planet, i.e. lexical system affecting the English language may also be concurrently conceptualized as the Moon. In conclusion, the knowledge we gain from astrophysics is that, aside from huge celestial bodies revolving round the Sun, i.e. planets with the Moon as bearing the special status, we also have smaller bodies, called planetoids whose role in the system cannot by any means be underestimated.

6.7.2. Planetoids

Harwit (1973: 501) says the following about planetoids:

> Besides the nine planets we have listed, there are many more minor planets orbiting the sun. They are sometimes also called *planetoids* or *asteroids*. Most of them travel along paths lying between the orbits of Mars and Jupiter, a region known as the *asteroidal belt*. The largest asteroid is Ceres. Its radius is 350 km. Its mass is about one ten-thousandth that of the earth.

The correlation of what has been delivered in the quotation with the world of linguistics can be presented as follows: planetoids correspond a great deal with more or less 'sized' lexical systems as perspectivized from the Earth. These lexical systems are viewed as planetoids because their force of impact onto the shape of the Earth's surface is less intense due to factors such as volume and mass (see section 8.1). This view is possible on account of the aforementioned Ptolemy's model of the universe and the concurrent working of anthropocentricity principle (Dirven and Verspoor 2004), which says that human beings occupy the privileged position in the conceptualization of events, the fact which is, in turn, reflected in language. If humans see themselves as privileged, the place where they live, i.e. the Earth may also be viewed as privileged by way of metonymic extension PLACE for PEOPLE. Altogether, planetoids have their due, yet subservient role in the

[15] The proximity of the Moon with respect to the Earth and the ensuing impact of the former on the latter has also been scientifically proven and stated in the Law of Universal Gravity. Newton, as already mentioned earlier, observed that the force of gravity is responsible for regulation of sea tides on the Earth. It has also been scientifically proven that the full Moon has a negative impact on people and statistically there is higher incidence of crime recorded in the area where the full Moon is observed. This shows us to what extent human naïve conception of the Moon and its influence intertwines with some scientific facts corroborating some degree of impact from the Moon.

solar system (i.e. the entire network of languages) with English as the centre – the lingua franca.[16]

6.7.3. Meteors and meteorites

Finally, a few words must be devoted to the discussion of the place of meteors and meteorites within the emerging model of lexical assimilation. Harwit (1973: 501) describes meteors in the following way:

> Many of the smaller known asteroids have diameters of the order of a kilometer. These objects number in the thousands and there must be many more orbiting masses that are too small to have been observed. Among these are bodies that might only be a few meters in diameter or smaller. From time to time, some of these approach the earth and survive the journey through the atmosphere. Such an object that actually impacts on the earth's surface is called a *meteorite*. Meteorites are studied with great interest because they are a direct means of learning about the physical and chemical history of at least a small class of extraterrestrial solar system objects. […] Even smaller than the meteorites are grains of dust that also circle the sun along orbits similar to those of planets. From time to time a grain of dust may enter the atmosphere. Much of it may burn through heat generated by its penetration into the atmosphere, and the particle becomes luminous through combustion and can be observed as a *meteor*, historically called a *shooting star*.

As Harwit (1973: 502) concludes: "We should notice that in talking about planets, asteroids, meteorites, meteors, and micro-meteoric dust grains we are enumerating different-sized members of an otherwise homogeneous group. The major known difference between these objects is their size."

What the above presented excerpt illustrates about the functioning of meteors and meteorites appears to be fully compatible with the model of foreign lexical assimilation that we attempt to sketch in this chapter. Thousands of meteors invading the atmosphere of the Earth may be thus compared to foreign lexical items that enter the lexical system of a given language. The number of people seeing the meteor from the Earth and the number of places from which the meteor can be seen are then correlative of x/y ratio (see section 6.1) recorded in the BNC, and thus of the resulting significance of the lexical item to the target lexical system. It is also consistent that the aforementioned significance of a foreign lexical item must be correlated with the mass of the meteor, as the greater the mass, the greater the chances for the meteor to reach the surface of the Earth and mark some more durable trace on it.

[16] As Crystal (2007a: 29) puts it: "Lingua francas have an obvious and important role in facilitating international communication; but even if one language does, through some process of linguistic evolution, become the world's lingua franca – a status which most people feel is likely to be held by English, it does not follow that this must be at the expense of other languages." See also some adversary viewpoint on the matter in Graddol (1997, after Crystal 2007a: 29).

The correlation between "time," "mass," and the force of impact that a meteor may have on the Earth is directly transposed upon the correlations manifested in language in which the significance of a foreign lexical item is viewed as stemming from its statistics of frequency in the corpus which, as may be hypothesized, is also correlated with the period of time in which a particular lexical item 'stays' in the target lexical system. Once the meteor reaches the ground, it means that it was big enough not to completely melt due to forces of resistance (air-friction). Linguistically, then, it already implicates some degree of assimilation in the system. However, if the meteorite can be said to have been fully recognized as alien element of the landscape depends on whether the meteorite is subject to further exploration by humans. In the same way, the degree of entrenchment of a lexical item, once it has already been noted in the world of discourse of a target language, depends if this lexical unit is further put to use. If so, the lexical item gains the status of a borrowing, if not, it will bear the status of a foreign word and phrase until, due to non-extensive usage, its significance will gradually diminish with the resulting disappearance of the item from the target language map.

The identification of meteorites within the framework of the adopted model of assimilation is thus possible. What is more, it is also possible to strictly demarcate between the set of lexical items bearing the status of borrowings and the set bearing the status of a foreign word of phrase, on the one hand, and the set of 'unknown' words and phrases, on the other. This is argued to be feasible once the criterion of CRAC is adopted in such analysis (more about this issue in section 8.6). For the present moment, we may postulate the following binary sets that illustrate far reaching analogies between the physical and the linguistic world:

- Meteor – any word or phrase of a foreign origin entering the target lexical system.
- Underexplored meteorite – a foreign word or phrase (a preadapted unit).
- Well-explored meteorite – a borrowing (a fully adapted unit).
- Asteroidal matter (a group of barely noticeable meteors) – short-lived 'unknown' words and phrases in the target lexical system.
- Exploration of meteorites with a view to learning more about the origins of the solar system – investigating foreign words to learn more about the origins of donor languages and their contributions within the network of languages in the world.
- Compositional homogeneity of planetary objects – the fundamental homogeneity of natural languages.[17]

A detailed explication of the just introduced correspondences between foreign words and the astrophysical world, together with justification of the presented typology, is found in Chapter 9 (see especially Figs. 48–51).

[17] See, for instance, attempts at (re)creating the universal grammar of all languages in the world, particularly noticeable in generative-transformational paradigm of linguistic research.

6.8. A closer look at force

The discussion upon the significance of force as central to modeling language structurization is not a new trend in language research. The two main cognitive linguists who investigated the concept of force in language are Johnson (1987) and Talmy (1985, 1988, 2002). The former discussed force as represented in cognition by means of FORCE schemas and showed its relevance to language structure, whereas the latter offered a comprehensive descriptive framework where force interaction has not only been noted in language structure at all levels, but also extended on to a discoursal, psychological, and social dimension.

6.8.1. Mark Johnson

Let us first look in more detail at the insights offered by Johnson (1987). First and foremost, he views force as a "dimension of experience." He says:

> I shall examine a second, every present dimension of our experience, that of forceful interaction. The previously described schemata for CONTAINMENT gave prominence to the limitation, restriction, and channelling of forces. By paying more attention to our experience of force as such, we uncover new considerations that did not arise in the analysis of boundedness. These considerations include motion, directness of action, degree of intensity, and structure of causal interaction (including notions of both agency and patienthood, for animate and inanimate things alike).
>
> (Johnson 1987: 41–42)

Later on the same pages, he concludes his general remarks connected with the presence of FORCE schema in the semantics of grammar of a natural language by referring to the experiential basis of force:

> In order to survive as organisms, we must interact with our environment. All such causal interaction requires the exertion of force, either as we act upon other objects, as we are acted upon by them. Therefore, in our efforts at comprehending our experience, structures of force come to play a central role. Since our experience is held together by forceful activity, our web of meanings is connected by the structures of such activity.
>
> (Johnson 1987: 42)

Johnson (1987: 42), then, specifies the details of his enterprise. His major goal is to investigate in a systematic way how the image schema of FORCE permeates the conventional imagery of language, thus posing limitations or constraints upon our inference patterns in reasoning. The main task formulated by Johnson is to "identify the image schemata for some of the more experientially important force structures that bear on semantics and the structure of our conceptual systems" (1987: 43). It is important to remark, that Johnson, as a cognitive linguist, primar-

ily deals with force interaction from the non-expert point of view. This stance is clear in the following quotation:

> I usually pay no attention to the wind, unless it is so strong that it resists my progress as I walk. Only then do I become aware of its force. Likewise, gravity is a force so pervasive that I am seldom aware of it. But we need only encounter a hill in our daily stroll to feel the existence of this force, as if we are suddenly being pulled back.
>
> (Johnson 1987: 42)

Force is then discussed by Johnson (1987: 43–44) as having gestalt structure whose typical features are enumerated below:

- Force is experienced through interaction.
- Our experience of force usually involves the movement of some object (mass) through space in some direction.
- There is typically a single path of motion.
- Forces have origins or sources.
- Forces have degrees of power or intensity.
- There is always a structure or sequence of causality as experienced via force interaction.

The above-presented features make up what Johnson (1987: 44) referred to as "gestalt structure for force." The notion of "gestalt structure" is conceived of "as some organized and unified whole within our experience and understanding that manifests a repeatable pattern or structure." All image-schemas are henceforth referred to by Johnson as possessing gestalt structure, which he considers a defining property of any image schema (Johnson 1987: 44). Of focal interest to the scholar is the investigation of the relationship between force and meaning, which he sees as follows:

> One of the main claims of this book is that meaning (both in the broad sense that I am using the term and in its more narrow sense, as linguistic meaning) is often carried by gestalt structures of this sort. […] We need to explore more concretely how forceful bodily experiences give rise to image-schematic structures of meaning that can be transformed, extended, and elaborated into domains of meaning that are not strictly tied to the body (such as social interactions, rational argument, and moral deliberation).
>
> (Johnson 1987: 44–45)

Finally, he analyses image schemata that illustrate seven of the most frequent structures pertinent to force that can be identified as active in our experience (Johnson 1987: 45).

Compulsion

Experiential basis: being moved by external forces, such as wind, water, physical objects, and other people. This force is regarded as irresistible.

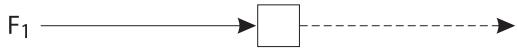


Fig. 27. *Compulsion* (Johnson 1987: 45)

Here the dark arrow represents an actual force vector and the broken arrow denotes a potential force vector or trajectory.

Blockage

Experiential basis: "In our attempts to interact forcefully with objects and persons in our environment, we often encounter obstacles that block or resist our force […]. The relevant gestalt can be represented as a force vector encountering a barrier and then taking any number of possible directions" (Johnson 1987: 45).

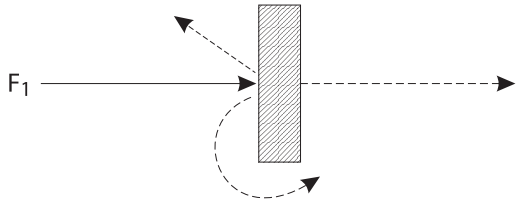


Fig. 28. *Blockage* (Johnson 1987: 46)

Counterforce

Experiential basis: "Football linemen are most familiar with this force gestalt. Here two equally strong, nasty, and determined force centres collide face-to-face" (Johnson 1987: 46).

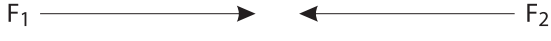


Fig. 29. *Counterforce* (Johnson 1987: 46)

Diversion

Experiential basis: "Rowing a boat at some angle oblique to the wind, you know that without compensation in your rowing, your initial force vector is lost before you know it" (Johnson 1987: 46).

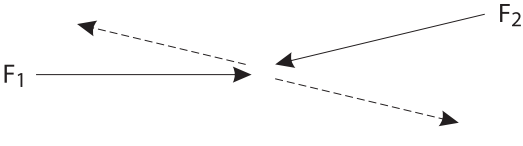


Fig. 30. *Diversion* (Johnson 1987: 46)

Removal of restraint

Experiential basis: "When the door is opened, we are free to come into the room. When the fence is taken away, the dog can visit its canine neighbours, if it so chooses" (Johnson 1987: 46).

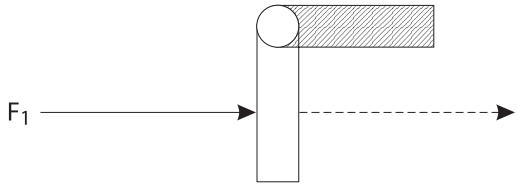


Fig. 31. *Removal of restraint* (Johnson 1987: 47)

Enablement

Experiential basis: "If you choose to focus on your acts of manipulation and movement, you can become aware of a felt sense of power (or lack of power) to perform some action. You can sense that you have the power to pick up the baby, the groceries, and the broom but not to lift the front end of your car" (Johnson 1987: 47).

Fig. 32. *Enablement* (Johnson 1987: 47)

Attraction

Experiential basis: "A magnet draws a piece of steel toward itself, a vacuum cleaner pulls dirt into itself, and the earth pulls back down when we jump. […] The force is not gravitational, in the standard sense, but it is a kind of gravitation toward an object. As such, it shares the same underlying ATTRACTION schema" (Johnson 1987: 47–48).

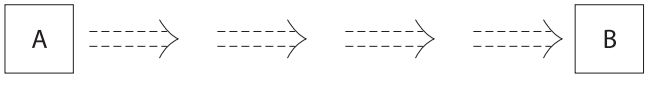


Fig. 33. *Attraction* (Johnson 1987: 47)

The remaining section on force in Johnson's book (1987: 49–57) is dedicated to the analysis of force gestalts in the root senses of modal verbs, which he borrows from a comprehensive study of modals by Eve Sweetser (1982, 1984). Later on he proceeds to the analysis how the structure of speech acts is characterised

by force image-schemas (1987: 57–63).[18] Summing up, we may say that Johnson's considerations over force involve identification of patterns of forceful interaction in the environment and establishing correlations between these patterns and language structure. His illustrations of force seem to reflect naïve view of force conceptualisation and rather ignore expert sources of knowledge about force. For example, the patterns of attraction, counterforce or blockage are discussed as derived from experiences of humans interacting with the environment. Rightly so, but the experiential basis as provided by Johnson appears to be 'grounded' deeper than presented above. It appears that the source of patterning of different types of forces presented by Johnson originates from our experience of gravity and the related action of centripetal and centrifugal force as fundamental forces of action and reaction discussed in Newton's 3rd Law of Motion (see sections 6.2 and 6.8).

The present book appears to overcome the aforementioned constraints in trying to account for both expert and non-expert knowledge about force as contributing to the integrated conceptualisation of lexical assimilation process without which, as the claim goes, the final picture would be 'impoverished' as founded only on one source of knowledge, whether expert or non-expert.

6.8.2. Leonard Talmy

Leonard Talmy, similarly to Johnson, has occupied a very eminent place in the history of linguistic research on force since the 80s of the 20th century, when he delivered his famous papers (1985, 1988) on the presence of force in language and cognition. This conception has evolved into a coherent model-theoretic framework called Force Dynamics: "Force Dynamics gained a good deal of attention in cognitive linguistics due to its claims of psychological plausibility and

[18] See Libura (2000: 167–175) for the modification and elaboration of Johnson's (1987) basic typology of image-schema patterns. An important part of the discussion is the axiological parameter of FORCE image-schema as discussed by Krzeszowski (1997). Krzeszowski postulates that the PLUS /MINUS schema is built in the structure of image-schematic pattern as proposed by Johnson. In his discussion of FORCE, he claims that FORCE is assigned a positive valuation in language provided it is correlated with the concept of CONTROL as excessive FORCE exerted on the object or basically uncontrollable FORCE are valuated MINUS. Krzeszowski (1997: 128) provides an experiential basis for FORCE, which is primarily vital energy, whereas secondarily, types of experiences as "biological, solar, hydric, vental, electric, nuclear, and psychic energy necessary to act." He then proceeds on to the enumeration of various lexical correlates of FORCE schema. Thus he lists positively charged expressions: *a forceful leader; a vigorous campaign, an energetic supporter,* and the like, as well as negatively charged ones: *weak, passive, powerless, an idle brain is the devil's work,* etc. (after Krzeszowski 1997: 128). His treatment of FORCE schema is, however, by no means central within the typological framework of preconceptual-image schemata discussed, which may be explicated through his perspectivization of image-schema discussion in respect of the aforementioned axiological relevance.

the elegance with which it generalizes ideas not usually considered in the same context."[19] The success that Talmy's Force Dynamics enjoyed was because the semantic category of Force Dynamics not only appeared to be "one of the preeminent conceptual organizing categories in language" (2002: 461), but also its structurizing significance has been noted outside the realm of language *per se*, i.e. to psychological, or social interactions as well as discourse (patterns of argumentation). Again, similarly to Johnson, the main question that Talmy (2002: 409) posed in his inquiry was *how* "entities interact with force" rather than *why* such interactions are possible at all. It seems that the answer to this query lies again in shifting the focus of attention more towards non-expert rather than expert type of knowledge in explicating the nature of forceful interaction. Talmy (2002: 409) states: "A semantic category that has previously been neglected in linguistic study is that of force dynamics – how entities interact with respect to force. Included here is the exertion of force, resistance to such a force, the overcoming of such a resistance, blockage of the expression of force, removal of such blockage, and the like."

A significant and pioneering, to the best of my knowledge, was Talmy's reference to the science of physics as an important discipline from which linguists may borrow in interdisciplinary cognitive-linguistic research. Talmy justifies his reflections upon the details of the contemporary physics in the following way: "The linguistic system, in fact, shows close parallels with the conceptual systems for force interaction both in naive physics and psychology, and in early science, as well as in casual treatments of modern science – though it is often at variance with rigorous modern science" (Talmy 2002: 410). However, his final conclusions regarding the category of force appear to be in striking correspondence with the assumptions made in the present book regarding the postulated metacognitive model of lexical assimilation processes, where the integration of both naïve and scientific model we reason about the universe and hence forceful interactions constitutes the theoretic foundation of the present study (see section 10.1.1).

Talmy (202: 411) proposes, then, how the semantic category of Force Dynamics is applicable to language and to psychology and how mental and naïve models may successfully contribute to the organization of our thinking about these fields via the category of Force Dynamics. Talmy proceeds to the illustration of these correspondences saying:

> In historical perspective, developed concepts of force interactions are of course not novel, in particular, for physical phenomena, long the study of disciplines like physics. Outside the physical, perhaps the most familiar application is that of Freud to the psyche with such psychodynamic concepts as libido and drives, repression and resistance, id-superego conflict, and tension-reduction model for restoring equilibrium.
>
> (Talmy 2002: 410)

[19] www.en.wikipedia.org/wiki/Force_Dynamics (ED:07/08).

Talmy, in a consistent way, presents the correspondences between the world of language and other non-linguistic domains by coming up with the terms *psychodynamics* and *sociodynamics*. The former "generalizes notions of physical pushing, blocking, and the like to the framing of such concepts as wanting and refraining" (Talmy 2002: 430), whereas the latter designates an extension of "interpsychological force interactions between sentinent entities" to force interactions of social provenance (Talmy 2002: 438).

It must be reminded that Force Dynamics has been labelled a semantic category; therefore, its application has been postulated to be primarily in language.[20] In consequence, Talmy does not seem to be much concerned with the investigation of forces within their primary physical mode of interaction, but rather focuses on the linguistic manifestation of force, treating other non-linguistic manifestations (whether psychological or sociological) as metaphorical extensions of Force Dynamics. Talmy's Force Dynamics is, then, primarily applicable to the characterization of language structure, and only of secondary if any interest at all are Talmy's inquiries into the physical nature of Force Dynamics. This differentiates Talmy's position from the one adopted in the present book where answering the question how lexical assimilation processes are structured is one, albeit not central concern of the research (Chapter 7). Equally, if not even more important is answering the query *why* lexical assimilation processes can be modelled the way they are suggested in this monograph (see Chapters 8 and 9). To sum up, Talmy gives in this respect an insightful and inspiring answer *how* things come in interactions with force, however, little insight is offered as to *why* such interactions occur at all.

Let us briefly outline Talmy's research on force in language as this part of his enterprise constitutes, accordingly to the remarks above, the major part of his investigation. Talmy (2002: 410), in discussing the history of his research upon force, mentions his own works from 1976 and 1981, Whorf's (1941) discussion on force opposition, Heider's (1958) insights into force and modality, Gee and Kegl's (1982) reference of force in American Sign Language, Sweetser's (1982, 1984) illuminating study of epistemicity of modals within force-dynamics framework, as well as other studies of force as elements of the adopted theoretical frameworks, e.g. Pinker (1989, 1997), Jackendoff (1990) and Brandt (1992).

To the list of scholars dealing with the conception of force by Talmy, it is important to add Achard (1996), Boye (2001), and Vandenberghe (2002) as far as the investigation of modal verbs and force are concerned. Also the usage of force dynamics in lexical semantics (Deane 1992, Da Silva 2003) and morphosyntactical analysis (Chun & Zubin 1990, Langacker 1999: 352–54) cannot be ignored.[21]

[20] It must be reminded that Force Dynamics is regarded by Talmy as privileged semantic category aside from such well recognized ones as number, aspect, mood, evidentiality (2002: 411), so its mode of existence is primarily linguistic.

[21] Source: www.en.wikipedia.org/wiki/Force_Dynamics (ED: 07/08).

Worth quoting is also Walter de Mulder's (2007) integrated account of force dynamics in linguistic research. Moreover, Krzeszowski's (1997) 'exculpation' of Conduit Metaphor in claiming its consonance with the third law of thermodynamics must also be noted as a significant contribution to the study of force in the context of axiological parameterization in the processes of metaphorization.

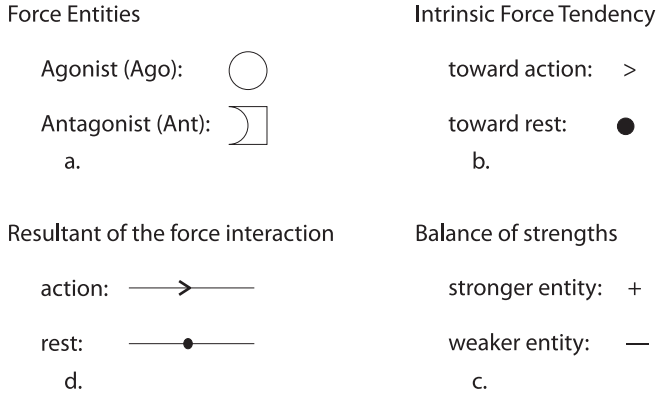


Fig. 34. *Agonist/Antagonist and force* (Talmy 2002: 414)

The following is a brief summary of basic force-dynamic distinctions and their applications in language. We begin with the basic distinction between the Agonist and the Antagonist that Talmy (2002: 413) illustrates in the following way: "Borrowing the terms from physiology where they refer to the opposing members of certain muscle pairs, I call the focal force entity the Agonist and the force element that opposes it the Antagonist."

In his application of force dynamics category to modals, Talmy (2002: 440) stresses a multicontextual analysis of modals conducted in linguistic research so far, however, he notices that all these analyses missed what appears central to the investigation of modality that is the opposition between forces. As for the more complex processes of discourse, Talmy also remarks that force dynamics plays a significant role, especially in the process of argumentation:

> Force dynamics functions extensively in the domain of discourse, and pre-eminently so in the process of *argumentation*. This is the rhetoric of persuasion and includes efforts to exhort, to convince, and to logically demonstrate. The process involves the deployment of points to argue for and against conflicting positions. In a force-dynamic understanding of 'argument space,' each such point can in turn oppose or reinforce another point and overcome or be overcome by it; each successive resultant of these encounters can move the current argument state closer to or further from one of the opposing conclusions.
>
> (Talmy 2002: 452)

Finally Talmy discusses a set of correlations or, put it better, a set of lacking correlations between the physical world and the linguistic world with regard to the category of Force Dynamics. He lists the following aspects:

> • The conception of privilege not existent in physics in case when two objects are equivalent. This is not the case in which equivalent objects either Agonist or Antagonist is conferred on a privileged status.
> • The conception of intrinsic force tendency of an object in language – questioned by modern physics, correlative of earlier scientific theories.
> • The concept of greater relative strength represented in diagrams with a plus sign. In one application of this conception, a *stronger* Antagonist is required so as to be able to block an Agonist with tendency toward motion and hold it in stationary in place. So natural is this linguistic, and perhaps also commonsense, conception that it may have escaped special attention during our exposition. Yet it is at variance with one of the more familiar principles of physics, that two interacting objects – including two objects in contact at zero velocity – must be exerting *equal* force against each other.
>
> (Talmy 2002: 456)

The whole of the above-quoted passage that directly relates to Newton's 3rd Law of Motion can, however, be deemed as falsely grounded in argumentation. The false premises on which Talmy's reasoning is based appears to originate in the wrong interpretation of the aforementioned law of motion, strictly the conception of the equal force exerted in the contact between two objects in contact. Thus, the concept of relative strength as discussed within the Force-Dynamic framework is, contrary to Talmy's claims, not at odds with what we know from scientific physics. The reason is that the strength of the object is not the resultant of one force being stronger than the other, but rather the action of other forces on that object. Hence the sum value of forces acting on that object is bigger that the value of the force(s) acting on the opposite object.

Other examples of lacking correspondences are manifested, in Talmy's words, in "some of the most basis force-dynamic concepts." Thus:

> Blocking and letting, resistance and overcoming – have no principled counterpart in physics. For their viability, these concepts depend on the ascription of entityhood to a conceptually delimited portion of the spatiotemporal continuum, and on the notion of an entity's having an intrinsic tendency toward motion or rest. […] These concepts of blocking and letting vanish, however, under physics' fine-structural perspective of individual particles and forces in local interaction.
>
> (Talmy 2002: 458)

It appears that Talmy's criticism of the lack of correspondence between the notions developed within the Force-Dynamic framework and the science of physics comes from the research perspective taken by the scholar, where force interactions constituting the linguistic Force Dynamics are analysed as to their degree of isomorphism with forces which are subject to a scientific investigation in physics.

Adopting such stance leads to some pessimism connected with the recorded missing counterparts of Force-Dynamic distinctions in the physical world.

The present book differs in that respect from what Talmy suggests, in that the postulated lack of isomorphism between the physical and non-physical forces is substituted for the postulate of search for far-reaching analogies between the world of physics and the world of language as subsumed under the already quoted conceptual metaphor LANGUAGE LAWS ARE PHYSICAL LAWS. Certainly, the proposed metaphorical conceptualisation should not be interpreted in terms of formula *x* is *y*, but rather *x* is like *y* (see section 8.1). Such a viewpoint helps us integrate the discussed lexical assimilation processes as founded on the analogy with the working of centripetal and centrifugal forces, without the need to reflect on, for instance, if Talmy's concepts of resistance, letting, blocking or overcoming are *sensu stricto* represented in nature. The concepts are rather seen as representing phenomena *analogous* to the ones observed in nature in that the concepts of blocking and resistance are both to be viewed as corresponding to a centrifugal force, whereas the concepts of letting and overcoming are analysed as corresponding to a centripetal force (see section 6.2).

6.8.3. Force and lexical assimilation

Concluding the discussion upon the presence of force in linguistic research, we are in position to enumerate a set of conceptual metaphors that will, hopefully, elucidate the nature of conceptual correspondences between the physical forces and their non-physical manifestations. It is important to notice how the fundamental opposition of centripetal vs. centrifugal force permeates the below-presented instantiations of the general metaphor LANGUAGE LAWS ARE PHYSICAL LAWS. The following metaphors are confined to 'force' aspect of the aforementioned correspondences as this has been argued to be central to our conception of lexical assimilation:

- Physical force as CENTRIPETAL (CENTRIFUGAL) FORCE OF GEOGRAPHICAL PROXIMITY (DISTANCE) between languages.
- Physical force as CENTRIPETAL (CENTRIFUGAL) FORCE OF CULTURAL PROXIMITY (DISTANCE) between languages.

Since the language has been with much consent regarded by linguists as a social phenomenon since the early 20s of the 20th century, it is legitimate to propose another metaphor strictly bound up with the emerging model of lexical assimilation where SOCIAL LAWS ARE PHYSICAL LAWS. This metaphor is instantiated by the following:

- Physical force as SOCIAL FORCE OF REJECTION (centrifugal) OR ACCEPTANCE (centripetal).

• Physical force as INSTITUTIONAL FORCE (media, academies, authorities) regulating the inflow of alien lexemes into the target system, etc. (may be both centripetal or centrifugal, depending on circumstances).

It is, thus, on the basis of these 4 fundamental linguistic manifestations of centripetal and centrifugal forces that the conceptualisation of processes of foreign lexical assimilation is based. This issue is also discussed in Chapter 9 (sections 9.1 and 9.2) and in section 10.1.1 in which the integration of all the above-findings is presented.

6.9. Astrophysical concepts and language concepts. Towards consolidation

This section constitutes the summary of the most significant discovered analogies presented in sections 6.1–6.8. They are of a great relevance to the theoretic framework of analysis adopted for the forthcoming analysis. For ease of reference, all of the following analogies are displayed in the form of Tab. 12, which contains 3 parts and represents the discovered correspondences between the physical and linguistic world.[22] Part 1 discusses ontological correspondences between the world of physics and language; Part 2 focuses on correspondences of action (processual) type rather than entity (substantial) type, therefore, they are called epistemic. Finally, Part III treats about the axiological (valuative) implications between physical events and their perceived linguistic counterparts. The axiological correspondences have not yet been explicitly evoked in the present study, but have been addressed numerously in the preceding sections, where the discussion on the nature of centrifugal force was presented. In the Tab. 12 presented below, some elaboration on the axiological treatment of foreign lexical assimilation process is manifested. It is in line with the adopted hybrid model of lexical assimilation that parallels the two views of the universe, i.e. the naïve geocentric and the expert heliocentric (see sections 9.2.1 and 9.2.2).

The naïve geocentric view is addressed in that it helps us conceive of the English language as the centre of the universe, whereas the expert view is given attention as it aids in acknowledging a significant role of solar energy (see Fig. 59, Chapter 9) in structuring the shape of the solar system. This energy may, in consequence, be perceived as positive by the inhabitants of the Earth (native English speakers), provided the influence from other languages is seen as enriching (nutritious). Conversely, the solar energy may be perceived as negative when the influence from other languages is seen as detrimental. This phenomenon has

[22] The presented correspondences in Tab. 12 have been made by the author of this book.

its analog in a human being's excessive exposure to UV radiation, i.e. indiscriminate usage of foreign words and phrases. Consistently, then, contamination of the Earth by foreign influences is commonly viewed as a direct cause of the greenhouse effect, as a result of which the protective shield of ozone layer diminishes. This leads to even the greater exposure of the Earth to detrimental radiation, which in the long run may pose the threat to its identity. The observed facts from the physical world have also their counterparts in the domain of language and linguistics.

Prior to presenting summary findings of the aforediscussed analogies, let us dwell on the crucial distinction between the Earth (English) and planets/planetoids (donor languages). These correspondences appear to be well experientially motivated. Firstly, the planet Earth as well as other planets has similar ontological status in astrophysics as stable landmarks within the solar system. Planetoids, being considerably smaller bodies, are assigned more peripheral status. This is in line with our conception of the English language and its special status among the languages of the world (lingua franca), similarly to the human conception of the planet Earth as having special status among the planets of the solar system. The other eighth planets thus correspond to major donor languages that provide substantial lexical input into the target system of English, whereas planetoids correspond to a significant number of minor languages whose individual lexical contributions into the system of the English language are relatively smaller. Certainly, as the perspective taken in the present research on foreign words and phrases is 'Earth-centred,' therefore, our valuations of major/minor status among languages is also to be regarded as relative to this anthropocentric (Earthly) viewpoint. What thus emerges from the above-presented considerations is the hybrid nature (planet-planetoid) of languages. The hybridism is perceived relative to the direction of investigation.

Tab. 12. *Correspondences between physical and linguistic phenomena*[23]

CONCEPTUAL WORLD	
PHYSICAL WORLD ◄	► LINGUISTIC WORLD

PART I

Ontological correspondences	
The Earth	The English language
Planets	Major donor languages

[23] The discussed correspondences (all of these are mine, M.K.) are tentatively claimed to be neurolinguistically based, where the lexemes are viewed as verbal physical stimuli activating a particular bundle of neural connections in the brain (see Ahlsén 2006). Of course, definitely more neurolinguistic and psychological evidence is needed to validate the postulated metadiscursive metaphor LANGUAGE LAWS ARE PHYSICAL LAWS on these grounds (see Preface).

Tab. 12 – cont.

The Moon	The donor language that affects the English language the most
Planetoids	Minor donor languages
Meteor	Lexeme
Meteor falling into the atmosphere	Incorporation of a foreign word or phrase into the target lexical system
Number of individuals observing the meteor	Number of occurrences of the lexeme in the corpus of the target language
Number of places where the meteor was noticed	Number of texts in which the lexeme appears in the corpus of the target language
Mass[24] of a meteor	CRACn1 value of the lexeme as the average of two variables (x – occurrences, y – texts)
Mass of a planet/planetoid	Number of lexemes times CRACn2 value
Meteor crater	The visual/auditory trace of the assimilation of a lexeme in the target lexical system
Salience of the meteor	Salience of the lexeme[25]
Well-explored meteorite	A borrowing
Underexplored meteorite	A foreign word
A shooting star	An 'unknown'[26] foreign lexeme

PART II

Epistemic correspondences	
Interplanetary gravitation	Network of interaction among languages with English as reference point
The law of universal gravitation. All systems of entities possessing a mass attract one another. The greater the mass of systems the stronger the attraction	The greater the number of lexemes making up systems (source and target) of two languages the greater the attraction between them
Gravitational field	The sphere of the lexical system where the appearance of foreign words and phrases is noted

[24] It should be emphasised that the category of mass is to be conceived as a dynamic, emergent, ontological category rather than presented to language users as an a priori entity. What is more, the mass of a meteor is correlated with the number of people noting the entity. Thus, the greater the mass, the greater the actual force of gravity, viewed as the need by the speech community to implement a non-native unit into the landscape of their own lexical system.

[25] A function of the average value calculated on the basis of the number of occurrences and the number of texts in which a given lexeme appears.

[26] Cf. LD and its competitive metaphor for the lack of assimilation of foreign units: "yet many of the terms Latin has contributed still seem unassimilated, like memorials graven in stone" (p. xi).

Part II of Tab. 12 – cont.

First Newton's Law: If an entity is at rest or its velocity and direction of movement is constant, then the resultant force acting upon this entity equals zero	If there is no need to borrow new lexemes into the target system, it naturally follows that assimilation process does not occur
Second Newton's Law: The acceleration of an object is directly proportional to the net force acting on the object, is in the direction of the net force, and is inversely proportional to the mass of the object	Successful assimilation of a foreign lexical unit will be viewed as a resultant force of prevailing forces of incorporation acting on that unit, on the one hand, and of negligible forces of resistance, on the other (see also section 9.1.2). Thus, the greater CRACn1 value of a particular lemma, the greater the force of incorporation acting on that unit relative to the target lexical system
Third Newton's Law: If A exerts its force upon B, then B exerts the same force upon A, but with the opposite vector	The two manifestations of force schema, i.e. (NO BLOCKAGE) and (BLOCKAGE) which correspond to respective centripetal and centrifugal forces, regulate the working of lexical assimilation process in that NO BLOCKAGE relates to the 'force' of acceptance of a given lexical unit in the target lexical system, whereas BLOCKAGE referes to its rejection from that system
Centripetal force	Action – (active/dynamic) process of incorporating a foreign lexical item in the target language
Centrifugal force	Reaction – (conservative/stabilizing) process of assimilating or rejecting a foreign lexical item in the target system
Resultant force	End-result of the interaction between foregin word or phrase and the target lexical system, i.e. the word or phrase's assimilation or rejection
Gravitational force equals centripetal force (Second Kepler's Law). Gravity force as centripetal force	Incorporation (appearance) of foreign lexical items in the target system
Work as function of interaction between force(s) and the affected entity	Assimilation process as function of the interaction of forces between the foreign word or phrase and the target lexical system
Kinetic energy/potential energy of an entity. The greater the mass and velocity of an entity the greater its kinetic energy	Potential (assimilatory) power of the lexeme. The greater the CRACn1 value of the lexeme, the greater its assimilatory potential
Electromagnetic (gravity) force acting between two systems and their elements	Value of incorporation power acting upon the lexeme conditioned by the distance between the poles of the source and target lexical systems
Weight of an entity in respect of gravity (electromagnetic) force acting between this entity and another one	Actual (assimilatory) value of the lexeme related to an interaction between two lexical systems[27]

[27] The postulated regularity that comes from this epistemic correspondence looks as follows: the greater the attraction between the two lexical systems, the higher the likelihood that lexemes (foreign words and phrases) pertinent to the source (home) system will become adapted more quickly in the target (host) lexical system.

Tab. 12 – cont.

PART III

Axiological correspondences	
Force of resistance (centrifugal) prevailing	Negative attitude of language users to foreign words and phrases
Force of gravity (centripetal) prevailing	Positive attitude of language users to foreign words and phrases
Contamination of the Earth	Perceived detrimental effects caused by the exposure of the English language to other languages[28]
Radiation from the Sun	Source of interacting energy between languages (the process of external change)
Protection of the ozone layer	Protective measures taken by institutional centres in an attempt at protecting English from excessive foreign influence
The greenhouse effect	Perceived threat to the identity of the English language as a result of the interaction with other languages

[28] See also Mańczak-Wohlfeld (2006) on an interesting discussion concerning the influence of English on other languages which is often regarded as highly pernicious (Claude Hagège's lectures).

Chapter 7
EXPLORING PLANETS AND PLANETOIDS

7.0. Classification and procedure

This section discusses classificatory criteria for the characterization of the semantic architecture of foreign words and phrases. One of the more complex methodological problems relates to the categorisation criteria in the process of eliciting semantic fields/domains for the analysis. The overriding principle, according to which the domains in question are selected, relates to the cognitive salience of a category. This means that some of the categories may appear unequal with regard to the degree of specificity in relation to one another. For example the category of "artist" is postulated as a classificatory domain on a par with the category of "writer" despite discernible discrepancy as regards the level of specificity between the two terms. Thus, although the category of "artist" is more schematic than the category of "writer," they are suggested at the same level of the hierarchy in classification of foreign words and phrases on account of the similar salience (entrenchment) of the two terms in English.

In view of the above assumptions, the procedure implemented in the analysis can be characterized as follows: A group of over 2000 specific foreign proper names are subsumed under different superordinate categories, on the basis of some commonality of meaning. Superordinate categories, in turn, identified on account of the aforementioned cognitive salience principle, serve a convenient tool of classification due to their collecting and highlighting function (see Ungerer and Schmid 1996: 109). Thus, when the meaning of a particular foreign word or phrase can be related to a superordinate category, this superordinate term is naturally used as the name of a given domain in the analysis. For instance, a personal proper name designating "a mathematician" can be related at the same time to a more schematic category of "scientist." Certainly, then, for the sake of clarity, the latter rather than the former has been selected as the name of the domain characterizing a given set of different lemmas with the related meaning. If, on the other hand, no such superordinate category can be found, then the basic-level concept as the name of a given domain is suggested instead. Such is the case with the category of "physician" or "politician," which has no salient superordinate counterparts to which they may potentially refer. The situation, when the basic-level term functions as a category label instead of a superordinate one, is connected with the so-called "gaps" in categorization networks, which are nothing extraordinary for any natural language.

For example, Dirven and Verspoor (2004: 39–40) discuss a hierarchical taxonomy for the superordinate conceptual domain called "article of dress" in which

there is a "lexical gap" at the basic level of categorization. In other words, there is no superordinate category that might conventionally subsume a set of specific items of clothing under the category at the same degree of salience as other categories.

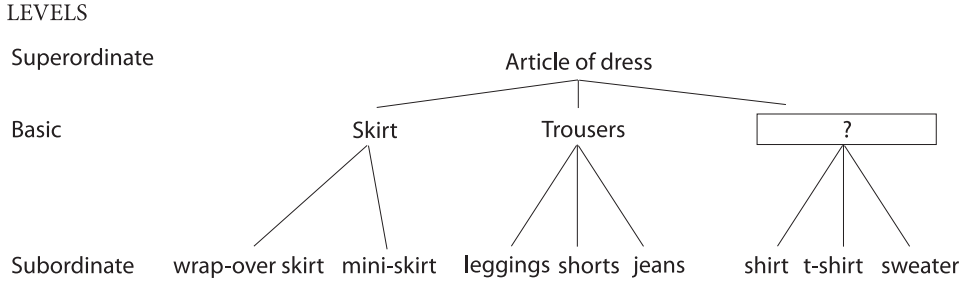


Fig. 35. *Hierarchical taxonomy of 'article of dress'* (Dirven and Verspoor 2004: 40)

To sum up, following the guidelines proposed by Mańczak-Wohlfeld (1995), a selection of more controversial superordinate categories (marked in italics) is presented along with sample enumeration of more specific (basic-level) categories identified in the course of data exploration and used to characterize meanings of particular words and phrases (see Glossary of foreign words and phrases at the end of the book):

- *Artist* – composer; painter; sculptor; musician (including conductors); actor; dancer; fashion designer; film director; tightrope walker,
- *Scientist* – psychologist; physicist; mathematician; biologist; chemist; bacteriologist; psychiatrist; astronomer; zoologist; anthropologist; botanist; anatomist; mineralogist; geographer; ornithologist; computer scientist; cartographer; logician,
- *Writer* – poet; prose writer; playwright; novelist,
- *Scholar* – linguist; philosopher; theologian; educationalist; historian; philologist; editor; printer,
- *Politician* – diplomat; politician; emperor,
- *Sportsman* – football player; golf player; tennis player; formula one driver; chess player,
- *Engineer* – architect; plumber; construction worker; industrialist;
- *Royal family* – historical dynasty; king,
- *Religious activist* – bishop; priest; religious thinker; prelate; pope; archbishop.

The procedure of searching for superordinate categories has also proved a bit of challenge in the case of non-personal proper names as well as place names. Again the principle of cognitive salience led us, in the case of place names, to create categories like "city" which are well cognitively entrenched, and at the same time, propose a few more schematic categories such as "province/region," "caves/rocks" or "mountain/valley." The typographic sign of slash (/) informs us of the contiguity

relationship between the juxtaposed categories. Because of this conceptual affinity between the juxtaposed domains, they have been subsumed under one category on a par with less schematic categories, such as the aforementioned category of "city."

Relatively least problematic turned out the elicitation of domains for the characterization of common words and phrases. This is because of the existent literature in the subject matter, which contributed a great deal to establishing fields or domains into which common words and phrases can be framed in the course of similar analyses. Thus, in discussing English common words and phrases in the Polish language, Fisiak (1970) proposes 11 semantic groups, into which he classified 721 anglicisms. Still, for each category presented, the scholar makes further subdivisions:

1) Sport,
2) Sea, ships, sailors,
3) Trade/economy,
4) Science and technology,
5) Transport, the automotive industry, travel,
6) Man and society,
7) Food, drink and meals,
8) Fashion, clothes, cosmetics, ornaments,
9) Culture and arts,
10) Political institutions and life,
11) Agriculture.

(Mańczak-Wohlfeld 1995: 66–67)

Mańczak-Wohlfeld (1995: 68–72) introduces some further specification to the above presented classification. For example, she adds „cybernetics," „computers," "gardening," "geology," "media," "photography," and many more. Altogether, Mańczak-Wohlfeld discusses 45 semantic fields (1995: 68–72). Most of them have been adopted as 'classificatory' domains for the analysis in the present chapter, but also some new categories have been added. The list of domains proposed for the analysis of foreign proper names and phrases as well as common words and phrases is presented below:

PROPER NAMES

Major subfields: place names, personal names, non-personal names

Place names[29]: Avenue; Caves/Rocks; City; City and River; City and Caves/Rocks; City and Province/Region; City and Lake; City and Mountain; Country; Country/River; Country and City;

[29] As one may notice, the category of place names abounds with polysemous, "artificial" subcategories, which however have been adopted for the analysis in order to avoid terminological overspecification. It is to be reminded that the slash mark "/" does not indicate polysemy, but rather constitutes an instance of conceptually bound heterogeneous unit. The psychologically real status of such category is corroborated by the frequent affinity of the two categories in lexicological classifications (see the above-quoted Mańczak-Wohlfeld 1995 and Fisiak 1970).

Country and Province; Lake/Sea; Miscellaneous[30]; Mountain/Valley; Mountain and Province; Mountain and River; Mountain and Lake; Province/Region; River; River and Province.

Personal names: Adventurer/traveller/sailor; Artist; Banker/economist; Deity/prophet/saint; Engineer/architect/industrialist; Ethnic group; First name/surname; Historical figure/ruler; Inhabitant of a Town; Inventor; Literary/legendary character; Miscellaneous; Physician; Politician; Print maker; Religious activist; Royal family/nobleman; Scientist/scholar; Soldier; Sports activist/sportsman; Writer.

Non-personal names: Administrative/political term; Artistic Institution; Buildings/monuments; Car make; Cheese; Concentration camp; Dynasty/kingdom; Epic poem; Form of address; Holy/religious scriptures; Language/dialect; Manufacturing company; Miscellaneous; Newspaper/magazine; Dish; Political party; Perfume; Religion; Non-alcoholic drink; Opera-related terms; Philosophy; Place-related quality; Publishing/editorial company; Airline; Airport; Cultural movement; Currency; Type of disease; Educational institution; Festival; Honorary title; Type of horse; Type of outerspace project; Type of watch; Wine/spirits.

COMMON WORDS AND PHRASES

Major subgroups: *Person, abstract, thing.*

Domains identified: Biology and animals; Communication; Entertainment/Movie/Media; Fashion/Clothes; Food and drink; Geography/Meteorology; Geology/Mining/Agriculture; Law; Lifestyle/Social life; Location/Buildings; Materials/Objects/Technical Appliances; Measures; Medicine; Military/warfare; Miscellaneous; Money/Trade/Economy; Music and Arts; Philosophy; Politics; Religion; Science and Education; Social Groups/Professions/People; Sport; Tourism and Transport.

At this point, a word should be spared about the procedure applied in the characterization of foreign lexical contributions from particular donor languages (see section 7.1). The order of presentation is not chronological, but anticipates findings about planetary/planetoid status of donor languages discussed in Chapter 8. Thus, we begin the investigation by looking at the most significant languages from the viewpoint of shaping foreign lexical 'mass' of the English language. The further we proceed, the less 'influential' languages are discussed. Eventually, the analysis concludes with the discussion of highly peripheral foreign lexical systems relative to the target English system. As already indicated, the justification of the procedure is explicated in detail in Chapter 8 (see especially section 8.1).

The survey of donor languages starts with analysis of proper names. Proper names are analyzed with regard to the three subtypes: "place names," "personal names," "non-personal" names. Then, some further specification is provided, this time with regard to CRAC parameter (see section 6.1). Once proper names are analyzed, the analysis proceeds to the second fundamental subtype of foreign lexis, i.e. common words and phrases. Here, the analysis is analogical to that of proper names and phrases. Thus, first, common words and phrases are studied with respect to the

[30] See the notions of intercategorial and intracategorial prototypicality (Geeraerts 2006c: 101–104) as an interesting terminological pair for the understanding of the category "miscellaneous" in the present book (section 8.4.1).

three more specific subtypes: "person," "thing" and "abstract." Subsequently, more specification is provided in which case we discuss relevant set of words and phrases relative to CRAC parameter. Similarly to proper names, a consolidation of all the findings presented in the present chapter is offered in Chapter 8.

Finally, a few remarks are needed about the list of foreign words and phrases themselves. The lists are compiled on the basis of two editions of LPD, i.e. 1995 and 2004. Foreign words and phrases that have appeared in the 2004 edition of the dictionary are italicized and additionally marked in bold-face, whereas the remaining ones are just italicized. This should enable the reader to trace the influx of new 'alien' words and phrases during the time span between the 1995 and 2004 edition.

7.1. Planets and planetoids in close-up

7.1.1. French

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.506	277
Place only	0.336	184
Non-personal name only	0.131	72
Miscellaneous	0.025	14
Total: 547		

Personal names: *Abelard; Aimee; Althusser;* ***Anatole****; Andre; Anouilh; Ansermet; Antoine; Antoinette; Appolinaire; Argand; Ariane; Armand; Aznavour; Balzac; Bardot; Baudelaire; Beauvoir; Becquerel;* ***Belmondo****;* ***Bergson****; Berlioz; Bernhardt; Bernoulli; Binet;* ***Binoche****; Bizet; Bleriot; Blondin; Boileau; Bonaparte; Boudin; Bougainville; Bourbon; Braque;* ***Brigitte****; Broca;* ***Camus****; Candide;* ***Cantona****; Cardin; Carnot; Cartier;* ***Casaubon****; Cesar; Cezanne; Chabrier; Chabrol; Chagall; Chagrin; Champlain; Chanel; Chantal; Charlemagne; Chateaubriand;* ***Chirac****; Christophe; Claude;* ***Clouseau****; Clouzot; Cocteau; Coeur de Lion; Colette; Comte;* ***Corbusier****; Coriolis;* ***Corneille****; Corot; Coué; Couperin; Courreges; Cousteau; Cro-Magnon; Curie; Cuvier; Cyrano de Bergerac; Danton; Dartagnan; Daudet; De Bauvoir; De Broglie; De Gaulle; Debussy;* ***Degas****; Delacroix;* ***Delibes****;* ***Delors****;* ***Depardieu****;* ***Derrida****; Descartes; Diderot;* ***Didier****; Dion; Dior; Dominique; Dreyfus; Dubois; Duchamp; Dufy;* ***Dukas****; Dumas; Dupont;* ***Dupuytren****; Durkheim; Duvalier; Eiffel; Emile; Escoffier; Etienne; Eugenie; Faure; Fauve; Feydeau; Flaubert; Foch; Foucault; Fourier;* ***Fragonard****; Franck; Francois; Francoise; Fresnel; Gauguin; Gay; Lussac; Genet; Genevieve;* ***Georges****;* ***Gericault****; Gide; Giselle;* ***Godot****;* ***Goncourt****; Gounod; Grand Guignol; Grignard; Guillaume; Heloise; Henri; Hilaire; Honegger; Hulot; Ionesco; Jacques; Jean; Jules; Julienne; Justine;* ***La Fontaine****; Lagrange; Lalique; Lamarck; Laplace; Lavoisier; Le Corbusier; Lefebvre; Lesseps;* ***Levi-Strauss****; Lisle;* ***Lissajous****; Louis Quatorze; Louis Quinze; Louis Seize; Louis Treize; Maeterlinck;* ***Maginot****; Magritte; Maigret;* ***Mallarme****; Manet; Marat;* ***Marceau****; Marianne; Massenet;*

Matisse; Mazarin; Messiaen; Michel; Michele; Milhaud; Mirabeau; ***Mitterand****; Moliere; Monegasque; Monet; Monique; Montaigne; Montesquieu;* ***Monteux****; Montfort; Montgolfier; Necker; Ney; Odette;* ***Pagnol****; Pantagruel; Parmentier; Pascal;* ***Passy****; Pasteur; Pele;* ***Pelleas****;* ***Pelletier****; Peltier; Perrault; Perrin;* ***Petain****; Philippe; Piaf; Piaget; Picard;* ***Pierre****; Pissarro;* ***Poincare****;* ***Poirot****; Poisson; Pompidou; Poulenc; Proudhon; Proust; Rabelais; Cracine; Rainier; Rameau; Raoul; Ravel; Raynaud; Reaumur;* ***Remy****;* ***Renoir****;* ***Resnais****; Richelieu; Rimbaud;* ***Robbe-Grillet****; Robespierre; Rodin; Ronsard; Rousseau; Sabatier; Sade; Saint-Saens;* ***Sand****;* ***Santer****; Sartre; Satie; Saussure;* ***Savarin****; Schuman; Serge; Seurat;* ***Simenon****; Solvay; St Laurent; Stendhal; Taine; Talleyrand; Tanguy; Tartuffe;* ***Tati****; Teilhard de Chardin; Therese; Tocqueville; Tortelier; Toulouse-Lautrec; Trudeau; Truffaut; Utrillo;* ***Valery****; Valois; Verlaine; Verne;* ***Villon****; Voltaire; Vuitton; Watteau; Yves; Zola.*

Place names: *Abbeville; Abidjan; Aisne; Aix-en-Provence; Aix-la-Chapelle; Aix-les-Bains; Ajaccio; Alencon; Alsace; Amiens; Angouleme; Anjou; Antibes;* ***Aquitaine****; Ardeche; Ardennes; Arles; Armentieres; Arras; Artois; Aubusson; Auvergne; Avignon; Bale; Bangui; Bayeux; Bayonne; Beauvais; Berne; Besancon; Bethune; Beziers; Biarritz; Bordeaux; Boulogne; Brazzaville; Brest; Bruges; Bruxellles; Caen; Calais; Calvados; Camargue; Cambray; Cannes; Carcassonne; Carnac; Casablanca; Cassel; Cenis; Chablis; Chambourcy; Chamonix; Champs Elysees; Chantilly; Chartres; Cherbourg;* ***Cleves****; Cluny; Conakry; Cote d'Azur;* ***Cote d'Ivoire****; Crecy; Croix; Dakar; Deauville; Dieppe; Dijon; Domremy; Dordogne;* ***Dore****; Douai; Douala; Dunkirk; Eupen; Fontainbleau;* ***Gabon****; Graves; Grenoble; Guadeloupe; Havre;* ***Ile de France****;* ***Jura****;* ***Languedoc****;* ***Lascaux****; Lausanne; Laval;* ***Le Havre****; Le Mans; Liege; Lille; Limoges; Limousin; Loire; Lome; Lorraine; Lourdes; Louvain; Lucerne; Luxembourg; Lyon; Macon; Mahe; Malplaquet;* ***Marseilles****;* ***Mayotte****; Medoc; Metz; Meuse; Midi; Millet; Miquelon; Monaco; Mont Blanc;* ***Montmartre****; Montpellier;* ***Montreal****; Montreux; Moselle; Namur;* ***Nancy****; Nantes; Narbonne; Navarre;* ***Neuchatel****; Niamey; Nice; Niger;* ***Nimes****;* ***Nouakchott****;* ***Noumea****; Nuits-Saint-George;* ***Oise;*** *Oran; Orleans; Orly; Ostend; Paris;* ***Perigrod****; Perpignan; Picardy; Poitiers;* ***Pont l'Eveque****; Port-au-Prince; Principe; Provence;* ***Puy-de-Dome****; Quai d'Orsay; Rambouillet; Ramillies;* ***Reims****; Rennes; Reunion; Rheims; Rhone; Rouen;* ***Saint-Etienne****;* ***Sancerre****;* ***Saone****; Seine; Sevres; Simplon; Sion; Somme; St Bernard; St Cloud; St Denis; St Malo;* ***St Tropez****; Strasbourg;* ***Tignes****; Toulon; Toulouse; Tours; Valenciennes; Vaucluse; Verdun; Versailles; Vichy; Vincennes; Vosges; Vouvray; Yaounde; Ypres.*

Non-personal names: *Apache; Arc de Triomphe; Badoit; Bally; Balmain; Barsac; Bastille; Beaujolais;* ***Beaujolais Nouveau****; Bordeaux Mixture; Bordelaise; Bovary;* ***Cheyne-Stokes****; Chillon; Citroen; Cointreau;* ***Comme des Garcons****; Courvoisier; Croix de Guerre; Cuisenaire; Directoire; Dubonnet; Duquesne; Elysee; Entre-Deux-Mers; Fleur; Folies Bergere;* ***Franglais****;* ***Frontenac****; Gauloise;* ***Gitane****; Gobelin; Grand Marnier; Gruyere; Jacquard; La Boheme;* ***Larousse****;* ***Le Monde****; Louvre;* ***Madame****; Marseillaise;* ***Megane****; Mesdemoiselles; Messieurs;* ***Michelin****; Monsiuer;* ***Moulin Rouge****; Muscadet; Notre Dame;* ***Oberlin****;* ***Percheron****;* ***Pernod****; Perrier; Peugeot; Pinot Noir; Port Salut; Provencal; Quebecois; Renault; Roquefort; Sabena; Sandoz; Sauterne;* ***Semillon****; Sorbonne; Suchard; Sylphides;* ***Thermidor****; Tissot;* ***Tour de France****; Tuileries.*

Miscellaneous: *Armagnac; Barbizon; Beaumont; Braille; Breton; Camembert;* ***D'Entrecasteaux****;* ***Evian****; Lafayette;* ***Meniere****;* ***Mont Saint-Michel****; Monte Carlo; Nez Perce;* ***Villeneuve****.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Place and personal name	0.500	7
Personal name and non-personal name	0.214	3
Place and non-personal name	0.285	4
Total: 14 (548)		

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Artist	0.242	69
Writer	0.147	43
Scientist	0.137	39
First name	0.133	38
Miscellaneous[31]	0.123	35
Politician	0.070	20
Scholar	0.060	16
Royal family name	0.028	8
Literary character	0.025	7
Soldier	0.011	3
Sailor	0.007	2
Sportsman	0.007	2
Traveller	0.007	2
Total: 285[32]		

PLACE NAMES

Type	Occurrence	Lemmas
City	0.507	99
Province/Region	0.287	56
Miscellaneous[33]	0.143	28
City and Province/Region	0.097	19
Country	0.020	4
River	0.020	4
River and Province	0.015	3
Avenue	0.010	2
Mountain and Province	0.010	2

[31] Miscellaneous: more than one reference (inter- and intracategorially) or affiliation mentioned once only. This understanding of "miscellaneous" is applicable throughout in analogical contexts.

[32] The total number of lemmas analyzed goes beyond the one presented in the general frequency count. This is because "place names" specification comprises both inter and intra-category polysemy. In the general frequency table the intracategory polysemy is excluded, hence the figure is lower.

[33] Miscellaneous: more than two references inter or intracategorially. This understanding of "miscellaneous" is applicable throughout in analogical contexts.

Mountain/Valley	0.010	2
City and Mountain	0.005	1
Cave/Rocks	0.005	1
Country/River	0.005	1
Total: 195		

NON-PERSONAL NAMES

Domain	Occurrence	Lemmas
Miscellaneous[34]	0.480	37
Wine	0.103	8
Car	0.064	5
Castle/Palace	0.064	5
Cheese type	0.051	4
Honorific form	0.051	4
Academic institution	0.038	3
Mineral water	0.038	3
Type of disease	0.025	2
Language	0.025	2
Liqueur	0.025	2
Name of perfume	0.025	2
Total: 77		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 53)	Lemmas
Personal name only	0.486	69
Place only	0.415	59
Non-personal name only	0.070	10
Miscellaneous[35]	0.028	4
Total: 142		

[34] Miscellaneous: more than one reference inter- or intracategorially or domain recorded once only. This understanding of "miscellaneous" is applicable throughout in analogical contexts.

[35] Miscellaneous: reference to more than one category both inter- and intracategorially.

COMMON WORDS

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0.132	79
Thing	0.454	272
Abstract	0.413	247
Total: 598		

Person: ***Abbe**; agent provocateur; aide-de-camp; **arriviste**; artiste; auteur; **berceuse**; betenoire; bien-pensant; bonne femme; bonvivant; bourgeois; chanteuse; charge d'affaires; chasseur; chatelaine; coiffeur; colon; colporteur; concierge; **confrere**; cordon bleu; curé; danser; danseuse; debutante; demi-mondaine; demoiselle; dirigiste; doyen; doyenne; duchesse; eminence grise; enceinte; enfant terrible; entrepreneur; faineant; farceur; **faux amis**; **faux-naif**; femme; **femme fatale**; fils; gamin; gamine; garcon; gendarme; gourmand; gourmet; idiot savant; ingenue; jardiniere; jeunesse doree; **jongleur**; litterateur; mademoiselle; maitre d'hotel; mangetout; meuniere; midinetce; moi; parvenu; passant; pere; perruquier; petit bourgeois; pierrot; **pointilliste**; raconteur; rapporteur; rentier; repetiteur; restaurateur; sans-culotte; savant; seigneur; **tete-beche**; vichyssoise; voyeur*.

Thing: *A la carte; aide-memoire; **aiguille**; aperitif; **appellation controlee**; arete; arrondissement; assignat; atelier; au jus; au lait; au naturel; au vin; aubade; auberge; autoroute; baccarat; **baguette**; ballade; **bandeau**; barre; bearnaise; beau geste; beau monde; bechamel; **beignet**; bidet; bijouterie; billet-doux; bisque; bistro; blanquette; blanquette de vetu; blouson; boeuf; boeuf bourguignon; bombe; bon voyage; bouchee; bougi; bouillabaisse; bouillon; bouillon cube; boules; bourree; brie; brioche; broderie anglaise; bureau; bureau de change; cabernet sauvignon; cabochon; café; **cafeterie**; canard; capote; carnet; cassis; cassoulet; cayenne; centime; chaconne; chaise longue; chamber; champagne; champignon; chanson; chartreuse; chateau; chateaux; chaud-froid; chef d'oeuvre; chemindefer; **chevre**; chinon; cocotte; cognac; coiffure; coqauvin; coquille; cordon sanitarie; cortege; **coulis**; court-bouillon; creche; creme; **creme de la creme**; creme de menthe; **creme fraiche**; crepes; croissant; crouton; crudités; cuisine; cul-de-sac; culotte; danse; laube; dauphin; decollete; decoupage; demarche; demitasse; denier; digestif; eau; eau de cologne; eau de Nil; eau de vie; eclair; eclat; **ecraseur**; ecu; elan; embonpoint; embouchure; en brochette; en brosse; **en croute**; enjambement; entr'acte; entrechat; entrecote; entree; entremets; entrepot; entresol; epee; escargot; esprit d'escalier; **estaminet**; estragon; etagere; etui; exposé; **fabliau**; faience; fait accompli; farandole; filet mignon; fines herbes; fleur-de-lis; foie gras; foyer; frise; frisson; **fromage frais**; frottage; galere; gateau; gigot; gigue; gouache; goujon; grand prix; grisaille; hors-d'oeuvre; jabot; kepi; kir; langouste; langoustine; langue de chat; limacon; singerie; liqueur; lorgnette; mal de mer; manege; maquillage; maquis; marron; massif; materie; **merlot**; **millefeuille**; **mirepoix**; mistral; mitrailleuse; mot; motjuste; moue; musette; musique concrete; navarin; **neve**; noisette; nouvelle cuisine; objet d'art; objet trouve; **oeil-de-boeuf**; oeuvre; opera bouffe; paillette; **pain au chocolat**; papier-mache; papillon; papillote; **parfum**; parterre; pas; pas de deux; passe-partout; pastisz; pate; pate de fois gras; patos; **pavé**; peignoir; penillion; pension; per sienne; petanque; petit four; petit pois; **picot**; piece de resistance; pied-a-terre; pince-nez; pique; pissoir; piton; **plage**; **planchette**; plat du jour; plein-air; plie; pompadour; pompon; portiere; pot pourri; potage; pot-au-feu; potiche; **pourboire**; poussin; prie-dieu; prix fixe; quarto; quinze; rapport; **ratatouille**; **repousse**; **rocaille**; roman; rouge; rouge et noir; sabot; seance; sedan; soixante-neuf; son et lumiere; table d'hote; **tapenade**; telepherique; tic doloureux; timbale; timbre; toilette; tole; tulle; vignette; vin; vin blanc; vin du pays; vin ordinaire; vin rouge; vingt-et-un; voile; voit; vol-au-ven;wagon-lit*.

Abstract: *A gogo; a la Grecque; a la mode;* ***actualite****; adieu; affaire; allemande; ambience; amour-propre; ampere; ancien regime;* ***anis****; apercu;* ***apres****; arriere-pensee; art deco; art nouveau; assai; assez; au contraire; au courant; au fait; au fond; au gratin; au pair; au revoir; avant-garde; badinage; barré; battue; beau; beaux-arts; belle; belle epoque; belles-lettres; bestie; blasé; bon; bonne boucle; bourgeoisie; brut; c'est la vie; carte blanche; cause celebre; causerie; chacun a son gout; cherchez la femme; chez nous; chinoiserie; cinemaverite; collage; commeilfaut; concours; concours d'elegance; conde; consomme; conte; conte; contretemps; corvee; coulomb; coup d'etat; coup de theatre; coups;* ***couture****; cri de coeur; crime passionel; cru; cuisine minceur; debacie; declasse; decolletage; deja vu; deluxe; demimonde; demode; denoument;* ***de rigeur****; dernier cri; deshabille; detente; detrop;* ***diablerie****; dieu est mon dront; dirigisme; distingue; distrait; divertissement; dossier; double entendre; douceur; droit de seigneur; carte; elan vital; embarrass de richesses; embourgeoisement; en bloc; en famille; en fete; en garde; en passant; en route; en gag; enmasse; ennui; ensuite; entente; entourage; entre nous; esprit de corps;* ***etude****; explication; extraordinaire; faites vos jeux; farci; parouche; faute de mieux; faux pas; fete;* ***fete champetre****; fiancé; fin de siecle; fine; flambe; folie a deux; folie de grandeur; fondu; force majeure;* ***formidable****; fouette; franc; frappe; gaucherie;* ***genre****; gite; grand; grand mal;* ***habitue****; haute; hauteur; honi soit qui mal y pense; hors de combat; idee fixe; idee recu; impasse; ingression; insouciance; insouciant; je ne sais quoi; jete; jeu d'esprit; joie de vivre; julienne; laisser-faire; langue; lese-majeste; liaison; luxe; leyce; lyonnaise; madrilene; manque;* ***mariniere****; melange; menage; menage a trois; mesalliance; métier; mignon; milieu; mise-en-scene; mon veneris; montage; morne; mouille; naivete;* ***napolitaine****; negligee; nicoise; noblesse oblige; nom de guerre; nom de plume; nouveau; nouveau riche;* ***nouvelle vague****; nuance;* ***nul point****;* ***ordinaire****;* ***outre****; par excellence; partipris; passe; penchant; persiflage; petillant; petit mal; petit point; petite; pisa ller;* ***plus ca change****; portugaise; poste restante;* ***pret-a-porter****; protégé; quel; raison d'etre;* ***rapprochement****;* ***recherche****;* ***regime****; renaissance; rendezvous; repechage;* ***repondez s'il vous plait****; retrousse; risque; rite de passage;* ***roman a clef****;* ***roman fleuve****; s'il vous plait; saint;* ***sangfroid****; sans; sauve qui peut; savoir-faire; savoir-vivre;* ***se-tenant****; soi-disant; soignee; spiree; sommelier; soupcon; succes; succes d'estime; succes de scandale;* ***succes fou****;* ***tachisme****; tete-a-tete; ton; tupet;* ***tour de force****;* ***tout court****;* ***tout ensemble****; trompe l'oeil;* ***vacherin****; virement; voila; volte-face.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 33)	Lemmas
Person	0.116	12
Thing	0.378	39
Abstract	0.504	52
Total: 103		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Communication	0.319	191
Food and drink	0.135	81
Lifestyle/Social life	0.091	55
Social Groups/Professions/People	0.086	52
Materials/Objects	0.071	43
Music and arts	0.061	37
Location/Buildings	0.033	20

Biology and animals	0.026	17
Fashion/Clothes	0.025	15
Miscellaneous[36]	0.023	14
Geography/Meteorology	0.021	13
Entertainment/Movie/Media	0.016	10
Politics	0.011	7
Tourism and transport	0.011	7
Military/Warfare	0.010	6
Science and education	0.008	5
Stationery	0.008	5
Name of religion	0.006	4
Sea	0.006	4
Law	0.005	3
Money/Trade/Economy	0.005	3
Philosophy	0.005	3
Sport	0.003	2
Medicine	0.0009	1
Total: 598		

To sum up, in the case of French, the dominant subcategories within the domain of "proper names" are "personal names" and "place names." This tendency is also manifested in global perspective (see Tab. 17, section 8.4.2.1), where the two categories are seen as clearly prevalent over the category of "non-personal names." When it comes to the category of "personal names," the dominant fields relate to the category of "artist" and "writer," whereas in the case of the category "place name," the most frequently recurring domain relates to "city." Finally, the category of "non-personal names" abounds with terms referred to as "miscellaneous," which are characterized by non-recurrent meanings or intracategorial as well as intercategorial polysemy. CRACn2 value in the case of French proper names is estimated at 53. The dominant category, consistently with tendencies mentioned above, is the domain of "personal names." When it comes to French common words, these are characterized by the predominance of the categories "thing" and "abstract" over the category "person." This tendency is maintained when we analyse common words and phrases in question from the perspective of CRACn2, which is estimated at 33. Looking at semantic specification of French common nouns, the dominating fields are related to "communication" and "food and drink."

[36] Miscellaneous: more than one domain (both intra- and intercategorially) referred to. This understanding of "miscellaneous" is applicable throughout in analogical contexts.

7.1.2. German

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.486	201
Place only	0.309	128
Non-personal name only	0.126	52
Miscellaneous	0.080	33
Total: 414		

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Personal name and non-personal name	0.575	19
Place and personal name	0.242	8
Place and non-personal name	0.181	6
Total: 33 (414)		

Personal names: *Adenauer; Adler; Bach; Baedeker; Bauer;* ***Beckenbauer****; Becker; Beckmann; Beethoven; Bismarck; Blucher; Bohm; Boltzmann; Bonhoeffer; Bormann; Brahms; Brandt; Brecht; Bruch; Bruckner; Brunnhilde;* ***Bulow****; Bunsen; Buxtehude; Clausewitz;* ***Curtius****; Dieter; Dietrich; Durer; Durrenmatt; Eckhart; Ehrlich; Eichmann; Einstein; Elsa; Emil; Engelbert; Engels; Ernst; Euler;* ***Fassbinder;*** *Frankenstein; Franz; Freud; Friedrich;* ***Frisch****; Froebel;* ***Furtwangler****; Geissler; Gluck; Godel; Goebbels; Goering; Goethe; Gottfried; Gretchen; Gretel; Gunter;* ***Gunther****; Gutenberg; Habsburg; Handel; Hans; Hansel; Hapsburg; Haydn; Hegel; Heidegger; Heimlich; Heine; Helmholtz; Hermann;* ***Hilbert****; Himmler; Hindemith;* ***Hindenburg****; Hitler;* ***Hofmannsthal****; Hofmeister; Hohenzollern; Holbein;* ***Honecker;*** *Humboldt; Humperdinck; Immelmann; Isolde;* ***Joachim;*** *Jung; Jurgen; Kant; Karajan; Kekule; Kepler; Kirchhoff; Klaus; Klebs-Loffler; Klee; Klemperer; Koch; Kochel;* ***Kohl****; Kokoschka; Kreutzer; Krupp; Kurt; Langerhans; Lauda; Leibnitz;* ***Liebig****; Liszt; Loeb; Lohengrin; Lorelei; Lorenz; Ludwig; Mahler; Marlene;* ***Matthaus****; Mayer; Melanchton; Mendelssohn; Metternich; Mohs; Mollweide;* ***Mossbauer****; Mozart; Muller; Neanderthal; Neumann; Nibelung; Niebuhr; Niersteiner; Nietzsche; Olbers;* ***Ophuls****; Orff;* ***Petri****; Planck;* ***Ribbentrop****;* ***Richter****;* ***Richthofen****;* ***Riefenstahl****; Riemann; Rollei; Rommel; Sachs; Schiller; Schlegel; Schleicher; Schliemann; Schneider;* ***Schnitzler****; Schoenberg; Schopenhauer;* ***Schroder****; Schrodinger; Schubert; Schumacher; Schumann; Schwartz;* ***Schwartzenegger****; Schwartzschild; Schwarzkopf; Schweitzer; Seebeck; Siegfried; Sieglinde; Sigmund; Sigurd; Spengler; Spohr; Stein; Steiner;* ***Stich****; Stockhausen; Strauss; Struwwelpeter;* ***Suppe****; Tannhauser; Telemann; Thyssen;* ***Trubner****; Ulrich; Wagner; Waldemar; Walter; Wankel; Wassermann; Weber;* ***Webern****; Weismann; Weizmann; Werner; Wernicke;* ***Wiesenthal****; Wildenstein; Wolf; Wolff; Wolfgang; Zeiss; Ziegler; Zwingli.*

Place names: *Aachen;* ***Augsburg****; Baader-Meinhof; Bad Godesberg; Baden; Baden-Baden; Bamberg;* ***Basle****; Battenberg; Bayreuth;* ***Berchtesgaden****; Berlin; Bern; Bielefeld;* ***Bochum****; Bonn; Brandenburg; Bremen;* ***Bremerhaven****; Brenner; Brocken; Charlottenburg;* ***Chemnitz****; Coburg; Colditz; Cuxhaven; Dahrendorf; Danzig; Darmstadt; Davos;* ***Detmold****; Deutschland; Dortmund; Dresden; Duisburg;* ***Duisenberg****; Dusseldorf; Eifel; Eiger; Elbe; Ems;* ***Erfurt****; Essen; Frankfurt; Freiburg; Gotterdam-*

*menrung; **Gottingen**; Graz; **Grindelwald**; **Gstaad**; Hamburg; Hamelin; Hanover; Hartz; Heidelberg; **Heilbron**; Heisenberg; Holstein; Innsbruck; Interlaken; Jungfrau; **Kaiserslautern**; Kiel; Kitzbuhel; **Koblenz**; Konigsberg; Leipzig; Liechtenstein; Linz; Lubeck; **Luneberg**; Luthe; Main; **Mainz**; Mannheim; Marburg; Marienbad; Meclenburg; Meissen; Monchen-Gladbach; Munich; Munster; Nassau; **Neisse; Neubrandenburg; Niedersachsen**; Nordhein-Westfalen; Nuremberg; Oberammergau; Oberland; Oder; Offenbach; Oldenburg; Osnabruck; **Paderborn**; **Pilsen**; Potsdam; Reich; Rhine; Rhineland; **Rostock**; Ruhr; Saar; **Saarbrucken**; Saarland; Salzburg; **Schengen**; Schleswig; **Schwartzwald**; Spandau; St Moritz; Stuttgart; Taunus; Tilsit; Trier; **Tubingen**; Tyrol; **Vaduz**; Weimar; Weser; Wiesbaden; **Wittenberg**; Wuppertal; Wurttenberg; Wurzburg; Zermatt; Zugspitze; Zurich*.

Non-personal names: *Altdorfer; Bauhaus; Belsen; **Bertelsmann**; **Biedermeier**; **Bierkeller**; **Borussia**; Buchenwald; Bundesbank; Bundesrat; Bundestag; Bundeswehr; Ciba; **Deutsche Mark**; Frau; Gewurtzaminer; **Hohner**; Kultur; Leica; **Liebfraumilch**; Lowenbrau; Lufthansa; Luftwaffe; Luger; Meistersinger; Messerschmidt; **Nessler**; **Nibelungenlied**; Opel; Parsifal; **Parzival**; Piesporter; Plattdeutsch; **Poggenpohl**; Porsche; **Quellenforschung**; Ratskeller; Reichstag; **Ricoh**; Siemens; **Spatlese**; Sprechstimme; Stuka; **Trabant**; Ursprache; Volkswagen; **Vorsprung durch Technik**; Wartburg; Weltanschauung; Weltschmerz*.

Miscellaneous: *Alzheimer; Baume; Benz; Bingen; Boehm; Dachau; Daimler; Fraunhofer; Gotha; Hansa; Herr; Hesse; Klosters; Mauser; Mobius; Mond; Munchausen; Neckar; Perutz; Pilatus; Radetzky; Riesling; **St Gotthard**; Stern; Tirpitz; Volk; Waldheim; **Waldstein**; Weil; Weiss; **Winterthur**; Worms; Wotan*.

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.263	60
Artist	0.157	36
Scientist	0.140	32
First name	0.096	22
Politician	0.070	16
Scholar	0.061	14
Family name	0.053	12
Writer	0.035	8
Literary character	0.039	9
Sporstman	0.018	4
Royal family	0.018	4
Engineer	0.013	3
Industrialist	0.013	3
Physician	0.013	3
Soldier/Admiral	0.009	2
Total: 228		

PLACE NAMES

Type	Occurrence	Lemmas
City	0.754	107
Province/Region	0.099	14
Mountain/Valley	0.070	10
River	0.056	8
Country	0.014	2
River and Province	0.007	1
Total: 142		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.384	28
Manufacturing company	0.137	10
Car make	0.096	7
Aircraft	0.068	5
Wine	0.068	5
Building/Monument	0.055	4
Bank	0.027	2
Concentration camp	0.027	2
Epic poem	0.027	2
Honorific form	0.027	2
Language	0.027	2
Parliament	0.027	2
Type of watch	0.027	2
Total: 73		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 62)	Lemmas
Personal name only	0.600	60
Place only	0.280	28
Non-personal name only	0.008	8
Personal name and non-personal name	0.004	4
Total: 100		

COMMON WORDS AND PHRASES AND WORD LISTS

Type	Occurrence	Lemmas
Person	0.087	8
Thing	0.450	41
Abstract	0.461	42
Total: 91		

Person: *Doppelganger; fraulein; fuhrer; hausfrau; junker; kaiser;* ***ubermensch****; wunderkind.*

Thing: ***Apfelstrudel****; auslese; autobahn; erg;* ***bergschrund****; bierkeller; bratwurst; Brocken spectre; Bunsen burner; dachshund; edelweiss;* ***festschrift****; foehn;* ***fohn****; gauleiter; glockenspiel;* ***gluhwein****; gneiss; groschen; gymnasium; homburg; kirsch; knackwurst; kummel; lederhosen; lied; less; panzer; pfennig; pumpernickel; quark; sauerbraten; sauerkraut; skat; spitz; stollen; strudel;* ***torte****; wurst; zeppelin;* ***Wernicke's area****.*

Abstract: *Ablaut; ach; ach-laut; achtung; angst; anschluss; auf wiedersehen; bildungsroman; blitzkrieg; ding an sich; echt; ersatz; gemeinschaft; gesellschaft;* ***gestalt****; gestapo; gesundheit; herrenvolk; hertz; ich dien; ich-laut; lebensraum; leitmotif; mach; oflag; ohm; realpolitik;* ***roentgen****;* ***Rorschach test****; schadenfreude; schuss; sieg heil; sprachgefuhl; sprechgesang; stoss; sturm und drang;* ***verboten****; wanderlust; wiener; zeitgeist;* ***Zollner's lines****; zugzwang.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 17)	Lemmas
Person	0.153	4
Thing	0.384	10
Abstract	0.461	12
Total: 26		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Communication	0.219	20
Food and drink	0.142	13
Social groups/Professions/People	0.098	9
Science (physics, chemistry, electricity) and education	0.076	7
Military/Warfare	0.054	5
Music and arts	0.054	5
Geography and meteorology	0.054	5
Geology/Mining and agriculture	0.043	4
Lifestyle/Social life	0.032	3
Biology and animals	0.032	3
Politics	0.032	3

Materials/Objects/Technical appliances	0.032	3
Fashion /Clothes	0.021	2
Entertainment/Movie/Media	0.021	2
Money/Trade/Economy	0.021	2
Law	0.010	1
Tourism and transport	0.010	1
Philosophy	0.010	1
Locations and buildings	0.010	1
Medicine	0.010	1
Total: 91		

In the case of German, the dominant subcategories within the domain of "proper names" relate, as in French, to "personal names" and "place names." When it comes to the category of "personal names," the dominant fields involve the category of "artist" (again analogically to French) and "miscellaneous," whereas in the case of the category "place name," the most frequently recurring domain relates, again, to "city." Finally, the category of "non-personal names" abounds with terms referred to as "miscellaneous" as well as terms connected with industry, i.e. "manufacturing companies." CRACn2 value in the case of German proper names is estimated at 62. The dominant category, consistently with tendencies mentioned above, is the domain of "personal names." When it comes to German common words, these are characterized by the predominance of the categories "thing" and "abstract" over the category "person." This tendency is maintained when we analyze common words and phrases in question from the perspective of CRACn2, which is estimated at 17. Looking at semantic specification of German common nouns, the dominating fields are related to "communication" and "food and drink." As we can notice, in the vast majority of cases, the dominant tendencies in the lexical import from the French are repeated in the case of German.

7.1.3. Italian

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.533	112
Place only	0.333	70
Non-personal name only	0.119	25
Miscellaneous	0.014	3
Total: 210		

Personal names: *Albinoni; Amati;* ***Andrea****; Andrea del Sarto; Angelico;* ***Annigoni****;* ***Antonioni****; Ariosto; Armani; Arturo; Avogadro; Bartolommeo; Bellini; Bernini;* ***Bertolucci****; Boccaccio; Bodoni; Boticelli; Cagliostro; Canaletto; Cannizzaro; Caravaggio; Carpaccio; Caruso; Casanova; Cavour; Cellini; Cenci; Cherubini; Cimabue; Corelli; Correggio;* ***D'Annunzio****; da Vinci; Dallapicolla; Dante; Don; Giovanni;* ***Don Pasquale****; Donatello; Donizetti;* ***Eco****; Enrico;* ***Fellini****; Fermi; Ferrari; Fibonacci; Fra; Galileo; Garibaldi;* ***Gianni****; Gigli;* ***Gioconda****; Giotto; Giovanni; Giulietta; Giuseppe; Golgi; Guido; Lamborghini;* ***Lampedusa****;* ***Lombardo****; Lombroso; Loren; Lucia; Machiavelli;* ***Malpighi****; Marconi;* ***Mascagni****;* ***Mastroianni****;* ***Medici****; Menotti; Michelangelo; Modigliani; Montessori; Monteverdi; Mussolini;* ***Nessun dorma****; Orsini; Paganini; Pagliacci;* ***Pasolini****; Pavarotti;* ***Peano****;* ***Pergolesi****; Perugino;* ***Peruzzi****; Pestalozzi; Pirandello;* ***Prodi;*** *Puccini; Ricci;* ***Roberto****; Rocco; Rossini; Salieri; Savonarola; Scarlatti; Schiaparelli; Sergio;* ***Tasso****;* ***Tiepolo****; Tintoretto; Toscanini; Uccello;* ***Umberto****; Verdi; Veronese; Verrazano; Vespucci;* ***Vinci****; Visconti; Vivaldi; Zeffirelli.*

Place names: *Abruzzi; Agrigento; Amalfi; Ancona;* ***Anzio****;* ***Arno****; Ascona;* ***Asiago****; Assisi;* ***Asti****;* ***Bari****;* ***Bergamo****; Bologna; Brindisi; Cagliari; Calabria; Canossa; Capodimonte; Capri; Capua; Carrara; Castel Gandolfo; Como; Cremona; Elba; Ferrara; Frascati; Friuli; Garda; Genoa; Gorgonzola;* ***Lazio;*** *Lepanto; Locarno; Lugano; Marengo;* ***Marsala****; Messina; Modena; Monte; Monza;* ***Olbia****; Otranto; Padua;* ***Palermo****; Palestrina;* ***Pantelleria****; Parma; Perugia;* ***Piacenza****; Pisa; Po;* ***Rapallo****;* ***Ravenna****;* ***Reggio****;* ***Rialto****; Rimini;* ***Romagna****;* ***Rome****;* ***Salerno****;* ***San Remo****;* ***Scutari****; Siena; Sorrento; Stromboli;* ***Taranto****; Ticino; Tivoli; Verona.*

Non-personal names: *Aida;* ***Alitalia****;* ***Bel Paese****;* ***Borlotti****;* ***Campari****;* ***Cannelloni****;* ***Cavalleria****; Rusticana; Cosi Fan Tutte; Eurydice; Ghia; Lambrusco; Lancia; Maestro; Mozzarella; Olivetti; Panini; Pinocchio; Portofino; Rigoletto; Risorgimento;* ***Signora****;* ***Traviata****;* ***Turandot****;* ***Uffizi****;* ***Zabaglione****.*

Miscellaneous: *Corti;* ***Don Carlos****;* ***Este****.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Personal name and non-personal name	0.009	2
Place name and non-personal name	0.004	1
Total: 3 (210)		

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Artist	0.443	51
Miscellaneous	0.121	14
Writer	0.113	13
First name	0.095	11
Scientist	0.060	7
Politician	0.052	6
Physician	0.052	6

Personal names – cont.

Scholar	0.026	3
Traveller	0.017	2
Sportsman	0.017	2
Total: 115		

PLACE NAMES

Type	Occurrence	Lemmas
City	0.471	33
City and Province/Region	0.271	19
Province/Region	0.171	12
River	0.028	2
River and Province	0.014	1
Lake and Province	0.014	1
City and Lake	0.014	1
Miscellaneous	0.014	1
Total: 70		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.392	11
Opera-related terms	0.250	7
Food and drink	0.142	4
Car	0.071	2
Manufacturing company	0.071	2
Honorific term	0.071	2
Total: 28		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 45)	Lemmas
Personal name only	0.545	24
Place name only	0.386	17
Non-personal name only	0.068	3
CRAC: above 45		
Total: 44		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0.110	12
Thing	0.385	42
Abstract	0.504	55
Total: 109		

Person: *Bambino; cacciatore; carabinieri; cicisbeo; condottiere; contessa; donna; duce;* ***lollo rosso****; osso bucco; papa bile; paparazzo*.

Thing: ***Aioli****; arietta; autostrada;* ***calamari****; cassata; chianti;* ***ciabatta****; cinquecento; espresso; fantoccini;* ***farfalle****; fettuccine;* ***focaccia****;* ***fusilli****; gnocchi; grappa; intaglio; lasagna; linguini; lira; marinara;* ***mascarpone****; panettone;* ***penne****; pesto; pizza;* ***presa****; prosciutto; provolone; quattrocento; radicchio; risotta;* ***ripieno****; risotto;* ***saltimbocca****;* ***scagliola****;* ***scena****; sirocco; segno;* ***spumante****; trattoria; vaporetto*.

Abstract: *Accelerando; aggiornamento; al dente;* ***alla****; andante; andantino; arrivederci; basso pro fundo; bel; bel canto;* ***ben trovato****; chiaroscuro;* ***commedia dell'arte****;* ***da capo****; dal; dal segno; dolce farni ente; dolce Vita; doloroso;* ***finto****; gran turismo; grave;* ***macchiato****; maestoso;* ***molto****; morbidezza; non troppo; opera buffa; opera seria; palladio;* ***parlando****;* ***pentimento****; pieta; poco; rallentando; rilievo; scherzando; scherzo; secco; segue; semplice; sempre; senza;* ***sgrafitto****; simpatico; sotto voce;* ***spiccato****; stretto; terza rima; tessitura; toccata; troppo; tutti;* ***verismo****; volta*.

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 17)	Lemmas
Person	0.142	3
Thing	0.380	8
Abstract	0.476	10
Total: 21		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Communication	0.366	40
Food and drink	0.211	23
Social groups/Professions/People	0.100	11
Music and arts	0.100	11
Materials/Objects/Technical appliances	0.055	6
Geography and meteorology	0.027	3
Lifestyle/Social life	0.018	2
Tourism and transport	0.018	2
Locations and buildings	0.018349	2

Common words and phrases – cont.

Measures	0.018349	2
Science and education	0.009174	1
Military/Warfare	0.009174	1
Biology and animals	0.009174	1
Fashion /Clothes	0.009174	1
Money/Trade/Economy	0.009174	1
Lakes/Seas	0.009174	1
Religion	0.009174	1

In the case of Italian, the dominant subcategories within the domain of "proper names" are, similarly to French and German, "personal names" and "place names." When it comes to the category of "personal names," the dominant fields relate to the category of "artist" and "miscellaneous" (exactly as in the case of German), whereas in the case of the category "place name," the most frequently recurring domain relates to "city." Finally, the category of "non-personal names" is dominated by the subcategory of "miscellaneous" as well as opera-related terms. CRACn2 value in the case of Italian proper names amounts to 45. The dominant category is the domain of "personal names," which repeats the pattern noted in the case of French and German. When it comes to common words, these are characterized by the predominance of the categories "thing" and "abstract" over the category "person." This tendency is maintained when we analyze common words and phrases in question from the perspective of CRACn2, which is estimated at 17. Looking at semantic specification of Italian common nouns, the dominating fields are related to "communication" and "food and drink". In conclusion, in the vast majority of cases, the dominant tendencies observed in the lexical import from French and German are also discernible in the case of Italian.

7.1.4. Spanish

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.502	108
Personal name only	0.386	83
Non-personal name only	0.083	18
Miscellaneous	0.027	6
Total: 215		

Place names: *Aconcagua; Algeciras; Alicante;* ***Almeria****; Altamira; Altiplano; Andalusia; Antofagasta; Aragon; Arrecife; Avila; Baja; Barcelona;* ***Bilbao****; Bogota; Buenos Aires; Cadiz;* ***Cancun****;*

Caracas; Cartagena; Ceuta; ***Chiapas****; Ciudad; Colima; Cordoba; Cordova;* ***Corunna****; Costa Brava; Costa del Sol; Costa Rica; Cotopaxi; Cristobal; Cuba; Darien; Ebro; Ecuador; El Monte; El Paso; El Salvador; Fray Bentos; Galapagos; Granada; Guadalajara; Guadalcanal; Guadalquivir; Guantanamo; Guayaquil; Havana; Ibiza;* ***La Nina****; Lanzarote; Las Palmas; Lima;* ***Madrid****; Majorca; Malaga; Managua;* ***Manresa****; Meta; Mexicali; Mexico; Minorca; Monterey; Nicaragua; Oaxaca; Oviedo;* ***Palma****; Pamplona; Panama;* ***Paraguay****; Peru;* ***Picchu****; Plata;* ***Popocatepetl****; Puerto Rico; Quito; Rio; Rio Grande;* ***Rivera****;* ***Rosario****; Salamanca; Salvador;* ***San Jose****;* ***San Miguel****; San Sebastian;* ***Santa Maria****; Santander;* ***Santo Domingo****; Saragossa; Seville;* ***Sierra Leone****;* ***Sierra Madre****;* ***Tampico****;* ***Tarragona****; Tenerife; Tierra del Fuego; Tijuana; Titicaca; Toledo; Torremolinos;* ***Trujillo****; Uruguay;* ***Valencia****;* ***Valladolid****; Valparaiso; Vigo;* ***Yorba Linda****; Yucatan;* ***Zaragoza****.*

Personal names: ***Albeniz****; Alfonso; Allende; Alonzo; Alphonso; Alvarez; Angel;* ***Angeleno****; Astrias; Balboa; Ballesteros; Bolivar; Borges; Bunuel;* ***Caballe****; Carlos; Carmen; Casals; Castro; Cervantes;* ***Chavez****; Cortez; Dali; De Soto; Diaz; Diego; Dolores; Domingo; Don Quixote; Evita; Falla;* ***Fernandez****; Fernando; Fidel; Franco; Garcia;* ***Geraldo****; Gonzales; Goya;* ***Granados****; Guevara; Inez;* ***Jacinta****;* ***Jorge****; Jose; Juan; Juanita; Junipero; Leon;* ***Lorca****; Loyola; Martinez;* ***Mendez****; Miguel; Murillo;* ***Noriega****;* ***Ortega****;* ***Pablo****;* ***Pachuco****; Panza; Pedro; Pepe;* ***Pepita****; Peron; Picasso; Pinochet; Pizarro; Puerto Rican; Quixote;* ***Ramon****;* ***Ramos****; Raquel; Raul; Sancho; Sancho Panza; Soledad; Torquemada;* ***Vega****; Velasquez; Xavier; Ximenes; Zapata.*

Non-personal names: ***Alcazar****;* ***Alhambra****;* ***Aymara****;* ***Bacardi****;* ***Balenciaga****;* ***Chicana****;* ***Chicano****;* ***Escorial****; Falange;* ***Guarani****;* ***Jerez****;* ***Macarena****;* ***Prado****;* ***Rosinante****;* ***Santeria****;* ***Senor****;* ***Senorita****;* ***Tupamaro****.*

Miscellaneous: ***Guernica****; Lopez;* ***Rioja****;* ***Segovia****;* ***Serra****;* ***Valderrama****.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Place name and personal name	0.018	4
Place name and non-personal name	0.009	2
Total: 6 (215)		

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Artist	0.220	19
First name	0.186	16
Miscellaneous	0.197	17
Politician	0.116	10
Writer	0.081	7
Surname	0.081	7
Literary character	0.058	5
Adventurer	0.034	3
Sportsman	0.023	2
Total: 86		

PLACE NAMES

Type	Occurrence	Lemmas
City	0.417	48
Province/Region	0.191	22
City and Province/Region	0.165	19
Country	0.095	11
Mountain/Valley	0.043	5
River	0.034	4
Country and City	0.017	2
Lakes/Seas	0.017	2
City and River	0.008	1
City and Caves/Rocks	0.008	1
Total: 115		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.619	13
Buildings	0.190	4
Politics	0.095	2
Address forms	0.095	2
Total: 21		

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 75)	Lemmas
Place names	0.538	28
Personal names	0.442	23
Miscellaneous	0.019	1
Total: 52		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Thing	0.692	45
Abstract	0.184	12
Person	0.123	8
Total: 65		

Person: *Aficionado; **caballero**; caudillo; conquistador; conquistadores; infanta; infante; **mariachi**.*

Thing: *Aguardiente; alameda; amontillado; angostura; arroyo; barrio; bodega; bola; burrito; burro; **cabana**; **cajon**; centavo; **chihuahua**; chorizo; copita; cordillera; costa; **fino**; fijole; **gaz pacho**; guacamole; jalapeno; jojoba; mantilla; manzanilla; paella; pampa; parador; peseta; peso; peyote; pina colada; **piñata**; plaza; poncho; presidio; pulque; **quesadilla**; salsa; sangria; sierra; solera; **tapas**; tortilla.*

Abstract: *Adios; camarilla; **cojones**; corrida; hasta la vista; hasta manana; junta; machismo; **pelota**; pinta; pronunciamento; real.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 26)	Lemmas
Thing	0.777	7
Abstract	0.222	2
Person	0	0
Total: 9		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Food and drink	0.307	20
Communication	0.107	7
Social groups/Professions/People	0.076	5
Lifestyle/Social life	0.076	5
Military/Warfare	0.076	5
Biology and animals	0.061	4
Money/Trade/Economy	0.061	4
Geography and meteorology	0.046	3
Locations and buildings	0.046	3
Fashion/Clothes	0.030	2
Politics	0.030	2
Music and arts	0.030	2
Geology/Mining and agriculture	0.015	1
Tourism and transport	0.015	1
Entertainment/Movie/Media	0.015	1
Total: 65		

Summing up, the dominant subcategory within the domain of Spanish proper names is the category of "place names." This stands in contrast to the above-mentioned three most important donor languages (i.e. French, German, and Italian) where it was rather the category of "personal names" that prevailed

over the category of "place names." The dominant fields in the category of "place names" relate to "city," whereas in the case of the category "personal names," the most frequently recurring domain relates to "artist." Finally, the category of "non-personal names" is dominated by the domain of "miscellaneous." CRACn2 value in the case of Spanish proper names is estimated at 75. The dominant category is consistently that of "place names." When it comes to common words, these are characterized by the predominance of the category "thing" over the remaining categories. This tendency is maintained when we analyse common words and phrases in question from the perspective of CRACn2, which is calculated at 26. Looking at dominant semantic fields of Spanish common nouns, these are related to "communication" and "food and drink." Again, we may conclude that prevailing tendencies observed in the lexical import from the most significant 'planets' (i.e. French, German, Italian) can also be noted in the case of Spanish.

7.1.5. Welsh

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.553	124
Personal name only	0.294	66
Non-personal name	0.107	24
Miscellaneous		10
Total: 224		

Place names: *Aberaeron; Aberavon; Aberfan; Abergele; Abersoch; Abersychan; Aberystwyth; Afon; Amlwch; Arenig; Arfon; Bangor; Beddgelert; Bedwellty; Bews-y-Coed; Blaenau; Brynmawr; Bwlch; Cader Idris; Caerau; Caernarfon;* ***Caerphilly****; Caersws; Capel; Cleddau; Clwyd; Conwy; Corwen;* ***Creigiau****; Criccieth; Cynon; Deganwy;* ***Dolgellau****; Dyffryn; Eynon; Fach; Fan; Fawr; Felindre; Festiniog; Ffestiniog; Ffynnongroew; Fron; Froncysyllte; Gabalfa;* ***Gelligaer****;* ***Glanyrafon****; Glyder; Gwynedd;* ***Hafod****; Harlech; Hirwain; Ithon; Llanberis; Llandeilo; Llandovery; Llandrindod; Llandudno; Llanelli; Llanfairfechan; Llanfairpwll; Llangollen; Llanwrst; Llanwrtyd; Lleyn; Machynlleth; Maelor; Maendy;* ***Maentwrog****; Maesteg;* ***Mawddach****; Menai;* ***Merioneth****; Merthyr; Merthyr Tydfil; Nantffrancon; Nantgarw; Ogwr; Padarn;* ***Penarth****; Penmaemawr; Penrhyndeudraeth; Pentre; Penybont; Pen-y-groes; Plynlimon;* ***Pontardawe****; Pontypridd; Porth Dinilaen; Porthmadog;* ***Prescelly****; Prestatyn; Pwllheli; Rheidol; Rhinog; Rhiwbina;* ***Rhos****; Rhosllanerchrugog; Rhosneigr; Rhuddlan; Rhyd-ddu; Rhyl; Ruabon; Talacre; Talybont; Tal-y-llyn; Tawe; Teifi; Tonfanau; Tonypandy; Tonyrefail; Torfaen; Towy; Towyn; Trawsfynydd; Trefor; Tryfan; Tywyn; Waunfawr; Ynys; Ynysybwl; Ystalyfera; Ystrad; Ystradgynlais.*

Personal names: *Aeronwy; Alun;* ***Aneirin****; Angharad; Anuerin; Arwel; Arwyn; Cain; Caradog; Ceinwen; Ceredig; Cyfeiliog; Cynan; Dafydd; Dyfrig; Dylan; Dynevor; Ednyfed; Eifion; Eirian; Eirlys; Emrys; Emyr; Eurig;* ***Ffion****; Gaenor; Gareth; Geraint; Goronwy; Gruffydd; Gwenllian; Gwynfor;*

Heulwen; Huw; Hywel; Iestyn; Illtud; Ioan; Iolo; Iorwerth; Iwan; Lago; Leuan; Llywelyn; Mair; Meredydd; Merfyn; Morfudd; Myfanwy; Olwen; Owain; Pantycelyn; Rhian; Rhiannon; Rhianydd; Rhodri; Rhonwen; Rhydderch; Rhys; Seiriol; Sulwen; Sulwyn; Talfan; Taliesin; Teleri; Tudur.

Non-personal names: *Aled; Bettws; Betws-yn-Rhos; Blodwen; Ceredigion;* ***Coch****; Dovey; Dyfed;* ***Eglwys****; Gorsedd;* ***Gregynog****;* ***Hyder****;* ***Ifor****; Mabinogion; Nant; Penrhos;* ***Plaid Cymru****;* ***Pontllanfraith****; Pontypoll; Ty; Urdd Gobaith Cymru;* ***Wylfa****; Ynys-ddu;* ***Ystwyth****.*

Miscellaneous: *Buddig; Coed; Gorseinon; Graig; Gwaun-cae-Gurwen; Meirion; Meurig; Morfa;* ***Pontardulais****; Rhondda*.

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Personal name and non-personal name	0.013	3
Place name and personal name	0.004	1
Place name and non-personal name	0.031	7
Total: 10 (224)		

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.656	86
Province/Region	0.167	22
Mountain/Valley	0.068	9
River	0.053	7
City and Province/Region	0.030	4
River and Province	0.007	1
Cave/Rocks	0.007	1
Mountain and River	0.007	1
Total: 131		

PERSONAL NAMES

Type	Occurrence	Lemmas
First name	0.231	16
Miscellaneous	0.217	15
Literary or legendary character	0.173	12
Historical figure	0.101	7
Writer	0.086	6
Politician	0.072	5

Personal names – cont.

Prince	0.057	4
Saint	0.028	2
Artist	0.028	2
Total: 69		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.606	20
Buildings	0.090	3
Manufacturing company	0.090	3
Administrative/Political term	0.090	3
Type of educational institution	0.060	2
Kingdom	0.060	2
Total: 33		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 21)	Lemmas
Place name	0.540	33
Personal name	0.245	15
Non-personal name	0.180	11
Place and non-personal name	0.032	2
Total : 61		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Abstract	0.714	5
Thing	0.285	2
Person	0	0
Total: 7		

Thing: *Bach; Cwm.*

Abstract: *Ach-y-fi; Cymru am byth; cynghanedd; eisteddfod; iechyd da.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRAC above 9)	Lemmas
Person	0	0
Thing	0.500	1
Abstract	0.500	1
Total: 2		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Type	Occurrence	Lemmas
Communication	0.428	3
Geography and meteorology	0.142	1
Lifestyle/Social life	0.142	1
Materials/Objects/Technical appliances	0.142	1
Music and arts	0.142	1

To sum up, the prevalent subcategory within Welsh proper names is, similarly to Spanish, the category of "place names." The dominant fields in the category of "place names" relate to "city," whereas in the case of the category "personal names," the most frequently recurring domain relates this time to the category "first name" and "miscellaneous." Finally, the category of "non-personal names" is dominated by the domain of "miscellaneous," exactly in the same way as in the case of Spanish. CRACn2 value for Welsh proper names is estimated at only 21. The dominant category is consistently that of "place names." When it comes to common words, these are not numerous in English. The predominant category is "abstract." Worth noticing is extremely low CRACn2 value for common nouns, which is calculated at 9. With such scarce amount of data, it is hard to delineate any viable tendencies.

7.1.6. Russian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.520	77
Place name only	0.412	61
Non-personal name	0.040	6
Miscellaneous	0.027	4
Total: 148		

Personal names: *Andrei; Bakst;* ***Bakunin****; Boris; Borodin; Brezhnev; Brodsky; Chekhov; Eisenstein;* ***Feodor****;* ***Gagarin****;* ***Glazunof****; Glinka; Godunov;* ***Gogol****; Gorbachev; Gromyko; Igor;* ***Kandinsky****;* ***Karenina****;* ***Karpov****;* ***Kerenski****;* ***Khachaturian****; Khrushchev; Korsakoff; Lenin;* ***Lermontow****; Lysenko; Mendeleyev; Mikhail; Molotov;* ***Moussorgsky****; Mussorgsky;* ***Nekrasov****; Nesselrode; Nijinsky;* ***Nureyev****; Oistrakh; Oleg; Olga;* ***Ordzhonikidze****;* ***Ouspensky****; Pasternak; Pavlov; Pavlova; Petrograf; Potemkin;* ***Primakov****;* ***Prokofiew****; Przewalski; Pushkin; Raisa; Rasputin; Rimsky-Korsakov; Romanov; Rostropovich; Sakharov;* ***Salyut****; Scriabin; Sergei; Shostakovich;* ***Solzhenitsyn****; Stalin; Stanislavski; Stravinsky; Tchaikovsky; Tolstoy; Trubetzkoy; Turgenev; Ustinov; Vladimir;* ***Yeltsin****;* ***Yevtushenko****; Yuri;* ***Zhirinovsky****; Zhukov;* ***Zinoviev*** .

Place names: *Amur;* ***Ashkhabad****; Astrakhan; Azerbaijan; Azov;* ***Babi Yar****; Baikal; Birobidzhan; Bokhara; Borodino; Bug; Bukhara; Chernobyl;* ***Dnepropetrovsk****; Dnieper; Dniester;* ***Grozny****; Inkerman; Irkutsk; Kamczatka;* ***Katyn****; Kazan; Kiev; Kyzyl Kum; Lena; Leningrad; Moskva; Murmansk; Nagorno-Karabakh;* ***Nakhichevan****; Neva;* ***Nizhni Novogrod****;* ***Novaya Zemlya****; Novgorod; Novosibirsk; Ob; Odessa; Okhotsk; Omsk;* ***Onega****; Perm;* ***Petropavlovsk****;* ***Rostov****; Sakhalin;* ***Samara****;* ***Samarkand****;* ***Saratov****; Sevastopol;* ***Smolensk****;* ***St Petersburg****; Sverdlovsk;* ***Tashkent****; Tomsk;* ***Tuva****; Ural; Vladivostok; Volga;* ***Volgograd****; Yakutsk; Yenisei;* ***Yerevan****.*

Non-personal names: *Aeroflot; Bolshoi Ballet; Komsomol; Moskvich; Prawda; Soyuz.*

Miscellaneous: *Gorki; Kirov; Ilyushin, Vostok.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Place and personal name	0.006	1
Personal name and non-personal name	0.013	2
Place and non-personal name	0.006	1
Total: 4 (148)		

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Artist	0.227	18
Miscellaneous	0.227	18
Politician	0.202	16
First name	0.139	11
Writer	0.126	10
Scientist	0.050	4
Scholar	0.025	2
Total: 79		

PLACE NAMES

Type	Occurrence	Lemmas
City	0.625	40
River	0.125	8
Miscellaneous	0.078	5
Country	0.062	4
Province/Region	0.062	4
City and Province/Region	0.015	1
City and River	0.015	1
Mountain or Lake	0.015	1
Total: 64		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.666	6
Spaceflight programme	0.333	3
Total: 9		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRAC above 55)	Lemmas
Personal name only	0.666	22
Place name only	0.303	10
Miscellaneous	0.030	1
Total: 33		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0.133	2
Thing	0.466	7
Abstract	0.400	6
Total: 15		

Person: *Babushka;* ***Przewalski's horse****.*

Thing: *Balaclava; balalaika; chernozem; dacha; piroshki; samowar;* ***taiga****.*

Abstract: ***Glasnost****; kolkhoz;* ***Markov process****;* ***oblast****; perestroika; samizdat.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACm2 above 28)	Lemmas
Person	0	0
Thing	0.400	2
Abstract	0.600	3
Total: 5		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Social groups/Professions/People	0.133	2
Geology/Mining and agriculture	0.133	2
Communication	0.0666	1
Food and drink	0.0666	1
Science and education	0.0666	1
Music and arts	0.0666	1
Geography and meteorology	0.0666	1
Biology and animals	0.0666	1
Politics	0.0666	1
Materials/Objects/Technical appliances	0.0666	1
Fashion /Clothes	0.0666	1
Money/Trade/Economy	0.0666	1
Locations and buildings	0.0666	1
Total: 15		

In the case of Russian, it is the category of "personal names" that comes back to its 'reigning' status as the most salient domain. The dominant subfields in this category relate to "artist" and "miscellaneous," whereas in the case of the category "place names," the most frequently recurring domain relates invariably to the domain of "city." Finally, the category of "non-personal names" is dominated by the domain of "miscellaneous," exactly in the same way as in Welsh and Spanish. CRACn2 value for Russian proper names is estimated at 55, which is, in fact, close to CRACn3 value calculated for the entire category of proper names and phrases (see Chapter 8). The dominant category, under CRACn2 constraints, is consis-

tently that of "personal names." When it comes to common words, no marked tendency in the semantic architecture can be discerned. Despite this fact, CRACn2 value for Russian common nouns is relatively high and amounts to 28.[37]

7.1.7. Chinese

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.678	74
Personal name only	0.174	19
Miscellaneous	0.110	12
Non-personal name only	0.036	4
Total: 109		

Place names: ***Amoy; Anhui; Anshun; Anyang; Chengdu; Chongqing; Chungking; Dalian; Foochow; Fuzhou; Gansu; Guangdong; Guilin; Guizhou; Hainan; Hangchow; Hangzhou; Harbin; Hebei; Heilongjiang; Hong Kong; Hubei; Hunan; Jiangsu; Jiangxi; Jilin; Jinan; Kaohsiung; Kunming; Lhasa; Liaoning; Nanchang; Nanjing; Nanking; Nanning; Ningbo; Ningxia; Pingtung; Qingdao; Qinghai; Quemoy; Qufu; Shaanxi; Shandong; Shanghai; Shantou; Shantung; Shenyang; Shenzhen; Shijiazhuang; Sichuan; Suzhou; Swatow; Szechuan; Taichung; Taipei; Taiping; Taishan; Taiwan; Tangshan; Tiananmen Square; Tianjin; Tibet; Urumqi; Wuhan; Wuxi; Xanadu; Xiamen; Xian; Xinjiang; Yangtse; Yangzhou; Yunnan; Zhejiang; Zhuhai***.

Personal names: ***Chao Yuen Ren***; *Chou Enlai*; ***Deng Xiaoping***; ***Jiang Qing***; ***Jiang Zemin***; ***Ladhar Laotse***; ***Li Tai Po***; ***Lin Biao***; ***Manchu***; ***Mao Tsetung***; ***Sui***; ***Sun Yatsen***; ***Woo***; ***Xia***; ***Zhang Xueliang***; ***Zhou Enlai***.

Non-personal names: *I Ching; Kuomintang; Ming; Putonghua; Qing; Shang; Tao Te Ching.*

Miscellaneous: *Han; Jin; Liao; Matsu; Min; Qin; Tang; Wu; Xia; Xinhua; Yuan; Yue.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Place name and personal name and non-personal name	0.454	5
Place name and personal name	0.272	3
Place name and non-personal name	0.180	2
Personal name and non-personal name	0.090	1
Total: 11 (109)		

[37] For an insightful study of Russian foreign influence upon English, see Piotrowski and Podhajecka (2004: 241–252).

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.500	43
Province/Region	0.441	38
River	0.046	4
River and Province	0.011	1
Total: 86		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.375	9
Politician	0.333	8
Aboriginal tribes	0.125	3
Writer	0.083	2
Surname	0.083	2
Total: 24		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.470	8
Dialect	0.235	4
Dynasty	0.176	3
Classic texts	0.117	2
Total: 17		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 42)	Lemmas
Place name only	0.500	6
Miscellaneous	0.500	6
Total: 12		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Abstract	0.625	5
Thing	0.375	3
Person	0	0
Total: 8		

Abstract: ***Feng shui; guyu; kung fu; pinyin; tai chi.***

Thing: *Renminbi; shih-tzu; yuan.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRAC above 44)	Lemmas
Person	0	0
Thing	0.500	1
Abstract	0.500	1
Total: 2		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Music and arts	0.375	3
Communication	0.250	2
Money/Trade/Economy	0.250	2
Biology and animals	0.125	1
Total: 8		

Chinese foreign proper names differ from the already presented languages in that the category of "place names" clearly dominates over the other categories within the domain of "proper names." The dominant subfield in this category of "place names" is the same as in other so far discussed languages, i.e. "city." In the case of the category "personal names," the most frequently recurring domains are "miscellaneous" and "politician." Similarly, the category of "non-personal names" is dominated by words characterized by inter- and intracategory polysemy. CRACn2 value for Chinese proper names is estimated at 42. Here, the dominant categories are *ex aequo* "place names" and "miscellaneous." In the case of common words, no marked tendency in structuring of semantic fields can be observed. As in Russian, CRACn2 value for Chinese common nouns is relatively high and amounts to 42, that is, it equals CRACn2 calculated for proper names.

7.1.8. Hindi

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.404	19
Personal name only	0.276	13
Non-personal name only	0.255	12
Miscellaneous	0.063	3
Total: 47		

Place names: *Allahabad; Amritsar; Andhra Pradesh; Bihar; Brahmaputra; Deccan; Gujarat; Gwalior; Haryana; Himalaya; Indore; Jaipur; Jalalabad; Lucknow; Meerut; Patna; Pradesh; Sind; Srinagar*.

Personal names: *Agni; **Arjuna**; Asoka; Buddha; Gandhi; **Ganesh**; Maharishi; Nehru; Pathan; Rabindranath; Shiva; Sikh; **Siva***.

Non-personal names: *Taj Mahal; Mahabharata; **Maharashtra**; Ramajana; **Granth**; Marathi; Urdu; Tabla; Karma; Vedanta; Bhagwan; Mahayana*.

Miscellaneous: *Gupta; Gujarati; Sindhi*.

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Personal name and non-personal name	0.666	2
Place and personal name	0.333	1
Total: 3 (47)		

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.550	11
Province/Region	0.350	7
River	0.050	1
Mountain/Valley	0.050	1
Total: 20		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.375	6
Politician	0.187	3
Deity	0.187	3
Englightened individual	0.125	2
Ethnic group	0.125	2
Total: 16		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.428	6
Language	0.214	3
Epic poem	0.214	3
Name of religion	0.142	2
Total: 14		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 34)	Lemmas
Personal name	0.416	5
Place name	0.333	4
Non-personal name	0.250	3
Total: 12		

COMMON WORDS

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Thing	0.368	14
Person	0.315	12
Abstract	0.315	12
Total: 38		

Person: *Babu; **chela**; deva; drobi; guru; **hanuman**; jai; maharaja; **maharanee**; mahatma; pandit; swami.*

Thing: ***Basmati**; bhindi; chapatti; dak; **dhal**; **dhansak**; dhoti; **garam masala**; **lassi**; **puggree**; roti; **saree**; sari; sitar.*

Abstract: *Bhangra; dharma;* ***hatha****; lakh; nirwana;* ***puja****;* ***purda****; raj; sandhi; satyagraha; suttee; svarabhakti.*

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 16)	Lemmas
Thing	0.444	4
Abstract	0.333	3
Person	0.222	2
Total: 9		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Food and drink	0.210	8
Social groups/Professions/People	0.157	6
Fashion/Clothes	0.105	4
Communication	0.105	4
Philosophy	0.105	4
Lifestyle/Social life	0.078	3
Music and arts	0.078	3
Name of religion	0.052	2
Biology and animals	0.026	1
Politics	0.026	1
Measures	0.026	1
Entertainment/Movie/Media	0.026	1
Total: 38		

In the case of Hindi proper names, we observe some relative growth in significance of the domain of "non-personal names" relative to "place names" and "personal names." The dominant subfield in this category of "place names" is "city." In the case of the category "personal names," the prevailing domains are "miscellaneous" and "politician," exactly in the same way as in Chinese. Likewise, the category of "non-personal names" is dominated by "miscellaneous." CRACn2 value for Hindi proper names is estimated at 34. The dominant category, under CRACn2 constraints, is "personal names," which constitutes some novelty with regard to the already discussed languages in that domains prevalent within general semantic fields normally preserved their dominance under CRACn2. This is not the case in Hindi. In the case of common words, the category "thing" prevails both with reference to the general domain of common words as with regard to

the domain limited by CRACn2, which is estimated at 16. When it comes to the semantic specification of common words, the most dominant subfield relates to "food and drink."

7.1.9. Dutch

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.460	29
Personal name only	0.380	24
Non-personal name only	0.079	5
Miscellaneous	0.079	5
Total: 63		

Place names: *Amsterdam; **Antwerp**; Arnhem; Aruba; Breda; **El Enschede**; Europoort; Friesland; Groningen; Haarlem; Hague; Hilversum; Holland; Leiden; Maas; Maastricht; **Mechlin**; Nijmegen; Oudenaarde; Paramaribo; **Scheldt**; **Scheveningen**; **Schiedam**; Suriname; Utrecht; Zeebrugge; Zeeland; Zuider Zee; **Zutphen**.*

Personal names: *Breughel; Buys Ballot; **Cuyp**; **Dijkstra**; Escher; **Haitink**; Huygens; Lorentz; Mondrian; Oort; Oosterhuis; Rembrandt; Rubens; **Ruisdael**; **Ruud**; **Ruyter**; Spinoza; Steen; Teniers; **Tinbergen**; Van der Waals; Van Gogh; **Vermeer**; Zeeman.*

Non-personal names: *Bosch; Concertgebouw; De Stijl; Elzevier; Schiphol.*

Miscellaneous: *Brabant; Edam; Gouda; Hals; Hoboken.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Place name and non-personal name	0.046	3
Place name and personal name	0.031	2
Total: 5 (63)		

SPECIFICATION

PLACE NAMES:

Type	Occurrence	Lemmas
City and Province/Region	0.352	12
City	0.294	10
Province/Region	0.235	8

Place names – cont.

River	0.058	2
Miscellaneous	0.058	2
Total: 34		

PERSONAL NAMES

Type	Occurrence	Lemmas
Artist	0.500	13
Miscellaneous	0.269	7
Scholar/Scientist	0.230	6
Total: 26		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.875	7
Buildings	0.125	1
Total: 8		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 70)	Lemmas
Place name	0.545	6
Personal name	0.272	3
Non-personal name	0.181	2
Total: 11		

To sum up, Dutch proper names are characterized by the prevalence of the domain of "place names" and "personal names." The dominant subfield in "place names" is the hybrid category "city and province/region." This is exception to the so far presented languages in which case it was the category of "city" itself that was observed as most salient. In the case of the category "personal names," the prevailing domain is "artist," similarly to the biggest 'planets' presented so far. Similarly to Hindi and Chinese, the category of Dutch "non-personal names" is dominated by "miscellaneous." CRACn2 value for proper names goes beyond the average, (i.e. 54 [CRACn3]) and reaches 70. The dominant category under CRACn2 is "place names." Surprisingly, there are no Dutch common words and phrases recorded as contributive to the lexical system of English.

7.1.10. Japanese

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.542	19
Non-personal name only	0.371	13
Miscellaneous	0.085	3
Total: 35		

Place names: *Hiroshima; Hokkaido; Honshu; Iwo Jima; Kobe; Kyoto; Kyushu; Nagasaki; Nagoya; Narita; Okinawa; Orinoco; Osaka; Ryukyu; Sapporo; Satsuma; Shikoku; Tokyo; Yokohama.*

Non-personal names: *Bushido; Fujitsu; Mitsubishi; Nikkei; Nissan; Noh; Sanyo; Shinto; Sumitomo; Suzuki; Toshiba; Toyota; Yamaha.*

Miscellaneous: *Fuji; Kawasaki; Nippon.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Place name and non-personal name	0.666	2
Place name and personal name and non-personal name	0.333	1
Total: 3 (35)		

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.500	11
Province/Region	0.318	7
City and Province/Region	0.045	1
Country	0.045	1
River	0.045	1
Mountain and River	0.045	1
Total: 22		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Manufacturing company	0.500	8
Miscellaneous	0.375	6
Car	0.125	2
Total: 16		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 82)	Lemmas
Non-personal name only	0.600	6
Place name only	0.300	3
Place name and non-personal name	0.100	1
Total: 10		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Abstract	0.535	15
Thing	0.464	13
Person	0	0
Total: 28		

Abstract: *Aikido; hiragana; ikebana; kabuki; kana; kanji; karate; katakana; kendo; koan; origiami; siatsu; shogun; sumo; tanka.*

Thing: *Futon; kakemono; kimono; netsuke; sukiyaki; sushi; tatami; tempera; teriyaki; tofu; torii; tsunami; tsutsugamushi.*

COMMON WORDS AND CRAC

Type	Occurrence (CRACn2 above 19)	Lemmas
Person	0	0
Abstract	0.571	4
Thing	0.428	3
Total: 7		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Music and arts	0.321	9
Communication	0.178	5
Food and drink	0.178	5
Materials/Objects/Technical appliances	0.071	2
Medicine	0.071	2
Fashion /Clothes	0.035	1
Locations and buildings	0.035	1
Military/Warfare	0.035	1
Lakes/Seas	0.035	1
Sport	0.035	1
Total: 28		

In the case of Japanese proper names, the domain of "place names" prevails over the category of "non-personal names." Very intriguing is the lack of the category of "personal names." The dominant subcategory of "place names" is "city." The category of "non-personal names" is, in turn, dominated by the subfield of "manufacturing companies" similarly to German. CRACn2 value for Japanese proper names is estimated at 84. The dominant category, in view of CRACn2 constraints, is "non-personal names." In the case of common words, the categories of "abstract" and "thing" absolutely dominate, with the category of "person" nonexistent in the lexical system of English. CRACn2 value for Japanese is 19, and the dominant subfield relates to "music and arts."

7.1.11. Arabic

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.547	23
Personal name only	0.309	13
Non-personal name only	0.119	5
Miscellaneous	0.023	1
Total: 42		

Place names: *Abu Dhabi; Aden; Amman; Aqaba; Aswan;* ***Baalbek****; Baghdad; Bahrain; Basra; Beirut; Benghazi;* ***Damascus****; Dhahran; Iraq; Jedda; Khartoum; Kuwait; Omdurman;* ***Qatar****; Rabat;* ***Riyadh****;* ***Sanaa****; Saudi Arabia.*

Personal names: *Ali; **Arafat**; **Gaddafi**; Ibrahim; **Iqbal**; Mohammed; Nasser; **Osama bin Laden**; **Quaddafi**; **Sadat**; **Saddam**; Saud; Yasser.*

Non-personal names: ***Hadith**; Islam; Ismaili; Koran; Qur'an.*

Miscellaneous: *Muscat.*

Miscellanous (type specification):

Type	Occurrence	Lemmas
Place name and non-personal name	0.023	1
Total: 1 (42)		

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.791	19
Country	0.208	5
Total: 24		

PERSONAL NAMES

Type	Occurrence	Lemmas
Politician	0.461	6
Miscellaneous	0.230	3
Prophet	0.153	2
Muslim fundamentalist	0.153	2
Total: 13		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Name of religion	0.833	5
Miscellaneous	0.166	1
Total: 6		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRAC above 187)	Lemmas
Place name only	0.555	5
Personal name only	0.333	3

Non-personal name only	0.111	1
Total: 9		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0.166	1
Thing	0.166	1
Abstract	0.666	4
Total: 6		

Person: *Mujaheddin.*

Thing: ***Mecca.***

Abstract: ***Fatwa; halal****; intifada.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 76)	Lemmas
Person	0.333	1
Thing	0.333	1
Abstract	0.333	1
Total: 3		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Law	0.333	2
Military/Warfare	0.333	2
Food and drink	0.166	1
Locations and buildings	0.166	1
Total: 6		

In the case of Arabic, we observe the predominance of "place name" category. The most prototypical subdomain is "city." As far as the category of "personal names" is concerned, Arabic is characterized by the prevalence of "politician," whereas in the case of "non-personal names," we record religious terms in the majority. CRACn2 for proper names is relatively very high and amounts to 187 (see discussion in section 10.1.1). Arabic foreign common words and phrases are not numerous in English. It is, thus, hard to discuss even tentative tendencies. Notwithstanding this scarce lexical import, CRACn2 for this category is still proportionately high and approximates 76.

7.1.12. Portuguese

PROPER NAMES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.717	28
Personal name only	0.205	8
Non-personal name only	0.051	2
Miscellaneous	0.025	1
Total: 39		

Place names: *Algarve; Bahia; Beira; Belem; Belo Horizonte; Benguela; Brasilia; Cabinda; Coimbra; Copacabana; Douro; Faro; Funchal; Lisbon; Luanda; Macau; Madera; Mato Grosso; Oporto; Para; Parana; Pernambuco; Porto Alegre; Rio de Janeiro; Sao Paulo; Sao Tome; Setubal; Xingu.*

Personal names: *Camoens; Chagas; Dias; Gomes; Juninho; Rodriguez; Soares; Vasco da Gama.*

Non-personal names: *Estoril; Verde.*

Miscellaneous: *Santos.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Place name and personal name	0.0256	1
Total: 1 (39)		

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.482	14
Province/Region	0.275	8
City and Province/Region	0.137	4
River	0.068	2
River and Province	0.034	1
Total: 29		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.750	6
Politician	0.250	2
Total: 8		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Racing track	0.500	1
Wine	0.500	1
Total: 2		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 32)	Lemmas
Place name only	0.666	8
Personal name only	0.250	3
Non-personal name only	0.083	1
Total 12		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0.600	3
Abstract	0.400	2
Total: 5		

Thing: ***Cruzeiro****; jacana;* ***vinho verde****.*

Abstract: *auto-da-fe;* ***fado****.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 6)	Lemmas
Person	0	0
Thing	1.00	2
Abstract	0	0
Total: 2		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Biology and animals	0.200	1
Food and drink	0.200	1

Common words – cont.

Lifestyle/Social life	0.200	1
Money/Trade/Economy	0.200	1
Music and arts	0.200	1
Total: 5		

To sum up, Portuguese foreign proper names, similarly to the above-discussed Arabic, boast the predominance of "place names" category. The most salient fields within "place names," personal names" subcategories are "city," and "miscellaneous," respectively. CRACn2 value for proper names is 32, and again, given these constraints, the domain of "place names" is seen as most entrenched. Portuguese foreign common words and phrases, are, as in Arabic, rather underrepresented, compared to proper names. Worth noticing is the lack of the category of "person." CRACn2 value amounts to 6. There are no discernible tendencies of field preference in the category of common words and phrases as a whole.

7.1.13. Greek

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.714	15
Personal name only	0.190	4
Non-personal name only	0.095	2
Total: 21		

Place names: ***Aegina; Corfu; Delphi; Epidaurus; Heraklion; Paphos; Parnassus; Patras; Paxos; Piraeus; Salonica; Samos; Siros; Thera;*** *Thessalonica*.

Personal names: ***Papadopoulos; Papandreou; Seferis;*** *Theodorakis*.

Non-personal names: *Avgolemono;* ***Katharevousa***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.533	8
Province/Region	0.333	5
City and Province/Region	0.066	1
Mountain/Valley	0.066	1
Total: 15		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	1.00	4
Other	0	0
Total: 4		

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Artist	0.250	1
Politician	0.250	1
Sportsman and politician	0.250	1
Writer	0.250	1
Total: 4		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Language	0.500	1
Type of dish	0.500	1
Total: 2		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRAC above 17)	Lemmas
Place name	0.857	6
Personal name	0.142	1
Non-personal name	0	0
Total: 7		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	1.00	11
Abstract	0	0
Total: 11		

Thing: *Bouzouki;* ***feta****;* ***filo****;* ***gyro****;* ***moussaka****; ouzo;* ***retsina****;* ***souvlaki****; taramasalata; taverna;* ***tzatziki****.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 15)	Lemmas
Person	0	0
Thing	1.00	5
Abstract	0	0
Total: 5		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Food and drink	0.818	9
Locations and buildings	0.090	1
Music and arts	0.090	1
Total: 11		

In the case of Greek proper names, the domain of "place names" prevails over the remaining categories. The most salient subcategory of "place names," as in the vast majority of the above-discussed languages, relates to "city." The category of "personal names" is, in turn, dominated by the field of "miscellaneous" as in Portuguese. CRACn2 value for Greek proper names is relatively low and amounts to 17. The dominant category, in view of CRACn2 constraints, is "place names." This is in accordance with some tentative tendency noted so far. In the case of common words, the category of "abstract" absolutely dominates, with the categories of "person" and "thing" as non-existent in the lexical system of English. CRACn2 value for Greek common words and phrases is 15, and the dominant field relates, somewhat unsurprisingly, to "food and drink."

7.1.14. Polish[38]

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.500	10

[38] Podhajecka (2002, after Mańczak-Wohlfeld 2006: 7) in her study on the presence of Polish words in the English language lists about 19 lexical units (source: *The Oxford English Dictionary*, 2nd edition, 1989), which constitutes 0,004% of all lexical entries recorded. As Mańczak-Wohlfeld (2006: 8) emphasizes, most of these words belong within "exoticismis" like *Polack*, *sejm*, *mazurka*, *zloty*, which indicates a low degree of usefulness of these words, hence their low frequency in English.

Personal name only	0.400	8
Miscellaneous	0.100	2
Total: 20		

Place names: ***Bydgoszcz***; *Bialystok; Gdansk;* ***Katowice***; ***Krakow***; ***Lodz***; ***Rzeszow***; ***Szczecin***; ***Torun***; *Wroclaw*.

Personal names: ***Gorecki***; *Jan;* ***Jaruzelski***; *Paderewski;* ***Pilsudski***; ***Wajda***; ***Walesa***; ***Wojtyla***.

Miscellaneous: *Kosciusko;* ***Strzelecki***.

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Place name and personal name	0.090	2
Total: 2 (22)		

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.833	10
Mountain/Valley	0.166	2
Total: 12		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.400	4
Politician	0.300	3
Artist	0.300	3
Total: 10		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 59)	Lemmas
Personal name	1.00	2
Miscellaneous	0	0
Total: 2		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0.750	3
Abstract	0.250	1
Total: 4		

Thing: ***grosz***; ***kielbasa***; *zloty*.

Abstract: *mazurka*.

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 10)	Lemmas
Thing	1.00	1
Person	0	0
Abstract	0	0
Total: 1		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Money/Trade/Economy	0.500	2
Food and drink	0.250	1
Music and arts	0.250	1
Total: 4		

Polish foreign proper names are characterized by the absolute prevalence of "place names" and "personal names" categories. More specifically, the most salient field within the domain of "place names" is "city," whereas in the domain of "personal names" it is "miscellaneous." Under CRACn2 constraints, whose value is 59, proper names are represented by the category "personal names." Polish foreign common words and phrases are relatively infrequent, so it is more than risky to suggest any definite tendencies of preference. CRACn2 value is also rather low, and amounts to 10.

7.1.15. Danish

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.523	11
Personal name	0.428	9
Non-personal name	0.047	1
Total: 21		

Place names: ***Aalborg***; ***Aarhus***; *Bornholm; Copenhagen; Elsinore; Esbjerg;* ***Godthaab***; *Jutland;* ***Odense***; ***Roskilde***; *Skagerrak*.

Personal names: *Bering; Bohr; Brahe; Jensen;* ***Jespersen***; *Kierkegaard;* ***Rask***; *Tycho Brahe; Verner*.

Non-personal names: *Lego*.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.363	4
Province/Region	0.363	4
City and Province/Region	0.272	3
Total: 11		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.555	5
Scholar	0.222	2
Scientist	0.222	2
Total: 9		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Toy company	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 29)	Lemmas
Place name	0.250	2
Personal name	0.625	5
Non-personal name	0.125	1
Total: 8		

COMMON WORDS

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Thing	0.333	1
Abstract	0.666	2
Person	0	0
Total: 3		

Thing: *krone.*

Abstract: *Oersted;* ***Verner's law****.*

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 14)	Lemmas
Thing	1.00	1
Person	0	0
Abstract	0	0
Total: 1		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Science and education	0.666	2
Money/Trade/Economy	0.333	1
Total: 3		

In conclusion, Danish foreign proper names, analogically to Polish, display a marked tendency towards the use of "place names" and "personal names." Within the category of "place names" we observe the lack of dominance of "city." Instead, this category is viewed as equally salient as "province/region." "Non-personal" names are represented by the famous toy company *Lego*. Under CRACn 2 constraints, whose value is 29, "personal names" are predominant. Danish foreign common words are rather scarce in the lexical system of English, so it is unreasonable to discuss any marked tendencies. CRACn 2 value in the case of common words and phrases is 14.

7.1.16. Swedish

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.380	8
Place name only	0.333	7
Non-personal name only	0.142	3
Miscellaneous	0.142	3
Total: 21		

Personal names: *Birgitta; Bjorn; Borg; Lars;* ***Palme****; Sibelius;* ***Strindberg****; Swedenborg.*

Place names: *Aland;* ***Gothenburg****; Lund;* ***Norrkoping****;* ***Orebro****; Stockholm; Uppsala.*

Non-personal names: *Bofors;* ***Ericsson****; Krona.*

Miscellaneous: ***Hammarskjold****; Nobel; Berzelius.*

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Scientist	0.272	3
First name	0.272	3
Miscellaneous	0.272	3
Politician	0.181	2
Total: 11		

PLACE NAMES

Type	Occurrence	Lemmas
City	0.571	4
City and Province/Region	0.285	2
Province/Region	0.142	1
Total: 7		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.666	4
Manufacturing company	0.333	2
Total: 6		

MISCELLANEOUS

Type	Occurrence	Lemmas
Name and non-personal name	0.272	3
Total: 3 (11)		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 50)	Lemmas
Personal name only	0.428	3
Place name	0.285	2
Non-personal name only	0.142	1
Personal name and non-personal name	0.142	1
Total: 7		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0.750	3
Abstract	0.250	1
Total: 4		

Thing: ***Fartlek****; ore; smorgasbord.*

Abstract: *Angstrom.*

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 3)	Lemmas
Person	0	0
Thing	1.00	1
Abstract	0	0
Total: 1		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Food and drink	0.25	1
Measures	0.25	1
Money/Trade/Economy	0.25	1
Sport	0.25	1
Total: 4		

In the case of Swedish proper names, the predominant categories are, similarly to other languages discussed above, "personal names" and "proper names." As far as "personal names" are concerned, there is no observable preference, thus the categories of "scientist," "first name," and "miscellaneous" are seen as equally entrenched. The category of "place names" is dominated by "city," whereas the category of "non-personal names" is primarily represented by "miscellaneous." CRACn2 value for Swedish proper names comes near the absolute average (see the discussion of CRACn3 in Chapter 8), and amounts to 50. Under these constraints, the category of "personal names" is viewed as most salient. Swedish common words and phrases are infrequent relative to proper names in English. It is hard to delineate any tendencies, except maybe the general category of "common words and phrases," where the category of "thing" prevails over the other domains. CRACn2 value for this category is very low and is estimated at 3.

7.1.17. Hungarian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.578	11
Place name	0.263	5

General frequency count and word lists – cont.

Miscellaneous	0.105	2
Non-personal name only	0.052	1
Total: 19		

Personal names: ***Bartok***; *Biro; Dohnanyi; Esterhazy; Gabor; Kodaly; Lehar; Lukacs;* ***Petofi***; *Sandor; Solti*.

Place names: ***Balaton***; ***Budapest***; ***Pecs***; *Pest;* ***Szeged***.

Non-personal names: ***Tokay***.

Miscellaneous: ***Kaposi***; *Magyar*.

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Personal name and non-personal name	0.105	2
Other	0	0
Total: 2 (19)		

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.615	8
Artist	0.384	5
Total: 13		

PLACE NAMES

Type	Occurrence	Lemmas
City	0.600	3
City or Province/Region	0.200	1
Lake/Ocean/Sea	0.200	1
Total: 5		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Alcoholic drink	0.333	1
Miscellaneous	0.666	2
Total: 3		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 28)	Lemmas
Place name	0.500	1
Personal name	0.500	1
Total: 2		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0.666	2
Abstract	0.333	1
Total: 3		

Thing: *forint; vizsla.*

Abstract: ***czardasz.***

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 6)	Lemmas
Person	0	0
Thing	1.00	1
Abstract	0	0
Total: 1		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Biology and animals	0.333	1
Money/Trade/Economy	0.333	1
Music and arts	0.333	1

Hungarian foreign proper names are characterized by the predominance of "personal names." Here the subcategory of "miscellaneous" 'wins' over the other fields. When it comes to "place names," and "non-personal names," a noticeable tendency can be observable towards the use of "city" in the case of the former category, whereas in the case of the latter, no marked preference can be discerned due to the insufficient data. CRACn2 value for the category of proper names is 28, whereas for the category of common words and phrases it is only 6. Common words and phrases are again too infrequent to designate any tendency.

7.1.18. Irish

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.350	7
Personal name	0.350	7
Non-personal name	0.300	6
Total: 20		

Place names: *Balie Atha Cliath; Cobh; Dun Laoghaire; Eire; Gaeltacht; Laois; Ni.*

Personal names: ***Deirdre**; Eithne; O'Fiaich; Padraig; Sean; Taoiseach; **Tearlach**.*

Non-personal names: *Dail; **Fianna Fail**; Oireachtas; Saoirse; Seanad; Sinn Fein.*

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.428	3
Country	0.285	2
Province/Region	0.285	2
Total: 7		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.571	4
First name	0.428	3
Total: 7		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Politics	1.00	6
Other	0	0
Total: 6		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 80)	Lemmas
Place name	0.400	2
Personal name	0.400	2
Non-personal name	0.200	1
Total: 5		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0	0
Abstract	1.00	2
Total: 2		

Abstract: ***cead mile faille****; garda.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence CRACn2 above 40)	Lemmas
Abstract	1.00	1
Person	0	0
Thing	0	0
Total: 1		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Communication	0.500	1
Lifestyle/Social life	0.500	1
Total: 1		

Irish foreign proper names do not display any marked tendency of preference towards any of the above-listed subcategories. In the case of "place names" the most salient domain is that of "city," albeit not as salient as in other languages discussed so far. In the case of "personal names" as well as "non personal names," the primarily activated fields relate to "miscellaneous" and "politics," respectively. CRACn2 value for Irish foreign proper names is relatively high, i.e. 80, whereas for common words and phrases it is 40, although it is hard to present here tendencies at all, given just two lemmas fulfilling the criteria of "foreignness" discussed in section 0.2.

7.1.19. Norwegian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.500	8
Personal name	0.437	7
Non-personal name	0.062	1
Total: 16		

Place names: *Bergen; **Lofoten**; **Longyearbyen**; Oslo; Stavanger; Svalbard; Trondheim; Utsira.*

Personal names: *Amundsen; Grieg; Haakon; Ibsen; Munch; Olaf; Roald.*

Non-personal names: *Bokmal.*

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.250	2
Province/Region	0.500	4
City or Province/Region	0.250	2
Total: 8		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.428	3
Artist	0.285	2
First name	0.285	2
Total: 7		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Language	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 38)	Lemmas
Place name	0.500	3
Personal name	0.500	3
Total: 6		

Summing up, Norwegian foreign proper names are characterized by almost equal contribution of "place names" and "personal names" subcategory. In the case of "place names," the most salient field, rather exceptionally, relates to "province/region" (see also section 7.1.9). Within the category of "personal names," the relatively most salient is the category of "miscellaneous." CRACn2 value for Norwegian is 38, but no marked tendencies of preference are noted. In this language we do not record any instances of foreign common words and phrases.

7.1.20. Czech

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.230	3
Personal name only	0.692	9
Non-personal name only	0.076	1
Total: 13		

Place names: *Brno; **Olomouc**; **Vltava**.*

Personal names: *Dvorak; **Havel**; Huss; Janacek; Kafka; **Martinu**; Smetana; **Vaclav**; **Zapotek**.*

Non-personal names: *Skoda.*

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.666	2
River	0.333	1
Total: 3		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.666	6
Artist	0.333	3
Total: 9		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Car	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 41)	Lemmas
Place name	0	0
Personal name	1.00	3
Non-personal name	0	0
Total: 3		

To sum up, Czech foreign proper names are characterized by the predominance of "personal names" over the remaining subcategories. In the case of "personal names" the most salient field is connected with "miscellaneous", whereas in the case of second most salient subcategory, i.e. "place names," it is "city." Czech proper names, as constrained by CRACn2 (41), display the absolute dominance of the category of "personal names."

7.1.21. Afrikaans

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.666	6
Non-personal name only	0.333	2
Place name only	0.166	1
Total: 9		

Personal names: *Botha; Coetzee;* ***de Klerk****; Kruger;* ***Verwoerd****; Voortrekker.*

Non-personal names: *Broederbond;* ***Taal****.*

Place names: *Witwatersrand.*

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Politician	0.666	4
Miscellaneous	0.333	2
Total: 6		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Fraternal organization promoting Afrikaners	0.500	1
Dialects	0.500	1
Total: 2		

PLACE NAMES

Type	Occurrence	Lemmas
Country and Province	1.00	1
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 61)	Lemmas
Place name only	0	0
Personal name only	1.00	2
Non-personal name only	0	0
Total: 2		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0.666	2
Abstract	0.333	1
Total: 3		

Thing: *Rand; sjambok.*

Abstract: *Verkrampte.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 30)	Lemmas
Person	0	0
Thing	1.00	1
Abstract	0	0
Total: 1		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Type	Occurrence	Lemmas
Materials/Objects/Technical appliances	0.3333	1
Money/Trade/Economy	0.3333	1
Politics	0.3333	1

In the case of Afrikaans proper names, the prevailing category, similarly to Czech, is that of "personal names." The most salient field within this category is that of "politician." When talking about remaining categories of "place names" and "non-personal names," no marked tendencies can be noted due to insufficient

data. Under CRACn2, whose value is 61, "personal names" absolutely prevail (as in Czech). Afrikaans foreign common nouns are represented by one lemma, so it is of course impossible to sketch any tendency of preference.

7.1.22. Korean

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.600	6
Non-personal name only	0.400	4
Personal name only	0	0
Total: 10		

Place names: ***Inchon***; ***Panmunjom***; ***Pusan***; *Pyongyang*; ***Seoul***; *Taegu*.

Non-personal names: ***Chaebol***; ***Daewoo***; *Hyundai*; ***Samsung***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	1.00	7
Other	0	0
Total: 7		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Manufacturing company	0.750	3
Miscellaneous	0.250	1
Total: 4		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 43)	Lemmas
Place name	0.500	2

Proper names and phrases and crac – cont.

Non-personal name	0.500	2
Total: 4		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0.500	1
Abstract	0.500	1
Total: 2		

Thing: *hangul.*

Abstract: *taekwondo.*

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 20)	Lemmas
Person	0	0
Thing	0	0
Abstract	1.00	1
Total: 1		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Communication	0.5	1
Sport	0.5	1
Total: 2		

In the case of Korean proper names, we revert to the dominating position (albeit not that clearly marked) of "place names." Within this category, the absolute predominance is given to "city." In the case of "non-personal names," the most salient is the domain designating manufacturing companies (analogically to the above-discussed Japanese). CRACn2 value for Korean proper names is 43. Similarly to Afrikaans, Korean foreign common nouns are represented by one lemma, so it is of course impossible to outline even a tentative tendency of preference.

7.1.23. Cantonese

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.333	2
Personal name only	0.333	2
Non-personal name only	0.333	2
Total: 6		

Place names: ***Kowloon; Wanchai***.

Personal names: ***Cheung; Wong***.

Non-personal names: ***Kai Tak; Hang Seng index***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
Province/Region	1.00	2
Total: 2		

PERSONAL NAMES

Type	Occurrence	Lemmas
Surname	1.00	2
Other	0	0
Total: 2		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Type of airport	0.500	1
Stock market	0.500	1
Total: 2		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 21)	Lemmas
Place name	0.333	1
Personal name	0.333	1
Non-personal name	0.333	1
Total: 3		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	1.000	2
Abstract	0	0
Total: 2		

Thing: ***dim sum; wok***.

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 21)	Lemmas
Person	0	0
Thing	1.000	1
Abstract	0	0
Total: 1		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Food and drink	0.500	1
Materials/Objects/Technical appliances	0.500	1
Total: 2		

Cantonese foreign proper names do not exhibit any tendencies of preference in relation to their basic subcategories. "Place names" are dominated by "province/region," whereas "personal names" are characterized by the prevalence of the field of "surname." CRACn2 value, in the case of Cantonese, amounts to 21. Common words and phrases are infrequent enough to discuss any tendencies.

7.1.24. Scottish-Gaelic

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.500	3
Personal name	0.333	2
Miscellaneous	0.166	1
Total: 6		

Place names: ***Bheinn***; ***Gaidhealtachd***; *Sgurr*.

Personal names: *Sassenach; Seumas.*

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
Province/Region	0.333	1
Mountain/Valley	0.666	2
Total: 3		

PERSONAL NAMES

Type	Occurrence	Lemmas
First name	0.666	2
Miscellaneous	0.333	1
Total: 3		

Miscellaneous: *Seonaid.*

Miscellaneous (type specification):

Type	Occurrence	Lemmas
Personal name and non-personal name	0.166	1
Total: 1 (6)		

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 13)	Lemmas
Place name	1.00	2

Proper names and phrases and CRAC – cont.

Other	0	0
Total: 2		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0	0
Abstract	1.00	2
Total: 2		

Abstract: *Pibroch; slainte*.

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 2)	Lemmas
Person	0	0
Thing	0	0
Abstract	1.00	2
Total: 2		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Communication	0.5	1
Music and arts	0.5	1
Total: 2		

In the case of Scottish-Gaelic proper names, the categories of "place names" and "personal names" are seen as prevailing. Thus, in the category of "place names," the most salient field relates to "mountain/valley." For the first time in our discussion, we observe other domain than "city" or "province/region" as most lexically entrenched. "Personal names" are primarily characterized by the sub-category of "first name." CRACn2 for Scottish-Gaelic proper names and phrases is 13, whereas for the highly underrepresented group of common words, the value of CRACn2 amounts to highly marginal 2.

7.1.25. Turkish

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.571	4
Personal name	0.285	2
Non-personal name	0.142	1
Total: 7		

Place names: ***Ankara***; *Dalaman*; *Istanbul*; ***Smyrna***.

Personal names: *Ataturk*; ***Mehemet***.

Non-personal names: *Doner*.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.500	2
City and Province/Region	0.500	2
Total: 4		

PERSONAL NAMES

Type	Occurrence	Lemmas
Politician	0.500	1
First name	0.500	1
Total: 2		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Food	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 51)	Lemmas
Place names	1.00	2
Other	0	0
Total: 2		

COMMON WORDS AND PHRASES

Dolma (type of dish) is the only Turkish common noun listed in LPD (2004 edition), so further specifications of the kind presented above seem not to be justified. As far as Turkish proper names are concerned we conclude that the category of "place names" prevails over "personal names" and "non-personal names." The most salient field within the category of "place names" is "city." The number of lemmas pertinent to the two other categories is too small to mark any tendencies. CRACn2 value for Turkish proper names is 51 and, under these constraints, we observe the dominance of the category "place names."

7.1.26. Serbo-Croatian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place names only	0.857	6
Miscellaneous	0.142	1
Total: 7		

Place names: *Herzegovina; **Pristina**; Sarajevo; **Tuzla**; **Vukovar**; Zagreb.*

Miscellaneous: *Mohorovicic.*

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.500	3

City and Province/Region	0.300	2
Province/Region	0.200	1
Total: 6		

MISCELLANEOUS

Type	Occurrence	Lemmas
Personal names and non-personal names	0.200	1
Other	0	0
Total: 1 (6)		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 67)	Lemmas
Place names	1.00	2
Other	0	0
Total: 2		

Summing up the statistics of the above-presented Serbo-Croatian, we observe the prevalence of the subcategory of "place names" within the general category of proper names and phrases. Other subcategories are underrepresented, so the discussion of tendencies is pointless. CRACn2 value for Serbo-Croatian proper names is 67, and given this prerequisite, we notice the predominance of the category "place names."

Serbo-Croatian itself is an interesting 'planetoid' because it is no longer existent as a self-contained entity on the contemporary world map of languages. This is due to political changes that occurred in the Balkans during the 1990s. Still, in LPD (1995 edition), this 'planetoid' is recorded, which is evidenced by the use of Serbo-Croatian language marker. Therefore, the present book also takes note of this fact. However, since the current discussion of foreign lexical input relates to the turn of the 21st century, a new 'updated' picture is also presented. According to this, Serbian and Croatian are recognized officially as separate languages, the fact that is naturally reflected in the forthcoming study (see sections 7.1.33 and 7.1.36, respectively).

7.1.27. Zulu

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name	0.500	2
Place name	0.250	1
Non-personal name	0.250	1
Total: 4		

Personal names: ***Buthelezi***; *Cetshwayo*.

Place names: ***Hluhluwe***.

Non-personal names: ***Inkatha***.

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Politician	1.00	2
Other	0	0
Total: 2		

PLACE NAMES

Type	Occurrence	Lemmas
Province/Region	1.00	1
Other	0	0
Total: 1		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Name of political party	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 48)	Lemmas
Personal name	0.500	1
Non-personal name	0.500	1
Total: 2		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	0.333	1
Abstract	0.666	2
Total: 3		

Thing: *impala*.

Abstract: *impi; indaba*.

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 5)	Lemmas
Person	0	0
Thing	1.00	1
Abstract	0	0
Total: 1		

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Biology and animals	0.333	1
Lifestyle/Social life	0.333	1
Military/Warfare	0.333	1
Total: 3		

Zulu proper names are not characterized by some clearly marked tendency of preference as far as their subcategories are concerned. Worth noting is the subcategory of "place names," in which the dominance of "politician" is observed. CRACn2 value for proper names amounts to 48, whereas in the case of even less numerous common words, the CRACn2 value is calculated at only 5. Due to insufficient lexical input, tendencies of preference can hardly be delineated.

7.1.28. Finnish

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Place name only	0.600	3
Personal name only	0.200	1
Non-personal name only	0.200	1
Total : 5		

Place names: *Helsinki; Lahti; Turku*.

Personal names: *Aalto*.

Non-personal names: ***Kalevala***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	1.00	3
Other	0	0
Total: 3		

PERSONAL NAMES

Type	Occurrence	Lemmas
Architect	1.00	1
Other	0	0
Total: 1		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Epic poem	1.00	1

Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 41)	Lemmas
Place name	1.00	1
Personal name	0	0
Non-personal name	0	0
Total: 1		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	1.00	1
Abstract	0	0
Total: 1		

Thing: *sauna*.

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 195)	Lemmas
Person	0	0
Thing	1.00	1
Abstract	0	0
Total: 1		

COMMON WORDS AND PHRASES – FURTHER SPECIFICATION

Type	Occurrence	Lemmas
Locations and buildings	1.00	1
Other	0	0
Total: 1		

In the case of Finnish proper names, we notice the preponderance of "place names." More specifically, the most salient field within "place names" is "city."

CRACn2 value for Finnish proper names is 41 and it is only one lemma that fulfills these constraints. Finnish common words are instantiated by only one, however, quite popular lemma. This means that CRACn1 value (195) becomes also CRACn2 for this category.

7.1.29. Hebrew

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Non-personal name only	0.666	3
Personal name only	0.200	1
Miscellaneous	0.200	1
Total: 5		

Non-personal names: *Chanukah; Hanukah; Torah.*

Personal names: *Chaim.*

Miscellaneous: *Eretz.*

Type	Occurrence	Lemmas
Place and non-personal name	0.200	1
Total: 1 (5)		

SPECIFICATION

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Jewish type of festival	0.333	1
Religious book	0.333	1
Miscellaneous	0.333	1
Total: 3		

PERSONAL NAMES

Type	Occurrence	Lemmas
First name	1.00	1
Other	0	0
Total: 1		

PLACE NAMES

Type	Occurrence	Lemmas
Country	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 19)	Lemmas
Personal name only	0.500	1
Non-personal name only	0.500	1
Place name only	0	0
Total: 2		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	1.00	1
Abstract	0	0
Total: 1		

Thing: *tallith*.

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

No frequency statistics is indicated since *tallith* is not represented in the 2000 world edition of BNC.

COMMON WORDS – FURTHER SPECIFICATION

Type	Occurrence	Lemmas
Fashion /Clothes	1.000	1
Other	0	0

In conclusion, Hebrew proper names are marked with the prevalence of the subcategory of "non-personal names." However, there is not any clearly delineated

tendency observed, as regards the salience of particular fields pertinent to that domain. CRACn2 value for Hebrew proper names is estimated at 19, whereas no such value can be calculated for common words and phrases.

7.1.30. Romanian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.833	5
Non-personal name	0.166	1
Total: 6		

Place names: *Cluj; Constanta; Oradea; Ploesti; Timisoara.*

Non-personal names: *Vlad.*

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.400	2
Province/Region	0.400	2
City and Province/Region	0.400	1
Total: 5		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Dynasty	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 15)	Lemmas
Place name	1.00	2

Miscellaneous	0	0
Total: 2		

In the case of Romanian proper names, we observe the dominance of the category "place names." Within "place names" we notice an equal contribution in respect of the degree of salience of the category "city" and "province/region." CRACn2 value for this category is 15. There are no common words and phrases listed for this language.

7.1.31. Catalan

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.600	3
Personal name only	0.400	2
Total: 5		

Place names: *Andorra; Lloret;* ***Montserrat***.

Personal names: ***Gaudi; Samaranch***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.333	1
Province/Region	0.333	1
Country	0.333	1
Total: 3		

PERSONAL NAMES

Type	Occurrence	Lemmas
Architect	0.500	1
Olympic games activist	0.500	1
Total: 2		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 22)	Lemmas
Place name	1.00	2
Other	0	0
Total: 2		

In the case of Catalan proper names, we notice, again, the prevalence of the category "place names." Within "place names" we do not, however, observe any preference regarding the salience of pertinent fields. CRACn2 2 value for Catalan proper names is 15.

7.1.32. Icelandic

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.800	4
Personal name only	0.200	1
Non-personal name only	0	0
Total: 5		

Place names: *Hekla; Keflavik; Surtsey; Reykjavik*.

Personal names: ***Bjork***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.500	2
Province/Region	0.250	1
Mountain/Valley	0.250	1
Total: 4		

PERSONAL NAMES

Type	Occurrence	Lemmas
Artist	1.00	1

Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 13)	Lemmas
Place name	1.00	1
Other	0	0
Total: 1		

Icelandic proper names are characterized by the preponderance of "place names" over the remaining two subcategories. It is again "city" that is assigned the status of the most salient field within "place names." CRACn2 value for Icelandic proper names is calculated at 13.

7.1.33. Serbian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name	0.600	3
Place name	0.400	2
Total: 5		

Personal names: ***Karadzic; Milosevic; Radovan***.

Place names: ***Pale; Subotica***.

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Politician	0.666	2
Miscellaneous	0.333	1
Total: 3		

PLACE NAMES

Type	Occurrence	Lemmas
Province/Region	0.500	1
City and Province/Region	0.500	1
Total: 2		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACN2 above 36)	Lemmas
Personal name	1.00	3
Other	0	0
Total: 3		

Serbian proper names, contrary to the above-discussed Icelandic, exhibit the prevalence of "personal names" over "place names." Within the subcategory of "personal names," we note the field of "politician" as most salient. This is maybe due to the reasons presented in section 7.1.26. CRACn2 value for Serbian proper names is 36 and the absolutely dominant subcategory here is again that of "personal names."

7.1.34. Xhosa

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.500	2
Place name only	0.250	1
Miscellaneous	0.250	1
Total: 4		

Personal names: *Mandela;* ***Thabo***.

Place names: ***Soweto***.

Miscellaneous: ***Xhosa***.

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Politician	0.666	2
Miscellaneous	0.333	1
Total: 3		

PLACE NAMES

Type	Occurrence	Lemmas
Province/Region	1.00	1
Total: 1		

MISCELLANEOUS

Type	Occurrence	Lemmas
Personal name and non-personal name	1.00	1
Other	0	0

PROPER NAMES AND PHRASES AND CRAC – FURTHER SPECIFICATION

Type	Occurrence (CRACn2 above 114)	Lemmas
Personal name	1.00	1
Other	0	0
Total: 1		

In the case of Xhosa proper names, we observe that the category "personal names" prevails over the others. Within this subcategory it is the field of "politician" that emerges as primarily activated. CRACn2 value for Xhosa proper names is relatively high and amounts to 114.

7.1.35. Albanian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Personal name only	0.666	2
Place name only	0.333	1
Total: 3		

Personal names: ***Hoxha; Zog***.

Place names: *Tirana*.

SPECIFICATION

PERSONAL NAMES

Type	Occurrence	Lemmas
Politician	1.00	2
Other	0	0
Total: 2		

PLACE NAMES

Type	Occurrence	Lemmas
City	1.00	1
Total: 1		

FURTHER SPECIFICATION

Type	Occurrence (CRACn2 above 38)	Lemmas
Place name	1.00	1
Other	0	0
Total: 1		

Analogically to Xhosa, Albanian proper names are characterized by the prevalence of the category "personal names." Within this subcategory it is again the field of "politician" that can be described as most salient. CRACn2 value for Albanian proper names is, however, much lower than that of Xhosa, and approximates 38.

7.1.36. Croatian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.666	2
Personal name	0.333	1
Total: 3		

Place names: ***Osijek; Rijeka***.

Personal names: ***Stepinac***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	1.00	2
Total: 2	0	0

PERSONAL NAMES

Type	Occurrence	Lemmas
Priest	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 42)	Lemmas
Place name	1.00	2
Other	0	0
Total: 2		

As far as Croatian proper names are concerned, we observe the statistical prevalence of the subcategory "place names." Within this subcategory we note what may be called 'traditional' from the perspective of the hitherto conducted analysis, i.e. the salience of "city." CRACn2 value for Croatian proper names is 42, and it is again the subcategory of "place names" that absolutely dominates.

7.1.37. Malay

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	1.00	2
Other	0	0
Total: 2		

Place names: *Perak;* ***Perlis***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
Province/Region	1.00	2
Total: 2		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 1)	Lemmas
Place name	1.00	2
Other	0	0
Total: 2		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	1.00	1
Abstract	0	0
Total: 1		

Thing: ***sambal***.

COMMON WORDS – FURTHER SPECIFICATION

Type	Occurrence	Lemmas
Food and drink	1.0	1
Total: 1		

What can be concluded about Malay, is the salience of the subcategory "place names" within "proper names." Other generalizations are risky due to insufficient lexical input from that language. For statistical consistence, we note that CRACn2 value for Malay proper names is infinitesimal and amounts to 1.

7.1.38. Punjabi

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Person	0	0
Thing	1.00	3
Abstract	0	0
Total: 3		

Thing: *Kara; kirpan; kuccha.*

COMMON WORDS – FURTHER SPECIFICATION

Domain	Occurrence	Lemmas
Locations/Buildings	0.333	1
Materials/Objects	0.333	1
Military/Warfare	0.333	1
Total: 3		

Punjabi foreign words in English are somewhat unique in that there are no proper names and phrases recorded for that language. In the category of "common words and phrases," we observe the absolute prevalence of subcategory of "thing." However, when we look at the semantic specification of this subcategory, there is no clearly marked tendency manifested as regards the salience of pertinent fields. No CRACn2 value is calculated for Malay common words since all of the identified lemmas are not represented in the BNC.

7.1.39. Bulgarian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	1.00	2
Other	0	0
Total: 2		

Place names: ***Plovdiv; Sofia.***

SPECIFICATION

Type	Occurrence	Lemmas
City	0.500	1
City and Province/Region	0.500	1
Total: 2		

FURTHER SPECIFICATION

PROPER NAMES AND CRAC

Type	Occurrence (CRACn2 above 65)	Lemmas
Place name only	1.00	1
Other	0	0
Total: 1		

Bulgarian proper names are characterized by the prevalence of the subcategory of "place names." CRACn2 value for proper name is 65. Otherwise, no other marked tendencies can be noted.

7.1.40. Ndebele

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.500	1
Personal name	0.500	1
Other	0	0
Total: 2		

Place names: *Bulawayo*.

Personal names: ***Nkomo***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City and Province/Region	1.00	1
Total: 1		

PERSONAL NAMES

Type	Occurrence	Lemmas
Politician	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 26)	Lemmas
Place name	1.00	1
Other	0	0
Total: 1		

Not much can be said about Ndebele proper names. Hardly any generalization can be made about them due to insufficient lexical input. The only fact that can be noted about that language relates to CRACn2 value which amounts to 26.

7.1.41. Persian

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	0.500	1
Non-personal name	0.500	1
Total: 2		

Place names: *Qom.*

Non-personal names: *Bahai.*

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City or River	1.00	1
Other	0	0
Total: 1		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Name of religion	1.00	1
Other	0	0
Total: 1		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 8)	Lemmas
Place name	1.00	1
Other	0	0
Total: 1		

Similarly to Ndebele, hardly anything can be conclusively asserted about Persian proper names. Hardly any generalization can be made about them due to insufficient lexical input. The only fact that can be noted about that language relates to CRACn2 value which amounts to 26.

7.1.42. Vietnamese

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name	1.00	2
Other	0	0
Total: 2		

Place names: ***Haiphong; Hanoi***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	1.00	2
Total: 2		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 58)	Lemmas
Place name	1.00	1
Other	0	0
Total: 1		

In the case of Vietnamese proper names, we observe an absolute dominance of the subcategory of "place names." Within this subcategory, it is the field of "city" that emerges as most salient. CRACn2 value for proper names in general amounts to 58.

7.1.43. Other planetoids

This section presents donor languages, whose lexical input for the English language can be called infinitesimal. Due to this fact, this section is confined to mere presentation of some basic statistical facts connected with this heterogeneous group of languages and, therefore, no generalizations are made at the end of the section as the contribution from each language in respect of either the category "proper names" or "common words" does not go beyond two lemmas. The discussion of this group of languages is given a detailed treatment in Kuźniak (2008: 423–438). As regards the category of "proper names" the lexical input from following languages is analyzed: Akan; Bengali; Esperanto; Estonian; Farsi; Ibo; Khmer; Latin; Latvian; Lithuanian; Macedonian; Maori; Slovak; Slovene; Swahili; Thai; Tswana; Twi; Urdu; Yoruba. In the case of the category of "common words," Maori and Sanskrit are investigated.

PROPER NAMES AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Place name only	0.600	12
Personal name only	0.150	3
Non-personal name only	0.100	2
Miscellaneous	0.150	3
Total 20		

Place names: *Tallin; **Kaunas**; Bratislava; Mombasa; Botswana; Lahore; Ibadan; **Phuket**; **Ljubljana**; **Skopje**; **Riga**; **Phnom Penh**.*

Personal names: *Tagore; Caesar;* ***Rafsanjani***.

Non-personal names: *Twi; Esperanto*.

Miscellaneous: ***Akan; Ibo; Maori***.

SPECIFICATION

PLACE NAMES

Type	Occurrence	Lemmas
City	0.733	11
Province/Region	0.200	3
Country	0.066	1
Total: 15		

PERSONAL NAMES

Type	Occurrence	Lemmas
Miscellaneous	0.666	4
Political leader	0.333	2
Total: 6		

NON-PERSONAL NAMES

Type	Occurrence	Lemmas
Language	0.800	4
Miscellaneous	0.200	1
Total: 5		

MISCELLANEOUS

Type	Occurrence	Lemmas
Place name and personal name and non-personal name	0.090	2
Personal name and non-personal name	0.045	1
Total: 3 (21)		

FURTHER SPECIFICATION

PROPER NAMES AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 57)	Lemmas
Place name	0.625	5
Personal name	0.375	3
Total: 8		

COMMON WORDS AND PHRASES

GENERAL FREQUENCY COUNT AND WORD LISTS

Type	Occurrence	Lemmas
Thing	0.500	1
Abstract	0.500	1
Total: 2		

Thing: *ngaio*.

Abstract: *om*.

SPECIFICATION

COMMON WORDS AND PHRASES AND CRAC

Type	Occurrence (CRACn2 above 9)	Lemmas
Thing	1.00	1
Other	0	0
Total: 1		

PART IV
GETTING TO THE GIST

Overview

PART IV is the last large section of the book. It consolidates all the findings from previous parts into one coherent, unified descriptive framework. Chapter 8 constitutes the summary of facts established in Chapter 7 and offers some new insights that emerge from the adoption of broader perspective in the study of foreign lexical influence upon the English language. In addition, concepts of mass and force which were introduced in Chapter 6 are once more taken into consideration and elaborated on in view of the results obtained in Chapter 7. Chapter 8 concludes with word recognition test, the role of which is to verify, via the survey, some major tendencies of foreign lexical structuring discussed throughout the chapter. Finally, Chapters 9 and 10 provide the detailed characterization to the central claim laid out in section 0.2. The main concern of Chapter 9 is, thus, to postulate a tentative model of foreign lexical assimilation as well as presentation of a unified framework of the conception of foreign word as a lexical category. Chapter 10, on the other hand, relates the findings discussed in the book to some further potential avenues of research, such as correlations between the presented model of foreign lexical assimilation and the Rutherford-Bohr's model of atom, or the increasingly popular in IT science, the concept of scale-free networks.

Chapter 8
PLANETS, PLANETOIDS AND THE EARTH

8.0. Tendencies in structuring of planetoid–planetary mass

8.1. Mass and volume revisited

In section 6.1, the entire mass of a planet was linguistically correlated with the following formula:

$$M_1 = (\text{Ln} \times \text{CRACn2}\ [y]),$$
$$M_2 = (\text{Ln} \times \text{CRACn2}\ [y]),$$
$$M = M_1 + M_2,$$

where M_1 relates to the value obtained on the basis of the formula that reads: the number of proper names or phrases (Ln) times (×) CRACn2 value (y) calculated for the relevant set. It is to be reminded that CRACn2 value is the averaged set value calculated on the basis of CRACn1 values of a set of lemmas pertinent to a particular category, whether "proper names" or "common words." M_2, on the other hand, relates to the value obtained on the basis of the formula that reads: the number of common words or phrases (Ln) times (×) CRACn2 value (y) for the relevant set. M stands for the sum total of M_1 and M_2. The significance in calculating M of a given donor language is underlain by the following hypothesis:

The greater the mass, the greater the chances there are for a set of lemmas pertinent to that language to become adapted into the target lexical system. Therefore the greater the mass, the greater the chances that a given donor language plays more central role in structuring the target lexical system.

Thus, M value may function as a valid indicator of status assignment of a given donor language as planet (central) or planetoid (peripheral) in structuring the lexical system of the target language. The hypothesis, albeit tentative one, is in accordance with scientific facts discussed in physics, especially the ones concerning force, energy, and motion (see sections 6.2–6.3 and 6.5).

Even a more significant criterion from the cognitive-linguistic viewpoint is the criterion of volume (V) in assigning the status of planet or planetoid to a particular donor language. V is the value obtained as the sum total of lemmas ("proper" and "common") pertinent to a given donor language. The resulting formula, thus, reads as follows:

$$V = V_1 + V_2,$$

where V_1 stands for the total of proper names and phrases of a given language, and V_2 stands for the total of common words and phrases belonging to a given donor language.

Taking V as fundamental criterion adopted in order to designate which language functions as a 'planet' and which language functions as a 'planetoid' is debatable. This criterion is, however, tentatively selected as default in assigning the status of a language on two major grounds. Firstly, it essentially correlates with the following regularity:

The greater the volume of a given language, the greater the mass of a planet, and hence the importance of a donor lexical system to the target language.

Secondly, it is consonant with the folk view of magnitude of entities according to which MORE IS UP (see Krzeszowski 1997) and, therefore, BIGGER entities are seen as MORE important. This leads us to formulating the following naïve-view of the world hypothesis about lexical assimilation processes:

The more words a given donor language 'lends' to a target lexical system, the more important it is to that system.

However, BIGNESS does not directly need to designate HEAVINESS of an object, hence the fundamental distinction between volume, mass, and weight discussed in scientific physics (see section 6.1). If the above-presented naïve view were to hold true, it would stand in direct opposition to the Fourth Newton's Law of Dynamics, where $F = mg$ (force of gravity is a resultant of mass and gravitational constant), that is the heavier the object, the greater the force of gravity. The linguistic correlation of this physical law is:

The heavier a given language lexical system is (the greater the mass it possesses), the more assimilatory potential this language has in the target lexical system.

What follows is that V may serve as a convenient tool of organizing donor languages in order of their importance (and this is actually done in the book), however since it does not need to correlate with M, it cannot by any means function as the only criterion of evaluation of the role a given language plays in the process of lexical assimilation in respect of the target system. Tab. 13 presents V and M values calculated for donor languages of the English language.

Tab. 13. *V and M values of donor languages*

Language	V_1	V_2	V	M_1	M_2	M
French	547	598	**1146**	28991	19734	**48725**
German	414	91	**505**	25668	1547	**27215**
Italian	210	109	**319**	9450	1853	**11303**

Spanish	215	65	**280**	16125	1690	**17815**
Welsh	224	7	**231**	4704	63	**4767**
Russian	148	15	**163**	8140	420	**8560**
Chinese	109	8	**117**	4469	352	**4821**
Hindi	47	38	**85**	1598	608	**2206**
Dutch	63	0	**63**	4410	0	**4410**
Japanese	35	28	**63**	2870	532	**3402**
Arabic	42	6	**48**	7854	456	**8310**
Portuguese	39	5	**44**	1248	30	**1278**
Danish	21	3	**24**	609	42	**651**
Greek	21	11	**32**	357	165	**522**
Swedish	21	4	**25**	1050	12	**1062**
Polish	20	4	**24**	1260	40	**1300**
Hungarian	19	3	**22**	532	18	**550**
Irish	20	2	**22**	1600	80	**1680**
Norwegian	16	0	**16**	608	0	**608**
Czech	13	0	**13**	91	0	**91**
Afrikaans	9	3	**12**	549	90	**639**
Korean	10	2	**12**	430	40	**470**
Cantonese	6	2	**8**	100	66	**166**
Scottish-Gaelic	6	2	**8**	78	4	**82**
Turkish	7	1	**8**	357	2	**359**
Serbo-Croatian	7	0	**7**	469	0	**469**
Zulu	4	3	**7**	192	15	**207**
Finnish	5	1	**6**	205	195	**400**
Hebrew	5	1	**6**	95	0	**95**
Romanian	6	0	**6**	90	0	**90**
Catalan	5	0	**5**	110	0	**110**
Icelandic	5	0	**5**	65	0	**65**
Serbian	5	0	**5**	180	0	**180**
Xhosa	4	0	**4**	456	0	**456**
Albanian	3	0	**3**	114	0	**114**
Croatian	3	0	**3**	126	0	**126**
Malay	2	1	**3**	2	1	**3**
Punjabi	0	3	**3**	0	0	**0**
Bulgarian	2	0	**2**	130	0	**130**
Maori	1	1	**2**	94.5	14.5	**109**
Ndebele	2	0	**2**	52	0	**52**

Tab. 13 – cont.

Persian	2	0	**2**	16	0	**16**
Vietnamese	2	0	**2**	116	0	**116**
Akan	1	0	**1**	1.5	0	**1.5**
Bengali	1	0	**1**	21.5	0	**21.5**
Esperanto	1	0	**1**	25.5	0	**25.5**
Estonian	1	0	**1**	3	0	**3**
Farsi	1	0	**1**	110	0	**110**
Ibo	1	0	**1**	11	0	**11**
Khmer	1	0	**1**	122	0	**122**
Latin	1	0	**1**	305.5	0	**305.5**
Latvian	1	0	**1**	73.5	0	**73.5**
Lithuanian	1	0	**1**	4	0	**4**
Macedonian	1	0	**1**	12.5	0	**12.5**
Sanskrit	0	1	**1**	0	4	**4**
Slovak	1	0	**1**	62	0	**62**
Slovene	1	0	**1**	24.5	0	**24.5**
Swahili	1	0	**1**	37.5	0	**37.5**
Thai	1	0	**1**	20	0	**20**
Tswana	1	0	**1**	151	0	**151**
Twi	1	0	**1**	0	0	**0**
Urdu	1	0	**1**	61	0	**61**
Yoruba	1	0	**1**	12	0	**12**

The eight planets that, thus, can be identified on the basis of *V* criterion are the following: French, German, Italian, Spanish, Welsh, Russian, Chinese and Hindi.

The striking fact is the difference in *M* between the top eight languages: the range of *M* magnitude varies from nearly 2200 to 49,000. The two leading languages (French and German) are, however, more orderly in terms of *M*/*V* ratio. Namely, the biggest language in terms of *V,* i.e. French is also the heaviest one in terms of *M*. This also applies relative to German, which is roughly twice smaller than French, and thus, consistently possessing half the *M* value of French. This regularity does no longer apply to languages with much lower *V* values, e.g. Welsh and Russian, in which case the higher *V* value of Welsh in relation to Russian does not entail the greater *M* value. We can, therefore, draw the conclusion that although Welsh is *perceived* as bigger than Russian, its lexical assimilatory potential (*M*) is *actually* half as big as Russian's. We, thus, may expect Russian lexical items to become more permanently assimilated into the target lexical system of English than Welsh words.

Our remarks are based on the following formulae:

1) Statement (a) *X is perceived as Y* is qualitatively different from statement (b) *X is Y* in the scientific view of the world.

2) Statement (a) *X is perceived as Y* is qualitatively isomorphic with statement (b) *X is Y* in the folk view of the world.

As section 9.2 shows, the two aforementioned views of the world, i.e. folk and naïve, offer parallel and complementary frameworks for the conception of foreign lexicalisation processes. Hence both formulae 1) and 2) are equally valid for the considerations over the nature and mechanisms of lexical assimilation processes that we attempt to model in the book, yet for two fundamentally different reasons.

Formula 1) is taken into serious consideration for the reason of our trying to explicate which donor languages bear actual lexical assimilatory significance to the English language, whereas formula 2) is taken into account for the reason of compliance with the nonetheless weighty principle in cognitive-linguistic studies, whereby it is the human conceptualisation (perception) of things that matters rather than their actual status (see Lakoff and Johnson's [1980] notion of experientialism). On account of formula 2), then, the classification of languages into more significant (planet-like) and less significant (planetoid-like) appears to be validated.

In accordance with the principle of natural categorisation (Rosch 1977), and the resulting division of the conceptual world into three levels: superordinate, basic, subordinate, we may consistently divide the donor languages into three main categories: planets, major planetoids, and minor planetoids, given *V* criterion into consideration. The following Tab. 14 illustrates this tentative division:

Tab. 14. *Status of languages in respect of V criterion*

Status	Language	*V*
PLANETS perceived as 8 biggest celestial bodies in the solar system revolving round the ninth planet, i.e. the Earth	French	1145
	German	505
	Italian	319
	Spanish	280
	Welsh	231
	Russian	163
	Chinese	117
	Hindi	85
MAJOR PLANETOIDS making up >1% of the global lexical assimilation incidence on the planet Earth.	Dutch	63
	Japanese	63
	Arabic	48
	Portuguese	44

Tab. 14 – cont.

MINOR PLANETOIDS <1% of the global lexical assimilation incidence on the planet Earth (English)	Greek	32
	Swedish	25
	Danish	24
	Polish	24
	Hungarian	22
	Irish	22
	Norwegian	16
	Czech	13
	Afrikaans	12
	Korean	12
	Cantonese	8
	Scottish-Gaelic	8
	Turkish	8
	Serbo-Croatian	7
	Zulu	7
	Finnish	6
	Hebrew	6
	Romanian	6
	Catalan	5
	Icelandic	5
	Serbian	5
	Xhosa	4
	Albanian	3
	Croatian	3
	Malay	3
	Punjabi	3
	Bulgarian	2
	Maori	2
	Ndebele	2
	Persian	2
	Vietnamese	2
	Akan	1
	Bengali	1
	Esperanto	1
	Estonian	1
	Farsi	1
	Ibo	1

MINOR PLANETOIDS <1% of the global lexical assimilation incidence on the planet Earth (English)	Khmer	1
	Latin	1
	Latvian	1
	Lithuanian	1
	Macedonian	1
	Sanskrit	1
	Slovak	1
	Slovene	1
	Swahili	1
	Thai	1
	Tswana	1
	Twi	1
	Urdu	1
	Yoruba	1

8.2. Characterization of tendencies in mass structuring of planets and selected planetoids

Taking the criterion of *V* into account, the following regularities in the internal composition of donor lexical systems can be observed. In the case of characterization of minor planetoids, only a selected group of them is presented on the basis of order of alphabet principle.

8.2.1. Planets

French, German, Italian, Spanish, Welsh, Russian, Chinese, Hindi

8.2.1.1. Characterization of tendencies in proper names. Personal names

First, one notices that out of 17 semantic fields relative to which the subcategory of foreign personal names was analysed, only 9 fields appear as dominant in the lexical systems of 8 major donor languages (planets). Out of 9 fields, 6 appear more than once with the category "miscellaneous" as occurring 7 times out of 8. The category "miscellaneous" is the most outstanding one as it is the most copious of all, comprising not only individual instances of names of professions but also indicative of inter- and intracategory polysemy. The predominant fields or elements constituting the mass structure of the identified planets are subsumed under Fig. 36:

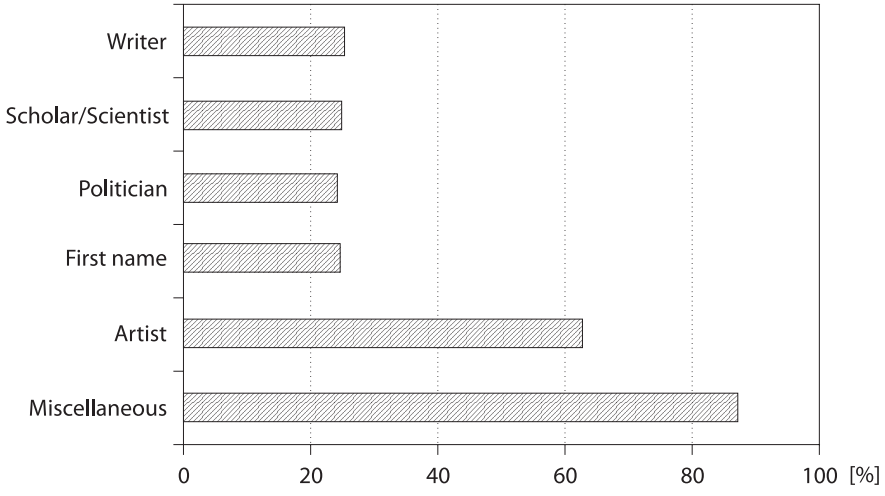


Fig. 36. *Dominant semantic fields: Personal names – planets*

8.2.1.2. Characterization of tendencies in proper names. Place names

In the case of foreign proper place names, there is also noticeable uniformity as far as dominant fields structuring the lexical system of planet-like donor languages. Out of 24 identified semantic fields, only 6 appear as recurring in the internal composition of the discussed source lexical systems. These are the following:

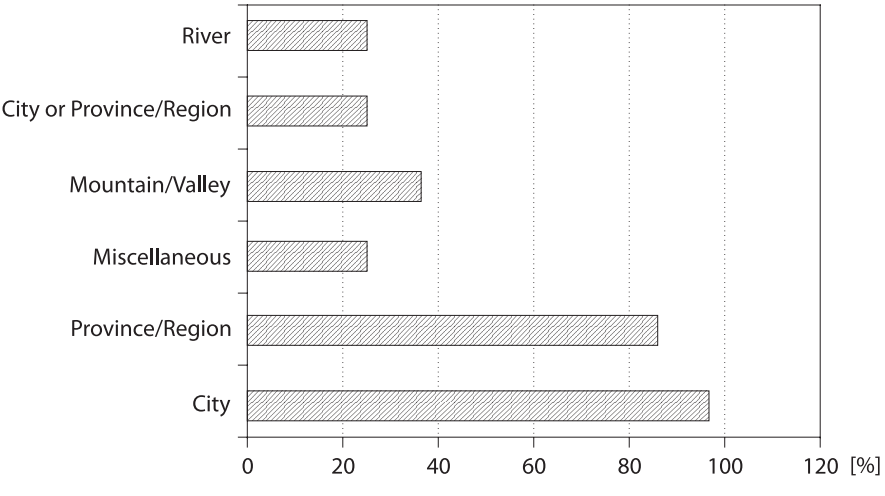


Fig. 37. *Dominant semantic fields: Place names – planets*

As it turns out the "city" component is an inevitable element in the composition of the place name section in lexical systems of major donor languages. Again,

of interest are the fields inherently polysemous in nature. Since the phenomenon of polysemy appears to prevail in this subcategory, the field "miscellaneous" is understood to comprise here more than two references (both inter- or intracategorially). Hence, for instance, the category "city and province/region" is treated here not as strictly polysemous despite the fact that it relates to more than one (but not more than two) references.

8.2.1.3. Characterization of tendencies in proper names. Non-personal names

The subcategory of foreign non-personal names is the most heterogeneous so far in that there is not any clearly dominating and internally homogeneous field structuring lexical systems of major donor languages. Again, there is quite a lot of uniformity in this diversity, that is, out of 33 fields identified for this subgroup, 11 categories emerge as dominant in the structures of source lexical systems of major languages, out of which 5 categories recur on more than one planet, i.e. in more than one lexical system. These are:

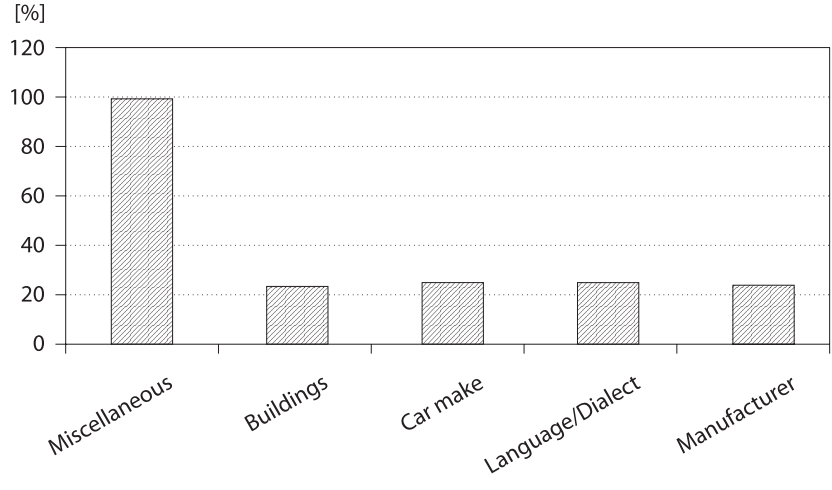


Fig. 38. *Dominant semantic fields: Non-personal names – planets*

8.2.1.4. Characterization of tendencies. Common words

Similarly to the subcategory of place names and personal names, the section of common words and phrases exhibits a high uniformity in structuring the internal composition of major donor languages. Out of 25 fields identified for this section of vocabulary only 8 emerge as dominant, out of which only 4 recur in more than one lexical systems. These are:

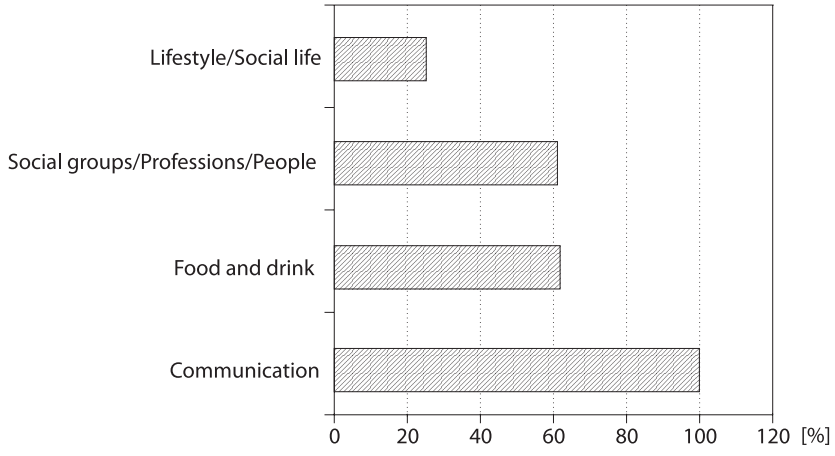


Fig. 39. *Dominant semantic fields: Common words – planets*

As one may notice, major source lexical systems are composed of items designating communicative functions, i.e. they are groups of fixed phrases or adjectives used for the description of other categories. As it turns out, this sort of foreign non-proper vocabulary constitutes the part and parcel of this section of donor lexical systems.

8.2.2. Major planetoids

Dutch, Arabic, Japanese, Portuguese

8.2.2.1. Characterization of tendencies in proper names. Personal names

As the analysis of tendencies in structurization of the "personal names" subcategory within the group of so called major planetoids shows, there is not any

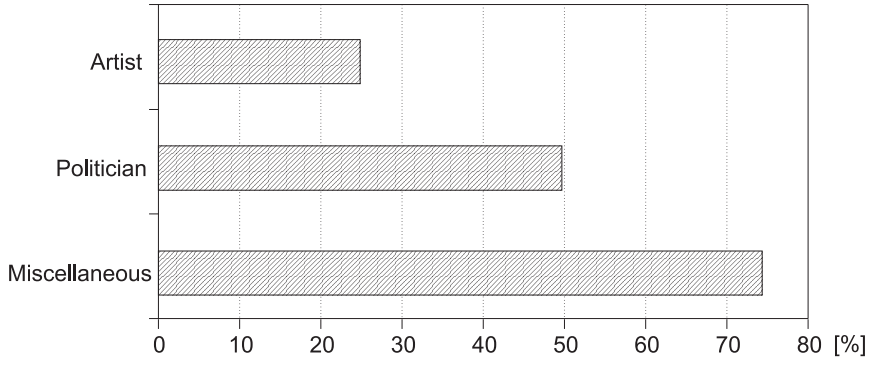


Fig. 40. *Dominant semantic fields: Personal names – major planetoids*

substantial difference in the selection of dominant fields constituting the lexical systems of these donor languages in comparison with the aforementioned planets. The dominant fields can be seen in Fig. 40.

8.2.2.2. Characterization of tendencies in proper names. Place names

Very similar results to the ones discussed in section 8.2.1.2 are obtained while characterizing the tendencies in the structurization of "place name" section of major planetoids. Here are the results:

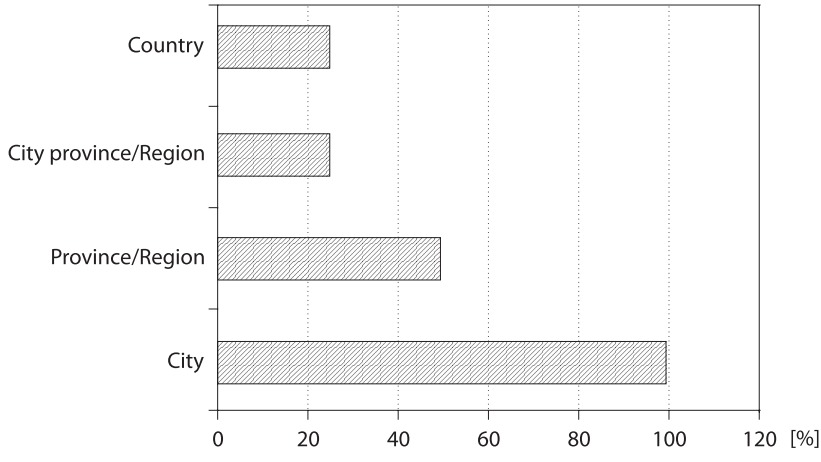


Fig. 41. *Dominant semantic fields: Place names – major planetoids*

It appears then, that "city" component of the "place name" section is the essential (prototypical) element characterizing this area of donor lexical systems. Other elements identified above also recur on other planets or planetoids, which indicates unequivocally that the most significant elements shaping the substantial part of mass structure of celestial bodies (donor lexical systems) influencing the Earth (English) are composed of the same matter.

8.2.2.3. Characterization of tendencies in proper names. Non-personal names

The subcategory of foreign non-personal names within the set of major planetoids displays exactly the same tendencies as in the case of the above-discussed planets. There is quite a lot of diversity in uniformity observed. The major planetoids possess relatively smaller masses in comparison to planets, which makes the characterization of tendencies more tentative. We may, for once, observe 'empty slots' in structuring of some languages, e.g. Japanese. We, thus,

predict that the smaller the masses and volumes of planetoids, the higher the 'risk' of encountering 'areas of emptiness' in a given subsection of the lexical system. Still, however, we identify here most of elements that have recurred in the mass structure of planets, i.e. "miscellaneous," "buildings," and "manufacturing company." Other categories such as "name of religion," "racing track" or "wine" are more incidental than illustrative of any tendencies within this section of vocabulary. The reason for this is relatively small portion of data on the basis of which the elicitation of categories was conducted. A good example is Portuguese, where categories of "wine" or "racing track" are seen as 'dominant' on account of being the only categories that the lexical system of this language features.

Therefore, Fig. 42 below shows only the recurrent category and its representation in the mass structure of major planetoids.

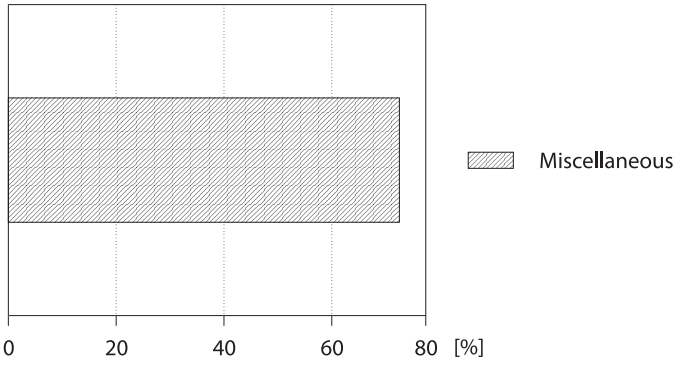


Fig. 42. *Dominant semantic fields: Non-personal name – major planetoids*

8.2.2.4. Characterization of tendencies. Common words

The "common" lexical stock pertinent to the systems of major planetoids is not worth diagraming below for the reason of not having a single dominant field recurring on any other major planetoids. Some important observation concerns the reiteration of elements characterizing the mass structure of planets with the mass structure of planetoids. These are "communication" and "food/drink" fields. Interesting is the lack of fields characterizing social life or those directly referring to humans. We, thus, predict that the tendency of 'deanthropomorphization' within the dominant fields will continue with diminishing importance of donor lexical systems influencing English. Similarly to the subsection of non-personal names discussed above, empty categories also tend to occur, e.g. in Dutch.

8.2.3. Selected minor planetoids

Afrikaans, Albanian, Bulgarian, Cantonese, Catalan, Czech, Croatian, Danish, Finnish, Greek

8.2.3.1. Characterization of tendencies in proper names. Personal names

After the analysis of the selection of 10 peripheral lexical systems within the domain of personal names, the following observations can be made: firstly, there is a remarkable growth in the number of 'empty slots' on the observed planetoids. That is, out of 20 possible spaces to be occupied (10 planetoids have, by and large, 2 spaces reserved to designate dominant fields within a given subcategory), 6 spaces were left 'non-accommodated.' However, the elements that do appear as structuring the mass of these minor planetoids are basically the same as in the case of major planetoids or planets. These are the following:

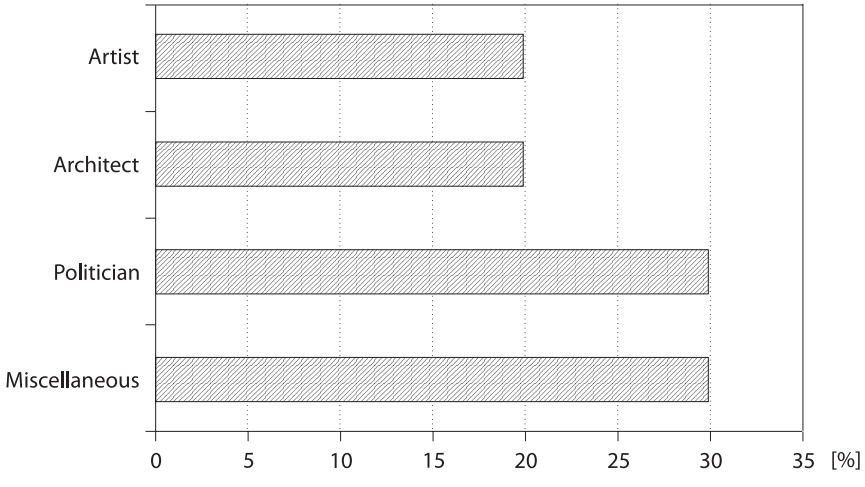


Fig. 43. *Dominant semantic fields: Personal names – minor planetoids*

8.2.3.2. Characterization of tendencies in proper names. Place names

No substantial change may be observed in the characterization of mass structure of minor planetoids in comparison to major planetoids or planets within the category of "place names." Analogically to the domain of "personal names," there is a significant number of 'empty slots' (5 out of 20). Again, the dominating field is that of "city" with other elements "province/region" recurring also on other minor planetoids. Even the elements that do not reappear within the discussed set have

already been identified as dominant on the so-called major planetoids or planets discussed above. These fields involve "city and province/region," "country," and "river." The dominant fields are, thus, the following:

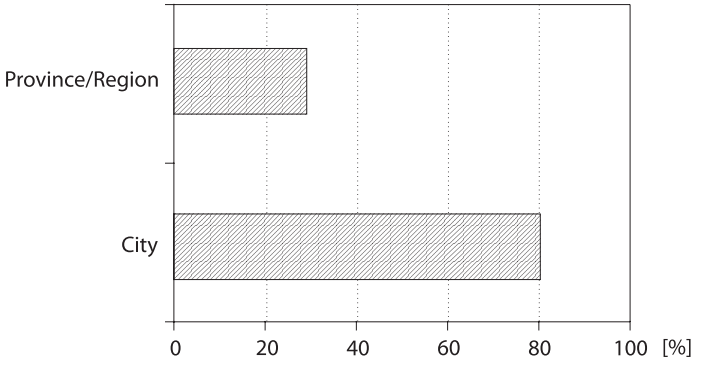


Fig. 44. *Dominant semantic fields: Personal names – minor planetoids*

8.2.3.3. Characterization of tendencies in proper names. Non-personal names

The domain of foreign non-personal names appears as the most varied of all domains regardless of the type of celestial bodies analysed (whether planets, major planetoids or the current minor planetoids). There are domains like "dialects/language," "epic poems," "name of dish," "car make" that have already been detected on the other planets or major planetoids. However, they do not constitute the dominant elements within the mass structure of minor planetoids in that they do not recur on them. Again, analogically to the aforementioned subcategories of "personal names" and "place names," this category of "non-personal names" also displays the cases of 'empty' field spaces (8 out of 20).

The recurring elements are the following:

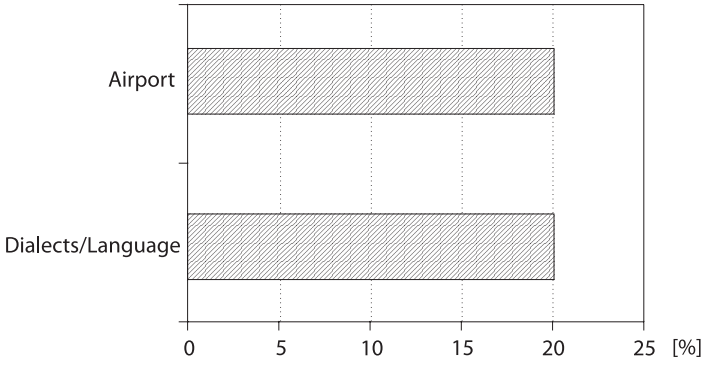


Fig. 45. *Dominant semantic fields: Non-personal names – minor planetoids*

Worth noticing is the diminishing contribution of the 'dominant' elements within the structure of minor planetoids relative to the contribution of elements structuring major planetoids or planets.

8.2.3.4. Characterization of tendencies. Common words

The first striking observation made while analyzing the domain of foreign common words within the set of minor planetoids, is the remarkable number of 'empty slots' in the spaces provided for the characterization of dominant fields. Out of 20 spaces available, there is 11 'holes' in the structure of the discussed lexical systems.

Otherwise, the same phenomenon of 'deanthropomorphisation' of foreign common words may be observed, i.e. out of 5 dominant fields identified for this domain, not a single one directly describes humans, as was the case in the analysis of a planetary structure of common words, where such fields as "social life/lifestyle," "social groups/professions/people" or "communication" were identified as prevalent. The dominant fields in this group are presented below:

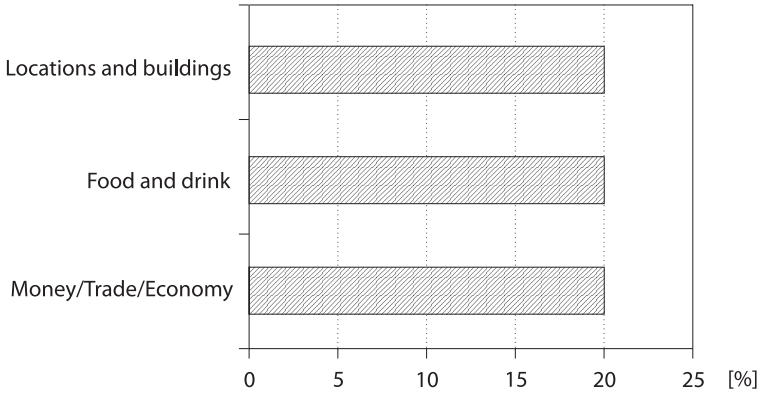


Fig. 46. *Dominant semantic fields: Common words – minor planetoids*

8.2.4. Planets and planetoids in compact view

As the title of this section suggests, we will confine ourselves to a compact presentation of the most outstanding regularities observed across planet–planetoid mass structurization.

Firstly, as far as the subcategory of "personal names" is concerned, the elements that make up the dominant portion of the mass of the discussed planets and planetoids are fundamentally composed of the same matter. Secondly, in the case of foreign proper place names, it appears that "city" component is

the essential (prototypical) element characterizing this area of donor lexical systems, regardless of the status of a donor system (i.e. whether it is planet-like or planetoid-like). Again, recurrence of elements making up the structure of planet–planetoid mass is noticeable, which leads to some broader generalization, i.e. that the most significant elements shaping the substantial part of mass structure of celestial bodies (i.e. donor lexical systems) affecting the Earth (English) are essentially composed of the same matter. Thirdly, in the case of non-personal names there may be observed a relative diversity in the selection of dominant fields structuring the planet–planetoid mass. This is paradoxically a 'unifying' feature characterizing this subcategory of donor lexical systems. Finally, in the case of foreign common words the tendency of 'deanthropomorphisation,' i.e. the relative insignificant role of fields directly descriptive either of individual human beings or groups they form, is visible to an extent that the tentative regularity (henceforth numbered for the convenience of cross-referencing) may be construed:

Regularity 1

'Deanthropomorphisation' within the dominant fields characterizing 'common' vocabulary will increase with the decreasing importance (measured in terms of *V*) of donor lexical systems influencing English.

Aside from the above-cited unifying tendencies in the structurization of particular subparts of vocabulary across major and minor donor lexical systems, some generalizations related to the overall mass structure and the volume factor can be made. We, thus, predict that the smaller the masses and volumes of a celestial body (whether planet or planetoid), the higher the 'risk' of lacking elements characterizing a given sub-section of the lexical system. Thus, another general regularity may be formulated:

Regularity 2

The number of 'holes' in the structure of lexical systems increase in direct proportion to the decreasing volume (or mass) of a given lexical system.

Another general observation is about the role of dominant fields in the structurization of planet–planetoid mass with regard to *V* parameter. Another tentative regularity may, thus, be stated as follows:

Regularity 3

The percentage of contribution of dominant elements characterizing the structure of donor lexical systems is inversely proportionate to the volume of a given lexical system.

The above observations on the noticeable uniformity of mass structurization across planet–planetoid system may lead us to think of yet another correspondence between the world of language and the world of physics. In the same way that astrophysicists explore the composition of the internal structure of other

celestial bodies, whether planetary or planetoid, in order to hypothesize about the changing nature of the universe, linguists may examine the structuring of lexical systems of other languages to learn about the history of external lexical change of a target language (see, e.g. Fisiak 2000).

Consequently, the knowledge we obtain from astrophysics and related to the way planets and planetoids behave in the universe may also account for the phenomenon of language death. As small pieces get detached from the planetoid under the forces of gravity and subsequently fall into the atmosphere of the Earth, a smaller language may find itself under "superstratal influence" from a more influential language due to geographical as well as cultural proximity and, thus, eventually disappear.[1]

8.2.4.1. Proper names and phrases

The following is the summary view of some basic statistics connected with the characterization of foreign "proper name" part of donor lexical systems characterizing the English language at the turn of the 21st century.

Tab. 15. *Proper names – general statistics*

Language	Items	Total	CRACn2	Mass	Frequency	Place	Personal name	Non-personal name	Miscellaneous[2]
French	547	2364	53	29044	0.2313	184	277	72	14
German	414	2364	62	25668	0.1751	128	201	52	33
Welsh	224	2364	21	4704	0.0947	124	66	24	10
Spanish	215	2364	75	16125	0.0909	108	83	18	6
Italian	210	2364	45	9450	0.0888	70	112	25	3
Russian	148	2364	55	8140	0.0626	61	77	6	4
Chinese	109	2364	41	4469	0.0461	74	19	4	12
Dutch	63	2364	70	4480	0.0266	29	24	5	5
Hindi	47	2364	34	1598	0.0198	19	13	12	3
Arabic	42	2364	187	7854	0.0177	23	13	5	1
Portuguese	39	2364	32	1248	0.0164	28	8	2	1

[1] The notions of substratal and superstratal influence are discussed in Dirven and Verspoor (2004: 207).

[2] See Tab. 17, where the category "miscellaneous" is more specifically understood. The category "miscellaneous" in Tab. 15 comprises, thus, the instances of intercategorial polysemy and the so-called 'hybrid' polysemy, i.e. the cases of lemmas which are polysemous both intracategorially and intercategorially (see the category "miscellaneous 1+2" in Tab. 17).

Tab. 15 – cont.

Japanese	35	2364	82	2870	0.0148	19	0	13	3
Danish	21	2364	29	609	0.0088	11	9	1	0
Greek	21	2364	17	357	0.0088	15	4	2	0
Swedish	21	2364	50	1050	0.0088	7	8	3	3
Polish	20	2364	63	1300	0.0084	10	8	0	2
Irish	20	2364	80	1600	0.0084	7	7	6	0
Hungarian	19	2364	28	532	0.0080	5	11	1	2
Norwegian	16	2364	38	608	0.0067	8	7	1	0
Czech	13	2364	7	91	0.0054	3	9	1	0
Korean	10	2364	43	430	0.0042	6	0	4	0
Afrikaans	9	2364	61	549	0.0038	1	6	2	0
Serbo-Croatian	7	2364	67	469	0.0029	6	0	0	1
Turkish	7	2364	51	357	0.0029	4	2	1	0
Cantonese	6	2364	21	126	0.0025	2	2	2	0
Romanian	6	2364	15	90	0.0025	5	1	0	0
Scottish-Gaelic	6	2364	13	78	0.0025	3	2	1	0
Catalan	5	2364	22	110	0.0021	3	2	0	0
Finnish	5	2364	41	205	0.0021	3	1	1	0
Hebrew	5	2364	19	95	0.0021	0	1	3	1
Icelandic	5	2364	13	65	0.0021	4	1	0	0
Serbian	5	2364	36	180	0.0021	2	3	0	0
Xhosa	4	2364	114	456	0.0016	1	2	0	1
Zulu	4	2364	48	192	0.0016	1	2	1	0
Albanian	3	2364	38	114	0.0012	1	2	0	0
Croatian	3	2364	42	126	0.0012	2	1	0	0
Bulgarian	2	2364	65	130	0.0008	2	0	0	0
Malay	2	2364	1	2	0.0008	2	0	0	0
Ndebele	2	2364	26	52	0.0008	1	1	0	0
Persian	2	2364	8	16	0.0008	1	0	1	0
Vietnamese	2	2364	58	116	0.0008	2	0	0	0
Akan	1	2364	1.5	1.5	0.0004	1	0	0	0
Bengali	1	2364	21.5	21.5	0.0004	0	1	0	0
Esperanto	1	2364	25.5	25.5	0.0004	0	0	1	0
Estonian	1	2364	3	3	0.0004	1	0	0	0
Farsi	1	2364	110	110	0.0004	0	1	0	0
Ibo	1	2364	11	11	0.0004	1	0	0	0

Khmer	1	2364	122	122	0.0004	1	0	0	0
Latin	1	2364	305.5	305.5	0.0004	0	1	0	0
Latvian	1	2364	73.5	73.5	0.0004	1	0	0	0
Lithuanian	1	2364	4	4	0.0004	1	0	0	0
Macedonian	1	2364	12.5	12.5	0.0004	1	0	0	0
Maori	1	2364	94.5	94.5	0.0004	0	1	0	0
Slovak	1	2364	62	62	0.0004	1	0	0	0
Slovene	1	2364	24.5	24.5	0.0004	1	0	0	0
Swahili	1	2364	37.5	37.5	0.0004	1	0	0	0
Thai	1	2364	20	20	0.0004	1	0	0	0
Tswana	1	2364	151	151	0.0004	1	0	0	0
Twi	1	2364	0	0	0.0004	0	0	1	0
Urdu	1	2364	61	61	0.0004	1	0	0	0
Yoruba	1	2364	12	12	0.0004	1	0	0	0

8.2.4.2. Common words and phrases

Upon the summary characterization of the "proper name" component, it is a matter of consequence to apply the same procedure to foreign common words and phrases. Tab. 16, presented below, offers a 'bird's eye-view' of some basic statistics connected with "common" lexical input of donor lexical systems, characterizing the English language at the turn of the 21st century.

Tab. 16. *Common words – general statistics*

Language	Items	Total	CRACn2	Mass	Frequency	Person	Thing	Abstract
French	598	1018	33	19734	0.5874	79	272	247
Italian	109	1018	17	1853	0.1070	12	42	55
German	91	1018	17	1547	0.0893	8	41	42
Spanish	65	1018	26	1690	0.0638	8	45	12
Hindi	38	1018	16	608	0.0373	12	14	12
Japanese	28	1018	19	532	0.0275	0	13	15
Russian	15	1018	28	420	0.0147	2	7	6
Greek	11	1018	15	165	0.0108	0	11	0
Chinese	8	1018	44	352	0.0079	0	3	5
Welsh	7	1018	9	63	0.0069	0	2	5
Arabic	6	1018	76	456	0.0059	1	1	4

Tab. 16 – cont.

Portuguese	5	1018	6	30	0.0049	0	3	2
Polish	4	1018	10	40	0.0039	0	3	1
Swedish	4	1018	3	12	0.0039	0	3	1
Afrikaans	3	1018	30	90	0.0029	0	2	1
Danish	3	1018	14	42	0.0029	0	1	2
Hungarian	3	1018	6	18	0.0029	0	2	1
Punjabi	3	1018	0	0	0.0029	0	3	0
Zulu	3	1018	5	15	0.0029	0	1	2
Cantonese	2	1018	21	42	0.0019	0	2	0
Irish	2	1018	40	80	0.0019	0	0	2
Korean	2	1018	20	40	0.0019	0	1	1
Scottish-Gaelic	2	1018	2	4	0.0019	0	0	2
Finnish	1	1018	195	195	0.0009	0	1	0
Hebrew	1	1018	0	0	0.0009	0	1	0
Malay	1	1018	1	1	0.0009	0	1	0
Maori	1	1018	14.5	14.5	0.0009	0	1	0
Sanskrit	1	1018	4	4	0.0009	0	0	1
Turkish	1	1018	2	2	0.0009	0	1	0

8.3. A few more words on force and mass

Another interesting phenomenon that we might observe is related to the dynamics of geographical forces of gravity as combined with cultural forces of gravity. Both forces have been described in section 6.8.3 and argued to be equally significant in the characterization of lexical assimilation process of foreign words and phrases in English. When we analyze "proper name" part of donor lexical systems affecting the English language and relate the discussion about the centripetal nature of the aforementioned geographical and cultural forces to the concept of mass as measured in terms of CRACn2, some interesting correlations can be noted.

For example, when we look at such languages as Arabic, Japanese, Irish, Polish, Afrikaans, Serbo-Croatian, Xhosa, Bulgarian, Vietnamese, Farsi, Khmer, Latin, Latvian, Maori, Tswana or Urdu, we notice a huge discrepancy between their individual volumes and CRACn2 values in favor of the latter. All these languages exceed CRACn3 value, some of them like Arabic, quite considerably. All of these languages, except maybe for Irish, are not cognates of English, nor are they in direct geographical proximity, yet still they manifest quite high CRACn2

values. This suggests that the volume of the planetoid itself is not a good indicator whether words and phrases of the languages in question will become incorporated into the English lexical system and even permanently adapted. The solution to this intriguing puzzle is again obtained from the world of astrophysics, strictly the observation of falling meteorites. This observation, *nota bene* consonant with the folk view, is that the bigger the meteor, the greater the chances of it falling onto the surface of the Earth. We can easily transpose this observation to the world of linguistics and argue as follows:

Regularity 4

The lesser the mass of a planetoid and the greater the geographical distance from the Earth, the bigger the chance that individual meteors entering the atmosphere will eventually find their way onto the surface of the Earth due to their relatively greater masses. These relatively greater masses of particular meteorites are the result of greater force of resistance that they have to overcome in order to get to the surface of the Earth. Why? Because the necessity to reach the surface of the Earth as driven by cultural force of gravity entails a bigger mass of a meteor if this meteor is to neutralize reaction force of resistance caused by the aforesaid geographical distance between the planetoid and the Earth.

We may here observe the phenomenon of compensation where the relatively weak centripetal force of geographical gravity is compensated for the centripetal force of cultural gravity. Were it not for this compensation, the languages distant from English (with its centre in Great Britain) would not be able to influence the lexicon of English in any way. This is obviously not true. Conversely, if "the centripetal force of geographical proximity," (see Townend 2002),[3] is combined with the centripetal force of cultural proximity, both forces act with a twofold intensity on the planet or planetoid, causing it to have a great influence on the target language system. This is, naturally, concordant with the results discussed above, where French and German are seen as main, stimulating forces acting upon the lexical structure of the English language. The dynamics of both forces is additionally stimulated by the fact that the English language conceived of as the Earth has actually two main centres of force radiation, i.e. historically primary Great Britain and secondary the USA. Both centres are complementary in contributing to what is now regarded as the English language. Therefore, the analysis of foreign words and phrases, although based on the BNC, which is rather confined to the British 'radiation centre,'[4] does not ignore American sources, where we also search for the traces of foreign meteorites.

[3] Cf. LD and its formulation of culture influence as stimulated by lexical importations: "The contributions of loanwords from particular languages into English offer insights into cultural influence" (p. x).

[4] This may be regarded as one of the weakness of the present study; however, it is enforced by the rationale behind the scope of the research.

8.4. A glance at the structure of global planetary–planetoid mass

We are at the point of the analysis where summary of global planetary–planetoid mass, i.e. the overview of characterization of semantic fields of donor lexical systems affecting the English language, can be conducted. We will go through the presentation of these global results by adhering to the already introduced division of foreign vocabulary into proper names and phrases, on the one hand, and common words and phrases, on the other.

8.4.1. Polysemy

Let us firstly investigate the results we obtain after the analysis of all foreign proper names and phrases. The subdivision of this category is into subgroups: "place names," "personal names," "non-personal names" and the special category of "miscellaneous" – the term used to cover polysemous cases. Since the phenomenon of polysemy is nothing out of the ordinary in characterizing lexical system of any natural language, it has naturally been taken into account in the analysis.

One disadvantage of the characterization provided below relates to somewhat artificial differentiation between cases classified as polysemous or incidental, and therefore subsumed under the aforementioned category "miscellaneous," and the cases of so-called monosemous lexical items. The phrase "so-called" must be stressed at this moment since not an *a priori* assumption is made in the book that 'pure' monosemy exists at all. The study rather complies with the view held by cognitive linguists, according to which the phenomenon of polysemy in the structure of lexical items is rather maximized than minimized (see, e.g. Evans and Green 2006). Still, however, one of the goals set out at the introduction to this book is to reflect the average knowledge of vocabulary possessed by native speakers – the fact which requires some sort of idealization, i.e. simplification of presented data and, as a result, necessitates distinguishing between tendencies towards a higher and a lower degree of polysemy in the mind of the average language user.

An excellent tool for such demarcation is *Wikipedia* which appears to offer a consensus between the detailed encyclopedic expert view of the world, and naïve or simplified folk view. It was on the basis of this source that lexemes were categorized either as monosemous, i.e. displaying one salient sense correlated with the item, or as polysemous, in which case lexemes were shown to correlate with more than one identifiable sense. It appears then, that the phenomenon of polysemy is 'polysemous' in itself (see Geeraerts 2006c, Evans & Green 2006: Chapter 10) and may cover cases of intracategorial proliferation of meanings, i.e. differentiation of meanings within a particular category (whether "place names," "personal names"

or "non-personal names"), or it may cover the case of intercategory extension of meanings pertinent to one lexeme, i.e. differentiation of meanings dispersed across the categories of "place names," "personal names" and "non-personal names." The two instantiations of polysemy have been technically referred to as "miscellaneous 1" and "miscellaneous 2," respectively (see also section 7.1.1, Fts. 33–36).

However, as the analysis has shown, the picture may get additionally more complicated when we look at the cases of lexical items which are both polysemous intracategorially and intercategorially. These cases were subsumed under the hybrid category "miscellaneous 1 and 2." As already noted above, we do not claim that the picture that we get as a consequence of adopting these criteria of demarcation is 100% reliable, but the simplification appears to be an indispensable element of ordering the vocabulary to the extent that such idealization is possible, given the framework of cognitive-linguistic study. This idealization effect has hopefully been reduced by adopting *Wikipedia* as an external reference source on the one hand, and using BNC statistics of frequency of occurrence, on the other.

8.4.2. Anthropocentricity, egocentricity and symbolicity as structuring force-like principles

8.4.2.1. Proper names and phrases

As Tab. 17 suggests, foreign proper words and phrases are largely composed of "place names" and "personal names" displaying tendencies of being perceived as predominantly 'monosemous' in the minds of English native speakers. One might speculate about the causes of such state of affairs, but it seems that the attraction of this type of vocabulary is both in line with egocentricity principle and anthropocentricity principle (Dirven and Verspoor 2004: 5–8). The former is directly connected with "deictic orientation" and the tendency of humans to place the description of events from the viewpoint of "here and now." The latter is related to the perception of a human being by human beings as privileged entity in the description of events, where other non-human entities are also involved. As a consequence, then, foreign proper names and phrases (place names along with numerous non-personal names), as the dominant layer of foreign planet–planetoid mass structure, relate directly to this 'locative imperative' entailed in the egocentricity principle. Also the anthropocentricity principle is seen as almost equally represented as a force influencing a part of the English lexical system. This is observed in a huge number of foreign "personal names" identified in the course of the analysis. Summing up, we may say that these two principles do not only appear to permeate our everyday communication, but they also govern the metacognitive processes of lexical forces of gravity as further explicated in Chapter 9.

Tab. 17. *Planet–planetoid mass of foreign proper names and phrases*

Type	Occurrence	Lemmas
Place names only	0.422	999
Personal names only	0.418	989
Non-personal names only	0.114	271
Miscellaneous 1	0.038	90
Miscellaneous 2	0.039	94
Miscellaneous 1 and 2	0.006	15
Total: 2364		

Key:
a) Miscellaneous 1: Intercategorial polysemy
b) Miscellaneous 2: Intracategorial polysemy
c) Miscellaneous 1+2: Both inter- and intracategorial polysemy

After having a glimpse at the general structuring of planet–planetoid mass affecting the planet Earth, we may proceed to the more detailed specification of that cumulated 'foreign input' offered by donor lexical systems. As Tab. 18 shows, the 'active zones' (see Evans and Green 2006: 238–240) of different conceptual domains/fields identified in the subcategory of "place names" concentrate on locations which are most natural bearers of identification for humans. Here, the aforementioned 'locative imperative' as driven by egocentricity principle also carries deeper cognitive implications. These implications concern the so-called CONTAINER schema (see Krzeszowski 1997) and the primary experience of this concept by humans connected with the mother's womb and the consequent birth. No wonder, then, that domains such as "city," "province/region," "river," "mountain/valley" or "country" serve as the most frequently recurring in the specification of foreign lexical subpart of the English language.

However, the tentative claim is that the set of the aforementioned features will remain essentially unaltered in the examination of foreignisms affecting the lexical system of any other language that we adopt as target for our consideration. The basis for this claim is that these domains might serve as "matrix of domains" (see Lakoff 1987) characterizing the following complex cognitive model of LOCATION as grounded in the conceptual metonymy LOCATION FOR HUMAN IDENTIFICATION. Thus, we may observe how the principle of anthropocentricity, as related to the need of identification, perfectly intertwines with the egocentricity principle as related to the human need of conceptualizing other entities in terms of a three-dimensional CONTAINER, to eventually form the dominant element structurizing the lexical systems of donor languages in the target language.

Tab. 18. *Active zones in the planet–planetoid structuring of foreign proper place names*

Type	Occurrence	Lemmas
City	0.557	596
Province/Region	0.214	229
City and Province/Region	0.088	94
River	0.041	44
Mountain/Valley	0.031	34
Country	0.030	33
River and Province	0.006	7
Miscellaneous[5]	0.005	6
Cave/Rocks	0.002	3
City and River	0.002	3
Lake/Sea/Ocean	0.002	3
Avenue	0.001	2
City and Mountain	0.001	2
Country/River	0.001	2
Country and City	0.001	2
Mountain and Province	0.001	2
Mountain and River	0.001	2
City and Caves/Rocks	0.001	1
City and Lake	0.0009	1
Country and Province	0.0009	1
Mountain and Lake	0.0009	1
Total: 1069		

Tab. 19 offers specification of foreign lemmas pertinent to the category of "personal names" which characterize the structure of English. As we have already seen, the anthropocentricity principle is the one that 'obliges' humans to perceive themselves as 'privileged in the description of events' (see Dirven and Verspoor 2004: 6). The word "privileged" is crucial here as it conventionally bears an axiological positive charge associated with the meaning of this word. We, thus, predict that the most prototypical domains pertinent to that subcategory of proper names and phrases also display absolutely positive conventional valuation.[6] Not surprisingly, the most prominent category identified in the course of the analysis relates to one of the most socially valued professions, i.e. "artist," as this inherently

[5] Miscellaneous: more than two references both inter- and intracategorially.

[6] See Krzeszowski (1997) on the discussion of concepts of valuation in language.

heterogeneous category possesses a feature that uniquely distinguishes it from other categories. This feature relates to a creative form of activity, which, in turn, is connected with how people prototypically conceptualize the internal structure of events as "actions in which one entity acts upon another" (Dirven and Verspoor 2004: 10). This prototypical perception of event structure not only appears to motivate the grammatical structure of language (e.g. word order structure, see Dirven and Verspoor 2004: 10), but also the lexical structure in the processes of incorporating foreign words and phrases into a target lexical system.

Remarkable in this respect is the correlation observed between the frequency of personal names and their reference to prototypical action events. The correlation may be stated as follows:

Regularity 5

The frequency of foreign personal names incorporated into the target lexical system increases with the social perception of these names viewed as prototypically "creative" in the description of events.

Of course, regularities do not exclude exceptions, so a discernible odd one out in Tab. 19 provided below is the category of "inventor." This, however, can be easily explained by the fact that persons who might be labeled "inventors" and who actually get entrenched in the awareness of the average individual conceptualizer are rare on everyday basis, therefore the relative frequency of foreign items designating these form of activity is correspondingly scarce.

Tab. 19. *Active zones in the planet–planetoid structuring of foreign proper personal names*

Type	Occurrence	Lemmas
Artist	0.213	226
Miscellaneous[7]	0.177	187
First name/Surname	0.141	150
Scientist/Scholar	0.112	119
Politician	0.099	105
Writer	0.092	98
Literary/Legendary character	0.033	35
Royal family/Noblemen	0.017	19
Historical figure/Ruler	0.015	16
Sports activist/Sportsman	0.015	16
Adventurer/Traveller/Sailor	0.012	13
Physician	0.012	13
Religious activist	0.012	13

[7] More than one reference inter- and intracategorially or occupation/quality recorded once only.

Engineer/Architect/Industrialist	0.011	12
Deity/Prophet/Saint	0.008	9
Soldier	0.008	9
Ethnic group	0.003	4
Inventor	0.003	4
Inhabitant of a Town	0.002	3
Print maker	0.002	3
Banker/Economist	0.002	3
Total: 1057		

The subcategory of foreign "non-personal names" has already been alluded to in the discussion of 'locative imperative' above. Similarly to the aforementioned categories of "place names" and "personal names," the category "miscellaneous" appears to be a valid contributor to characterizing the mass of planets and planetoids structuring the foreign lexical input of the English language from the perspective of "non-personal names." Otherwise, the types of non-personal foreign words and phrases also exhibit the regularity where the most frequent categories correlate with more prototypical instantiations of CONTAINER schema, whereas the less frequent categories correlate with less prototypical instantiations of CONTAINER. The measure of prototypicality is again connected with anthropocentricity principle and involves the association between the types of CONTAINER and their use in everyday life of a human conceptualizer. Naturally, then, such domains as "manufacturing company," "buildings," "wines/spirits," "car" or "language," as highly entrenched in daily experiences of human conceptualizers living in Judeo-Christian culture, serve as dominant lexical foreign input in the expression of communicative needs by English native speakers.

Tab. 20. *Active zones in the planet–planetoid structuring of foreign proper non-personal names*

Type	Occurrence	Lemmas
Miscellaneous[8]	0.394	140
Manufacturing company	0.074	26
Buildings	0.057	20
Wine/Spirits	0.057	21
Type of car	0.045	16
Language	0.037	13
Administrative/Political term	0.023	8
Name of religion	0.023	8
Dynasty/Kingdom	0.020	7

[8] Inter- and intracategorial polysemy or domain recorded once only.

Tab. 20 – cont.

Form of address	0.020	7
Holy/Religious scriptures	0.020	6
Artistic institution	0.017	6
Name of dish	0.017	6
Type of educational institution	0.017	6
Epic poem	0.017	6
Name of political party	0.017	6
Type of airline	0.014	5
Cheese type	0.014	5
Opera-related terms	0.011	4
Philosophy/Philosophical term	0.011	4
Type of honorary title	0.008	3
Newspaper/Magazine	0.008	3
Type of outer-space project	0.008	3
Type of watch	0.008	3
Type of airport	0.006	2
Concentration camp	0.006	2
Type of currency	0.006	2
Type of disease	0.006	2
Type of festival	0.006	2
Non-alcoholic drink	0.006	2
Name of perfume	0.006	2
Place-related quality	0.006	2
Publishing/Editorial company	0.006	2
Type of horse	0.006	2
Type of cultural movement	0.006	2
Total: 354		

8.4.2.1.1. Proper names and phrases and CRAC

Interesting results can be observed when we compare types of subcategories making up the category of foreign proper names and phrases in the context of their relative contribution (i.e. assimilatory) potential measured in terms of CRACn2 value. It is to be reminded that CRACn2 value is obtained as the average of all CRACn1 values attributed to particular foreign proper names and phrases. The value thus calculated (in this case it is 54) serves as a landmark for designating which of the discussed lemmas are likely to enter the process of adaptation into the target system, and which are not. The principle entailed in this concept is that the higher the individual CRACn1 value (i.e. equal or above 54), the greater the

likelihood that a particular lemma will become a stable element in the lexicon of the target language. As we may conclude from the analysis of the data presented in Tab. 21, the number of lemmas pertinent to categories "personal names" or "place names" is relatively higher than the number of lemmas pertinent to other categories, especially the category of "miscellaneous," which itself maximizes polysemy phenomena. The two regularities, thus, appear to be here in action:

Regularity 6

The higher the CRACn1 value of a particular lemma, the greater the structuring force of egocentricity and anthropocentricity principles acting on it. In effect, a particular lemma will belong to either of two subcategories: "personal names" or "place names."

Regularity 7

The tendency for maximization of monosemous representation of a given lemma in the mind of a human conceptualizer increases proportionately to its CRACn1 value.

The tendency that we might, thus, observe is that a better entrenchment of a foreign proper word or phrase in the target lexical system is associated with the minimization of 'fuzziness effect' that every single lemma carries with it. In other words, the clearer the meaning of a lexical item, the more chances there are that this item will be preadapted and/or eventually fully integrated with the target lexicon, i.e. become recognized as part of the target lexicon by native speakers.[9]

Tab. 21. *The structure of planet–planetoid mass of foreign proper names and phrases and CRAC*

Type	Occurrence (CRACn3 above 54)	Lemmas
Personal names only	0.415	225
Place names only	0.400	217
Non-personal names only	0.079	43
Miscellaneous 2	0.061	33
Miscellaneous 1	0.030	16
Miscellaneous 1 and 2[10]	0.015	8
Total: 542		

8.4.2.2. Common words and phrases

In section 8.4.2.1 we offered a concise characterization of the structure of foreign proper names and phrases and accounted for the nature of the structuring processes relative to a larger cognitive framework, in which the principles of egocentricity and anthropocentricity appear as prevalent. According to Dirven and

[9] See the word-recognition test discussed in section 8.6.

[10] Miscellaneous 1 – intercategorial polysemy; Miscellaneous 2 – intracategorial polysemy.

Verspoor (2004), there are two more principles structuring language, i.e. iconicity principle and symbolicity principle. It seems that the latter is of particular significance in the description of common words and phrases. The reasons are twofold. Firstly, "common words and phrases" is a group of foreign lexical items distinct from proper names and phrases among others by the fact that they are the 'property' of a greater number of individuals, whether animate or inanimate. Thus, the bondage between the word and its referent (whether physical or conceptual) is much weaker than in the case of proper names. This naturally leads to the principles of egocentricity and anthropocentricity (both belong within indexicality principles according to Dirven and Verspoor 2004) to have lesser structuring impact on common words and phrases. Secondly, common words and phrases due to their entrenchment in their 'home' lexical systems (on account of working of symbolicity principle) exhibit lesser degree of assimilatory potential as measured collectively in terms of CRACn3 than proper names and phrases.

In actual fact, it turns out that CRACn3 value reaches the half of the value identified for proper names and amount only to 28 (see also Kuźniak 2009b). The principle that takes over the structuring role in the case of common words and phrases is 'the principle of symbolicity' (Dirven and Verspoor 2004: 12–13). As the authors claim: "The principle of symbolicity refers to the conventional pairing of form and meaning and is typically found in the word stock of a language" (2004: 12). Not surprisingly, the common words and phrases found their due place in many lexicological descriptions, being much more standardized units of language than proper names which were not infrequently relegated to the status of a second-rate phenomenon.[11] This book offers a different perspective in that we see proper foreign names and phrases as dominant in the foreign word stock structuring the lexical system of English in that they are based on principles which are experientially closer to the human conceptualizer.

Tab. 22. *Planet–planetoid mass of foreign common words and phrases*

Type	Occurrence	Lemmas
Person	0.120	123
Thing	0.467	476
Abstract	0.409	417
Miscellaneous[12]	0.001	2
Total: 1018		

As we found out from the discussion of interplanetary gravitation in Chapter 6, the closeness between the two celestial bodies entails a greater gravity force. Therefore, it appears absolutely natural that proper names and phrases possess a greater CRACn3 value as compared to common words and phrases. Coming

[11] See the discussion upon the linguistic research on proper names in Berezowski (2001).

[12] Intercategorial polysemy ("thing" and "abstract").

back to the latter, Tab. 22 illustrates the fundamental structuring levels (domains) of this group of foreign vocabulary and their relative contribution to the overall planet–planetoid mass.

Striking as it may seem, the category "miscellaneous" occupies a marginal structuring role in this stock of vocabulary. The reason is the aforementioned symbolicity principle along with the suggested entrenchment of common words and phrases in their 'home' lexical systems. The result is that a common foreign word or phrase, when 'transferred' on to the target lexical system, usually conveys one dominant meaning it originally possesses in its own lexical system. Possession of one prominent or central meaning itself appears to be fundamental to the organization of common words in the lexicon from semasiological and onomasiological perspective (see Dirven and Verspoor 2004: 26–46).

All things considered, common words and phrases primarily relate to the domains of "thing" and "abstract," and only secondarily do common words and phrases designate human beings. This occurs on account of the aforementioned lesser impact of egocentricity and anthropocentricity principles as overshadowed by the working of symbolicity principle.

When we take a closer look at the specification of domains subsumed under the three general subcategories, i.e. "person," "thing," "abstract," it appears self-evident that the most dominant elements constituting the "common" foreign planet–planetoid mass relate primarily to actions undertaken by humans rather than humans directly themselves. The two most dominant domains, i.e. "communication" and "food and drink" take up 43% of the overall volume. The former relates to the one of the absolutely fundamental actions undertaken by humans on everyday basis, i.e. linguistic communication, whereas the latter is associated with equally significant domain responsible for the physical survival of human beings. The close entrenchment of these two domains in human life ensures a higher degree of incorporation of the corresponding lemmas into the target system. How is this degree of incorporation (measured in terms of *V*) transposed on to the degree of assimilation (measured in terms of CRACn3) is presented in the following section.

Tab. 23. *Active zones in the planet–planetoid structuring of foreign common words and phrases*

Domain	Occurrence	Lemmas
Communication	0.272	278
Food and drink	0.161	165
Social groups/Professions/People	0.083	85
Music and arts	0.074	76
Lifestyle/Social life	0.071	73
Materials/Objects/Technical appliances	0.061	63
Location/Buildings	0.031	32
Biology and animals	0.030	30
Fashion/Clothes	0.026	27

Tab. 23 – cont.

Geography/Meteorology	0.025	26
Military/Warfare	0.021	22
Money/Trade/Economy	0.019	20
Miscellaneous[13]	0.018	19
Science and education	0.015	16
Politics	0.014	15
Entertainment/Movie/Media	0.013	14
Tourism and transport	0.009	10
Philosophy	0.007	8
Name of religion	0.006	7
Geology/Mining/Agriculture	0.006	7
Lakes/Seas	0.005	6
Law	0.005	6
Sport	0.004	5
Medicine	0.003	4
Measures	0.003	4
Total: 1018		

8.4.2.2.1. Common words and phrases and CRAC

When we analyze (see Tab. 24) the assimilatory potential of common words and phrases measured in terms of CRACn3 value, we might observe analogical processes occurring to the ones we noticed while investigating proper names and phrases. The two regularities should be reiterated here, one slightly modified (Regularity 8); the other – unchanged (Regularity 9).

Regularity 8

The higher the CRACn1 value of a particular lemma, the greater the structuring force of symbolicity principle acting on it at the expense of forces of egocentricity and anthropocentricity. In effect, a particular lemma will belong to either of the two subcategories "thing" or "abstract."

Regularity 9

The tendency for maximization of monosemous representation of a given lemma in the mind of a human conceptualiser increases proportionately to its CRACn1 value.

It is also extremely regular that the two dominant domains identified in the course of the global analysis of common words and phrases, i.e. "thing" and "abstract" preserve the same relative value of contribution to the entire volume of

[13] More than one cross-domain reference.

common words and phrases, whether viewed globally (ca. 87%) or under CRAC constraints (ca. 89%). Very intriguing is the preservation of the same percentage of contribution at lower levels of analysis, that is when we investigate the two dominant subdomains (i.e. "communication" and "food and drink") pertinent to the categories "abstract" and "thing," respectively. Not only do they preserve the status of 'leader' in the case of investigating assimilatory potential of lemmas as equaling or above CRACn3 28, but also preserve the same percentage (ca. 43%) contribution to the overall volume of stock identified as above that 28 value! It must also be noted that the same regularity is observed in the case of proper names and phrases, where the relative contribution of the categories of "personal names" and "place names" viewed globally and under CRAC constraints, is almost equal and amounts to 81%. This leads us to formulate the regularity which may be stated as follows:

Regularity 10

The relative quantitative contribution of the dominant elements structuring either proper names or common words to *V* of their respective stocks viewed globally (i.e. without any CRAC constraints) corresponds (with negligible percentage of difference) to the relative quantitative contribution of these dominant elements to *V* constrained by CRACn3 value.

Tab. 24. *The structure of planet–planetoid mass of foreign common words and phrases and CRAC*

Type	Occurrence (CRACn3 above 28)	Lemmas
Person	0.108	21
Thing	0.423	78
Abstract	0.467	85
Total: 184		

8.4.2.3. Anthropocentricity, egocentricity, symbolicity forces. Consolidation

Summing up the discussion on the anthropocentricity, egocentricity as well as symbolicity principles as relevant in the characterization of the internal composition of "proper" and "common" foreign lexical input we may, by way of illustration, draw a diagram in which the aforementioned principles will be viewed as correlated with particular 'layers' of mass, i.e. with particular subsystems of the foreign lexical input. A prominent, but not overtly discussed place, is occupied by the principle of iconicity as discussed by Dirven and Verspoor (2004).

The principle of iconicity, according to the authors, is further subdivided into three subprinciples, i.e. principle of distance, sequential order and principle

of quantity. These principles were not discussed for the reason of their being implicitly addressed throughout the book. This implicit address relates to all the discovered analogies between the working of lexical assimilation processes and the working of physical laws. So, for instance, the principle of quantity 'the more form, the more meaning' can easily be transposed onto the correlation between the volume and mass of the planet or planetoid with the meaningful contribution of it to the target lexical system. The principle of distance, for a change, accounts for the analogy between what we know from astrophysics as interplanetary gravity forces and the dynamics of centripetal geographical and cultural forces regulating the cross-language influences. Finally, the principle of sequential order manifests itself on a most general level in explicating the basis for the entire analogical reasoning laid down in the book, i.e. the assumption that spatiotemporal phenomena occurring in the physical world are reflected in an iconic way in the spatiotemporal event occurring in linguistic world. The set of all these iconic correspondences have been discussed in section 6.9.

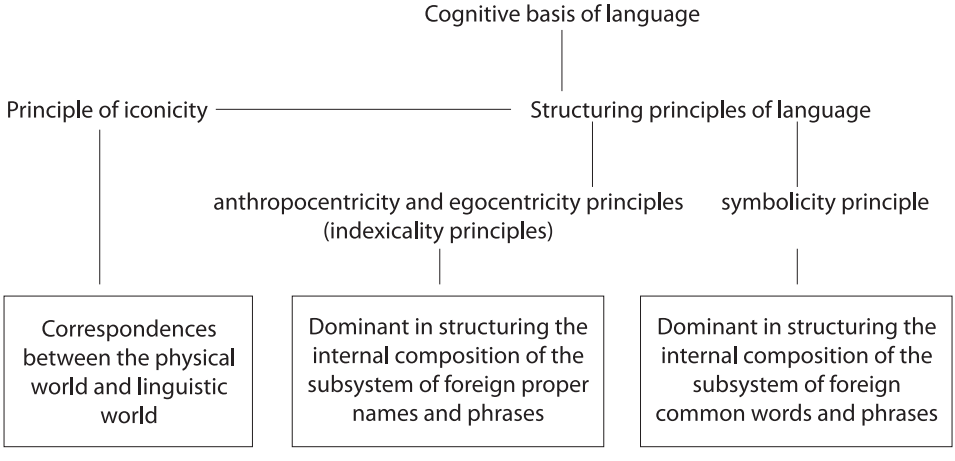


Fig. 47. *Structuring principles of language and subsystems of foreign word stock*

8.5. Identifying core meteorites

In this section, we present the set of prototypical foreign words and phrases across donor languages along the two major categories: proper names and common words. We recall from section 6.7.3 the working typology of foreign words and phrases into those fully adapted into the target lexical system of a given language, those being on way of becoming a part of the lexical landscape of the target language (preadapted foreign words or phrases) and those that do appear

within the system of the target language but shortly enough to mark any significant role in it. Let us recall our findings from section 6.7.3 for the clarity of presentation:

- Meteor – a word or phrase of a foreign origin entering the target lexical system.
- Well-explored meteorite – a borrowing (a fully adapted unit).
- Underexplored meteorite – a foreign word or phrase (a preadapted unit).
- Asteroidal matter – 'unknown' words and phrases in the target lexical system.

Given the validity of all the analogies between the world of physics and the world of language discussed in Chapter 8, we are able to first present a set of "underexplored meteorites" or prototypical foreign words and phrases (see Fig. 12, section 5.3). The criterion of elicitation is CRACn1 value as well as CRACn3 (see section 6.1), where CRACn3 means the value obtained as average of all CRACn1-s in a particular category, whether proper or common. On the basis of the study of all 'nominal' foreign words and phrases preselected for the research on the basis of phonological criteria (see Introduction), CRACn3 values for the group of proper names and common words have been calculated. In the case of the foreign proper name group the value amounts to approximately 54, whereas in the case of foreign common word group the value is 28. An interesting observation is that the foreign proper words and phrases are twice as much condensed in the target lexical system as common words. We already endeavoured to find some reason for this discrepancy (see section 8.4.2.2). At this moment we will confine ourselves to the presentation of lexical units whose CRACn1 values equal or approximate[14] 54 or 28 depending on whether they are classified as proper or common words, respectively.

Tab. 25. *Core underexplored meteorites (proper names and phrases)*

Word or phrase	Donor language	CRACn1
Cocteau	French	56
Mendelssohn	German	56
Umberto	Italian	56
Vinci	Italian	56
Prado	Spanish	56
Camus	French	55.5
Cartier	French	55.5

[14] Here the word "approximate" is to be understood as (+/-) two-point deviation from the reference CRACn3 value for a particular category.

Tab. 25 – cont.

Eiffel	French	55.5
Liszt	German	55.5
Bihar	Hindi	55.5
Schoenberg	German	55
Deutschland	German	55
Oleg	Russian	55
Gothenburg	Swedish	55
Valois	French	54.5
Papandreou	Greek	54.5
Vega	Spanish	54.5
Ming	Chinese	54
Poisson	French	54
Chagall	French	54
Caen	French	54
Montpellier	French	54
Fra	Italian	54
Assisi	Italian	54
Nagasaki	Japanese	54
Bjorn	Swedish	54
Kruger	Afrikaans	53.5
Armagnac	French	53.5
Nantes	French	53.5
Brecht	German	53.5
Hilbert	German	53.5
Roald	Norwegian	53.5
Lhasa	Chinese	53
Rodin	French	53
Havre	French	53
Ulrich	German	53
Rhineland	German	53
Sind	Hindi	53
Potsdam	German	52.5
Jiang Zemin	Chinese	52
Notre Dame	French	52
Freiburg	German	52
Sumitomo	Japanese	52
Lena	Russian	52

Tab. 26. *Core underexplored meteorites (common words and phrases)*

Word or phrase	Donor language	CRACn1
boeuf	French	30
chanson	French	30
volte-face	French	29.5
andante	Italian	29.5
seance	French	29.5
cayenne	French	29.5
pesto	Italian	29.5
gendarme	French	29.5
pas de deux	French	29
force majeure	French	29
ersatz	German	29
roman	French	29
cwm	Welsh	29
realpolitik	German	28.5
dance	French	28
artiste	French	28
extraordinaire	French	27.5
sedan	French	27.5
vignette	French	27.5
ennui	French	27
faux pas	French	27
tutti	Italian	26.5
gateau	French	26.5
taverna	Greek	26.5
fuhrer	German	26.5
cortege	French	26
mistral	French	26
sangria	Spanish	26
junker	German	26

Certainly, the borderline between the above-illustrated prototypical set of foreign words and phrases and the remaining word stock is fuzzy. It would be more than unreasonable to claim that words reaching CRACn1 values between, say 40–70 (in the case of proper names), and 20–40 (in the case of common words) cannot be qualified as "foreign" in the strict sense of the word discussed in the book. Prospective research directed towards approximating such fine-grained boundaries would be a challenge in itself and might serve as subject-matter for

another research. In the present study, one of the central goals is to rather demonstrate that what originally counts as a homogeneous group of foreign words and phrases, according to LPD, may not necessarily count as such. The following statistics presented in Tabs. 27 and 28 are quite telling in that respect. As a point of departure for the subdivision of 'nominal' foreign words and phrases, we adopted CRACn3 value.

Tab. 27. *Lexical dispersion with regard to CRACn3 (proper names and phrases)*

Type of vocabulary	Total	CRACn3	Above 56	52–56	Below 52
Proper names and phrases	2367	54	516	44	1807

Tab. 28. *Lexical dispersion with regard to CRACn3 (common words and phrases)*

Type of vocabulary	Total	CRACn3	Above 30	26–30	Below 26
Common words and phrases	1018	28	170	28	820

The data presented in Tabs. 27 and 28 show some striking regularity in that the percentage of words and phrases below CRACn3 value established for a particular group amounts to ca. 80% for both proper names and common nouns. It means that what is nominally referred to as foreign words and phrases in LPD is in the vast majority of cases instances of words and phrases whose status in the target lexicon may be deemed as at most near prototypical foreign, and at least 'unknown,' given the tertiary classification provided above. Around 20% of all vocabulary discussed may be classified as at least near prototypical foreign, and at most bearing the status of borrowings. If this division of nominal set of foreign words and phrases into "well explored meteorites," "meteorites" and "asteroidal matter" can be upheld, it must be subject to some further inquiry. This is the topic of the following section.

8.6. Word recognition test[15]

For the reasons of corroborating the psychological reality of the postulated tertiary division of the foreign word stock, a sample of 30 lexical items pertinent to proper name and common word group have been elicited for the word recognition test conducted among 100 adult native speakers of English. The words have been

[15] The tradition of vocabulary knowledge assessment goes back to the research by Seashore and Eckerson (1940) – "Seashore and Eckerson defined a 'word' as an item listed in the 1937 edition of Funk and Wagnall's *New Standard Dictionary of the English Language,* which contains approximately 450,000 entries. Of these, they reckoned that just under half, about 166,000 were 'basic words' such as *loyal*, and the remaining 204,000 or so were derivatives and compounds, such as *loyalism, loyalise, loyally,* and *Loyal Legion.* Obviously, it is impractical to test anyone on all the words in the dictionary, so a representative sample of the total needs to be obtained" – after Aitchison (1987: 6). Other significant research is that by Diller (1978, after Aitchison 1987: 7).

selected according to the CRACn1 criterion. That is, the total of 30 words has been divided into 3 subsets. Each subset contains a subset of 10 lemmas and is constituted on the following conditions:

Subset 1 CRACn1>CRACn3.
Subset 2 CRACn1 = CRACn3 value.
Subset 3 CRACn1<CRACn3.

A selection of 10 words for a particular subset is based on the order of alphabet principle and a specific CRACn1 value. Thus, Subset 1 comprises items with highest CRACn1 values (the top ten 'biggest' meteorites), Subset 2 contained 'meteorites proper' (CRACn1 value 54, or 28 depending on whether the proper or common word group, respectively). Finally, Subset 3 subsumes the ten 'smallest meteors' whose CRACn1 value amounts to 0. We, in effect, receive 3 subsets with CRACn1 values ranging from the biggest to the smallest. The hypothesis is, thus, the following: while being examined, the native speakers should indicate words belonging to Subset 1 as the ones they are most familiar with, i.e. they simply know the words belonging to it. They should then rate most of the words listed 1–10 a tick (v) mark. As regards Subset 2 (words listed 11–20), the respondents should be rather hesitant about the vocabulary, in that they should generally have some but not sound idea about the meanings of words they are confronted with. In this case the prevailing answer should be a question mark (?). Finally, with regard to Subset 3 (words listed 21–30), respondents are stipulated to have generally no knowledge about the meanings of words, as these are the words with the lowest CRACn1 values. In this case, the respondents are expected to put a hyphen mark (-) alongside each tested word. The answers provided are later translated into points in that a tick answer (v) is given 2 pts, (?) – 1 pt, and (-) – 0 pt. Of course, the format of the survey has been designed in such a way as to maximally conceal the aforementioned series of hypotheses. The respondents have merely been given a list of 30 words and requested to answer the survey as shown in the instruction below. The format of the survey was the following:

PART I[16]
(Proper names and phrases)

Instruction

Read through the list of words below, putting a tick (v) next to words you know, a question mark (?) next to words you are not sure about and a hyphen (-) next to words you simply do not know.

Subset	No.	Lemma	(v), (?), (-)
SUBSET 1	1	Paris	
	2	Rome	
	3	Iraq	
	4	Hong Kong	

[16] This research should be followed by the verification to what extent if at all the selected set of foreign words and phrases is represented in dictionaries of foreign words and phrases.

Part I – cont.

SUBSET 1	5	Berlin	
	6	Walter	
	7	Holland	
	8	Mexico	
	9	Hitler	
	10	Gorbachev	
SUBSET 2	11	Ming	
	12	Caen	
	13	Chagall	
	14	Montpellier	
	15	Poisson	
	16	Assisi	
	17	Fra	
	18	Nagasaki	
	19	Bjorn	
	20	Armagnac	
SUBSET 3[17]	21	Ystalyfera	
	22	Qur'an	
	23	Zatopek	
	24	Putonghua	
	25	Stepinac	
	26	Skagerrak	
	27	Buys Ballot	
	28	Cuisenaire	
	29	Gewurtzaminer	
	30	Katharevousa	

PART II
(Common words and phrases)

Instruction

Read through the list of words below, putting a tick (v) next to words you know, a question mark (?) next to words you are not sure about and a hyphen (-) next to words you simply do not know.

Subset	No.	Lemma	(v), (?), (-)
SUBSET 1	1	regime	
	2	champagne	
	3	bureau	
	4	liaison	

[17] Naturally, this column was absent from the original survey.

Part II – cont.

SUBSET 1	5	bourgeois	
	6	donna	
	7	grand prix	
	8	en route	
	9	dosta	
	10	genre	
SUBSET 2	11	pas de deux	
	12	force majeure	
	13	roman	
	14	ersatz	
	15	cwm	
	16	realpolitik	
	17	vignette	
	18	danse	
	19	ennui	
	20	sedan	
SUBSET 3	21	repetiteur	
	22	sans-culotte	
	23	cacciatore	
	24	macchiato	
	25	guyu	
	26	pinyin	
	27	arriere-pensee	
	28	au courant	
	29	betise	
	30	bonne bouche	

8.6.1. Test results

Tabs. 29 and 30 below present the results of the survey outlined in the preceding section, following which, the discussion of the emerging tendencies is provided.

Tab. 29. *Results of word recognition test (proper names and phrases)*

Subset	No.	Lemma	WRV	WRV per subset
SUBSET 1	1	Paris	2	1.97
	2	Rome	2	
	3	Iraq	2	

Tab. 29 – cont.

SUBSET 1	4	Hong Kong	2	1.97
	5	Berlin	2	
	6	Walter	1.79	
	7	Holland	2	
	8	Mexico	2	
	9	Hitler	1.99	
	10	Gorbachev	1.93	
SUBSET 2	11	Ming	1.84	1.52
	12	Caen	1.20	
	13	Chagall	1.13	
	14	Montpellier	1.82	
	15	Poisson	1.70	
	16	Assisi	1.72	
	17	Fra	0.68	
	18	Nagasaki	1.77	
	19	Bjorn	1.92	
	20	Armagnac	1.47	
SUBSET 3[18]	21	Ystalyfera	0.10	0.53
	22	Qur'an	1.54	
	23	Zatopek	0.62	
	24	Putonghua	0.21	
	25	Stepinac	0.17	
	26	Skagerrak	0.51	
	27	Buys Ballot	0.62	
	28	Cuisenaire	0.83	
	29	Gewurtzaminer	0.58	
	30	Katharevousa	0.68	

Tab. 30. *Results of words recognition test (common words and phrases)*

Subset	No.	Lemma	WRV	WRV per subset
SUBSET 1	1	regime	1.99	1.93
	2	champagne	1.99	
	3	bureau	1.99	
	4	liaison	1.95	
	5	bourgeois	1.79	

[18] Naturally, this column was absent from the original survey.

SUBSET 1	6	donna	1.81	1.93
	7	grand prix	2	
	8	en route	1.89	
	9	costa	2	
	10	genre	1.94	
SUBSET 2	11	pas de deux	1.56	1..34
	12	force majeure	1.48	
	13	roman	1.91	
	14	ersatz	1.12	
	15	cwm	0.61	
	16	realpolitik	1.15	
	17	vignette	1.42	
	18	danse	1.35	
	19	ennui	1.13	
	20	sedan	1.68	
SUBSET 3	21	repetiteur	1.23	0.81
	22	sans-culotte	1.08	
	23	cacciatore	0.87	
	24	macchiato	1.16	
	25	guyu	0.13	
	26	pinyin	0.39	
	27	arriere-pensee	0.68	
	28	au courant	0.87	
	29	betise	0.5	
	30	bonne bouche	1.17	

The results of the word recognition test (Tabs. 29 and 30) appear, on the whole, to confirm tendencies signaled in section 8.6. Thus, the postulated division of the word stock into three subcategories[19] definitely adds to the claim in favor of psychological reality of the results obtained in Chapter 7 and it is also compatible with the summary statistics presented in the present chapter.

Thus, we observe both with respect to proper name and common word subsets a declining curve of word recognition with the Subset 1 displaying the highest value of recognition, the Subset 2 a relatively lower value, and the Subset 3 the lowest value. These tendencies are, as already said above, in perfect compliance with the hypotheses stated in section 8.6, and generally are in tune with the assumptions laid down in Introduction to the book, and later elaborated on in

[19] See section 6.1 for some anticipatory observations about the validity of such tertiary division as supported by the discussion of CRAC value criteria.

Chapter 6 and the present Chapter 8. What is worth emphasizing, is that values obtained with regard to proper name subsets are relatively higher than values obtained from the analysis of common word subsets. This is again compatible with the CRAC statistics, whereby CRACn3 value for proper names has been calculated at 54, whereas CRACn3 value for common words at 28. In consequence, the higher average occurrences of foreign proper names in the BNC correlate with, the higher recognition statistics received through the word recognition test. Consequently, then, the lower average occurrences of common proper names in the BNC correlate with, the lower recognition values obtain as a result of the aforementioned recognition test.

Another interesting observation relates to the recognition values obtained for individual lemmas. We notice, on the whole, that the individual values of the discussed lemmas pertinent to Subset 1 are higher than the values of individual lemmas pertinent to Subset 2. Predictably, then, individual values of lemmas pertinent to Subset 2 are higher than the values of individual lemmas pertinent to Subset 3. We certainly notice some deviation from this general tendency, but these are individual cases (e.g., the case of relatively unfamiliar *Fra* from Subset 2 in comparison with relatively familiar *Qur'an* from Subset 3). These irregularities that apparently go against the expected tendencies may well be accounted for the fact of the constantly changing lexical system of English (see Introduction), which sees some of the "underexplored meteorites' (Subset 2) to be left to natural processes of oblivion (the case of *Fra*), whereas some of the 'meteors' (shooting stars) accumulate enough gravitational energy through time[20] to finally gain the status of the "underexplored meteorite" (foreign word proper) with chances of more stable adaptation in the system, provided the process of exploration continues.

[20] This process of status change from 'meteor' to 'meteorite' may be here compared to the situation in which a human being takes once an incidental picture of a meteor, then tucks the photo under a pile of documents in his or her shelf without showing it to friends. After some time, due to some accident, or for example because of the rising popularity of 'meteor' talks in the media, the person recalls about the picture, shows it to a couple of his or her friends, who scan it and spread it further via e-mail to other people.

Chapter 9
TOWARDS THE MODEL OF FOREIGN LEXICAL ASSIMILATION PROCESSES

9.0. Overview

This chapter seeks to integrate all of the findings presented in Chapters 6, 7 and 8 into one coherent framework. The goal will be to (re)construct the model of foreign lexicalization processes as underlain by the metadiscursive astrophysical metaphor LANGUAGE LAWS ARE PHYSICAL LAWS (see Preface and section 0.2). There will be two perspectives taken: the internal one, where we try to outline the mechanisms governing the processes of assimilation, and the external one, where we try to take a 'bird's eye view' of the lexical assimilation both from the vantage point of an expert (via the lenses of Copernicus' world), and from the vantage point of an average human conceptualizer (via the lenses of Ptolemy's world). We will eventually see how these two different perspectives complement one another in building a coherent metacognitive representation of foreign lexical assimilation.

9.1. Towards conceptualisation of the model (internal view)

In an attempt to (re)construct the model of foreign lexical assimilation, as viewed from the perspective of internal mechanisms governing its constitution, we have to recall the overriding conceptual metaphor LANGUAGE LAWS ARE PHYSICAL LAWS (see Preface, sections 0.2 and 6.8.3) that has laid the foundation for the present study. On account of this metaphor we have concluded that there exist far reaching correspondences between how words are subject to the processes of assimilation in the target system in respect of some fundamental astrophysical facts connected with the nature and organization of the solar system from the perspective of both: the expert and non-expert worldview.

We will see in section 9.1.1 how these two complementary worldviews are intertwined in effecting the coherent framework within which the model of foreign lexical assimilation can be grasped. However, before more specific framework is postulated we have to integrate into a coherent metaconceptual model the idea of a foreign lexical unit as discussed through Chapters 5–8.

Foreign words and phrases in English, 2009

9.1.1. The manifestations of the word in the conceptual metaspace

We start, then, with modeling the conception of foreign word or phrase accordingly with the tertiary division proposed in section 6.7.3. For convenience, this division is presented here again. It is, thus, to be reminded that we so far postulated three possible models of existence that a lexical unit pertinent to a donor language may have with respect to a target language:

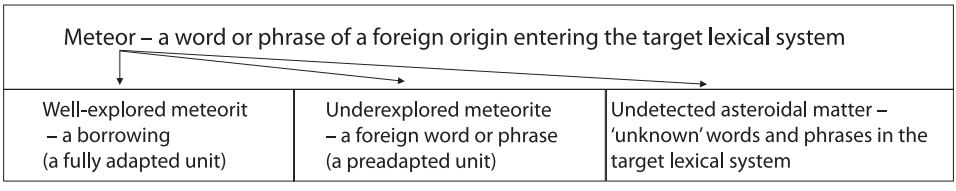


Fig. 48. *The word as a meteor*

In compliance with the underlying principal metaphor LANGUAGE LAWS ARE PHYSICAL LAWS, we are in a position to sketch the following three models of a foreign lexical unit presented in Figs. 49–51:

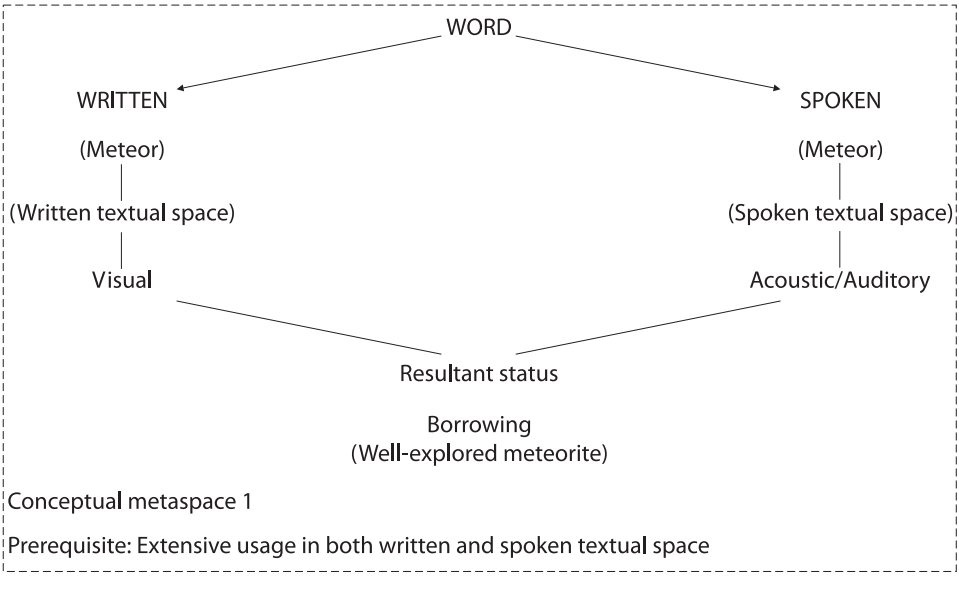


Fig. 49. *The conception of a fully adapted lexical unit in the target system*

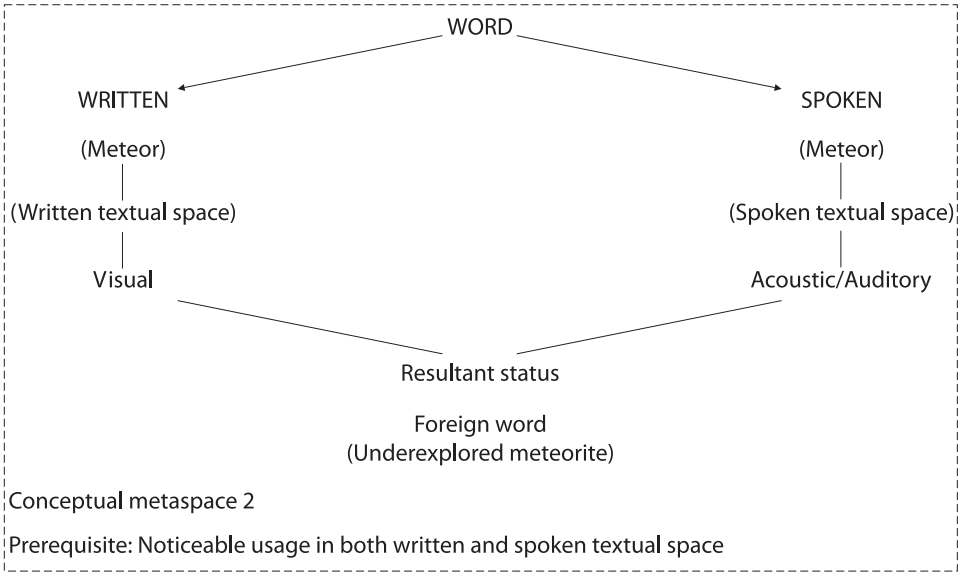


Fig. 50. *The conception of a preadapted lexical unit in the target system*

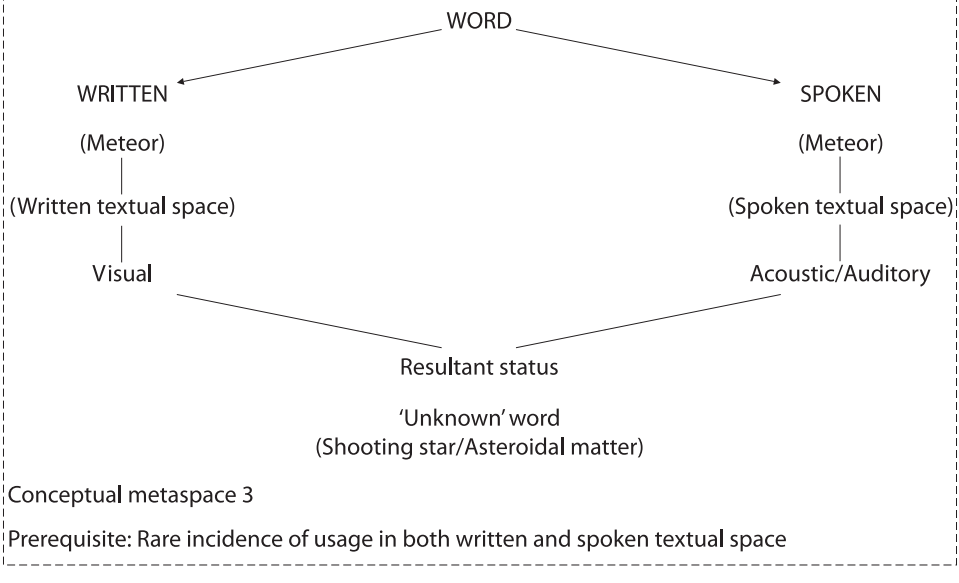


Fig. 51. *The conception of non-adapted lexical unit in the target system*

As illustrated above, the three modes of lexical existence relate naturally both to spoken and written texts. Depending on the frequency of occurrence in these texts, a particular unit is classified as "borrowing," "foreign word," and "un-

known" words along with the correspondent astrophycial 'labels,' i.e. "well-explored meteorite," "underexplored meteorite," and "shooting star/ asteroidal matter," respectively. The conceptual spaces, within which these modes of non-native lexis are discussed, are called "metatextual," as they refer to how these three types of 'words' may potentially be conceived of in metatextual universe of discourse.

9.1.2. Foreign lexical assimilation model (internal perspective)

Having been equipped with the three conceptual models of a foreign word or phrase, we are at a point where we can further explore the mechanisms upon which the assimilation of a foreign word or phrase is stipulated to be founded. It appears that a fundamental internal organizing principle governing the processes in question is CENTER/PERIPHERY opposition.

CENTER/PERIPHERY schema is inextricably related to the conception of CIRCLE. The conception of CIRCLE is again related to the conception of circular movement as we know it from our daily experience. A description of circular movement is found in the following:

> A point on the outside of a merry-go-round or turnable moves a greater distance in one complete rotation than a point on the inside. […] The speed of something moving along a circular path can be called tangential speed because the direction of motion is always tangent to the circle […]. Tangential speed, unlike rotational speed, depends on the distance from the axis. At the very center of the rotating platform, you have no speed at all, you merely rotate. But as you approach the edge of the platform you find yourself moving faster and faster. Tangential speed is directly proportional to distance from the axis.
>
> (Hewitt 1998: 49–50)

The conception of circle-like movement has also been discussed in Chapter 6, in which we have argued that this idea is constitutive of the new conception of the world as proposed by Copernicus. This new conception of the world was seen as revolutionary because it involved a complete reorganization of knowledge about the place of the Earth in the universe and how the planet Earth along with other planets behave with respect to the Sun. We have concluded in Chapter 6 (also Chapter 8) how this new Copernicus' model of the physical world may serve as adequate source domain for modeling foreign lexical assimilation processes. We, thus, see that it is a matter of consistency to postulate the aforementioned CENTER/PERIPHERY schema as foundational for the internal composition of the conceptual model.

This central CENTER/PERIPHERY schema is, as already signaled above, strictly bound up with the conception of MOVEMENT. The conception of MOVEMENT is then construed upon FORCE schema (see section 6.8), and the

working of two fundamental types of force, i.e. centripetal, and centrifugal. The relevance of these two forces for the emerging model has already been widely discussed in Chapter 6. Now, this overriding CENTER/PERIPHERY schema, along with fundamental FORCE schema instantiated by the aforementioned centripetal and centrifugal force, are postulated to be highly schematic and, thus, constitute a macro-level of a metaconceptual representation of the model. FORCE schema is argued to further interact with BLOCKAGE/NO BLOCKAGE schema. In this way, we can talk about INCORPORATION of a foreign lexical unit within the bounds of a foreign lexical system. All this is guaranteed by the working of centripetal force. On the other hand, once incorporated, the foreign lexical unit may be subject to adaptation phases (see Fig. 13, section 5.3), provided centrifugal force of resistance represented by NO BLOCKAGE schema is too weak to prevent the foreign word from marking the trace on the target system 'ground.' Conversely, once incorporated, the foreign lexical item may not be adapted into the system at all, as the centrifugal force of resistance associated with BLOCKAGE schema will be effective enough to form a kind of 'protective shield' over the target lexical system and, in effect, stop the foreign word from 'reaching the ground.' Centrifugal force of resistance (see section 6.8) might, thus, correspond to people's individual negative attitudes towards a foreign word or phrase. Alternatively, it may be represented e.g. by national language policies (ex. Ch. 7 in Crystal 2007c) aimed at protecting a native language from foreign influences. This macro-level can be diagrammed as follows:

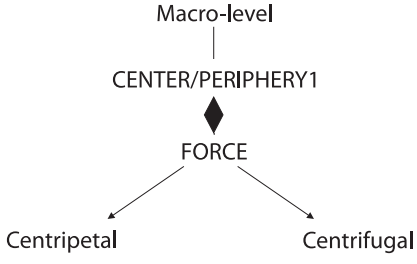


Fig. 52. *Macro-level representation of the lexical assimilation model*

The macro-level of metaconceptual representation is instantiated by more conceptually specific micro-level, which we have CENTER/PERIPHERY 2. This time the opposition is conceptually correlative of the central or peripheral status of a foreign lexical unit in relation to the target lexical system. The discussion of status implies hierarchization of entities. This will be the case here as well. That is, the more central the lexical unit, the more chances there are that it will become adapted into the lexical system of the target language. Conversely, the less central the lexical unit, the more chances there are that it will not become adapted into the lexical system of the target language.

The conception of centrality/peripherality as understood at this level is, thus, closely related to UP/DOWN schema, because the status of centrality/peripherality of an individual unit is measured in terms of CRACn1 value (or "mass" of an entity). We recall from Krzeszowski (1997) that MORE is UP, LESS is DOWN, so we further observe how the notion of centrality/peripherality at this level is subtly permeated with axiological-quantitative aspect, in which MORE is UP and it is fundamentally associated with PLUS pole of PLUS/MINUS schema, whereas LESS is DOWN and it is associated with MINUS pole of the schema. We, in effect, argue that the item that has the status of borrowing (see Fig. 49), i.e. is regarded as the most central foreign lexical unit (with the highest CRACn1), will consequently be viewed as 'welcome' in the target system. This appears logical because the very idea of "borrowing" entails that the target language speakers are in need of a foreign lexical item in order to express their communicative needs.

Consequently, less central will be the units called "foreign words proper" (see Fig. 50) whose status with respect to the target lexical system remains unresolved due to the individual CRACn1 values occupying the area somewhere halfway on the UP/DOWN vertical axis of quantity. This group of prototypical foreign words and phrases is also referred to as "meteorites proper." In the case of this group, it is a matter of speculation which of these 'meteorites' will further be explored and thus become a stable part of the lexical landscape of the target system, and which will disappear from the system through oblivion. It may, however, also be the case that such meteorites may, due to the psychosocial factors such as prestige or group identification, be neither further explored, nor left unexplored, but rather 'protected' or 'cherished' as a valuable thing by speakers of the recipient language. In this case, foreign words retain their native langue basic phonological characteristics, at the same time, however, preserving the traces of some degree of phonological adaptation to the target system caused by the action of assimilation forces.[21]

Finally, the least central, or the most peripheral will be the items (see Fig. 51) with the lowest CRACn1 values, profiling the area near the DOWN (MINUS) pole of UP/DOWN schema. The status of these lexical units will consistently be viewed as negative. We may, thus, call these 'meteors' – 'persona non-grata' or 'unwelcome' in the target system. As their CRACn1 value is relatively small (LESS is DOWN), they are pushed away from within the bounds of the target lexical system.

We may at this point integrate the two levels of conceptual representation of lexical assimilation process as viewed from the perspective of their internal make-up (see Fig. 53).

[21] Such tendencies may, for example, be observed in legal English where the preference for foreign as well as archaic terminology constitutes one of exponents of professionalism (see Brown and Rice 2007: 42)

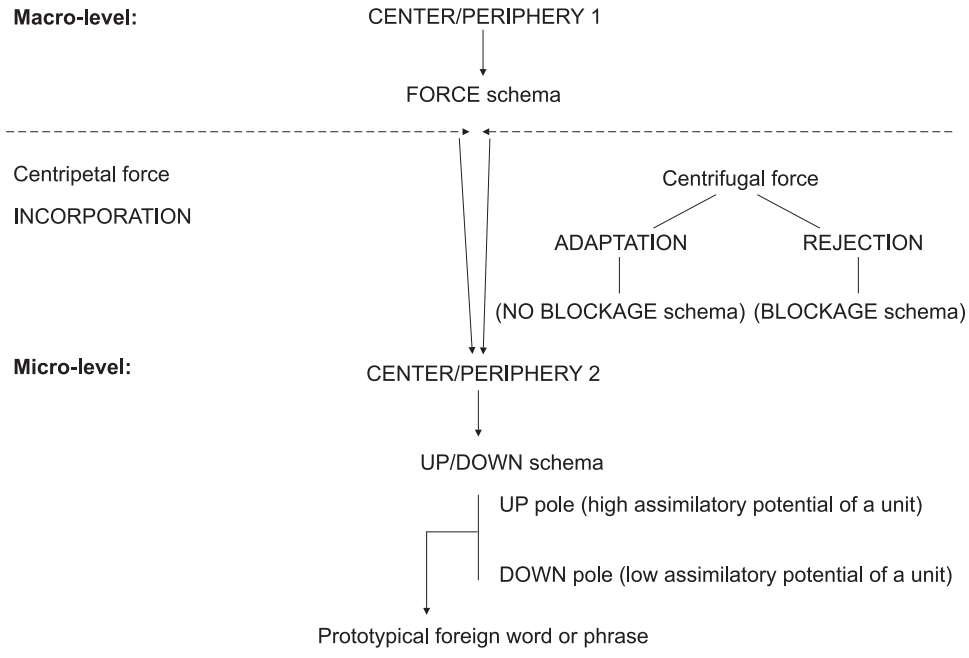


Fig. 53. *Correlation between macro- and microrepresentation in the internal model of foreign lexical assimilation*

The model of lexical assimilation which is presented above appears to be consonant with all of the findings presented in Chapters 6–8. The prototypicality status of foreign words and phrases implies the tertiary division of foreign words and phrases presented *inter alia* in section 9.1.1. The fact that these items occupy a midway area on UP/DOWN vertical axis is dictated by their CRACn1 values, which can be called 'prototypical' because they are equal or near equal to CRACn3 value calculated for a particular group of words. It is to be reminded that CRACn3 relates to the cumulative average relative count of all CRACn1 values of lemmas belonging either to "proper name" or "common word" category of foreign words and phrases. The very idea of averaging implies the search for some 'halfway' value. No wonder then, the prototypicality of a foreign unit is correlative of neither the top, nor the bottom area on the UP/DOWN schema, but rather the area somewhere in-between. This occurs because prototypicality is seen as derivative of CRACn3, i.e. the 'average' (halfway) mass of a given unit. Therefore, the relationship between the notion of prototypicality and CRACn3 value turns out to be natural.

Tentatively speaking, the above-delineated model appears to be neurolinguistically compatible with passive, indirect, responsive model, as well as active, direct search model discussed in section 1.3.1. The postulated compatibility of the model is dictated by the nature of processes governing the inclusion of new words into the target lexicon, whereby the discussed centripetal and centrifugal forces

correspond to the processes of accommodating or rejecting the neurological auditory or visual stimuli. Alternatively, the centripetal/centrifugal forces might be represented as cultural forces of social acceptance or rejection of alien lexical items in the target system. All in all, establishing more definite conclusions in relation to the neurolinguistic basis of the postulated model would be to claim that the metaphor LANGUAGE LAWS ARE PHYSICAL LAWS is to be interpreted literally, i.e. that it is constitutive rather than explicatory. This book is, however, far from making such a claim (see Preface).

9.2. Towards conceptualisation of the model (external view)

9.2.1. Integrating the two fundamental views of the universe

In this section we attempt to outline the external model of foreign lexical assimilation processes by investigating the two apparently conflicting frameworks within which the model can be encapsulated. These two frameworks have already been introduced in section 6.7. We have referred to them as Ptolemy's geocentric model of the world and Copernicus' heliocentric model. That these two models are mutually not exclusive but rather complementary in the proposed metacognitive framework of foreign lexical assimilation constitutes one of the central claims made in the book.

Fig. 54. *Ptolemy's conception of the universe* (Ptolemy, *Almagest*)

The anthropocentricity of the model of the physical world with the Earth as the centre of universe relates in a direct way to the ancient model advocated by Ptolemy as presented by Demiański (2008: 12).

In this model the Earth occupies the central position in the universe with the Sun, planets, and other stars revolving round it. It appears that this geocentric perspective is in accordance with the folk view of the world shared by the average human conceptualizer. Therefore, it cannot be rejected in cognitive-semantic considerations in which, according to the model of conceptual world (Dirven and Verspoor 2004: 14), human conceptualizer and his/her experienced world are the starting point for processes of categorization, be it linguistic or non-linguistic ones (see Fig. 55 below).

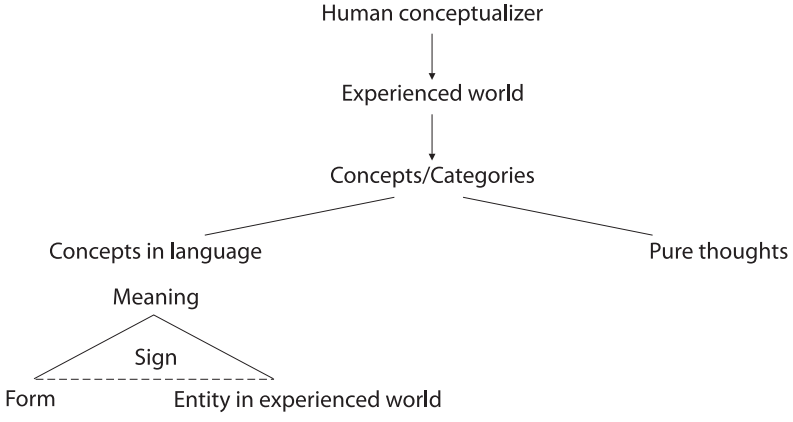


Fig. 55. *Model of the conceptual world (Dirven and Verspoor 2004: 14)*

The set of central correspondences presented in section 6.9 should, however, be seen as a compromise between Ptolemy's geocentric view of the universe and Copernicus' heliocentric view, where the Earth no longer occupies a privileged position in the universe and the Sun that takes over the role of the central celestial body of our planetary system. A systematic reference to Newton's laws of gravitation and other facts recognised by physical science and discussed in section 6.9 are rather the fruit of heliocentric view (see Fig. 56, after Demiański 2008: 12).

It appears that the process of lexical assimilation that we endeavour to model is a perfect juncture point for the two above-presented apparently conflicting philosophies. The model of lexical assimilation process will both respect the egocentricity view in that the English language is metaconceptually represented as the Earth, the centre of the planetary system, with other celestial bodies (planetoids) revolving round it. On the other hand, the laws of physics discussed in Chapter 6 are in accordance with the heliocentricity of the universe. This heliocentric view corresponds to the descriptive component of the lexical assimilation processes model (see Fig. 57) in which we try to answer the question about how things *are like*, rather than how things *are perceived to be like* in which case it is Ptolemy's geocentric model that appears as more adequate framework. We, thus, return to

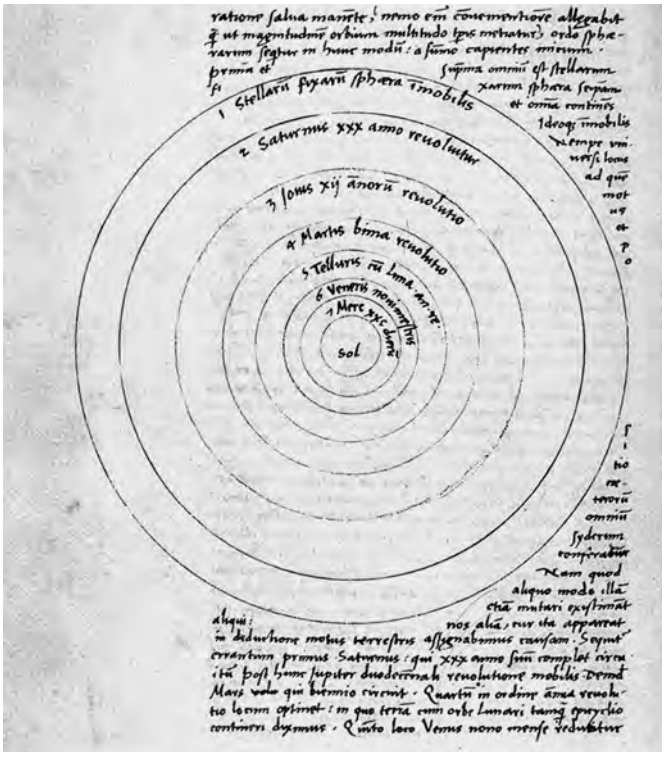


Fig. 56. *Copernicus' conception of the universe* (Nicolaus Copernicus, *De revolutionibus orbium coelestium*, Jagiellonian Library, Cracow)

the already discussed crucial qualitative difference implicated by the two seemingly isomorphic formulae: *x is y* vs. *x is like y* (see section 8.1). The former formula underlies the descriptive component of the externally perspectivized model of lexical assimilation which relates to Copernicus' model of the world, whereas the latter underpins the evaluative component that relates to Ptolemy's folk view of the world in which human emotions and feelings involving perception of foreign lexical units are taken into account. What needs to be emphasised is that the evaluative component of the model of lexical assimilation is actually reduced in this book to a sample collection (see section 4.3) of lexical evidence from lexicologists describing the phenomenon of lexical assimilation. Because the model of assimilation proposed in this monograph is essentially metacognitive or metadiscursive, it rather excludes 'average' language users from getting a conscious access to the model. The exception may be the cases of live TV programmes, when an eminent authority responds to the linguistic queries and puzzles asked by viewers during online conversations.[22]

[22] The example of such TV coverage is a very popular series entitled *Prof. Miodek odpowiada*, in which the Polish linguist, Jan Miodek, attempts to sort out various linguistic puzzles prompted by viewers. Naturally, then, the conversations conducted in this series stimulate the viewers

9.2.2. Lexical assimilation and two models of the universe

Very schematically, then, we may now attempt to illustrate the model of lexical assimilation via the reference to the two models of the universe and the two types of knowledge discussed above.

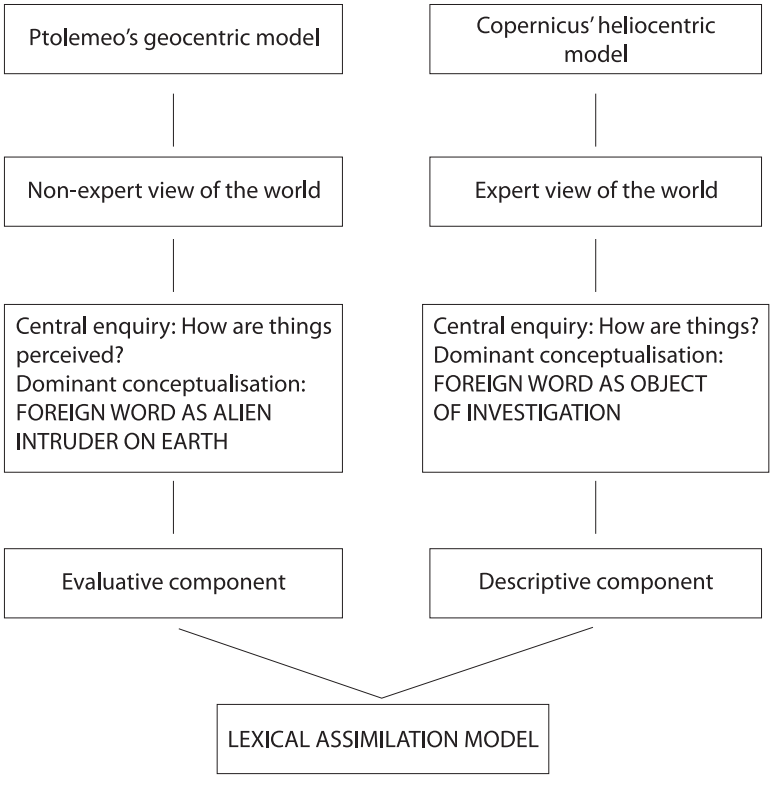


Fig. 57. *Models of the universe, types of knowledge and foreign lexical assimilation processes*

The integration of the two views is possible when we outline the model of foreign lexical assimilation as conjoining the expert and non-expert knowledge, which hopefully reflects our daily intuitions about how physical phenomena occur in the universe. It appears that, although *de facto* mutually exclusive, the two aforementioned views of the world can eventually appear to be rather complementary and *actually* exist parallel in the metalinguistic consciousness of the human conceptualizer. That this may be the case can be observed via introspection, where we have two types of knowledge about how the Earth behaves relative to the Sun. One type of knowledge, the so-called naïve or folk view of the world, makes

towards metacognitive reflections upon the nature of language and its functioning in a speech community.

us think of the Sun as revolving round the Earth, which is linguistically codified in fixed phrases, where we commonly refer to the Sun *rising* and the Sun *setting* over the horizon. On the other hand, we know, via the process of education at school, that it is the Earth rather than the Sun that revolves round the Earth and that it is the former rather than latter that constitutes the centre of our solar (not Earthly) system.

In the external model of foreign lexical assimilation that we attempt to sketch below, we reflect both the naïve and expert model of the world. The former is addressed as it is consonant with anthropocentricity as well as egocentricity principles (Dirven and Verspoor 2004: 6–8), where both principles focus our attention on a human language user as an Earthly creature living *hic et nunc*. The latter is addressed as it serves as a good illustration of evolutionary mechanisms governing the processes of change affecting all natural languages without prioritizing one language over the other. This, in turn, constitutes a perfect illustration of mechanisms governing the processes of foreign lexical assimilation that shape the internal composition of the model and are viewed as analogous to quite a few fundamental facts established in the science of astrophysics (see Chapter 6).

We now see how this vision of the integrated models of the world underpinning the metacognitve composition of the metaphor LANGUAGE LAWS ARE PHYSICAL LAWS interacts with the model of foreign lexical assimilation as viewed externally. As we conclude from Fig. 58, the folk view helps us understand the target language system, i.e. English as the centre of the universe with other languages being attracted to it. This carries, however, opposite axiological implications. On the one hand, the implications conveyed by this view are negative, because in the same way as human beings are occasionally informed of a potential crash of an alien asteroid with the Earth (the famous Armageddon), the humans may perceive other languages and, thus, their lexical units as potential carriers of danger to their home ground (see section 3.1). On the other hand, the implications entailed in this view may, however, be positive or at least neutral.

It is certainly true that the view of the world, whereby the English language is equated with the Earth does not need to lead to the aforementioned pessimism connected with interplanetary collisions. We may observe many people being truly enthusiastic about astronomical investigations. On a more linguistic level of consideration, such people could probably be compared to individuals potentially treating instances of foreignisms in their own language as objects of intriguing investigation rather than off-hand dismissal.

It must finally be reminded at this point that the folk view with English as the centre is of course relativistic, as it is dependent on which language is selected as a target language. We can, thus, easily imagine German, Polish, Swahili or any other language being perspectivized as the Earth because the goal of the research will be to investigate foreign words and phrases in this particular language, and so on. Nevertheless, positing the role of English as the 'prototypical' Earth appears

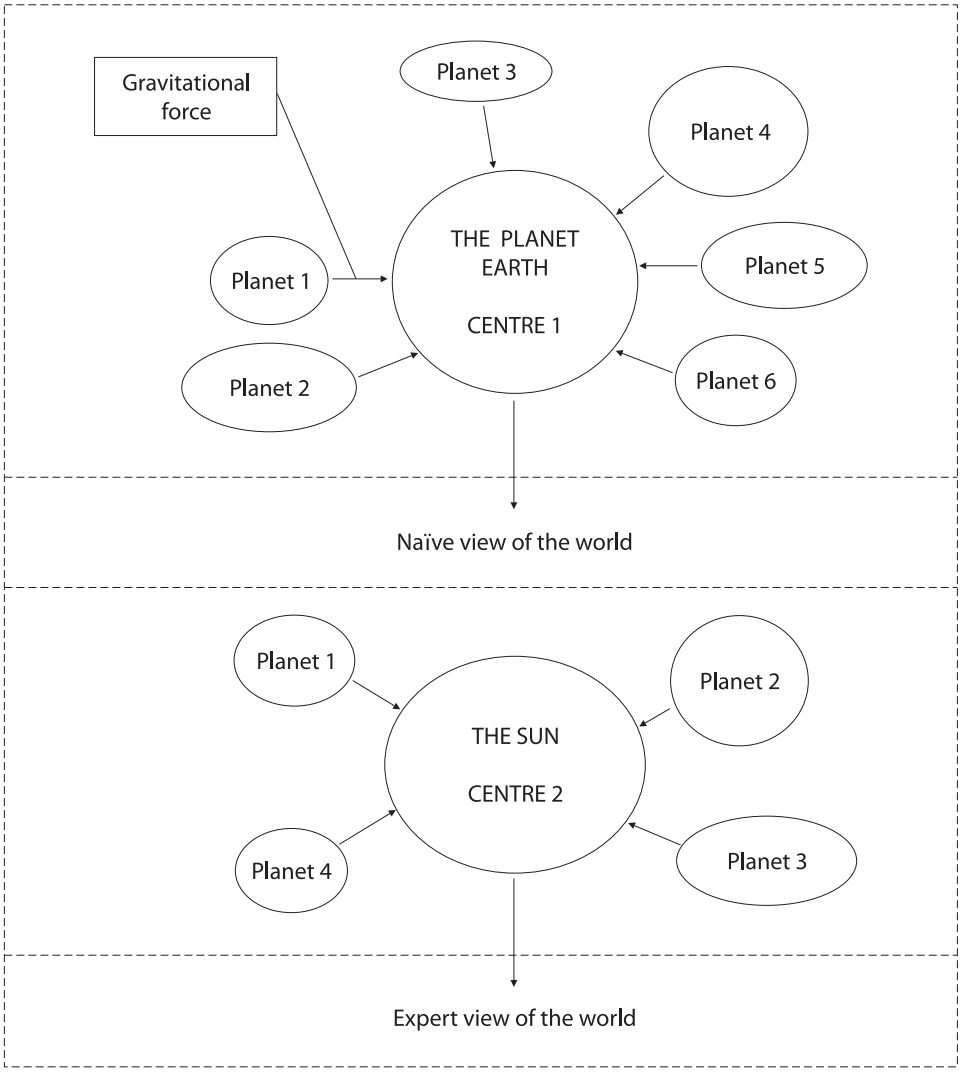


Fig. 58. *Co-existing models of the world*

somewhat natural, given its status of "the lingua franca" of contemporary communication in the world (see sections 6.7.2 and 10.2).

In contrast, under the expert view of the world, we conclude that English is one of many other languages in the universe driven by some central force of change, and remains to the same extent subject to evolutionary change as other languages. This evolutionary fact is perfectly corresponding to the heliocentric view of the world with the Sun viewed as the centre and other planets viewed as affected by the impact of its incessant force.

Another aspect connected with the placement of the Sun in the centre involves the solar radiation and the resulting positive as well as negative scientific facts connected with it. These are intermingled with the folk view of language change, whereby these centrally driven (powered) processes of change may lead to further development of a language (positive radiation), or they may cause a language to die as a result of negative effects of radiation caused by the loss of a 'protective shield' over that language (see section 6.9). For the reason of the aforementioned intermingling of folk and expert knowledge connected with the prospective evaluation of effects of language change, as it may be both the subject of popular (although rare, see Ft. 22 in section 9.2.1) and scientific debate. This issue has been subsumed under the category "expert" and "folk" in Fig. 58.

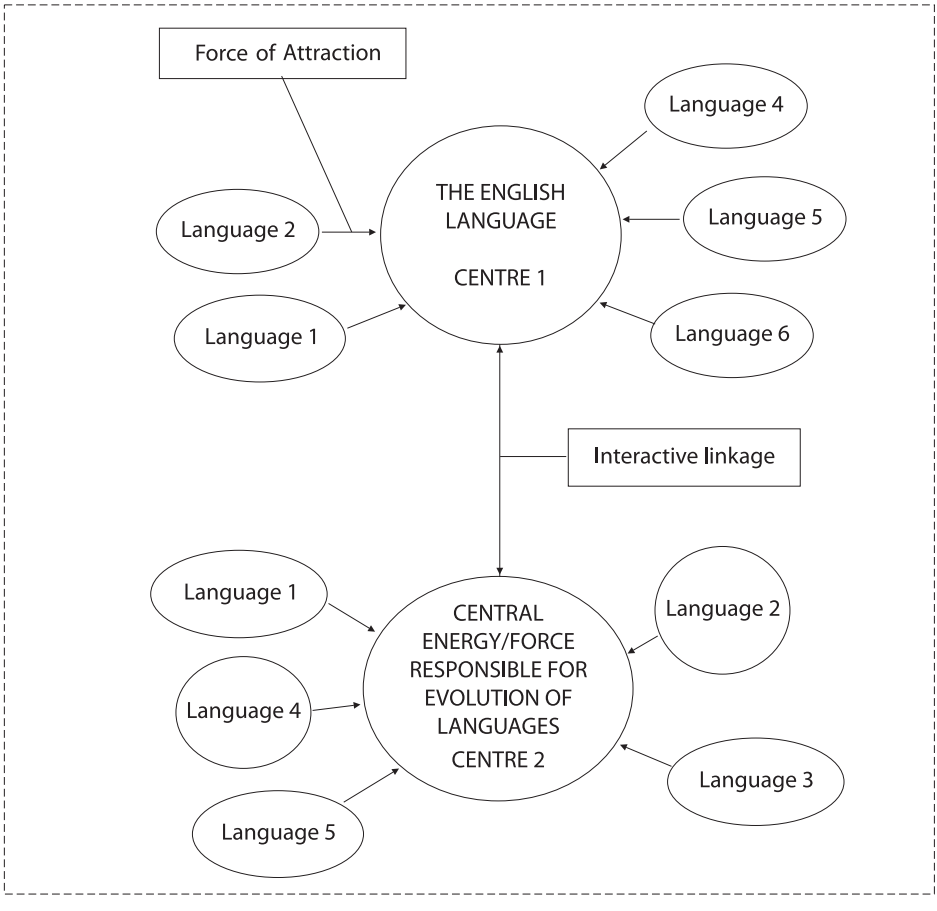


Fig. 59. *The external metalinguistic model of foreign lexical assimilation*

To sum up, the external model of foreign lexical assimilation that accounts for the two parallel views of the world results in the emerging two centres: one

with English as the centre of the world, and the other one that postulates the existence of some central energy that becomes the source of language change (see Fig. 59). Both perspectives can only be manifested at a metadiscursive level of human communication, hence the inherently postulated meta-discursive character of the general metaphor LANGUAGE LAWS ARE PHYSICAL LAWS.

As a conclusion to this section and the previous one, we summarise some selected major factors regulating the metaconceptualisation of foreign lexical assimilation processes from both internal and external perspective (see Tab. 31). More detailed list of correspondences has been presented in section 6.9.

Tab. 31. *Foreign assimilation processes in compact view*

Foreign lexical assimilation processes		
View	Folk	Expert
Internal	Our daily experiences connected with ATTRACTION/REPULSION and UP/DOWN scale with the ensuing regularities (after Krzeszowski 1997): MORE is UP, LESS is DOWN Thus, ATTRACTION is UP, REPULSION is DOWN	Facts connected with circular movement and the ensuing action of centripetal and centrifugal forces
	The 'bigger' the foreign lexical unit, the more chances it will become adapted into the target language system. The 'smaller' the unit, the more chances it will become rejected by the target system	LANGUAGE LAWS ARE PHYSICAL LAWS Some fundamental metaconceptualisations: Physical force as CENTRIPETAL (CENTRIFUGAL) FORCE OF GEOGRAPHICAL PROXIMITY (DISTANCE) between languages Physical force as CENTRIPETAL (CENTRIFUGAL) FORCE OF CULTURAL PROXIMITY (DISTANCE) between languages Physical force as SOCIAL FORCE OF REJECTION (centrifugal) OR ACCEPTANCE (centripetal) Physical force as INSTITUTIONAL FORCE (media, academies, authorities). This force may be both centripetal or centrifugal, depending on circumstances
External	The Earth as the centre of the universe with other celestial bodies subsidiary to it	The Sun as the centre with other planets revolving round it and subject to its influence
		Facts connected with solar radiation
	The English language as occupying the central position in the network system of languages in the world	All languages are equal and subject to the same central force of evolutionary change
	Positive/negative effects of language change	

Chapter 10
SUMMARY AND CONCLUSION

10.0. Overview

Chapter 10, the concluding part of the book, offers a brief summary of the contents of Parts 1–4 of this study (section 10.1.1) and suggests some further prospects in the interdisciplinary research upon foreign lexis (section 10.2). These concern, for example, the integration of the model proposed in the monograph with other theories such as the model of atom (Rutherford-Bohr) or recent developments in IT science (the notion of scale-free networks).

10.1. Summary of Parts 1–4

Part 1 (Introduction and Chapters 1–2) places the study of foreign lexical assimilation within the external framework underlain by the metadiscursive metaphor LANGUAGE LAWS ARE PHYSICAL LAWS. It also looks at the problematics of lexical assimilation studies within the field of lexicology as such and reflects upon the ensuing major concerns. Part 1 also provides an outline history of lexical semantics, and introduces us to the latest proposals offered by Cognitive and Corpus Linguistics in lexicological research.

Part 2 (Chapters 3–5) starts with a short Chapter 3 that discusses what has been said so far, asserted or feared about foreign words and phrases by native speakers of English. Chapter 4 gives an outline history of research upon foreign lexis. It looks at most significant dictionaries and related publications of foreign words and phrases in search of theoretical digressions concerning the status of foreign words and phrases in language, their usage, and categorization criteria. Chapter 5 attempts at identifying sources of methodological disorder observable in the literature on non-native lexis, and then suggests some target typology of foreign words and phrases.

Part 3 (Chapters 6–7) constitutes a theoretic as well as analytical foundation of the book. Chapter 6 elaborates on the central claim stated in Introduction and examines in detail a set of multiple correspondences between the physical and the linguistic world. More specifically, we observe far-going correspondences between such phenomena as corpus frequency of a lemma (understood in this book as CRAC) and the assimilatory potential of that lexical unit. This somewhat trivial observation is explicated by discovering a set of non-trivial analogies between linguistic concepts relevant to the description of assimilation processes and such astrophysical concepts and phenomena as centripetal/centrifugal force, interplan-

etary gravitation, the organization of the solar system along with its most salient celestial bodies, the notion of motion and energy. All these correspondences are compacted into one coherent metacognitive model of description presented in Chapter 9. Chapter 7 is a detailed semantic investigation of source languages (referred to as 'planets' or 'planetoids') and their lexical import into the structure of English. The novelty of the proposal relates to the inclusion, aside from common words and phrases, of a great number of proper foreign words and phrases. The rationale behind this step is supported by the conception of "foreignness" adopted in the book as well as methodology proposed.

Part 4 (Chapters 8–10) integrates findings from Part 3 into one coherent metaconceptual model of lexical assimilation. Chapter 8 gives us a global perspective upon the discussed material presented in Chapter 7. It offers a characterization of tendencies in the semantic typology of foreign lexical units across the source languages listed. The chapter closes with the delivery of some compact view of source lexical systems and their lexical input with regard to the English language. The discussion terminates with the elicitation of prototypical set of foreign words and phrases in use in English at the turn of a new millennium. Chapter 9 concludes the investigation. In this chapter both internal and external models of foreign lexical assimilation are proposed and seen as integrated holistic metacognitve representation of lexical assimilation.

10.2. Integrated model of foreign word representation and assimilation. Conclusion

Foreign words and phrases have always been part of the lexical landscape of a given language. Attitudes to foreign sounding elements in one's native language certainly vary, but, as the research has hopefully demonstrated, the very idea of assimilation should be regarded as absolutely positive phenomenon (sections 9.1 and 9.2).[23] The positive nature of foreign assimilation processes is related to the fundamental working of centripetal force in our life, manifested in terms of force of gravity and potentially represented in metadiscursive talks as the profiling of the ATTRACTION pole of ATTRACTION/REPULSION schema to eventually yield a metacognitive metaphor LEXICAL ASSIMILATION IS ATTRACTION.

We have observed that foreign lexical assimilation processes are to a large extent subject to the analogous phenomena as the ones that abide in the physical world (see particularly Preface, sections 0.2 and 6.9). For the sake of clarity, let us recapitulate the central claim of this book along with the ensuing argumentation restricted to the phenomenon of lexical assimilation:

[23] For the discussion on absolute vs. actual valuation in language, see Krzeszowski (1997).

Language undergoes continual change. Change can be discussed in terms of motion.[24] Language can, thus, be described as being in constant motion. Motion is determined by forces. Forces that determine language change are analogical to physical forces. There are two fundamental types of forces: centripetal (centre-seeking) and centrifugal (away-from-the-centre) forces. These physical forces are present in circular motion. Therefore, language change can analogically be modeled as a circular motion. Languages being in constant motion can be compared to planets. Their mutual interactions are guaranteed by the centripetal force of interplanetary gravity. The atmosphere of a planet (e.g. the Earth [i.e. the English language]) is, however, the residue of opposite centrifugal-like forces. Once an alien body (e.g. a meteor), driven by the force of gravity, enters the Earth's atmosphere, it meets the resistance of densely accumulated air-particles that form the opposite (centrifugal-like) force(s) acting on this alien entity.

The consequence of the claim for the description of lexical assimilation processes is the following:

Successful (pre-)adaptation of a foreign lexical unit (a meteor[ite]) into the target language system (in this book English – the planet Earth) occurs if the sum of centripetal force(s) that governs the incorporation of a foreign word or phrase *into* the target lexical system prevails over the sum of centrifugal force(s) that conserves the system. Breaking through the protective barrier formed by centrifugal-like forces, thus, ensures completion of assimilation process by a foreign word or phrase and its ensuing entrenchment in the target system. Consistently, then, failure to overcome the barrier formed by centrifugal force(s) is analogical to the expulsion of a foreign word or phrase *from* the target system, i.e. its lack of adaptation in the English language.

The scope of correspondences as well as linguistic evidence (CRAC values of the discussed lemmas) attested in the British National Corpus has been noted to be broad enough to legitimize the validity of the principal metadiscursive metaphor LANGUAGE LAWS ARE PHYSICAL LAWS along with the ensuing argumentation. The advantage of the approach in which a search for some external framework is made to explicate how lexical assimilation processes can potentially be represented at metadiscursive level of lexicological description relates to a tentative claim about universality of the proposed model.[25] It is, thus, stipulated

[24] For the discussion of CHANGE is MOTION and CAUSES are FORCES metaphors, see Lakoff (1993: 225).

[25] This observation may lead us to treat the model of foreign lexical assimilation as a manifestation of lexical universals. This seems to be also observed by Rayevska (1979: 240), who states that "the linguistic evidence drawn from the nature of foreign borrowings in a given language presents a peculiar interest in the domain of the so-called lexical universals. Most developed modern languages have a common international fund of words which comprises scientific, technical and sociopolitical terminology, in the main." This commonality may safely be attributed to the universality of mechanisms of physical laws that, as the present book argues, exhibit far-reaching analogies to the laws abiding in the realm of language, strictly foreign lexical assimilation processes. If physical laws constitute a ground for the universalist claims advanced in the book, these are seen as compatible with a cognitive-linguistic approach to language, where it is the nature of human physique that represents a starting point for considerations upon a linguistic organization. As Evans and Green (2006: 54–55) conclude, all these "commonalities are explained by the existence of general cognitive principles shared by all humans, in addition to the fundamentally similar ex-

that any language system perspectivized as the target one, relative to which foreign lexical assimilation is investigated, is susceptible to the same mechanisms of incorporation, adaptation or rejection of foreign words and phrases depending which forces (see section. 6.8.3) prevail at a particular moment. As we have observed in Chapter 6, there may be four major manifestations of centripetal and centrifugal forces existing in language. If, for instance, the centripetal forces of geographical and cultural proximity combine in acting upon a foreign word or phrase, the greater the likelihood that this item will become adapted into the target lexicon. That the forces may be added in the mathematical sense of the word is brought to us by the science of physics, provided the forces act on one and the same entity. The conception of prevailing forces of centripetal or centrifugal force, thus, constitutes the foundation of the processes of foreign lexical assimilation[26] as understood in this book.

Moreover, the incorporation of Ptolemy's and Copernicus' model of the universe as a part of the external metacognitive framework within which foreign lexical assimilation processes are presented can be substantiated inasmuch as the two models both relate, albeit in different ways, to a human bodily experience as a starting point of conceptualization process. The former does so via the reference to anthropocentricity and egocentricity principles (see section 8.4.2.3) which give a human being a privileged position in (meta)conceptualization of the reality, whereas the latter does so in focusing on the mechanisms governing the processes constituting the perceptible composition of the solar system. As such, then, Copernicus' model gives priority to Newton's law of universal gravitation with the concept of force of gravity being one of the fundamental, primary (i.e. preconceptual) experiences of a human being (see Krzeszowski 1997: 130).

Another issue that has not as yet been integrated into a unified framework relates to possible ways of foreign lexical representation in the target lexical system, i.e. the routes that a foreign word or phrase may be modeled to take towards complete adaptation or annihilation in the target language. The ways towards complete adaptation have been extensively discussed in the literature on borrowings (e.g. Görlach 2007, see also sections 5.1 and 5.2.1) and have been subsumed under the three components: 1) graphemic (spelling adapta-

periences of the world also shared by all humans due to embodiment." For more discussion upon universals, see (Jackendoff 1983, Pinker 1994, Li and Gleitman 2002, Wierzbicka 1996).

[26] It must be reminded that the term assimilation as used in the book must not be confused with the notion of adaptation of a foreign lexical item in the target lexical system. "Adaptation" relates to the idea of complete integration of the unit with the system of the receiving language and is thus more suitable a term in the discussion of borrowings, whereas 'assimilation' is understood in the study as a cover term to encompass such notions as incorporation, preadaptation (prototypical foreignness), adaptation as well as non-adaptation, i.e. relegation of a foreign term out of the scope of the target lexical system (see section 9.1.1).

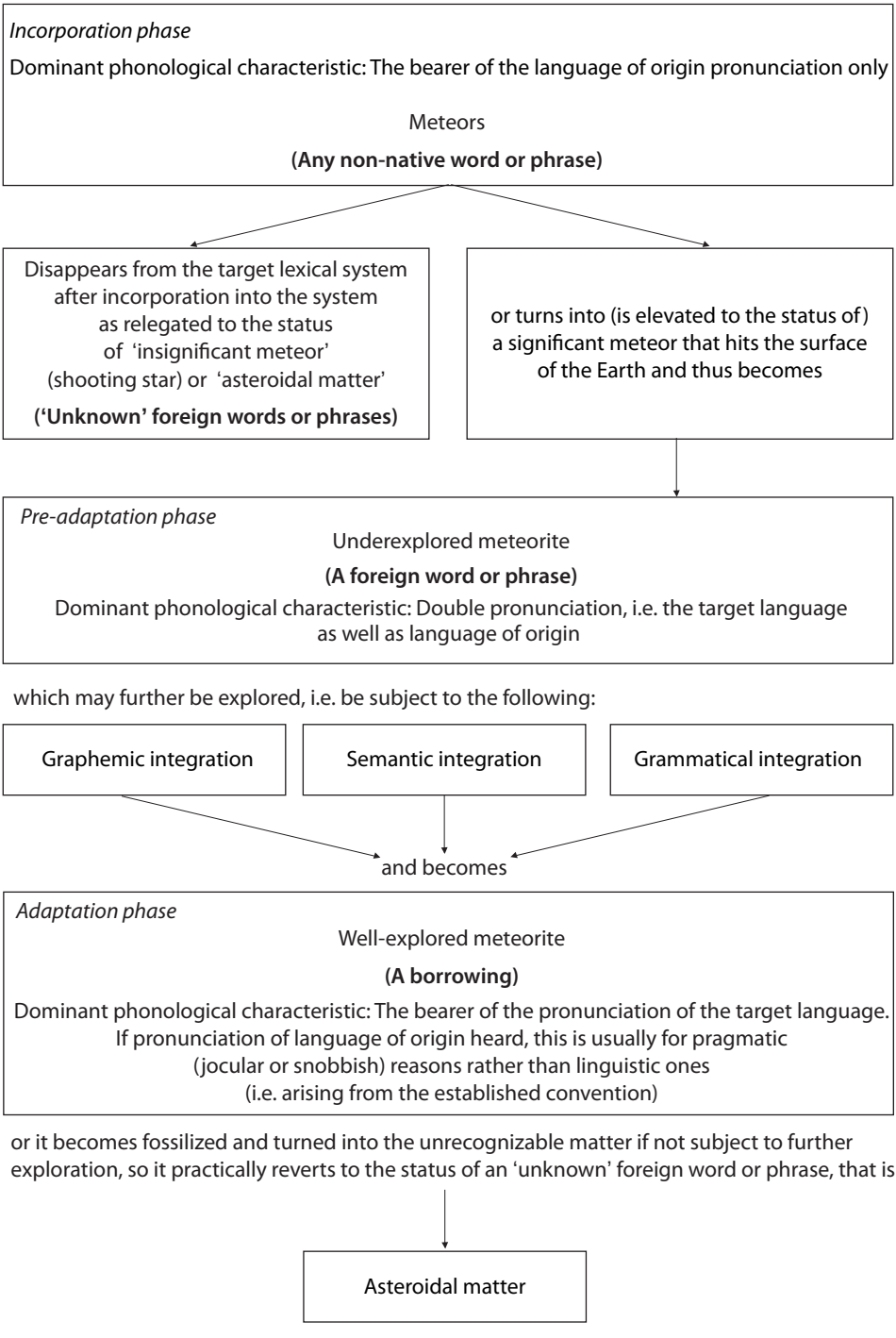


Fig. 60. *An integrated model of foreign word representation and assimilation in the target lexical system*

tion),[27] 2) semantic integration (adaptation), and 3) grammatical integration (both morphophonological and morphosyntactic). These stages, on account of the aforementioned extensive research already done in the field (e.g. Görlach, Mańczak-Wohlfeld, Fisiak) do not, thus, directly constitute the focal point of consideration in the present book (see Fig. 14 in section 5.3).

We are now at the stage when we may attempt at outlining this processual-propositional and metaconceptual model of foreign lexical representation in the target system (see Fig. 60). This model is an extension of Fig. 13 from Chapter 5 and involves the addition of elements of the descriptive framework developed in Chapter 6 (sections 6.7.3 and 6.9). We may, thus, see how the phonological criterion that constitutes the basis for data collection in the book intermingles with the graphemic criterion upon which, in turn, the BNC frequency data retrieval is hinged (see Fig. 1, section 0.1). Of course, it must be remembered that the borderlines between the identified stages are essentially fuzzy. In actual practice it is often difficult to determine *ad hoc* whether a given word or phrase counts as a borrowing or a foreign word.

As it is argued above these category assignment limitations can, however, be significantly overcome, once we adopt clear-cut data collection criterion (here phonological one) for the analysis, and provided that we apply objectivized tools which enable the prediction whether a particular lexical item carries the potential of being fully adapted to the target language system or not.

10.3. Alternative models of lexical assimilation

The question that remains open is whether the model of lexical assimilation as encapsulated within the framework of the astrophysical metaphor is the only adequate model of representation for the discussed processes. A cursory glance at other models operating in the physical science shows that a negative answer to the query may possibly be found. One such alternative conception is, for example, the model of atom structure by Rutherford-Bohr (see Fig. 61).

It appears, then, that the model of atom structure in its quintessence is consonant with Newton's laws of dynamics in that neutrons are modeled as particles orbiting round the nucleus in the same fashion as planets orbit round the Sun. As such, then, the model could successfully serve as a sound analogical framework within the domain of physical existence for the metaconceptual representation of foreign lexical assimilation proposed in the book. It is believed that more conceptions discussed originally in the realm of physics and the associated sciences

[27] McEnery & Wilson (2004: 43) notice that native speakers often adopt the strategy of omission of diacritic marks in writing when the issue of accent or non-Roman alphabet arises.

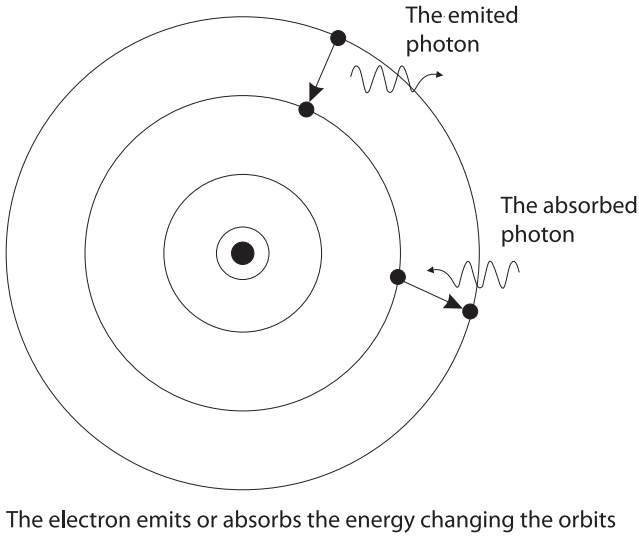


Fig. 61. *Model of atom structure by Rutherford-Bohr*[28]

could be helpful in modeling the processes of lexical assimilation. Such research, however, goes beyond the scope of the present study; nonetheless, one more especially intriguing example of such prospective avenue of investigation is discussed below.

Namely, the model of foreign lexical assimilation could tentatively be rendered through the reference to a quite popular theoretical framework developed in the contemporary cybernetics, i.e., so-called "scale-free network." A few quotations below taken from the famous article *Scale-free networks*[29] introduce us with the idea of scale-free network as such:

> Scientists have recently discovered that various complex systems have an underlying architecture governed by shared organizing principles. This insight has important implications for a host of applications, from drug development to Internet security […]. The brain is a network of nerve cells connected by axons, and cells themselves are networks of molecules connected by biochemical reactions. Societies, too, are networks of people linked by friendships, familial relationships and professional ties. On a larger scale, food webs and ecosystems can be represented as networks of species. And networks pervade technology: the Internet, power grids and transportations systems are but a few examples. Even the language we are using to convey thoughts to you is a network, made up of words connected by syntactic relationships. […] Yet despite the importance and pervasiveness of networks, scientists have had little understanding of their structure and properties. […] Over the past few years, investigators from a variety of fields have discovered that many networks – from the World Wide Web to a cell's metabolic system

[28] www.library.thinkquest.org (ED: 08/08).

[29] The article *Scale-Free Networks*. By: Barabasi, Albert-Laszlo, Bonabeau, Eric. *Scientific American,* 00368733, May 2003, Vol. 288, Issue 5 (www.web.ebscohost.com [ED: 04/08]).

> to actors in Hollywood – are dominated by a relatively small number of nodes that are connected to many other sites. Networks containing such important nodes, or hubs, tend to be what we call "scale-free" in the sense that some hubs have a seemingly unlimited number of links and no node is typical of the others. These networks also behave in certain predictable ways; for example, they are remarkably resistant to accidental failures but extremely vulnerable to coordinated attacks. […] Unexplained by previous network theories, hubs offer convincing proof that various complex systems have a strict architecture, ruled by fundamental laws – laws that appear to apply equally to cells, computers, languages and society.[30]

One of eminent examples of "scale-free network" is definitely the World Wide Web where "a few highly connected pages are essentially holding the World Wide Web together. More than 80% of the pages on the map had fewer than four links, but a small minority, less than 0.001% of all nodes, had more than 1000."[31] It is very interesting to observe that scale-free structures have been discovered in a wide range of other systems. For example scale-free networks have been recorded to exist in the realm of business, biology (cellular-metabolic network), the protein-interaction network as well as social networks. The last one can be exemplified by a "network of sexual relationships among people in Sweden."[32] This network followed what the researchers labelled a power law, that is, "although most individuals had only a few sexual partners during their lifetime, a few (the hubs) had hundreds."

Another interesting principle upon which the idea of "scale-free network" is hinged is the so-called "the rich get richer principle." This reads as follows:

> Most of us are familiar with only a tiny fraction of the full Web, and that subset tends to include the more connected sites because they are easier to find. By simply linking to those nodes, people exercise and reinforce a bias toward them. This process of "preferential attachment" occurs elsewhere. In Hollywood the more connected actors are more likely to be chosen for new roles. On the Internet the more connected routers, which typically have greater bandwidth, are more desirable for new users. In the U.S. biotech industry, well-established companies such as Genzyme tend to attract more alliances, which further increases their desirability for future partnerships. Likewise, the most cited articles in the scientific literature stimulate even more researchers to read and cite them, a phenomenon that noted sociologist Robert K. Merton called the Matthew effect, after a passage in the New Testament: "For unto every one that hath shall be given, and he shall have abundance." These two mechanisms – growth and preferential attachment – help to explain the existence of hubs: as new nodes appear, they tend to connect to the more connected sites, and these popular locations thus acquire more links over time than their less connected neighbours. And this "rich get richer" process will generally favour the early nodes, which are more likely to eventually become hubs.[33]

[30] *Ibidem*.
[31] *Ibidem*.
[32] *Ibidem*.
[33] *Ibidem*.

The question that remains to be resolved is to what extent language rules are subject to the same principle that constitutes scale-free networks illustrated above. The positive answer that may be provided is already hinted at in the above excerpt. It may further be claimed that the application of scale-free network to the external model of lexical assimilation as discussed in section 9.2 may also positively verify the assumption made in the quoted article. When we look at the network of languages with English postulated as CENTRE1, we notice that it can be successfully ascribed the status of a hub in the sense of the term discussed above. In examining volumes of planets or planetoids orbiting round the Earth, i.e. the English language (see section 8.1), we noticed that the degree of interaction of English with other languages is by no means the same, and can be claimed as largely conditioned by the action of four principal types of centripetal and centrifugal force discussed above. The action of these forces exerts, in turn, a direct impact on the overall volume and mass of the planet as indicated in the aforementioned section 8.1. In effect, we have a few languages (planets) that shape the overall foreign lexical composition of the contemporary English.

The role of other languages, although quite numerous, is actually marginal in that respect. We can, thus, see that the main donor languages 'planets' can successfully play the role of 'hubs' in the network system upon which the foreign lexical enrichment of English is based. As a reference, we look at Tab. 14 (section 8.1) on the status of languages in respect of *V* criterion. *V* criterion relates to the number of individual lemmas pertinent to a particular language recorded as foreign in the English language according to LPD. As a result, the more lemmas a given donor language boasts with respect to the target language (in this book the English language), the more connections it has with it. Quite arbitrarily, then, the languages referred to as 'planets' have been automatically assigned the status of 'hubs.' The situation can be observed from Fig. 62.

Postulating the English language as the central 'hub' in the network does not seem to stem only and exclusively from the English language-oriented perspective adopted in the book. It also appears to be consonant with the popular belief about the status of English as a lingua franca of the contemporary world (see, e.g. Strevens 1972, Crystal 2007a; see also section 6.7.2).

Interesting, however, might be the investigation of possible future development of the network with new languages gaining the status of 'hubs' while others losing or preserving it relative to the English language. It may be stipulated that the category of "mass" of a language elaborated on in the book (see, e.g. sections 6.1 and 8.1) may function, among others, as an interesting criterion in that respect. We say "among others" because another factor prompting the future development of the network is the inflow of new foreign words and phrases recorded in the 2004 edition of LPD in comparison with the 1996 version.

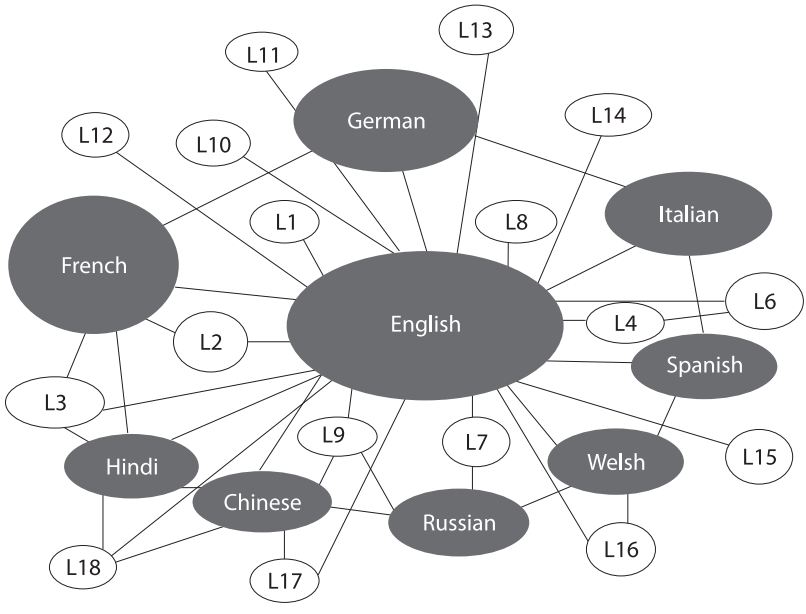


Fig. 62. *Scale-free network in the realm of languages*[34]

Coming back to the concept of "mass," we have observed so far that the volume of a planet does not need to correspond in a direct way to its mass. It is to be reminded that the notion of "mass" relates to the actual assimilatory potential of language calculated by the formula: The number of individual lemmas pertinent to a donor language times CRACn2 value. It is, thus, a significant indicator of the dynamicity of relationship of a given language in relation to English. The notion of 'dynamicity' must be carefully distinguished from the notion of 'intensity' in which case it is the criterion of *V* that gains significance, as it is connected with the number of lemmas associated with a given language. Let us, then, once more, refer to a selected portion of data presented in Tab. 13 in which we have some basic statistics presented involving the entire volume (*V*) and mass (*M*) of languages ('planets') affecting the English language:

Tab. 32. *Volume and mass of planets and planetoids recapitulated*

Language	V	M
French	1146	48778
German	505	27215

[34] Black areas indicate 'hubs' in the network, whereas other areas labelled as language 1 (L1), language 2 (L2), etc. are viewed as satellites with peripheral significance to the system. Of course due to editorial reason the number of 'satellite' languages is smaller than the actual number of languages identified as affecting the English language, i.e. 63.

Tab. 32 – cont.

Italian	319	11303
Spanish	280	17815
Welsh	231	4767
Russian	163	8560
Chinese	117	4821
Hindi	85	2206
Dutch	63	4410
Japanese	63	3402
Arabic	48	8310
Portuguese	44	1278
Greek	32	522
Polish	26	1338
Danish	25	680

We can, in effect, formulate at least three regularities that emerge from the interpretation of data presented above. The regularities read as follows:

Regularity 11

The lesser the discrepancy between *V* and *M* value as calculated for a particular language, the greater the chances that the language will preserve its status with regard to English in a scale-free network in the non-definite future.

Regularity 12

The greater the mass of a language relative to its volume, the greater the chances that a role of a given language will become more significant with respect to English in the non-definite future.

Regularity 13

The greater the volume of a language relative to its mass, the greater the chances that a role of a given language will become less significant with respect to English in the non-definite future.

On the basis of Regularity 11, we conclude that the status of French, German, Italian, and Spanish as 'hubs' seems to be constant with respects to English as their *V*/*M* ratios are considerably greater relative to other languages. On the basis of Regularity 12, we conclude that Welsh is in danger of losing its status of a 'hub' in the evolving network of influential languages relative to English due to a proportionally high discrepancy between *M* and *V* value in favour of the latter (Regularity 13).

Chinese, Russian, and especially Arabic appear to be on the opposite trend. For example, Arabic boasts a huge discrepancy between *V* value and *M* value in favor of the latter, which indicates an extreme dynamicity of assimilation with

the simultaneous lesser intensity when it comes to the number of individual words subject to adaptation. It may, thus, be claimed that the very dynamicity of the contact between Arabic and English observed at the turn of the 21st century may result in more intense relationship in the future, accordingly with the commonsense rule: dynamicity ensures intensity of the relationship in the long run. This rule can be intuitively felt to be operational in our daily personal or business contacts, whereby the energy we put in contacting one another usually effects in more frequent encounters (intensity) in the future. More specific examples may involve intimate contacts between people who are in love with each other, or on a less personal basis, economic contacts between countries.

Conversely, the lesser the dynamicity of contacts there is between two sides (in our research, the greater the discrepancy between *V* and *M* value in favor of the former relative to *V*/*M* ratios of other languages), the lesser their prospective intensity. This is the case of Welsh, which although still quite numerously represented in English in terms of individual lemmas (*V* value), is in danger of losing the status of an influential donor language (hub) as its mass is relatively smaller than, for example Arabic. The predictions for Welsh are, thus, pessimistic and it can be stipulated that this language will lose its significant status as a foreign lexical 'deliverer' in the non-definite future, unless some measures will be taken to prevent it. One such possible measure may be more extensive propagation of Welsh not only on its home territory but also outside.

A few words must also be spared regarding the native language spoken by the author of the present book, i.e. Polish. The Polish language is still relatively insignificant as a donor language for English, but prospects of the Polish language contribution to the lexis of English seem to be optimistic in that respect. We see that the discrepancy between *V* and *M* values is quite discernible in favour of *M*, which indicates that Polish may become more and more represented in English in the future. It appears that following the opening of the British job market to Polish people and the increasing number of Polish immigrants on the British Isles, the chances that the position of the Polish language as the average 'player' in delivering foreign lexis to English lexicon are higher. It is believed, however, that the confirmation of this prediction will only be possible no sooner than after examining new editions of LPD issued in the next decade of the 21st century.

The astrophysical metaphoric model of foreign lexical assimilation as proposed in the book, along with the postulated repertoire of tools, such as, e.g. CRACn (1, 2 and 3), the concept of *V*, the concept of *M*, may offer some alternative proposals of investigating foreign lexis of a given language as based on a clear phonological criterion. One of the potential advantages behind this approach is a tentative predictive power of the postulated model. Establishing CRACn1 and CRACn3 values, thus, allows us to more or less clearly demarcate between lexical items carrying higher chances of adaptation into the target lexicon (borrowings) and those whose status is undecided (prototypical foreign words) or decided in

the sense that the status may be deemed as accidental in the lexicon with no great chance of adaptation ('unknown' words and phrases). It is in the description of this central stage of the process (i.e. prototypicality of foreignness and the resulting process of a foreign lexical incorporation into the target lexical system) that the novelty of the proposal can be ascribed to.

The predictive power of the model has also been extrapolated from the analysis of individual lemmas onto a more general ground as the discussion in this chapter has hopefully demonstrated. Thus, on the basis of interacting categories *V* and *M*, we were able to determine the current status of a donor language and make tentative predictions regarding the relative significance of this language with respect to English in the future. Such predictions can tentatively be postulated on account of our discussion of scale-free network, within the framework of which the advanced model of foreign lexical assimilation appears to be validated. It is believed that this predictive potential behind the model along with its explanatory power based on the discussed multifarious analogies between the world of physics and the world of linguistics (section 6.9) has helped the reader of this book at least to minimally uncover some of the so far unresolved queries connected with the description of foreign lexis in the English language.

REFERENCES

Achard, M. 1996. 'French modals and speaker control.' In: A. Goldberg (ed.), *Conceptual Structure, Discourse and Language*. Stanford, CA.: CSL & I, 1–15.

Adamska-Sałaciak, A. 1996. *Language Change in the Works of Kruszewski, Baudouin de Courtenay and Rozwadowski*. Poznań: MOTIVEX.

Ahlsén, E. 2006. *Introduction to Neurolinguistics.* Amsterdam, Philadelphia: John Benjamins Publishing Company.

Aitchison, J. 1994. *Language Joyriding.* Oxford: Clarendon Press.

Aitchison, J. 1996 [1987]. *Words in the Mind. An Introduction to the Mental Lexicon.* 2nd edition. Oxford, Cambridge: Blackwell Publishers.

Alexieva, N. 2007. 'Bulgarian.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 241–260.

Algeo, J. 1973. *On Defining the Proper Name.* Gainesville: University of Florida Press.

Allerton, D.J. 1987. 'The linguistic and sociolinguistic status of proper names.' *Journal of Pragmatics 11*, 61–92.

Allwood, J. 2003. 'Meaning potentials and context: some consequences for the analysis of variation in meaning.' In: H. Cuyckens, R. Dirven, J.R. Taylor (eds.), *Cognitive Approaches to Lexical Semantics.* Berlin, New York: Mouton de Gruyter, 29–65.

Allwood, J., Andersson, L.G. & Dähl, Ö. 1977. *Logic in Linguistics.* Cambridge: Cambridge University Press.

Altenberg, B. & Granger, S. 2002. 'Recent trends in cross-linguistic lexical studies.' In: B. Altenberg (ed.), *Lexis in Contrast. Corpus-based Approaches.* Philadelphia: John Benjamins Publishing Company, 3–48.

Apresjan, J.D. 2008 [2000]. 'On the language of explications and semantic primitives.' In: P. Hanks (ed.), *Lexicology*, vol. III. London, New York: Routledge, 14–32.

Aristotle. 1949. 'Categoriae et Liber De Interpretatione.' In: L. Minio-Paluello (ed.), Oxford: Oxford University Press.

Aristotle. 2008. 'Meaning and essence. Excerpts from Aristotle's writings.' In: P. Hanks, *Lexicology*, vol. I. London and New York: Routledge.

Aston, G., Burnard, L. 1998. *The BNC Handbook. Exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.

Atkins, B.T.S., Kegl, J. & Levin, B. 1988. 'Anatomy of a verb entry: from linguistic theory to lexicographic practice.' *International Journal of Lexicography 1* (2), 84–126.

Atkins, B.T.S., Levin, B. & Zampolli, A. 1994. 'Computational approaches to the lexicon: an overview.' In: B.T.S. Atkins & A. Zampolli, *Computational Approaches to the Lexicon*. Oxford and New York: Oxford University Press, 17–45.

Atkins, B.T.S. & Levin, B. 1995. 'Building on a corpus: a linguistic and lexicographical look at some near synonyms.' *International Journal of Lexicography 8* (2), 85–114.

Ayto, J. 2007 [1999]. *A Century of New Words.* Oxford: Oxford University Press.

Baldinger, K. 1980. *Semantic Theory. Towards a Modern Semantics.* Oxford: Basil Blackwell.

Battarbee, K. 2007. 'Finnish.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 261–276.

Bauer, L., Nation, P. 1993. 'Word families.' *International Journal of Lexicography 6* (4), 253–279.

Bauer, L. 1998. *Vocabulary.* Florence, KY: Routledge.

Becker, H. 1948. *Der Sprachbund.* Leipzig: Humboldt Bücherei Gerhard Mindt.

Bencini, G. & Goldberg, A.E. 2000. 'The contribution of argument structure constructions to sentence meaning.' *Journal of Memory and Language 43* (4), 640–651.

Berezowski, L. 2001. *Articles and Proper Names*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Berglund, Y. 1997. 'Future in present-day English: corpus-based evidence on the rivalry of expressions.' *ICAME Journal 21*, 7–20.
Berlin, B. 1978. 'Ethnobiological classification.' In: E. Rosch & B.B. Lloyd (eds.), *Cognition and Categorization.* Hillsdale, NJ: Lawrence Erlbaum, 9–26.
Berteloot, A. & van der Sijs, N. 2007. 'Dutch.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 37–56.
Betz, W. 1949. *Deutsch und Lateinisch: die Lehnbildungen der altochdeutschen Benediktinerregel.* Bonn: H. Bouvier.
Biber, D., Conrad, S., Reppen, R. 1998. *Corpus Linguistics. Investigating Language Structure and Use.* Cambridge: Cambridge University Press.
Bliss, A.J. 1966. *Dictionary of Foreign Words and Phrases in Current English.* London, New York: Routledge.
Bloomfield, L. 1933. *Language.* New York: Holt, Rinehart and Winston.
Boas, F. 1940. *Race, Language, and Culture.* New York: Macmillan.
Boese, H. 1988. *Common Allusions and Foreign Terms.* Redlands, California: Simplicity Press.
Bogusławski, A. 2008 [1970]. 'On semantic primitives and meaningfulness.' In: P. Hanks (ed.), *Lexicology*, vol. III. London, New York: Routledge, 3–13.
Bolinger, D. 2008 [1965]. 'The Atomization of Meaning.' *Language 41*(4): 555–573.
Boye, K. 2001. 'The Force-Dynamic core meaning of Danish modal verbs.' *Acta Linguistica Hafniensia 33*, 19–66.
Braisby, N. 2008 [1990]. 'Situating word meaning.' In: P. Hanks (ed.), *Lexicology*, vol. V. London, New York: Routledge, 194–220.
Brandt, P.A. 1992. *La charpente modal du sens: pour une simio-linguistique morphogenitique et dynamique.* Amsterdam: Benjamins.
Bréal, M. 1897. *Essai de Sémantique.* Paris: Hachette.
Brown, G.D., Rice, S. 2007. *Professional English in Use Law.* Cambridge: Cambridge University Press.
Brown, G. & Yule, G. 1983. *Discourse Analysis*. Cambridge, New York: Cambridge University Press.
Brugman, C. & Lakoff, G. 2006. 'Cognitive topology and lexical networks.' In: D. Geeraerts (ed.), *Cognitive Linguistics: Basic Readings.* Berlin, New York: Mouton de Gruyter, 109–140.
Bühler, K. 1965 [1934]. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Stuttgart: UTB.
Burton-Roberts, N. 1986. *Analysing Sentences. An Introduction to English Syntax.* London, New York: Longman.
Busse, U. & Görlach, M. 2007 [2002]. 'German.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 12–36.
Buttler, D., Kurkowska, H., Satkiewicz, H. 1973. *Kultura języka.* Warszawa: PWN.
Bybee, J.L. 2006. *Frequency of Use and Organization of Language.* Oxford: Oxford University Press.
Bybee, J.L. & Eddington, D. 2006. 'A usage-based exemplar model approach to Spanish verbs of 'becoming.' *Language 82*, 323–355.
Carnap, R. *Introduction to Semantics.* Cambridge, Mass.: Harvard University Press.
Chomsky, N. 1957. *Syntactic Structures.* The Hague: Mouton.
Chomsky, N. 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT.
Chomsky, N. & Halle, M. 1968. *The Sound Pattern of English.* New York: Harper and Row.
Chun, S.A. & Zubin, D.A. 1990. 'Experiential vs. agentive constructions in Korean narrative.' *Proceedings of the Berkeley Linguistics Society 16*, 81–93.
Cienkowski, W. 1964. 'Ogólne założenia metodologiczne badania zapożyczeń leksykalnych.' *Poradnik Językowy*, 417–429.
Clyne, M. 1987. 'History of research on language contact.' In: H. von Urlich Ammon, N. Dittmar and K.J. Mattheier (eds.), *Sociolinguisitcs – Sociolinguistik,* vol. 1. Berlin: Walter de Gruyter, 452–459.

Coates, R. 2006. 'Properhood.' *Language 82*, 356–382.
Collins, A. & Loftus, E. 1975. 'A spreading-activation theory of semantic processing.' *Psychological Review 82*, 407–428.
Collins, A. & Quillian, M. 1969. 'Retrieval time from semantic memory.' *Journal of Verbal Learning and Verbal Behavior 8*, 240–247.
Constantinescu, I., Popovici, V. & Ştefânescu, A. 2007. 'Romanian.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 168–194.
Coseriu, E. 1970. 'Die lexematischen Strukturen.' In: E. Coseriu, *Sprache, Syrukture, und Funktionen.* Tübingen: Narr, 159–179.
Coseriu, E. & Geckeler, H. 1981. *Trends in Structural Semantics.* Tübingen: Narr.
Croft, W. 1999. 'Some contributions of typology to cognitive linguistics, and vice versa'. In: T. Janssen and G. Redeker (eds.), *Cognitive Linguistics: Foundations, Scope, and Methodology*. (Cognitive Linguistic Research 15.) Berlin, New York: Mouton de Gruyter, 61–93.
Croft, W. & Wood, E. 2000. 'Construal operations in linguistics and artificial intelligence.' In: L. Albertazzi (ed.), *Meaning and Cognition.* Amsterdam, Philadelphia: Benjamins, 51–78.
Cruse, D.A. 1986. *Lexical Semantics.* Cambridge: Cambridge University Press.
Cruse, D.A. 2000. *Meaning in Language: An Introduction to Semantics and Pragmatics.* Oxford: Oxford University Press.
Crystal, D. 2003 [1995]. *The Cambridge Encyclopaedia of the English Language.* Cambridge: Cambridge University Press.
Crystal, D. 2005. *The Stories of English.* London, New York: Penguin Books.
Crystal, D. 2007a [2000]. *Language Death.* Cambridge: Cambridge University Press.
Crystal, D. 2007b. *Words, Words, Words.* Oxford: Oxford University Press.
Crystal, D. 2007c. *The Fight for English.* Oxford: Oxford University Press.
Darbyshire, A.E. 1967. *A Description of English.* London: Edward Arnold Publishers Ltd.
Deane, P.D. 1992. 'Polysemy as the consequence of internal conceptual complexity: the case of over.' In: *Proceedings of the Eastern States Conference on Linguistics (ESCOL)* 9, 32–43.
De Haan, P. 1992. 'The optimum corpus sample size?' In: G. Leitner (ed.), *New Dimensions in English Language Corpora.* Berlin: Mouton de Gruyter, 3–19.
Demiański, M. 2008. 'O początku i ewolucji wszechświata.' In: E. Dobierzewska-Mozrzymas, J. Gajda-Krynicka, A. Jezierski (eds.), *Ku interdyscyplinarności… Różne oblicza rzeczywistości.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, 11–23.
De Mulder, W. 2007. 'Force dynamics.' In: D. Geeraerts & H. Cuyckens, *The Oxford Handbook of Cognitive Linguistics.* Oxford: Oxford University Press, 294–313.
Deroy, L. 1956. *L'emprunt linguistique.* Les Belles Lettres.
Diller, K.C. 1978. *The Language Teaching Controversy.* Rowley, Mass.: Newbury House.
Dirven, R., Goosens, L., Putseys. Y., Vorlat, E. 1982. *The Scene of Linguistic Action and Its Perspectivization by* Speak*,* Talk*,* Say*, and* Tell*.* Amsterdam, Philadelphia: Benjamins.
Dirven, R. & Taylor, J. 1988. 'The conceptualisation of vertical space in English: the case of *tall*.' In: B. Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics.* Amsterdam, Philadelphia: Benjamins, 379–402.
Dirven, R. & Verspoor, M. 2004. *Cognitive Exploration of Language and Linguistics.* Amsterdam, Philadelphia: John Benjamins Publishing Company.
Di Sciullo, A.M. & Williams, E. 1987. *On the Definition of Word.* Cambridge, Mass.: MIT Press.
Divjak, D. 2006. 'Ways of intending: delineating and structuring near-synonyms.' In: S.Th. Gries, A. Stefanowitsch (eds.), *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis.* Berlin, New York: Mouton de Gruyter, 19–56.
Dixon, R.M.W. & Aikhenvald, A.Y. 2007. 'Word – a typological framework.' In: R.M.W. Dixon & A.Y. Aikhenvald (eds.), *Word. A Cross-Linguistic Typology.* Cambridge: Cambridge University Press, 1–41.

Donnellan, K. 1972. 'Proper names and identifying descriptions.' In: D. Davidson & G. Harman (eds.), *Semantics of Natural Language*. Dordrecht: Reidel, 356–379.
Doroszewski, W. 1938. *Język polski w Stanach Zjednoczonych*. A.P. Warszawa: Prace Towarzystwa Naukowego Warszawskiego.
Dummett, M. 1973. *Frege: Philosophy of Language*. London: Duckworth.
Dustoor, P.E. 1968. *The World of Words*. Bombay, Calcutta: Asia Publishing House.
Evans, V., Green, M. 2006. *An Introduction to Cognitive Linguistics.* Edinburgh: Edinburgh University Press.
Faber, P. & Mairal Usón, R. 1997. 'Towards a typology of predicate schemata in a Functional-Lexematic Model.' In: G. Wojtak (ed.), *Toward a Functional Lexicology*. Frankfurt am Main: Peter Lang, 11–36.
Farkas, J. & Kniezsa, V. 2007. 'Hungarian.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 277–290.
Fauconnier, G. & Turner, M. 1996. 'Blending as a central process of grammar.' In: A.E. Goldberg (ed.), *Conceptual Structure, Discourse, and Language*. CA: CSLI Publications, 113–131.
Fay, D. & Cutler, A. 1977. 'Malapropisms and the structure of the mental lexicon.' *Linguistic Inquiry 8*, 505–520.
Filipović, R. 2007. 'Croatian.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 229–240.
Fillmore, Ch. 1977. 'Scenes-and-frames semantics.' In: A. Zampolli (ed.), *Linguistic Structures Processing.* Amsterdam: North-Holland Publishing Company, 55–81.
Fillmore, Ch. 1982. 'Toward a descriptive framework of spatial deixis.' In: R.J. Jarvella, & W. Klein (eds.), *Speech, Place, and Action. Studies in Deixis and Related Topics.* Chichester: John Wiley, 31–59.
Fillmore, Ch. 2008 [1975]. 'An alternative to checklist theories of meaning.' In: P. Hanks (ed.), *Lexicology*, vol. IV. London, New York: Routledge, 219–226.
Finkenstaedt, T. & Wolff, D. 1973. *Ordered Profusion: Studies in Dictionaries and the English Lexicon.* Heidelberg: Carl Winter.
Fisiak, J. 1970. 'The semantics of English loanwords in Polish.' *SAP 2*, 41–49.
Fisiak, J. 1985a. *Wstęp do współczesnych teorii lingwistycznych.* Warszawa: Wydawnictwa Szkolne i Pedagogiczne.
Fisiak, J. 1985b. 'A note on the adaptation of English loanwords in Polish.' *ITL Review of Applied Linguistics 67–68,* 69–75.
Fisiak, J. 1986. 'The word-formation of English loanwords in Polish.' In: W. Viereck and W.D. Bald (eds.), *English in Contact with Other Languages: Studies in Honour of Broder Carstensen on the Occasion of his 60th Birthday*. Budapest: Akadémiai Kiadó, 253–63.
Fisiak, J. 2000. *An Outline History of English. Volume One. External History.* 3rd edition. Poznań: Wydawnictwo Poznańskie.
Fodor, J., Fodor, J.A. & Garrett, M. 1975. 'The psychological unreality of semantic representations.' *Linguistic Inquiry 6*, 515–531.
Forster, K. 1976. 'Accessing the mental lexicon.' In: T. Wales & E. Walker (eds.), *New Approaches to Language Mechanisms.* Amsterdam: North Holland, 139–174.
Foster, B. 1981. *The Changing English Language.* London and Baringstoke: The Macmillan Press LTD.
Francis, N. 1992. 'Language corpora B.C.' In: J. Svartvik, *Directions in Corpus Linguistics.* Berlin: Mouton de Gruyter, 17–32.
Fraurud, K. 1996. 'Cognitive ontology and NP form.' In: T. Fretheim & J. Gundel (eds.), *Reference and Referent Accessibility.* Amsterdam: John Benjamins, 65–87.
Frege, G. 1892. 'On sense and reference.' In: F. Zabeh (ed., *et al.*), *Readings in Semantics.* Urbana: University of Illinois Press, 117–140.

Furdal, A. 2000. *Językoznawstwo otwarte*. Wrocław–Warszawa–Kraków: Zakład Narodowy im. Ossolińskich – Wydawnictwo.
Geach, P.T. 1962. *Reference and Generality.* Ithaca, N.Y.: Cornell University Press.
Gee, J. & Kegl, J. 1982. 'Semantic perspicuity and the locative hypothesis.' In: *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society.* Berkeley, Calif.: Berkeley Linguistics Society.
Geeraerts, D. 1985. *Paradigm and Paradox. Explorations into a Paradigmatic Theory of Meaning and Its Epistemological Background*. Leuven: Universitaire Pers.
Geeraerts, D. 1988. 'Where does prototypicality come from?' In: B. Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics.* Amsterdam, Philadelphia: John Benjamins, 207–229.
Geeraerts, D. 1989. 'Introduction: prospects and problems of prototype theory.' *Linguistics 27*, 587–612.
Geeraerts, D. 2006a. 'A rough guide to Cognitive Linguistics.' In: D. Geeraerts (ed.), *Cognitive Linguistics: Basic Readings.* Berlin, New York: Mouton de Gruyter, 1–28.
Geeraerts, D. 2006b. 'Methodology in Cognitive Linguistics.' In: G. Kristiansen (ed., *et al.*), *Cognitive Linguistics: Current Applications and Future Perspectives.* Berlin, New York: Mouton de Gruyter, 21–49.
Geeraerts, D. 2006c. *Words and Other Wonders. Papers on Lexical and Semantic Topics.* Berlin, New York: Mouton de Gruyter.
Geeraerts, D., Grondelaers, S., Bakema, P. 1994. *The Structure of Lexical Variation. Meaning, Naming, and Context.* Berlin, New York: Mouton de Gruyter.
Geeraerts, D., Grondelaers, S., Speelman, D. 1999. *Convergentie en divergentie in de Nederlandse woordenschat.* Amsterdam: Meertens Instituut.
Gibbs, R., Beitel, M., Harrington, P., Sanders, P. 1994. 'Taking a stand on the meanings of stand: bodily experience as motivation for polysemy.' *Journal of Semantics 11*, 231–251.
Gilquin, G. 2006. 'The place of prototypicality in corpus linguistics: causation in the hot seat.' In: S.Th. Gries, A. Stefanowitsch (eds.), *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis.* Berlin, New York: Mouton de Gruyter, 159–191.
Givón, T. 1986. 'Prototypes: between Plato and Wittgenstein.' In: C. Craig (ed.), *Noun Classes and Categorization.* Amsterdam, Philadelphia: John Benjamins, 77–102.
Goddard, C. 1998. *Semantic Analysis. A Practical Introduction.* Oxford: Oxford University Press.
Goddard, C. 2005. 'Lexico-semantic universals: a critical overview.' *Linguistic Typology* 5 (1), 1–66.
González, F.R. 2007. 'Spanish.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 128–150.
Gooden, P. 2005. *Faux Pas? A No-Nonsense Guide to Words and Phrases from Other Languages.* New York: Walker & Company.
Goodenough 2008 [1956]. 'Componential Analysis and the Study of Meaning.' In: P. Hanks (ed.), *Lexicology*, vol. II. London, New York: Routledge, 233–258.
Gordon, W.T. 1982. *A History of Semantics.* Amsterdam, Philadelphia: John Benjamins Publishing Company.
Görlach, M. 2003. *English Words Abroad.* Amsterdam, Philadelphia: John Benjamins Publishing Company.
Görlach, M. (ed.). 2007 [2002]. *English in Europe*. Oxford: Oxford University Press.
Görlach, M. 2007. 'Introduction.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 1–12.
Götz, H.P. 1998. *Fizyka bez problemów.* Warszawa: Edukacja MUZA SA.
Graddol, D. 1997. *The Future of English.* London: The British Council.
Graedler, A.L. 2007. 'Norwegian.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 57–81.

Grice, H.P. 1967. 'Logic and Conversation.' In: B. Cole, J.L. Morgan (eds.), *Syntax and Semantics. Speech Acts.* New York–London: Academic Press, 41–58.
Grice, H.P. 1989. *Studies in the Way of Words.* Cambridge, Mass.: Harvard University Press.
Gries, S.Th. 2006a. 'Corpus-based methods and cognitive semantics: the many senses of *to run*.' In: S.Th. Gries, A. Stefanowitsch (eds.), *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis.* Berlin, New York: Mouton de Gruyter, 57–99.
Gries, S.Th. 2006b. 'Introduction.' In: S.Th. Gries, A. Stefanowitsch (eds.), *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis.* Berlin, New York: Mouton de Gruyter, 1–17.
Gries, S.Th., Stefanowitsch, A. (eds.). 2006. *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis.* Berlin, New York: Mouton de Gruyter.
Grybosiowa, A. 2000. 'Percepcja sygnałow obcości w pokoleniu średnim Polaków.' *Poradnik Językowy*, 46–49.
Grzegorczykowa, R., Laskowski, R., Wróbel, H. 1984. *Gramatyka współczesnego języka polskiego.* Warszawa: PWN.
Grzenia, J. 1998. *Słownik nazw własnych.* Warszawa: Wydawnictwo Naukowe PWN.
Halliday, M.A.K. 1970. 'Language structure and language function.' In: J. Lyons (ed.), *New Horizons in Linguistics*. Harmondsworth, Middlesex: Penguin Books, 140–165.
Halliday, M.A.K., Teubert W., Yallop, C., Cermakova, A. 2005. *Lexicology and Corpus Linguistics. An Introduction*. London, New York: Continuum.
Hanks, P. 1996. 'Contextual dependency and lexical sets.' *International Journal of Corpus Linguistics 1* (1), 75–98.
Hanks, P. 2008. *Lexicology. Critical Concepts in Linguistics*, vols. I–VI. London, New York: Routledge.
Hansack, E. 2004. 'Das Wesen des Namens.' In: A. Brendler & S. Brendler (eds.), *Namenarten und ihre Erforschung. Ein Lehrbuch für das Studium der Onomastik* (Festschrift Karlheinz Hengst). Hamburg: Baar, 51–65.
Harley, H. 2006. *English Words. A Linguistic Introduction.* Oxford: Blackwell Publishing.
Harvalik, M. 2005. 'Common nouns or proper names?' In: D. Brozović-Rončević & E. Caffarelli (eds.), *Naming the World. From Common Nouns to Proper Names* (Proceedings International Symposium, Zadar 2004; RIOn International Series 1). Roma: Soc. Ed. Romana, 15–22.
Harwit, M. 1973. *Astrophysical Concepts.* New York, London: John Wiley & Sons.
Haspelmath, M. 1998. 'How young is standard average European?' *Language Sciences 20* (3), 271–287.
Haugen, E. 1950a. 'Problems of bilingualism.' *Lingua 2,* 271–290.
Haugen, E. 1950b. 'The analysis of linguistic borrowing.' *Language 26*, 210–31.
Haugen, E. 1953. *The Norwegian Language in America: A Study in Bilingual Behaviour*, vol. 1: *The Bilingual Community*; vol. II: *The American Dialects of Norwegian*. Bloomington: Indiana University Press (Reprinted 1969).
Haugen, E. 1968. *Bilingualism in the Americas: A Bibliography and Research Guide.* Alabama: American Dialect Society.
Heider, E.R. 1971. 'Focal color areas and the development of color names.' *Developmental Psychology 4*, 447–455.
Heider, E.R. 1972. 'Universals in color naming and memory.' *Journal of Experimental Psychology 93*, 10–20.
Heider, F. 1958. *The Psychology of Interpersonal Relations.* New York: Wiley.
Heine, B. & Kuteva, T. 2006. *The Changing Languages of Europe.* Oxford: Oxford University Press.
Hewitt, P. G. 1998. *Conceptual Physics.* Reading, Menlo Park, New York: Addison Wesley Longman, Inc.
Hockett, Ch.F. 1958. *A Course in Modern Linguistics.* New York: Macmillan.
Hoey, M. 2005. *Lexical Priming. A New Theory of Words and Language.* London, New York: Routledge.

Hoffmann, R.R. 1989. 'Some ambiguities in the study of ambiguity.' In: D.S. Gorfein (ed.), *Resolving Semantic Ambiguity.* New York: Springer, 204–222.
Hope, T.E. 1963. 'Loanwords as cultural and lexical symbols.' *Archivum Linguisticum 14* (2), 111–122 and *15* (1), 29–42.
Hudson, R. 1984. *Word Grammar.* Oxford: Basil Blackwell.
Humbley, J. 2007. 'French.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 108–127.
Ipsen, G. 1924. 'Der alte Orient und die Indogermanen'. In: J. Friedrich (ed.), *Stand und Aufgaben der Sprachwissenschaft. Festschrift für Streitberg*. Heidelberg: Winter, 200–237.
Itkonen, E. 2003. *What Is Language? A Study in the Philosophy of Linguistics.* Turku: Åbo Akademis tryckeri.
Jackendoff, R. 1983. *Semantics and Cognition*. Cambridge, MA: MIT Press.
Jackendoff, R. 1990. *Semantic Structures.* Cambridge, Mass.: MIT Press.
Jackendoff, R. 2002a. *Foundations of Language.* Oxford: Oxford University Press.
Jackendoff, R. 2002b. 'What is in the Lexicon?' In: S. Nootebom, F. Weerman, F. Wijnen (eds.), *Storage and Computation in the Language Faculty.* Dordrecht: Kluwer, 23–58.
Jackson, H. 1988. *Words and Their Meaning.* London, New York: Longman.
Jackson, H. & Zé Amvela, E. 2007 [2000]. *Words, Meaning and Vocabulary. An Introduction to Modern English Lexicology.* London, New York: Continuum.
Jacot de Boinod, A. 2006. *The Meaning of Tingo and Other Extraordinary Words from Around The World.* New York: The Penguin Books.
James, C. 1980. *Contrastive Analysis*. London: Longman.
Jenkins, J.J. 1970. 'The 1952 Minnesota word association norms.' In: L. Postman & G. Keppel, *Norms of Word Associations.* New York: Academic Press, 1–38.
Jespersen, O. 1924. *The Philosophy of Grammar. With a New Introduction and Index by James D. McCawley.* Chicago, London: The University of Chicago Press.
Johnson, M. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* Chicago: University of Chicago Press.
Johnson, S. 1852. *A Dictionary of The English Language*, 2nd edition. London (1st – 1755).
Johnson-Laird, P. 1982. 'The mental representation of the meaning of words.' *Cognition 25*, 189–211.
Jolles, A. 1934. 'Antke Bedeutungsfelder'. In: Beträge zur Geschichte der deutschen Sprache und Literatur 58, 97–109.
Jonansson, S. 1998. 'On the role of corpora in cross-linguistic research.' In: S. Jonansson & S. Oksefjell (eds.), *Corpora and Cross-Linguistic Research.* Amsterdam and Atlanta: Rodopi, 1–24.
Jonasson, K. 1994. *Le Nom Propre. Constructions et Interprétations.* Louvain-la-Neuve.
Jorgensen, J.C. 1990. 'The psychological reality of word senses'. *Journal of Psycholinguistic Research 19* (3): 167–190.
Juilland, A. & Chang-Rodriguez, E. 1964. *Frequency Dictionary of Spanish Words.* The Hague: Mouton.
Katamba, F. 2006. *English Words.* London: New York: Routledge.
Katz, J.J. 2001. 'The end of Millianism: multiple bearers, improper names, and compositional meaning.' *Journal of Philosophy 98*, 137–166.
Katz, J. J. & Fodor, J.A. 1963. 'The structure of a semantic theory.' *Language 39,* 170–210.
Katz, J.J. & Nagel, R. 1974. 'Meaning postulates and semantic theory.' *Foundations of Language 11*, 311–340.
Kaufman, E. 1939. 'Der Fragenkreis ums Fremdwort.' *Journal of English and Germanic Philology 38*, 42–63.
Kemmer, S. 2001. 'Causative constructions and cognitive models: the English *make* causative.' In: *The First Seoul International Conference on Discourse and Cognitive Linguistics: Perspectives for the 21st Century.* Seoul: Discourse and Cognitive Linguistics Society of Korea, 803–846.
Kempson, R. 1977. *Semantic Theory.* Cambridge: Cambridge University Press.

Kennedy, G. 1992. 'Preferred ways of putting things.' In: J. Svartvik (ed.), *Directions in Corpus Linguistics.* Berlin: Mouton de Gruyter, 335–373.

King, G. 2001. *Foreign Phrases.* Glasgow: Harper Collins Publishers.

Kittay, E.F. & Lehrer, A. 1992. 'Introduction.' In: A. Lehrer, E.F. Kittay, *Frames, Fileds, and Contrasts. New Essays in Semantic and Lexical Organisation.* Hillsdale, N.J: Lawrence Erlbaum, 1–18.

Kleiber, G. 2004. 'Peut-on sauver un sens de dénomination pour les noms propres?' In: J.CH. Verstraete (ed.), *Grounding and headedness in the noun phrase. Studies for Willy Van Langendonck.* Special Issue of *Functions of Language 11* (1), 115–145.

Knowles, G. 1996. 'Corpora, databases and the organisation of linguistic data.' In: J. Thomas & M. Short. *Using Corpora for Language Research.* London, New York: Longman, 36–53.

Ködderitzsch, R. & Görlach, M. 2007. 'Albanian.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 291–300.

Kowalik, H. 1999. 'Język pod ochroną.' *Przegląd Tygodniowy 30*, 4–5.

Kowalski-Glikman, J. 2008. 'Mit teorii wszystkiego.' In: E. Dobierzewska-Mozrzymas, J. Gajda-Krynicka, A. Jezierski (eds.), *Ku interdyscyplinarności… Różne oblicza rzeczywistości.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, 39–46.

Kripke, S. 1972. 'Naming and necessity.' In: D. Davidson & G. Harman (eds.), *Semantics of Natural Language.* Dordrecht: Reidel, 253–355.

Kuryłowicz, J. 1956. 'La position linguistique du nom propre.' *Onomastica 2*, 1–14.

Krzeszowski, T.P. 1990. *Contrasting Languages: Scope of Contrastive Linguistics (Trends in Linguistics: Studies & Monographs)*. Berlin, New York: Mouton de Gruyter.

Krzeszowski, T.P. 1997. *Angels and Devils in Hell. Elements of Axiology in Semantics.* Warszawa: Wydawnictwo Energeia.

Kuźniak, M. 2007. 'A few general remarks upon a place and function of foreign words and phrases in the English lexicon.' *Anglica Wratislaviensia 45*, 107–118.

Kuźniak, M. 2008. 'Peripheral foreign language corpora in English. Preliminary study.' In: K. Rasulić and I. Trbojević Milošević, *ELLSAC Proceedings. English Language and Literature Studies: Structures across Cultures*, vol. I. Belgrade: Faculty of Philology, University of Belgrade, 423–438.

Kuźniak, M. 2009a. 'Kilka słów w obronie pleonazmów.' Katowice (in print).

Kuźniak, M. 2009b. 'Issues in lexicology: In search of THE boundary where fuzziness appears to prevail.' In: Z. Wąsik & J. Zaprucki (eds.), *Słowo jako wyraz duchowości człowieka*. Zeszyty Wydziału Humanistycznego III. Jelenia Góra: Kolegium Karkonoskie w Jeleniej Górze, 99–110.

Kvaran, G. & Svavarsdóttir, Á. 2007. 'Icelandic.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 82–109.

Labov, W. 1973. 'The boundaries of words and their meanings.' In: C.J.N. Bailey & R.W. Shuy (eds.), *New Ways of Analyzing Variation in English*. Washington, DC: Georgetown University Press, 340–373.

Lakoff, G. & Johnson, M. 1980. *Metaphors We Live by*. Chicago: The University of Chicago Press.

Lakoff, G & Johnson, M. 1999. *Philosophy in the Flesh. The Embodied Mind and Its Challenge to Western Thought.* New York: Basic Books.

Lakoff, G. 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind.* Chicago & London: University of Chicago Press.

Lakoff, G. 1990. 'The invariance hypothesis: is abstract reason based in image-schemas?' *Cognitive Linguistics 1* (1), 39–74.

Lakoff, G. 1991. 'Cognitive linguistics versus generative linguistics: how commitments influence results.' *Language and Communication 11* (1/2), 53–62.

Lakoff, G. 1993. 'The contemporary theory of metaphor.' In: A. Orthony, *Metaphor and Thought*. Cambridge: Cambridge University Press, 202–252.

Langacker, R.W. 1983. *Foundations of Cognitive Grammar.* I. *Orientation,* II: *Semantic Structure.* Trier: Linguistic Agency of the University of Trier.

Langacker, R.W. 1987, 1991. *Foundations of Cognitive Grammar*, vol. I–II*: Theoretical Prerequisites*. Stanford: Stanford University Press.
Langacker, R.W. 1990. *Concept, Image, and Symbol. The Cognitive Basis of Grammar.* Berlin, New York: Mouton de Gruyter.
Langacker, R.W. 1999. 'Grammar and Conceptualization.' *Cognitive Linguistics Research*, vol. 14. Berlin, New York: Mouton de Gruyter.
Langacker, R.W. 2005 'Construction grammars: cognitive, radical, and less so.' In: F.J. Ruiz de Mendoza Ibáñez & M.S. Peña Cervel (eds.), *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interaction.* Berlin: Mouton de Gruyter, 101–159.
Langacker, R.W. 1999. 'Reference point constructions.' In: R.W. Langacker (ed.), *Grammar and Conceptualisation.* Berlin, New York: Mouton de Gruyter, 171–202.
La Palme Reyes, M., MacNamara, J., Reyes, G.E., Zolfaghari, H. (ed.). 1993. 'Proper names and how they are learned.' *Memory 1* (4), 433–455.
Leech, G. 1966. *English in Advertising: A Linguistic Study of Advertising in Great Britain*. London: Longman.
Lehmann, A. & Martin-Berthet, F. 1997. *Introduction à la Lexicologie: Sémantique et Morphologie.* Dunod.
Lehrer, A. 1974. *Semantic Fields and Lexical Structure*. Amsterdam, New York: North Holland Publishing Co.
Leibniz, G.W. 2008. 'Excerpts from table of definitions.' In: P. Hanks (ed.), *Lexicology*, vol I. London, New York: Routledge, 166–176.
Leisi, E. 1985. *Das heutige Englisch.* 7 Aufl., Heidelberg: Winter.
Leonardi, A. 2001. *Siła i energia.* Warszawa: Dom Wydawniczy Bellona.
Lewandowska-Tomaszczyk, B. 1985. 'On semantic change in a dynamic model of language.' In: J. Fisiak (ed.), *Historical Semantics. Historical Word-Formation*. Berlin: Mouton de Gruyter, 297–323.
Lewandowska-Tomaszczyk, B. 1986. 'Cognitive basis for dynamic semantics.' *Studia Anglica Posnaniensia 19*, 107–117.
Lewandowska-Tomaszczyk, B. 1991a. 'Cognitive and interactional conditioning of semantic change.' In: W. Kellerman, M.D. Morrissey (eds), *Diachrony within Synchrony. Language History and Cognition.* Frankfurt am Main: Peter Lang, 229–250.
Lewandowska-Tomaszczyk, B. 1991b. 'Semantic change and diachronic methodology.' In: N. Boretzky, W. Enninger, B. Jessing, Th. Stolz (eds.), *Beitrage zum 8. Bochum-Essener Kolloqium uber 'Sprachwandel und seine Prinzipien.'* Bochum: Universitätsverlag Dr. N. Brockmeyer, 10–19.
Lewandowska-Tomaszczyk, B. (ed.). 1993. *Topics in English Lexicology. Folia Linguistica 30.* Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
Lewandowska-Tomaszczyk, B. 2005. *Podstawy Językoznawstwa Korpusowego.* Łodź: Wydawnictwo Uniwersytetu Łódzkiego.
Lewandowska-Tomaszczyk, B. 2006. 'Metody empiryczne i korpusowe w językoznawstwie kognitywnym.' In: P. Stalmaszczyk, *Metodologie językoznawstwa. Podstawy teoretyczne.* Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 251–281.
Lewis, M. 1993. *The Lexical Approach.* Hove: Language Teaching Publications.
Li, P. & Gleitman, L. 2002. 'Turning the tables: language and spatial reasoning.' *Cognition 83*, 265–94.
Libura, A. 2000. *Wyobraźnia w języku. Leksykalne korelaty schematów wyobrażeniowych CENTRUM–PERYFERIE i SIŁY.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
Lipka, L. 2002. *English Lexicology.* Tübingen: Gunter Narr Verlag.
Lounsbury, F.G. 2008 [1964]. 'The structural analysis of kinship semantics.' In: P. Hanks (ed.), *Lexicology*, vol. II. London, New York: Routledge, 259–280.
Loveday, L.J. 1996. *Language Contact in Japan: A Socio-Linguistic History.* Oxford: Clarendon Press.

Lyons, J. 1963. *Structural Semantics.* Oxford: Blackwell.
Lyons, J. 1968. *Introduction to Theoretical Linguistics.* Cambridge: Cambridge University Press.
Lyons, J. 1977. *Semantics*. Cambridge: Cambridge University Press.
Lyons, J. 2008 [1969]. 'Semantic structure.' In: P. Hanks (ed.), *Lexicology*, vol. II. London, New York: Routledge, 194–230.
Lyons, Ch. 1999. *Definiteness.* Cambridge: Cambridge University Press.
Lyovin, A.V. 1997. *An Introduction to the Languages of the World.* Oxford: Oxford University Press.
McArthur, T. (ed.). 1992. *The Oxford Companion to the English Language.* Oxford: Oxford University Press.
McClelland, J. & Elman, J. 1986. 'The TRACE model of speech perception.' *Cognitive Psychology 18*, 1–86.
McEnery, T. & Wilson, A., 2004. *Corpus Linguistics.* Edinburgh: Edinburgh University Press.
Mańczak, W. 1980. 'Laws of analogy.' In: J. Fisiak (ed.), *Historical Morphology.* Berlin: Mouton, 283–288.
Mańczak, W. 1992. *De la préhistoire des peuples indo-européens.* Kraków: Uniwersytet Jagielloński.
Mańczak-Wohlfeld, E. 1994. *Angielskie elementy leksykalne w języku polskim.* Kraków: Universitas.
Mańczak-Wohlfeld, E. 1995. *Tendencje rozwojowe współczesnych zapożyczeń angielskich w języku polskim.* Kraków: Universitas.
Mańczak-Wohlfeld, E. 2006. *Angielsko-polskie kontakty językowe.* Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
Mańczak-Wohlfeld, E. 2007. 'Polish.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 213–228.
Marchand, C.H. 1960. *The Categories and Types of Present-day English Word-Formation.* Wiesbaden.
Marmaridou, S.S.A. 1989. 'Proper names in communication.' *Journal of Linguistics 25*, 355–372.
Marslen-Wilson, W. & Tyler, L. 1980. 'The temporal structure of language understanding.' *Cognition 8*, 1–71.
Maximova, T. 2007. 'Russian.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 195–212.
Mel'čuk, I. 2006. 'The explanatory combinatorial dictionary.' In: G. Sica (ed.), *Open Problems in Linguistics and Lexicography.* Monza: Italiy: Polimetrica, 225–355.
Metclaff, A. 1999. *The World in So Many Words.* Boston: New York: Houghton Mifflin Company.
Meyer, Ch. 1991. 'A corpus-based study of apposition in English.' In: K. Aijmer & B. Altenberg (eds.), *English Corpus Linguistics.* London: Longman, 166–181.
Mikołajczak-Matyja, N. 2008. *Hierarchiczna struktura leksykonu umysłowego. Relacje semantyczne w leksykonie widzących i niewidomych użytkowników języka.* Poznań: Wydawnictwo Naukowe UAM.
Mill, J.S. 1843. *A System of Logic.* London: Longman.
Miller, G. & Fellbaum, Ch. 1991. 'Semantic networks of English.' *Cognition 41*, 197–229.
Milroy, L. 1987. *Language and Social Networks.* 2nd edition. Oxford: Basil Blackwell.
Miodek, J. 1971. 'Spór o *handicap*.' *PorJ* 257–259.
Miodek, J. 1980. 'A klasa i test-mecz.' *JP 60*, 180–181.
Morawski, W. 1985. 'Angielskie cytaty w tekstach wybranych pamiętników emigrantów przesyłanych ze Stanów Zjednoczonych na konkurs w 1936 r.' *Poradnik Językowy*, 98–111.
Morawski, W. 1992. *Angielskie elementy leksykalne w języku polskich emigrantów w Stanach Zjednoczonych i w Kanadzie.* Kraków: PAN, Instytut Języka Polskiego.
Morton, J. 1969. 'Interaction of information in word recognition.' *Psychological Review 76*, 165–178.
Muysken, P. 1981. 'Creole tense/mood/aspect systems: the unmarked case?' In: P. Muysken (ed.), *Generative Studies on Creole Languages*. Dordrecht: Foris, 181–199.
Müller, F.M. 1875. *Lectures on the Science of Language,* vol. 1. New York: Scribner, Armstrong.

Napoli, E. 1997. 'Names, indexicals, and identity statements.' In: W. Küne, A. Newen, and M. Anduschus (eds.), *Direct Reference, Indexicality, and Propositional Attitudes.* Stanford: CSLI Publications, 185–211.

Nerlich, B. 1992. *Semantic Theories in Europe 1830–1930. From Etymology to Contextuality.* Amsterdam, Philadelphia: John Benjamins.

Nordquist, D. 2004. 'Comparing elicited data and corpora: what the mismatch reveals about the lexicon.' In: M. Achard & S. Kemmer (eds.), *Language, culture, and mind.* Stanford, CA: CSLI.

Nosofsky, R.M. 1988. 'Similarity, frequency, and category representation.' *Journal of Experimental Psychology: Learning, Memory and Cognition 14*, 54–65.

Nuyts, J. 2005. 'Brothers in arms? On the relations between Cognitive and Functional Linguistics.' In: F.J. Ruiz de Mendoza Ibáñez & M.S. Peña Cervel (eds.), *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interaction.* Berlin: Mouton de Gruyter, 69–100.

Ooi, V.B.Y. 1998. *Computer Corpus Lexicography.* Edinburgh: Edinburgh University Press.

Otwinowska-Kasztelanic, A. 2000. *A Study of the Lexico-Semantic and Grammatical Influence of English on the Polish of the Younger Generation of Poles (19–35 Years of Age)*. Warszawa: Wydawnictwo Akademickie Dialog.

Palmer, H. 1933. *Second Interim Report on English Collocations.* Tokyo: Institute for Research in English Teaching.

Palmer, F.R. 1988. *Semantics.* 2nd edition. Cambridge: Cambridge University Press.

Palmer, S. & Low., E. 1998. *Words Borrowed from Other Languages.* Harlow: Longman.

Paul, H. 1880. *Prinzipien der Sprachgeschichte.* Halle: Max Neimeyer.

Phillips, B.S. 2001. 'Lexical diffusion, lexical frequency, and lexical analysis.' In: J.L. Byebee and P. Hopper *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, 123–126.

Phillipson, R. 1996. *Linguistic Imperialism.* 3rd edition. Oxford: Oxford University Press.

Pinker, S. 1989. *Learnability and Cognition.* Cambridge, Mass.: MIT Press.

Pinker, S. 1994. *The Language Instinct*. New York: Morrow.

Pinker, S. 1997. *How the Mind Works.* New York: Norton.

Piotrowski, T. 1986. 'The development of the general monolingual dictionary in Poland.' In: R.R.K. Hartmann, *The History of Lexicography.* Amsterdam/Philadelphia: John Benjamins Publishing Co.

Piotrowski, T. 1994. *Z zagadnień leksykografii.* Warszawa: Wydawnictwo Naukowe PWN.

Piotrowski, T. 2001. *Zrozumieć leksykografię.* Warszawa: Wydawnictwo Naukowe PWN.

Piotrowski, T. 1998. 'O anglicyzmach – a może germanizmach? – w języku polskim.' *Język Polski 78*, (3–4), 271–273.

Piotrowski, T. 2003. 'Językoznawstwo korpusowe: wprowadzenie do problematyki'. In: S. Gajda (ed.), *Językoznawstwo polskie. Stan i perspektywy*. Warszawa, Opole: PAN, Uniwersytet Opolski, 143–154.

Piotrowski, T. 2005. 'Psudoanglicyzmy czy neoanglicyzmy? O współczesnych zapożyczeniach z języka angielskiego.' In: M. Balowski and W. Chlebda (eds.), *Ogród nauk filologicznych. Księga jubileuszowa poświęcona Profesorowi Stanisławowi Kochmanowi.* Opole: Wydawnictwo Uniwersytetu Opolskiego, 503–510.

Piotrowski, T. & Podhajecka, M. 2004. 'Russianisms in English (OED – BNC – LDOCE).' In: B. Lewandowska-Tomaszczyk, *Practical Applications in Language and Computers PALC 2003*. Frankfurt: Peter Lang, 241–252.

Podhajecka, M. 2002. 'Zapożyczenia polskie w języku angielskim na materiale Oxford English Dictionary (OED), *Język Polski 82*, 330–337.

Porzig, W. 1934. 'Wassenhafte Bedeutungsbeziehungen.' *Beiträge zur Geschichte der Deutschen Sprache und Literatur 58*, 70–97.

Porzig, W. 2008 [1934]. 'Intrinsic meaning relations.' In: P. Hanks (ed.), *Lexicology*, vol. II. London, New York: Routledge, 3–21.
Pottier, B. 1964. 'Vers une sémantique moderne.' *Travaux de Linguistique et de Littérature 2*, 107–137.
Pulcini, V. 2007. 'Italian.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 151–167.
Pustejovsky, J. 1991. 'The Generative Lexicon.' *Computational Linguistics 17* (4), 409–441.
Putnam, H. 1975. 'Is semantics possible?' In: *Mind, Language, and Reality. Philosophical Papers II.* Cambridge: Cambridge University Press, 215–271.
Quine, W.V.O. 1951. 'Two dogmas of empiricism.' *Philosophical Review 60*, 20–43.
Quine, W.V.O. 1960. *Word and Object.* Cambridge, Mass.: MIT Press.
Quirk, R. & Greenbaum, S. 1973. *A University Grammar of English.* London: Longman.
Quirk, R. & Stein, G. 1996. 'Sipping a cocktail of corpora.' In: J. Thomas & M. Short (eds.), *Using Corpora for Language Research.* London and New York: Longman, 27–35.
Radden, G. 1992. 'The cognitive approach to natural language.' In: M. Pütz (ed.), *Thirty Years of Linguistic Evolution. Studies in Honour of René Dirven on the Occasion of His Sixtieth Birthday.* Amsterdam, Philadelphia: John Benjamins, 513–541.
Radden, G. & Dirven, R. 2007. *Cognitive English Grammar.* Amsterdam, Philadelphia: John Benjamins Publishing Company.
Raukko, J. 1999. 'An 'intersubjective' method for cognitive-semantic research on polysemy: the case of *get*.' In: M.K. Hiraga, Ch. Sinha and S. Wilcox (eds.), *Cultural, Psychological and Typological Issues in Cognitive Linguistics.* Amsterdam, Philadelphia: John Benjamins, 87–105.
Rayevska, N.M. 1979. *English Lexicology.* Kiev: Head Publishing House.
Recanati, F. 1997. *Direct Reference: From Language to Thought.* Oxford: Blackwell.
Reddy, M. 1979. 'The conduit metaphor – a case of frame conflict in our language about language.' In: A. Ortony (ed.), *Metaphor and Thought.* Cambridge: Cambridge University Press, 284–384.
Rice, S. 1996. 'Prepositional prototypes.' In: M. Pütz & R. Dirven (eds.), *The Construal of Space in Language and Thought.* Berlin, New York: Mouton de Gruyter, 135–165.
Roland, D. & Jurafsky, D. 2002. 'Verb sense and verb subcategorization probabilities.' In: P. Merlo & S. Stevenson (eds.), *The Lexical Basis of Sentence Processing: Formal, Computational, and Experimental Issues.* Amsterdam, Philadelphia: John Benjamins, 325–346.
Ropa, A. 1974. 'O najnowszych zapożyczeniach.' *Poradnik Językowy*, 518–526.
Rosch, E. 1973. 'Natural categories.' *Cognitive Psychology 4*, 328–350.
Rosch, E. 1975. 'Cognitive representations of semantic categories.' *Journal of Experimental Psychology 104*, 192–233.
Rosch, E. 1977. 'Human categorization.' In: N. Warren (ed.), *Studies in cross-cultural psychology*, vol. I. London: Academic Press, 1–49.
Rosch, E. & Mervis, C.B. 1975. 'Family resemblances: studies in the internal structure of categories'. *Cognitive Psychology 7*, 573–605.
Rudzka-Ostyn, B. 1988. 'Semantic extensions into the domain of verbal communication'. In: B. Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics.* Amsterdam, Philadelphia: Benjamins, 507–554.
Ruhl, Ch. 1989. *On Monosemy. A Study in Linguistic Semantics.* Albany: State University of New York Press.
Rusiecki, J. 1980. 'Latent Bilingualism.' *Papers and Studies in Contrastive Linguistics*, vol. 12, 81–98.
Russell, B. 1918. 'Descriptions.' In: J. Rosenberg & Ch. Travis (eds.), *Readings in the Philosophy of Language,* Englewood Cliffs: Prentice Hall, 166–175.
Rybicka-Nowacka, 1976. *Losy wyrazów obcych w języku polskim.* Warszawa: PWN.
Sandra, D. & Rice, S. 1995. 'Network analyses of prepositional meaning: mirroring whose mind – the linguist's or the language user's?' *Cognitive Linguistics 6* (1): 89–130.
Sapir, E. 1921. *Language.* New York: Harcourt, Brace and World.

Saussure, F. de, 1916. *Cours de linguistique générale.* Paris: Payot.
Scheler, M. 1977. *Der Englische Wortschatz.* Berlin: Schmidt.
Schlauch, M. 1959. *The English Language in Modern Times.* Warszawa: Państwowe Wydawnictwo Naukowe.
Schmid, H.J. 2000. *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition.* Berlin, New York: Mouton de Gruyter.
Schmidt, J. 1872. *Die Verwantschaftsverhältnisse der indogermanischen Sprachen.* Weimar: H. Bohlau.
Schuchardt, H. 1884. *Slawo-deutsches und Slawo-italienisches.* Graz: Leuschner and Lubensky.
Schulze, R. 1988. 'A short story of *down*.' In: W. Hüllen & R. Schulze (eds.), *Understanding the Lexicon: Meaning, Sense, and World Knowledge in Lexical Semantics.* Tübingen: Max Niemeyer Verlag.
Searle, J.L. 1958. 'Proper names.' *Mind 67*, 166–173.
Searle, J. 1969. *Speech Acts.* Cambridge: Cambridge University Press.
Seashore, R.H. & Eckerson, L.D. 1940. 'The measurement of individual differences in general English vocabularies.' *Journal of Educational Psychology 31*, 14–38.
Sedelow, S. & Sedelow, W. 1969. 'Categories and procedures for content analysis in the humanities.' In: G. Gerbner (ed., *et al.*), *The Analysis of Communication Content.* New York: John Wiley.
Seiler, F. 1907–13. *Die Entwicklung der deutschen Kulture im Spiegel des Lehnwortes.* 4 vols. Halle: Buchhandlung des Waisenhauses.
Seppänen, A. 1974. *Proper Names in English: A Study in Semantics and Syntax.* Tampere: Tampere University.
Silva, A.S. da. 2003. 'Image schemas and category coherence: the case of the Portuguese verb deixar.' In: H. Cuyckens, R. Dirven and J. Taylor (eds.), *Cognitive Approaches to Lexical Semantics*, 281–322.
Sinclair, J. 1985. 'Lexicographic evidence.' In: R.Ilson (ed.), *Dictionaries, Lexicography, and Language Learning. ELT Documents 120.* Oxford: The British Council and Pergamon Press, 81–94.
Sinclair, J. (ed.). 1987. *Looking up.* London: Collins.
Sinclair, J. 1991. *Corpus, Concordance, and Collocation*. Oxford: Oxford University Press.
Sinclair, J. 1996. 'An international project in multilingual lexicography.' In: J. Sinclair, J. Payne, and C. Pérez Hérnandez (eds.), *Corpus to Corpus: A Study of Translation Equivalence.* Special issue of *International Journal of Lexicography*, 179–196.
Sloat, C. 1969. 'Proper nouns in English.' *Language 45* (1), 26–30.
Solomon, R.L. & Howes, D.H. 1951. 'Word frequency, personal values and visual duration thresholds.' *Psychological Review 58*, 256–270.
Speelman, D., Grondelaers, S. and Geeraerts, D. 2003. 'Profile-based linguistic uniformity as a generic method for comparing linguistic varieties.' *Computers and the Humanities* 37, 317–337.
Sperber, H. 1923. *Einführung in die Bedeutungslehre.* Bonn: Schroeder.
Stathi, E. 2007. 'Modern Greek.' In: M. Görlach (ed.), *English in Europe*. Oxford: Oxford University Press, 301–330.
Stefanowitsch, A. 2001. 'Constructing causation: a construction grammar approach to analytic causatives'. PH.D. dissertation, Rice University. [After Gilquin 2006: 189].
Stefanowitsch, A. & Gries, S.Th. 2003. 'Collostructions: investigating the interaction between words and constructions.' *International Journal of Corpus Linguistics 8* (2): 209–243.
Stern, G. 1931. *Meaning and Change of Meaning.* Götoborg: Elanders.
Strang, B.M.H. 1968 [1962]. *Modern English Structure.* 2nd edition. London: Edward Arnold.
Strawson, P.F. 1950. 'On referring.' *Mind* 59, 320–344.
Strevens, P. 1972. *British and American English.* London: Collier-Macmillan Publishers.
Stubbs, M. 1995. 'Collocations and semantic profiles: on the cause of the trouble with quantitative studies'. *Functions of Language 2*, 23–55.

Stubbs, M. 2002. 'Two quantitative methods of studying phraseology in English.' *International Journal of Corpus Linguistics 7* (2), 215–244.
Summers, D. 1996. 'Computer lexicography: the importance of representativeness in relation to frequency.' In: J. Thomas & M. Short (eds.), *Using Corpora for Language Research.* London and New York: Longman, 260–266.
Svartvik, J. 1996. 'Corpora are becoming a mainstream.' In: J. Thomas & M. Short, *Using Corpora for Language Research.* London, New York: Longman, 3–13.
Sweetser, E. 1982. 'A proposal for uniting deontic and epistemic modals.' In: *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society.* Berkeley, Calif.: Berkeley Linguistics Society.
Sweetser, E. 1984. *Semantic Structure and Semantic Change: English Perception Verbs in an Indo-European Context.* Trier: Linguistic Agency of the University of Trier.
Szymanek, B. 1989. *Introduction to Morphological Analysis.* Warszawa: PWN.
Talmy, L. 1978. 'Figure and ground in complex sentences.' In: J. Greenberg *et al.* (eds.), *Universals of Human Language*, vol. 4. Stanford: Stanford University Press, 627–649.
Talmy, L. 1985. 'Force dynamics in language and thought.' In: W.H. Eilfort, P.D. Kroeber, & K.L. Peterson (eds.), *Papers from the Parasession on Causatives and Agentivity at the Twenty-First Regional Meeting of the Chicago Linguisitc Society.* Chicago: Chicago Linguistic Society, 293–337.
Talmy, L. (1988). 'Force dynamics in language and cognition.' In: *Cognitive Science 12* (1), 49–100.
Talmy, L. (2002). *Toward a Cognitive Semantics,* vol. I: *Concept structuring systems*. Cambridge: MIT Press.
Tarski, A. 1936. 'Der Wahrheitsbegriff in den formalisierten Sprachen'. *Studia Philosophica I*, 261–405. Translated and reprinted as 'The concept of truth in formalised languages' in: Tarski, A. 1956, *Logic, Semantics, and Mathematics.* Oxford: Clarendon, 152–278.
Taylor, J. 1989. *Linguistic Categorisation. Prototypes in Linguistic Theory.* Oxford: Clarendon Press.
Taylor, J. 1995. *Linguistic Categorization. Prototypes in Linguistic Theory.* Oxford: Oxford University Press.
Taylor, J.R. 2003. 'Meaning and context.' In: R. Cuyckens *et al. Motivation in Language.* Amsterdam, Philadelphia: John Benjamins Publishing Company, 27–48.
Taylor, J., Cuyckens, H., Dirven, R. 2003. 'New directions in cognitive lexical semantic research.' In: H. Cuyckens, R. Dirven, J.R. Taylor (eds.), *Cognitive Approaches to Lexical Semantics.* Berlin, New York: Mouton de Gruyter, 1–28.
Thomas, G. 1975. 'The calque – an international trend in the lexical development of the literary languages of eighteenth-century Europe.' *Germano-Slavica 6*, 21–41.
Thomas, J., Short, M. (ed.). 1996. *Using Corpora for Language Research.* London, New York: Longman.
Thomason, S.G. & Kaufman, T. 1988. *Language Contact, Creolisation and Genetic Linguistics.* Berkeley: University of California Press.
Thorndike, E. 1921. *A Teacher's Wordbook.* New York: Columbia Teachers College.
Townend, M. 2002. *Language and History in Viking Age England: Linguistic Relations between Speakers of Old Norse & Old English*. Turnhout: Brepols Publishers.
Traugott, E.C. & Dasher, R.B. 2002. *Regularity in Semantic Change.* Cambridge: Cambridge University Press.
Trier, J. 1931. *Der Deutsche Wortschatz im Sinnbezirk des Verstandes. Die Geschichte eines Sprachlichen Feldes.* Heidelberg: Winter.
Trier, J. 2008 [1934]. 'The linguistic field. An investigation.' In: P. Hanks (ed.), *Lexicology*, vol. II. London and New York: Routledge, 22–44.
Tummers, J., Heylen, K., Geeraerts, D. 2005. 'Usage-based approaches in Cognitive Linguistics: A technical state of the art.' *Corpus Linguistics and Linguistic Theory 1*, 225–261.
Ullmann, S. 1957. *The Principles of Semantics.* Glasgow: Glasgow University Publications.
Ungerer, F. Schmid, H.J. 1996. *Introduction to Cognitive Linguistics*. Harlow: Longman.

Van Langendonck, W. 2007. *Theory and Typology of Proper Names.* Berlin, New York: Mouton de Gruyter.
Vandeloise, C. 1990. 'Representation, prototypes, and centrality.' In: S.L. Tsohatzidis (ed.), *Meanings and Prototypes. Studies in Linguistic Categorization.* London: New York: Routledge, 403–437.
Vandenberghe, W. 2002. 'Instigative setting-constructions: Force Dynamic research on 'new' types of agency.' In: *Leuvense Bijdragen 90* (4), 365–390.
Viberg, Å. 1996. 'Cross-linguistic lexicology. The case of English *go and Swedish gå.*' In: K. Aijmer, B. Altenberg, M. Johansson, *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguisitc Studies.* Lund: Lund University Press, 151–182.
Weinreich, U. 1953. *Languages in Contact: Findings and Problems.* New York: Linguistic Circle of New York. (Reprinted 1968, The Hague: Mouton)
Weinreich, U. 1964. *Languages in Contact.* The Hague: Mouton & Co. Publishers.
Weisgerber, L. 1927. 'Die Bedeutungslehre: ein Irrweg der Sprachwissenschaft?' *Germanisch-Romanische Monatsschrift 15*, 161–183.
West, M. 1953. *A General Service List of English Words.* London: Longman.
Whitcut, J. 1996. *Book of Exotic Words.* London, New York: Penguin Books.
Whitney, W.D. 1881. 'On mixture of language.' *Transactions of the American Philosophical Association 12*, 5–26.
Whorf, B.L. 1956. 'Languages and logic.' In: *Language, Thought, and Reality.* Cambridge, Mass.: MIT Press. (Originally published 1941)
Whorf, B.L. 1956b. *Language, thought and reality*. *Selected writings of Benjamin Lee Whorf.* Edited by J.B Carroll. Cambridge, MA: MIT Press.
Wierzbicka, A. 1985. *Lexicography and Conceptual Analysis.* Ann Arbor, MI: Karoma.
Wierzbicka, A. 1991. 'Semantic rules know no exceptions.' *Studies in Language 15*, 371–398.
Wierzbicka, A. 1996. *Semantic: Primes and Universals*. Oxford: Oxford University Press.
Wildgen, W. 2000. 'The history and future of field semantics. From Giordano Bruno to dynamic semantics.' In: L. Albertazzi (ed.), *Meaning and Cognition: A Multidisciplinary Approach.* Amsterdam: John Benjamins, 203–226.
Winford, D. 2005. *An Introduction to Contact Linguistics.* Oxford: Blackwell Publishing.
Wittgenstein, L. 1953. *Philosophical Investigations*. New York: The MacMillan Company.
Wojtak, G. (ed.). 1997. *Toward a Functional Lexicology.* Frankfurt am Main: Peter Lang.
Wundt, W. 1900. *Völkerpsychologie I: Die Sprache.* Leipzig: Englemann.
Zernik, U (ed.). 1991. *Lexical Acquisition: Exploiting On-line Resources to Build a Lexicon.* Hillsdale, NJ: Lawrence Erlbaum.
Zipf, G.K. 1935. *The Psychobiology of Language.* Boston: Houghton Mifflin.

PRIMARY INTERNET SOURCES:

BNC – British National Corpus at: www.natcorp.ox.ac.uk
www.natcorp.ox.ac.uk/corpus/index.xml (ED: 09/2007)
www.en.wikipedia.org (ED: 06-08/2007)

DICTIONARIES AND ENCYCLOPEDIAS:

BDFWP – *The Browser's Dictionary of Foreign Words and Phrases.* (Eds. M. Varchaver & F.L. Moore, 2006 [2001]). New York, Toronto: John Wiley & Sons, Inc.
CDFE – *A Concise Dictionary of Foreign Expressions*. 2nd edition. (Ed. B.A. Phythian, 1984 [1982]). Totowa, New Jersey: Barnes & Noble Books.
CDFWP – *Chambers Dictionary of Foreign Words and Phrases.* (Ed. R. Fergusson, 1995). Edinburgh: Chambers.

CFWP – *Cassell's Foreign Words* & *Phrases.* (Ed. A. Room, 2000). London: Cassell & Co.
DBW – *Dictionary of Borrowed Words.* (Eds. L. Urdang & F.R. Abate, 1991). New York: Wynwood Press.
DCFWP – *A Dictionary of Classical* & *Foreign Words* & *Phrases.* 2nd reprint. (Eds. J. Buchanan-Brown, J. Cang, J. Crawley, B. Galushka, B. McCabe, G. Parsons, C. Steiger, K. Willimas 1982 [1981]). London: Kogan Page.
DEA – *A Dictionary of European Anglicisms.* (Ed. M. Görlach, 2005 [2001]). Oxford, New York: Oxford University Press.
DFE – *A Dictionary of Foreign Expressions commonly (and not so commonly used in English).* (Ed. T. Tuleja, 1990). New York: MacMillan Publishing Company.
DFPA – *Dictionary of Foreign Phrases and Abbreviations.* 3rd edition. (Ed. K. Guinach 1990 [1965]). New York: The H.W. Wilson Company.
DFT – *Dictionary of Foreign Terms.* The Apollo Edition. (Ed. C.O.S. Mawson, 1963 [1934]). New York: Thomas Y. Crowell Company.
DFTEL – *The Dictionary of Foreign Terms in the English Language.* (Ed. D. Carroll, 1973). New York: The Hawthorn Books, INC.
DFWP – *Dictionary of Foreign Words and Phrases.* (Ed. M. Newmark, 1986 [1950]). New York: Philosophical Library.
DFWPCE – *A Dictionary of Foreign Words and Phrases in Current English.* (Ed. J.A. Bliss, 1972 [1966]). London: Routledge & Kegan Paul.
DOFT – *Dictionary of Foreign Terms.* (Eds. M. Pei & S. Ramondino, 1974). New York: Dell Publishing Co., Inc.
EJO – *Encyklopedia językoznawstwa ogólnego*. Wyd. II poprawione i uzupełnione. (Eds. K. Polański, M. Jurkowski, S. Karolak, R. Laskowski, A. Lewicki, Z. Saloni, 1999 [1993]). Wrocław: Ossolineum.
FFDFWP – *The Facts on File Dictionary of Foreign Words and Phrases.* (Ed. M.H. Manser, 2002). New York: Checkmark Books.
HDFT – *The Harper Dictionary of Foreign Terms.* 3rd Edition. (Ed. E. Ehrlich, 1987 [1975]). London, New York: Harper & Row Publishers.
LD – *Loanwords Dictionary.* 1st edition. (Eds. L. Urdang & F.R. Abate, 1988). Detroit, Michigan: The Gale Research Company.
LDOCE – *Longman Dictionary of Contemporary English.* (Eds. D. Summers, M. Rundell, *et al.*, 1995). Harlow: Longman.
LPD – *Longman Pronunciation Dictionary*. (Ed. J. Wells, 1995). Harlow: Longman.
ODFWP I – *Oxford Dictionary of Foreign Words and Phrases.* (Ed. J. Speake, 2005 [1997]). Oxford: Oxford UP.
ODFWP II – *From Bonbon to Cha-cha. Oxford Dictionary of Foreign Words and Phrases.* 2nd edition. (Ed. A. Delahunty, 2008). Oxford: Oxford University Press.
PDGWUE – *A Popular Dictionary of German Words Used in English.* (Ed. R.D. Knapp, 2005). Capitola, California: RobbsBooks. Com.
WDFE – *World Dictionary of Foreign Expressions. A Resource for Readers and Writers.* (Ed. T.J. Sienkiewicz, 1999). Wauconda, Illinois: Bolchazy-Carducci Publishers Inc.
WDFWE – *The Wordsworth Dictionary of Foreign Words in English.* 2nd edition. (Ed. J. Ayto, 1995 [1991]). Hertfordshire: Wordsworth Reference.

OTHER INTERNET SOURCES

Internet source (1): www.answers.com/topic/lexicology (ED: 2007/02/24)
Internet source (2): www.rincondelvago.com/lexicology (ED: 2007/02/24)

GLOSSARY OF FOREIGN WORDS AND PHRASES

Proper words and phrases

Lemma	Language	Solutions	Texts	CRACn1	Meaning
Botha	Afrikaans	171	56	113.5	politician or common surname
Broederbond	Afrikaans	1	4	2.5	social organisation
Coetzee	Afrikaans	87	11	49	writer
de Klerk	Afrikaans	521	98	309.5	politician
Kruger	Afrikaans	84	23	53.5	Boer leader or common family name
Taal	Afrikaans	5	2	3.5	language
Verwoerd	Afrikaans	8	4	6	politician
Voortrekker	Afrikaans	1	1	1	ethnic group
Witwatersrand	Afrikaans	15	13	14	mountain/valley
Akan	Akan	2	1	1.5	lowland; ethnic group; language
Hoxha	Albanian	38	17	27.5	politician or architect
Tirana	Albanian	14	42	28	city
Zog	Albanian	21	9	15	king
Abu Dhabi	Arabic	123	44	83.5	city
Aden	Arabic	167	53	110	city
Ali	Arabic	1311	248	779.5	Islamic leader or boxer
Amman	Arabic	125	46	85.5	city
Aqaba	Arabic	23	13	18	city
Arafat	Arabic	288	71	179.5	politician
Aswan	Arabic	95	25	60	city
Baalbek	Arabic	12	4	8	city
Baghdad	Arabic	582	132	357	city
Bahrain	Arabic	165	78	121.5	country
Basra	Arabic	57	25	41	province/region
Beirut	Arabic	656	139	397.5	city
Benghazi	Arabic	76	15	45.5	city
Damascus	Arabic	262	99	180.5	city
Dhahran	Arabic	11	8	9.5	city
Gaddafi	Arabic	66	23	44.5	politician
Hadith	Arabic	12	7	9.5	customs and traditions
Ibrahim	Arabic	289	78	183.5	prophet
Iqbal	Arabic	23	18	20.5	writer

Iraq	Arabic	3169	25	1597	city
Islam	Arabic	681	19	350	religion
Ismaili	Arabic	5	5	5	religion
Jedda	Arabic	3	1	2	country/state
Khartoum	Arabic	14	51	32.5	city
Koran	Arabic	96	53	74.5	holy/religious scriptures
Kuwait	Arabic	1629	189	909	city
Mohammed	Arabic	581	163	372	prophet
Muscat	Arabic	58	33	45.5	city; wine
Nasser	Arabic	139	56	97.5	politician
Omdurman	Arabic	17	15	16	city
Osama bin Laden	Arabic	0	0	0	religious activist
Qatar	Arabic	111	44	77.5	country/state
Quaddafi	Arabic	0	0	0	politician
Qur'an	Arabic	0	0	0	holy/religious scriptures
Rabat	Arabic	7	33	20	city
Riyadh	Arabic	87	41	64	city
Sadat	Arabic	11	36	23.5	politician
Saddam	Arabic	926	141	533.5	first name (of a politician)
Sanaa	Arabic	0	0	0	city
Saud	Arabic	28	18	23	royal family
Saudi Arabia	Arabic	843	197	520	country/state
Yasser	Arabic	63	38	50.5	politician
Tagore	Bengali	3	13	8	family of intellectuals
Plovdiv	Bulgarian	4	4	4	city
Sofia	Bulgarian	168	87	127.5	city
Cheung	Cantonese	13	9	11	common surname
Hang Seng index	Cantonese	28	25	26.5	stock market
Kai Tak	Cantonese	3	3	3	airport
Kowloon	Cantonese	41	2	21.5	province/region
Wanchai	Cantonese	2	2	2	province/region
Wong	Cantonese	73	41	57	common surname
Andorra	Catalan	49	22	35.5	country/state
Gaudi	Catalan	2	11	6.5	engineer (architect)
Lloret	Catalan	1	2	1.5	city
Montserrat	Catalan	63	29	46	province/region
Samaranch	Catalan	14	5	9.5	sports activist
Amoy	Chinese	3	3	3	city
Anhui	Chinese	4	2	3	province/region

Anshun	Chinese	0	0	0	city
Anyang	Chinese	0	0	0	city
Chao Yuen Ren	Chinese	0	0	0	scholar
Chengdu	Chinese	16	11	13.5	city
Chongqing	Chinese	3	2	2.5	province/region
Chou En-lai	Chinese	6	4	5	politician
Chungking	Chinese	8	3	5.5	province/region
Dalian	Chinese	4	14	9	city
Deng Xiaoping	Chinese	93	5	49	politician
Foochow	Chinese	0	0	0	city
Fuzhou	Chinese	0	0	0	city
Gansu	Chinese	9	7	8	province/region
Guangdong	Chinese	38	23	30.5	city
Guilin	Chinese	6	6	6	city
Guizhou	Chinese	5	3	4	province/region
Hainan	Chinese	7	6	6.5	province/region
Han	Chinese	18	52	35	river; ethnic group; dynasty; Chinese language
Hangchow	Chinese	2	1	1.5	city
Hangzhou	Chinese	1	1	1	city
Harbin	Chinese	11	7	9	city
Hebei	Chinese	7	5	6	province/region
Heilongjiang	Chinese	3	3	3	province/region
Hong Kong	Chinese	273	523	398	city; province/region
Hubei	Chinese	7	5	6	province/region
Hunan	Chinese	14	6	10	province/region
I Ching	Chinese	14	7	10.5	holy/religious scriptures
Jiang Qing	Chinese	16	5	10.5	politician
Jiang Zemin	Chinese	71	33	52	politician
Jiangsu	Chinese	11	7	9	province/region
Jiangxi	Chinese	4	3	3.5	province/region
Jilin	Chinese	2	2	2	city
Jin	Chinese	38	18	28	city; historical figure; dynastic kingdom; subdivision of spoken Chinese
Jinan	Chinese	7	5	6	city
Kaohsiung	Chinese	19	6	12.5	city
Kunming	Chinese	6	6	6	city
Kuomintang	Chinese	93	26	59.5	political party
Ladhar Lao-tse	Chinese	0	0	0	writer

Lhasa	Chinese	77	29	53	province/region
Li Tai Po	Chinese	0	0	0	writer
Liao	Chinese	3	3	3	province/region; common surname; dynasty
Liaoning	Chinese	9	6	7.5	province/region
Lin Biao	Chinese	3	2	2.5	politician
Manchu	Chinese	6	9	7.5	ethnic group
Mao Tsetung	Chinese	1	1	1	politician
Matsu	Chinese	8	6	7	province/region; deity
Min	Chinese	13	11	12	province/region; river; language
Ming	Chinese	71	37	54	dynasty
Nanchang	Chinese	5	3	4	city
Nanjing	Chinese	2	11	6.5	city
Nanking	Chinese	64	15	39.5	city
Nanning	Chinese	2	2	2	city
Ningbo	Chinese	1	1	1	city
Ningxia	Chinese	1	1	1	province/region
Pingtung	Chinese	0	0	0	city
Putonghua	Chinese	0	0	0	language
Qin	Chinese	17	7	12	city; dynasty
Qing	Chinese	62	17	39.5	dynasty
Qingdao	Chinese	2	2	2	city
Qinghai	Chinese	2	2	2	city
Quemoy	Chinese	7	6	6.5	city
Qufu	Chinese	0	0	0	city
Shaanxi	Chinese	6	6	6	province/region
Shandong	Chinese	14	9	11.5	province/region
Shang	Chinese	28	7	17.5	dynasty
Shanghai	Chinese	23	113	68	city
Shantou	Chinese	0	0	0	city
Shantung	Chinese	1	1	1	city
Shenyang	Chinese	5	4	4.5	city
Shenzhen	Chinese	29	16	22.5	city
Shijiazhuang	Chinese	0	0	0	city
Sichuan	Chinese	27	16	21.5	city
Sui	Chinese	41	36	38.5	dynasty or nationality
Sun Yatsen	Chinese	0	0	0	politician
Suzhou	Chinese	3	3	3	city
Swatow	Chinese	0	0	0	city

Szechuan	Chinese	8	7	7.5	city
Taichung	Chinese	0	0	0	city
Taipei	Chinese	95	47	71	city
Taiping	Chinese	0	0	0	province/region
Taishan	Chinese	0	0	0	city
Taiwan	Chinese	799	239	519	province/region
Tang	Chinese	151	99	125	city; dynasty; transliteration of family names; drink
Tangshan	Chinese	5	4	4.5	city
Tao Te Ching	Chinese	3	3	3	holy/religious scriptures
Tiananmen Square	Chinese	162	79	120.5	province/region
Tianjin	Chinese	13	6	9.5	city
Tibet	Chinese	251	99	175	province/region
Urumqi	Chinese	2	2	2	city
Woo	Chinese	288	171	229.5	common family name or director of martial arts films
Wu	Chinese	137	41	89	province/region; scientist or emperor or common family name; subdivision of spoken languages; one of historical kingdoms
Wuhan	Chinese	1	6	3.5	city
Wuxi	Chinese	3	3	3	city
Xanadu	Chinese	27	2	14.5	city
Xia	Chinese	4	4	4	common surname or dynasty philosophy
Xiamen	Chinese	2	2	2	city
Xian	Chinese	19	13	16	city
Xinhua	Chinese	85	44	64.5	city; media agency or film production company
Xinjiang	Chinese	42	14	28	city
Yangtse	Chinese	3	3	3	river
Yangzhou	Chinese	0	0	0	city
Yuan	Chinese	574	49	311.5	river; dynasty or family name political term
Yue	Chinese	5	4	4.5	city; ethnic group; language
Yunnan	Chinese	22	13	17.5	province/region
Zhang Xueliang	Chinese	1	1	1	politician
Zhejiang	Chinese	3	3	3	province/region
Zhou Enlai	Chinese	8	5	6.5	politician
Zhuhai	Chinese	3	3	3	city

Osijek	Croatian	17	13	15	city
Rijeka	Croatian	9	5	7	city
Stepinac	Croatian	0	0	0	religious activist
Brno	Czech	44	17	30.5	city
Dvorak	Czech	35	22	28.5	artist or common surname
Havel	Czech	313	72	192.5	politician
Huss	Czech	17	7	12	religious activist
Janacek	Czech	17	7	12	artist
Kafka	Czech	99	44	71.5	writer
Martinu	Czech	2	2	2	artist
Olomouc	Czech	16	1	8.5	city
Skoda	Czech	5	29	17	car
Smetana	Czech	26	12	19	artist or common surname
Vaclav	Czech	131	72	101.5	first name
Vltava	Czech	25	7	16	river
Zapotek	Czech	0	0	0	sportsman
Aalborg	Danish	13	4	8.5	city
Aarhus	Danish	12	1	6.5	city
Bering	Danish	43	21	32	navigator
Bohr	Danish	55	14	34.5	scientist
Bornholm	Danish	1	4	2.5	province/region
Brahe	Danish	16	5	10.5	astrologer or scholar
Copenhagen	Danish	38	159	98.5	city
Elsinore	Danish	7	5	6	province/region
Esbjerg	Danish	3	2	2.5	province/region
Godthaab	Danish	1	1	1	city
Jensen	Danish	19	51	35	scholar; scientist or silversmith
Jespersen	Danish	67	8	37.5	scholar
Jutland	Danish	43	19	31	province/region
Kierkegaard	Danish	43	15	29	scholar
Lego	Danish	142	6	74	manufacturer
Odense	Danish	7	6	6.5	province/region
Rask	Danish	13	3	8	scholar
Roskilde	Danish	14	6	10	province/region
Skagerrak	Danish	0	0	0	province/region
Tycho Brahe	Danish	7	3	5	nobleman
Verner	Danish	1	1	1	scholar
Amsterdam	Dutch	862	296	579	city; province/region
Antwerp	Dutch	215	84	149.5	city; province/region

Arnhem	Dutch	5	29	17	city
Aruba	Dutch	21	8	14.5	province/region
Bosch	Dutch	132	63	97.5	manufacturer
Brabant	Dutch	43	19	31	province/region; administrative term
Breda	Dutch	13	9	11	city
Breughel	Dutch	21	17	19	artist
Buys Ballot	Dutch	0	0	0	scientist
Concertgebouw	Dutch	15	6	10.5	type of building
Cuyp	Dutch	9	7	8	artist
De Stijl	Dutch	11	7	9	artistic movement
Dijkstra	Dutch	6	4	5	scientist
Edam	Dutch	52	13	32.5	city; province/region; cheese
El Enschede	Dutch	0	0	0	city; province/region
Elzevier	Dutch	0	0	0	publishing/editiorial company
Escher	Dutch	14	1	7.5	graphic artist or Swiss politician; geologician
Europoort	Dutch	1	1	1	province/region
Friesland	Dutch	1	3	2	city; province/region
Gouda	Dutch	16	9	12.5	city; province/region; cheese
Groningen	Dutch	44	2	23	city; province/region
Haarlem	Dutch	18	14	16	city
Hague	Dutch	591	169	380	city; province/region
Haitink	Dutch	14	4	9	artist
Hals	Dutch	15	14	14.5	city; artist
Hilversum	Dutch	1	3	2	city
Hoboken	Dutch	4	4	4	city; province/region; scholar
Holland	Dutch	1543	63	803	country/state; province/region
Huygens	Dutch	5	2	3.5	writer or scientist
Leiden	Dutch	41	34	37.5	city; province/region
Lorentz	Dutch	42	1	21.5	scientist
Maas	Dutch	14	11	12.5	river
Maastricht	Dutch	1243	165	704	city; province/region
Mechlin	Dutch	0	0	0	city; province/region
Mondrian	Dutch	36	2	19	artist
Nijmegen	Dutch	2	11	6.5	city; province/region
Oort	Dutch	4	1	2.5	scientist
Oosterhuis	Dutch	3	2	2.5	writer
Oudenaarde	Dutch	1	1	1	province/region

Paramaribo	Dutch	13	4	8.5	city
Rembrandt	Dutch	263	82	172.5	artist
Rubens	Dutch	119	57	88	artist
Ruisdael	Dutch	7	5	6	artist
Ruud	Dutch	55	43	49	first name
Ruyter	Dutch	2	1	1.5	soldier
Scheldt	Dutch	12	6	9	river
Scheveningen	Dutch	29	6	17.5	city
Schiedam	Dutch	1	1	1	city
Schiphol	Dutch	29	7	18	airport
Spinoza	Dutch	64	18	41	scholar
Steen	Dutch	272	38	155	artist
Surinam	Dutch	52	24	38	country/state
Teniers	Dutch	4	3	3.5	artist
Tinbergen	Dutch	37	9	23	scientist
Utrecht	Dutch	97	64	80.5	city; province/region
Van der Waals	Dutch	34	13	23.5	scientist
Van Gogh	Dutch	242	91	166.5	artist
Vermeer	Dutch	37	25	31	artist
Zeebrugge	Dutch	64	35	49.5	city
Zeeland	Dutch	11	7	9	province/region
Zeeman	Dutch	11	5	8	scientist
Zuider Zee	Dutch	2	2	2	province/region
Zutphen	Dutch	2	2	2	city
Esperanto	Esperanto	34	17	25.5	language
Tallin	Estonian	3	3	3	city
Rafsanjani	Farsi	172	48	110	politician
Aalto	Finnish	1	1	1	engineer
Helsinki	Finnish	26	114	70	city
Kalevala	Finnish	6	3	4.5	epic poem
Lahti	Finnish	3	3	3	city
Turku	Finnish	18	9	13.5	city
Abbeville	French	15	1	8	city
Abelard	French	47	22	34.5	first name
Abidjan	French	43	19	31	city
Aimee	French	2	7	4.5	artist
Aisne	French	17	4	10.5	province/region
Aix-en-Provence	French	12	11	11.5	city
Aix-la-Chapelle	French	9	8	8.5	city

Aix-les-Bains	French	0	0	0	city
Ajaccio	French	9	6	7.5	city
Alencon	French	6	2	4	city
Alsace	French	111	37	74	province/region
Althusser	French	436	23	229.5	scholar
Amiens	French	73	5	39	city
Anatole	French	9	9	9	first name
Andre	French	238	116	177	first name
Angouleme	French	1	1	1	city; province/region
Anjou	French	147	32	89.5	province/region
Anouilh	French	6	6	6	writer
Ansermet	French	6	3	4.5	artist
Antibes	French	27	14	20.5	city
Antoine	French	158	63	110.5	first name
Antoinette	French	235	37	136	first name
Apache	French	128	41	84.5	free software
Appolinaire	French	1	1	1	writer
Aquitaine	French	544	39	291.5	province/region
Arc de Triomphe	French	17	17	17	monument
Ardeche	French	2	2	2	province/region
Ardennes	French	18	14	16	province/region
Argand	French	1	8	4.5	scientist
Ariane	French	7	25	16	first name
Arles	French	52	21	36.5	city
Armagnac	French	9	17	13	province/region; alcohol/spirits
Armand	French	63	27	45	first name
Armentieres	French	5	3	4	city; province/region
Arras	French	24	18	21	city; province/region
Artois	French	68	3	35.5	city; province/region
Aubusson	French	17	11	14	city; province/region
Auvergne	French	156	22	89	province/region
Avignon	French	78	41	59.5	city
Aznavour	French	7	4	5.5	artist
Badoit	French	1	1	1	drink (non-alhoholic)
Bale	French	0	0	0	city; province/region
Bally	French	2	2	2	tradesmark
Balmain	French	25	13	19	perfume
Balzac	French	49	29	39	writer
Bangui	French	1	7	4	city

Barbizon	French	14	7	10.5	city; province/region; school of artists
Bardot	French	59	29	44	artist
Barsac	French	3	3	3	wine
Bastille	French	53	36	44.5	type of building
Baudelaire	French	0	0	0	writer
Bayeux	French	47	24	35.5	city
Bayonne	French	99	2	50.5	city
Beaujolais	French	45	3	24	wine
Beaujolais nouveau	French	0	0	0	wine
Beaumont	French	249	111	180	city; province/region; French scientist or American city
Beauvais	French	13	9	11	city; province/region
Beauvoir	French	66	25	45.5	writer
Becquerel	French	4	2	3	scientist
Belmondo	French	4	4	4	artist
Bergson	French	47	11	29	scholar
Berlioz	French	61	2	31.5	artist
Berne	French	14	62	38	city
Bernhardt	French	27	16	21.5	artist
Bernoulli	French	8	8	8	scientist
Besancon	French	0	0	0	city
Bethune	French	0	0	0	city
Beziers	French	0	0	0	city
Biarritz	French	95	27	61	city
Binet	French	1	3	2	scientist
Binoche	French	23	14	18.5	artist
Bizet	French	25	19	22	artist
Bleriot	French	5	5	5	engineer
Blondin	French	5	5	5	artist
Boileau	French	2	2	2	writer
Bonaparte	French	83	39	61	politician
Bordeaux	French	299	126	212.5	city
Bordeaux mixture	French	5	3	4	dish
Bordelaise	French	2	2	2	langauge
Boudin	French	4	3	3.5	artist
Bougainville	French	0	0	0	sailor
Boulogne	French	146	82	114	province/region
Bourbon	French	131	71	101	royal family

Bovary	French	6	1	3.5	novel
Braille	French	28	23	25.5	scholar; language
Braque	French	22	12	17	artist
Brazzaville	French	27	18	22.5	city
Brest	French	7	29	18	city
Breton	French	144	69	106.5	writer; langauge
Brigitte	French	79	48	63.5	first name
Broca	French	57	7	32	artist
Bruges	French	188	65	126.5	city
Bruxellles	French	6	5	5.5	city
Caen	French	72	36	54	city
Calais	French	239	116	177.5	city
Calvados	French	14	8	11	province/region
Camargue	French	24	8	16	province/region
Cambray	French	1	1	1	city; province/region
Camembert	French	39	21	30	province/region; cheese
Camus	French	78	33	55.5	writer
Candide	French	15	12	13.5	first name
Cannes	French	138	77	107.5	city
Cantona	French	319	26	172.5	sportsman
Carcassonne	French	17	9	13	city
Cardin	French	11	9	10	artist
Carnac	French	4	4	4	city; province/region
Carnot	French	8	2	5	scientist
Cartier	French	75	36	55.5	sailor
Casablanca	French	73	56	64.5	city
Casaubon	French	25	9	17	scholar
Cassel	French	19	1	10	city; province/region
Cenis	French	2	2	2	city; mountain
Cesar	French	41	3	22	artist
Cezanne	French	43	22	32.5	artist
Chablis	French	4	27	15.5	city
Chabrier	French	3	2	2.5	artist
Chabrol	French	2	2	2	artist
Chagall	French	77	31	54	artist
Chagrin	French	5	5	5	artist
Chambourcy	French	2	2	2	city; province/region
Chamonix	French	65	21	43	city
Champlain	French	5	5	5	traveller

Champs Elysees	French	12	1	6.5	city; avenue/square
Chanel	French	121	58	89.5	artist
Chantal	French	79	1	40	writer
Chantilly	French	38	22	30	city
Charlemagne	French	3	43	23	literary character
Chartres	French	46	3	24.5	city
Chateaubriand	French	11	9	10	writer
Cherbourg	French	69	32	50.5	city
Cheyne-Stokes	French	1	1	1	disease
Chillon	French	1	1	1	type of building
Chirac	French	98	4	51	politician
Christophe	French	39	19	29	first name
Citroen	French	126	41	83.5	car
Claude	French	235	132	183.5	first name
Cleves	French	11	1	6	city
Clouseau	French	5	5	5	literary character
Clouzot	French	2	1	1.5	artist
Cluny	French	52	2	27	city
Cocteau	French	76	36	56	writer
Coeur de Lion	French	11	8	9.5	artist
Cointreau	French	44	23	33.5	alcohol/spirits
Colette	French	84	38	61	writer or sportswoman
Comme des Garcons	French	1	1	1	perfume
Comte	French	121	49	85	scholar
Conakry	French	11	7	9	city
Corbusier	French	33	16	24.5	artist
Coriolis	French	16	1	8.5	scientist
Corneille	French	7	5	6	writer or artist
Corot	French	17	12	14.5	artist
Cote d'Azur	French	17	16	16.5	province/region
Cote d'Ivoire	French	2	17	9.5	country/state; city
Coué	French	1	1	1	writer
Couperin	French	12	7	9.5	artist
Courreges	French	3	3	3	artist
Courvoisier	French	6	6	6	manufacturer
Cousteau	French	28	15	21.5	traveller
Crecy	French	8	7	7.5	city; province/region
Croix	French	28	21	24.5	city; province/region
Croix de Guerre	French	1	1	1	badge

Cro-Magnon	French	7	4	5.5	early humans
Cuisenaire	French	0	0	0	pedagogical aids
Curie	French	36	27	31.5	scientist
Cuvier	French	7	9	8	scientist
Cyrano de Bergerac	French	16	11	13.5	writer
D'Entrecasteaux	French	2	2	2	province/region; sailor
Dakar	French	6	24	15	city
Danton	French	3	3	3	politician
Dartagnan	French	4	4	4	nobleman
Daudet	French	4	3	3.5	writer
De Bauvoir	French	0	0	0	writer
De Broglie	French	19	6	12.5	scientist
De Gaulle	French	1	81	41	politician
Deauville	French	26	18	22	province/region
Debussy	French	99	28	63.5	artist
Degas	French	13	38	25.5	artist
Delacroix	French	41	24	32.5	artist
Delibes	French	5	5	5	artist
Delors	French	458	119	288.5	politician
Depardieu	French	59	27	43	artist
Derrida	French	259	26	142.5	scholar
Descartes	French	19	55	37	scholar
Diderot	French	43	18	30.5	writer
Didier	French	169	47	108	first name
Dieppe	French	84	44	64	city
Dijon	French	91	47	69	city
Dion	French	55	21	38	artist
Dior	French	92	42	67	artist
Directoire	French	4	4	4	furniture
Dominique	French	73	45	59	first name
Domremy	French	0	0	0	province/region
Dordogne	French	52	26	39	river; province/region
Dore	French	63	31	47	city
Douai	French	6	6	6	city; province/region
Douala	French	13	9	11	city
Dreyfus	French	31	23	27	soldier
Dubois	French	7	27	17	artist
Dubonnet	French	2	2	2	aperitif
Duchamp	French	46	18	32	artist

Dufy	French	15	12	13.5	artist
Dukas	French	7	4	5.5	artist
Dumas	French	126	58	92	scientist
Dunkirk	French	169	86	127.5	city
Dupont	French	12	39	25.5	gardener
Dupuytren	French	2	1	1.5	physician
Duquesne	French	3	2	2.5	educational institution
Durkheim	French	252	36	144	scholar
Duvalier	French	26	17	21.5	politician
Eiffel	French	66	45	55.5	engineer
Elysee	French	14	1	7.5	type of building
Emile	French	84	59	71.5	first name
Entre-Deux-Mers	French	4	3	3.5	wine
Escoffier	French	26	4	15	restaurateur
Etienne	French	91	52	71.5	scholar; writer or politician
Eugenie	French	62	22	42	first name
Eupen	French	1	1	1	city
Evian	French	37	22	29.5	city; province/region; drink
Faure	French	28	2	15	politician or artist
Fauve	French	14	7	10.5	artist
Feydeau	French	11	8	9.5	artist
Flaubert	French	242	23	132.5	writer
Fleur	French	41	28	34.5	artistic institution
Foch	French	42	1	21.5	soldier
Folies Bergere	French	1	1	1	artistic institution
Fontainbleau	French	3	3	3	city; province/region
Foucault	French	389	44	216.5	scholar or scientist
Fourier	French	81	2	41.5	scholar or politician
Fragonard	French	17	11	14	artist
Franck	French	52	3	27.5	artist
Francois	French	138	89	113.5	first name
Francoise	French	16	1	8.5	first name
Franglais	French	9	8	8.5	language
Fresnel	French	2	2	2	scientist
Frontenac	French	18	1	9.5	type of building
Gabon	French	12	51	31.5	country/state; city
Gauguin	French	95	33	64	artist
Gauloises	French	14	12	13	cigarettes
Gay-Lussac	French	4	3	3.5	scientist

Genet	French	82	12	47	writer
Genevieve	French	51	28	39.5	first name
Georges	French	21	135	78	first name
Gericault	French	8	6	7	artist
Gide	French	17	18	17.5	writer
Giselle	French	56	11	33.5	first name
Gitane	French	7	4	5.5	manufacturer
Gobelin	French	2	2	2	ornamental material
Godot	French	11	8	9.5	literary character
Goncourt	French	9	6	7.5	writer
Gounod	French	12	6	9	artist
Grand Guignol	French	7	7	7	theatrical figure
Grand Marnier	French	17	13	15	alcohol/spirits
Graves	French	1	1	1	city; province/region
Grenoble	French	73	49	61	city
Grignard	French	3	3	3	scientist
Gruyere	French	3	3	3	cheese
Guadeloupe	French	33	17	25	city; province/region
Guillaume	French	1	36	18.5	artist or scientist
Havre	French	63	43	53	city
Heloise	French	11	7	9	writer
Henri	French	31	135	83	first name or footballer
Hilaire	French	17	31	24	writer
Honegger	French	19	7	13	artist
Hulot	French	1	1	1	first name
Ile de France	French	9	8	8.5	city; province/region
Ionesco	French	4	4	4	writer
Jacquard	French	172	24	98	spinning device
Jacques	French	963	319	641	first name
Jean	French	2	2	2	first name
Jules	French	195	89	142	common surname
Julienne	French	13	12	12.5	first name
Jura	French	72	29	50.5	city; mountain; province/region
Justine	French	67	16	41.5	first name
La Boheme	French	7	7	7	libretto or circle of scholars
La Fontaine	French	36	18	27	lawyer or politician
Lafayette	French	57	24	40.5	city; mountain; lake
Lagrange	French	15	7	11	scientist
Lalique	French	7	6	6.5	jeweller

Lalo	French	1	1	1	artist
Lamarck	French	5	11	8	scientist
Languedoc	French	25	13	19	city; province/region
Laplace	French	82	13	47.5	scientist
Larousse	French	1	8	4.5	publishing/editorial company
Lascaux	French	23	9	16	city; caves/rocks
Lausanne	French	11	58	34.5	city
Laval	French	17	13	15	city; province/region
Lavoisier	French	31	12	21.5	scientist
Le Corbusier	French	3	14	8.5	artist
Le Havre	French	57	41	49	city
Le Mans	French	13	4	8.5	city
Le Monde	French	246	89	167.5	newspaper
Lefebvre	French	99	14	56.5	religious activist
Lesseps	French	2	2	2	politician
Levi-Strauss	French	25	1	13	scientist
Liege	French	68	33	50.5	city
Lille	French	11	48	29.5	city
Limoges	French	112	29	70.5	city
Limousin	French	117	2	59.5	city; province/region
Lisle	French	1	7	4	first name
Lissajous	French	0	0	0	scientist
Loire	French	124	55	89.5	city; river
Lome	French	22	3	12.5	city
Lorraine	French	256	127	191.5	city; river
Louis Quatorze	French	1	1	1	royal family
Louis Quinze	French	3	3	3	royal family
Louis Seize	French	0	0	0	royal family
Louis Treize	French	0	0	0	royal family
Lourdes	French	131	48	89.5	city
Louvain	French	18	12	15	city
Louvre	French	23	64	43.5	type of building
Lucerne	French	53	25	39	city
Luxembourg	French	57	22	39.5	country/state; city
Lyon	French	574	21	297.5	city
Macon	French	29	8	18.5	city; province/region
Madame	French	977	194	585.5	form of address
Maeterlinck	French	7	5	6	writer
Maginot	French	15	1	8	politician

Magritte	French	12	29	20.5	artist
Mahe	French	2	3	2.5	city
Maigret	French	27	12	19.5	first name
Mallarme	French	2	2	2	writer
Malplaquet	French	3	2	2.5	city
Manet	French	59	36	47.5	artist
Marat	French	11	9	10	politician
Marceau	French	7	6	6.5	artist
Marianne	French	265	77	171	first name
Marseillaise	French	28	18	23	national anthem
Marseilles	French	22	57	39.5	city
Massenet	French	3	2	2.5	artist
Matisse	French	265	45	155	artist
Mayotte	French	12	7	9.5	city; province/region
Mazarin	French	14	8	11	religious activist
Medoc	French	75	6	40.5	city; province/region
Megane	French	0	0	0	car
Meniere	French	0	0	0	physician; disease
Mesdemoiselles	French	2	2	2	form of address
Messiaen	French	31	11	21	artist
Messieurs	French	15	12	13.5	form of address
Metz	French	45	25	35	city
Meuse	French	54	13	33.5	city; river; province/region
Michel	French	55	222	138.5	first name
Michele	French	325	48	186.5	first name
Michelin	French	132	63	97.5	car
Midi	French	5	3	4	city; province/region
Milhaud	French	13	11	12	artist
Millet	French	24	3	13.5	city
Miquelon	French	5	4	4.5	city; province/region
Mirabeau	French	9	4	6.5	politician
Mitterand	French	24	17	20.5	politician
Moliere	French	12	9	10.5	writer
Monaco	French	219	76	147.5	country/state; city
Monegasque	French	0	0	0	inhabitant of a town
Monet	French	155	59	107	artist
Monique	French	7	28	17.5	first name
Monsiuer	French	558	19	288.5	form of address
Mont Blanc	French	87	29	58	city; mountain

Montaigne	French	42	26	34	writer
Monte Carlo	French	117	76	96.5	city; province/region
Montesquieu	French	35	18	26.5	scholar
Monteux	French	7	4	5.5	artist
Montfort	French	94	28	61	royal family
Montgolfier	French	4	3	3.5	inventor
Montmartre	French	21	16	18.5	city; province/region
Montpellier	French	67	41	54	city; province/region
Montreal	French	427	191	309	city
Montreux	French	16	12	14	city
Mont-Saint-Michel	French	2	1	1.5	city
Moselle	French	27	13	20	city; province/region
Moulin Rouge	French	1	1	1	cabaret
Muscadet	French	1	9	5	wine
Namur	French	12	9	10.5	city; province/region
Nancy	French	562	223	392.5	city
Nantes	French	64	43	53.5	city
Narbonne	French	2	16	9	city
Navarre	French	9	26	17.5	city; province/region
Necker	French	1	1	1	banker
Neuchatel	French	11	7	9	city
Ney	French	5	5	5	soldier
Nez Perce	French	245	2	123.5	city; province/region ethnic group
Niamey	French	15	1	8	city
Nice	French	113	75	94	city
Niger	French	122	54	88	country/state; city; river
Nimes	French	18	13	15.5	city
Notre Dame	French	65	39	52	type of building
Nouakchott	French	16	11	13.5	city
Noumea	French	3	3	3	city
Nuits-Saint-George	French	0	0	0	city; province/region
Oberlin	French	3	3	3	educational institution
Odette	French	78	14	46	first name
Oise	French	7	4	5.5	city; river; province/region
Oran	French	13	1	7	city
Orleans	French	241	128	184.5	city
Orly	French	6	6	6	city; province/region
Ostend	French	51	26	38.5	city; province/region

Pagnol	French	3	2	2.5	writer or artist
Pantagruel	French	2	2	2	literary character
Paris	French	5938	2	3555	city
Parmentier	French	2	2	2	engineer
Pas de Basque	French	1	1	1	city; province/region
Pascal	French	27	89	58	scholar
Passy	French	11	3	7	scientist
Pasteur	French	51	29	40	scientist
Pele	French	53	25	39	sportsman
Pelleas	French	18	1	9.5	literary character
Pelletier	French	1	9	5	scientist
Peltier	French	8	5	6.5	scientist
Percheron	French	1	1	1	horse
Perigrod	French	0	0	0	city
Pernod	French	28	18	23	alcohol/spirits
Perpignan	French	26	12	19	city
Perrault	French	13	7	10	writer or doctor
Perrier	French	124	73	98.5	drink (non-alcoholic)
Perrin	French	66	31	48.5	artist or scientist
Petain	French	19	9	14	politician
Peugeot	French	326	18	172	car
Philippe	French	321	121	221	first name
Piaf	French	19	14	16.5	artist
Piaget	French	19	39	29	scientist
Picard	French	41	18	29.5	scientist
Picardy	French	38	32	35	city; province/region
Pierre	French	669	27	348	first name
Pinot Noir	French	52	19	35.5	wine
Pissarro	French	53	23	38	artist
Poincare	French	2	2	2	scholar
Poirot	French	56	23	39.5	literary character
Poisson	French	86	22	54	scientist
Poitiers	French	119	3	61	city
Pompidou	French	11	34	22.5	politician
Pont l'Eveque	French	2	2	2	city; province/region
Port Salut	French	13	4	8.5	cheese
Port-au-Prince	French	47	25	36	city
Poulenc	French	34	19	26.5	artist
Principe	French	23	17	20	city; province/region

Proudhon	French	4	4	4	politician
Proust	French	194	55	124.5	writer
Provencal	French	7	6	6.5	related to Provence
Provence	French	24	86	55	city; province/region
Puy-de-Dome	French	5	1	3	city; province/region
Quai d'Orsay	French	6	6	6	city; avenue/square
Quebecois	French	13	9	11	related to Quebec
Rabelais	French	16	14	15	writer
Racine	French	26	19	22.5	writer
Rainier	French	11	9	10	royal family
Rambouillet	French	6	5	5.5	city; province/region
Rameau	French	43	7	25	artist
Ramillies	French	7	2	4.5	city; province/region
Raoul	French	15	42	28.5	first name
Ravel	French	11	37	24	artist
Raynaud	French	7	4	5.5	physician
Reaumur	French	13	1	7	scientist
Reims	French	16	22	19	city
Remy	French	32	17	24.5	writer or road enginner
Renault	French	567	163	365	car
Rennes	French	33	24	28.5	city
Renoir	French	93	44	68.5	artist
Resnais	French	3	2	2.5	artist
Reunion	French	1	1	1	city; province/region
Rheims	French	71	16	43.5	city
Rhone	French	58	32	45	city; river
Richelieu	French	36	2	19	politician
Rimbaud	French	37	23	30	writer
Robbe-Grillet	French	54	2	28	writer
Robespierre	French	23	15	19	politician
Rodin	French	77	29	53	artist
Ronsard	French	7	3	5	writer
Roquefort	French	21	8	14.5	cheese
Rouen	French	15	63	39	city
Rousseau	French	251	83	167	scholar
Sabatier	French	9	7	8	scientist
Sabena	French	26	16	21	airline company
Sade	French	14	12	13	writer
Saint-Etienne	French	7	4	5.5	city

Saint-Saens	French	7	4	5.5	artist
Sancerre	French	15	1	8	city; province/region
Sand	French	1	1	1	writer
Sandoz	French	14	1	7.5	manufacturer
Santer	French	1	6	3.5	politician
Saone	French	3	3	3	city; river
Sartre	French	357	47	202	scholar
Satie	French	31	11	21	artist
Saussure	French	116	2	59	scholar
Sauterne	French	1	1	1	wine
Savarin	French	2	2	2	writer
Schuman	French	7	5	6	politician
Seine	French	158	82	120	city
Semillon	French	7	2	4.5	wine/type of grape
Serge	French	48	38	43	first name
Seurat	French	35	2	18.5	artist
Sevres	French	1	7	4	city; province/region
Simenon	French	4	1	2.5	writer
Simplon	French	11	9	10	city; province/region
Sion	French	35	24	29.5	city
Solvay	French	21	1	11	scientist
Somme	French	128	61	94.5	city; province/region
Sorbonne	French	55	43	49	educational institution
St Bernard	French	34	26	30	city
St Cloud	French	7	6	6.5	city
St Denis	French	66	19	42.5	city
St Laurent	French	18	16	17	artist
St Malo	French	19	12	15.5	city
St Tropez	French	35	18	26.5	city
Stendhal	French	19	13	16	writer
Strasbourg	French	424	177	300.5	city
Suchard	French	21	7	14	confectionary brand
Sylphides	French	2	4	3	ballet
Taine	French	5	3	4	scholar
Talleyrand	French	14	1	7.5	politician
Tanguy	French	4	3	3.5	artist
Tartuffe	French	5	3	4	literary character
Tati	French	3	3	3	artist
Teilhard de Chardin	French	5	4	4.5	scholar

Therese	French	31	21	26	first name
Thermidor	French	8	8	8	name of the month in a revolutionary calendar
Tignes	French	15	9	12	city; province/region
Tissot	French	14	8	11	watch
Tocqueville	French	28	11	19.5	politician
Tortelier	French	8	3	5.5	artist
Toulon	French	67	3	35	city
Toulouse	French	167	6	86.5	city
Toulouse-Lautrec	French	25	17	21	artist
Tour de France	French	5	29	17	sports competition
Tours	French	18	16	17	city
Trudeau	French	2	2	2	politician
Truffaut	French	1	7	4	artist
Tuileries	French	5	13	9	type of building
Utrillo	French	24	9	16.5	artist
Valenciennes	French	15	1	8	city; province/region
Valery	French	55	39	47	writer
Valois	French	87	22	54.5	royal family
Vaucluse	French	5	4	4.5	city; province/region
Verdun	French	288	17	152.5	city
Verlaine	French	12	12	12	writer
Verne	French	23	18	20.5	writer
Versailles	French	22	115	68.5	city
Vichy	French	123	47	85	city; province/region
Villeneuve	French	16	8	12	city; province/region; common surname
Villon	French	9	5	7	artist
Vincennes	French	31	15	23	city; province/region
Voltaire	French	11	56	33.5	scholar
Vosges	French	29	22	25.5	city; mountain; province/region
Vouvray	French	7	5	6	city; province/region
Vuitton	French	28	18	23	artist
Watteau	French	17	12	14.5	artist
Yaounde	French	1	1	1	city
Ypres	French	65	26	45.5	city; province/region
Yves	French	119	71	95	first name
Zola	French	56	31	43.5	writer
Aachen	German	126	36	81	city

Adenauer	German	151	15	83	politician
Adler	German	117	63	90	scientist
Altdorfer	German	3	1	2	planetoid
Alzheimer	German	28	81	54.5	scientist; disease
Augsburg	German	64	26	45	city
Baader- Meinhof	German	0	0	0	city
Bach	German	459	146	302.5	artist
Bad Godesberg	German	8	4	6	city
Baden	German	96	31	63.5	city
Baden-Baden	German	9	9	9	city
Baedeker	German	23	14	18.5	printmaker
Bamberg	German	18	11	14.5	city
Basle	German	98	44	71	city
Battenberg	German	12	11	11.5	city
Bauer	German	64	32	48	politician
Bauhaus	German	36	23	29.5	artistic institution
Baume	German	2	2	2	watch
Bayreuth	German	23	9	16	city
Beckenbauer	German	15	8	11.5	sportsman
Becker	German	291	11	151	sportsman
Beckmann	German	31	13	22	artist
Beethoven	German	457	148	302.5	artist
Belsen	German	29	25	27	concentration camp
Benz	German	31	22	26.5	engineer; car
Berchtesgaden	German	25	12	18.5	city
Berlin	German	2444	542	1493	city
Bern	German	63	2	32.5	city
Bertelsmann	German	1	9	5	media concern
Biedermeier	German	3	3	3	bourgeois lifestyle
Bielefeld	German	8	8	8	city
Bierkeller	German	5	4	4.5	entertainment club
Bingen	German	4	4	4	city; scholar
Bismarck	German	18	37	27.5	politician
Blucher	German	1	1	1	soldier
Bochum	German	13	9	11	city
Boehm	German	17	6	11.5	industrialist; manufacturer
Bohm	German	33	9	21	economist
Boltzmann	German	58	1	29.5	scientist
Bonhoeffer	German	25	13	19	scholar

Bonn	German	449	188	318.5	city
Bormann	German	17	7	12	politician
Borussia	German	23	12	17.5	football team
Brahms	German	174	46	110	artist
Brandenburg	German	87	45	66	city
Brandt	German	131	5	68	politician
Brecht	German	66	41	53.5	writer
Bremen	German	146	74	110	city
Bremerhaven	German	6	4	5	city
Brenner	German	55	19	37	province/region
Brocken	German	5	2	3.5	mountain/valley
Bruch	German	17	5	11	artist
Bruckner	German	47	11	29	historician; geographer; writer or artist
Brunnhilde	German	3	2	2.5	literary character
Buchenwald	German	1	1	1	concentration camp
Bulow	German	18	8	13	politician or artist
Bundesbank	German	257	11	134	banking institution
Bundesrat	German	7	25	16	administrative/political term
Bundestag	German	27	58	42.5	administrative/political term
Bundeswehr	German	23	12	17.5	military corps
Bunsen	German	34	27	30.5	scientist
Buxtehude	German	44	4	24	artist
Charlottenburg	German	15	7	11	city
Chemnitz	German	8	8	8	city
Ciba	German	1	8	4.5	manufacturer
Clausewitz	German	12	6	9	soldier
Coburg	German	92	24	58	city
Colditz	German	36	15	25.5	city
Curtius	German	24	3	13.5	scientist or scholar
Cuxhaven	German	2	2	2	city
Dachau	German	9	9	9	city; concentration camp
Dahrendorf	German	29	11	20	city
Daimler	German	132	47	89.5	engineer; car
Danzig	German	315	29	172	city
Darmstadt	German	25	2	13.5	city
Davos	German	63	16	39.5	city
Detmold	German	2	2	2	city
Deutsche Mark	German	2	8	5	currency

Deutschland	German	65	45	55	country/state
Dieter	German	217	54	135.5	first name
Dietrich	German	9	57	33	singer; artist or first name
Dortmund	German	57	38	47.5	city
Dresden	German	237	19	128	city
Duisburg	German	13	1	7	city
Duisenberg	German	0	0	0	city
Durer	German	15	11	13	artist
Durrenmatt	German	2	2	2	writer or novelist
Dusseldorf	German	71	44	57.5	city
Eckhart	German	5	3	4	scholar
Ehrlich	German	15	11	13	professor of law or medician
Eichmann	German	13	8	10.5	politician
Eifel	German	3	3	3	mountain/valley
Eiger	German	5	19	12	mountain/valley
Einstein	German	398	17	207.5	scientist
Elbe	German	56	32	44	river
Elsa	German	66	26	46	first name
Emil	German	111	48	79.5	first name
Ems	German	23	67	45	river
Engelbert	German	4	4	4	first name
Engels	German	575	68	321.5	scholar
Erfurt	German	12	11	11.5	city
Ernst	German	38	123	80.5	chemician or first name
Essen	German	49	31	40	city
Euler	German	21	1	11	scientist
Fassbinder	German	6	5	5.5	artist
Frankenstein	German	27	66	46.5	literary character
Frankfurt	German	568	257	412.5	city
Franz	German	34	118	76	first name
Frau	German	319	3	161	form of address
Fraunhofer	German	7	5	6	optician; educational institution
Freiburg	German	79	25	52	city
Freud	German	1559	2	780.5	scientist
Friedrich	German	125	83	104	first name
Frisch	German	37	13	25	writer
Froebel	German	15	12	13.5	scholar
Furtwangler	German	5	4	4.5	artist
Geissler	German	0	0	0	politician

Gewurtzaminer	German	0	0	0	wine
Gluck	German	29	18	23.5	artist
Godel	German	1	1	1	scientist
Goebbels	German	56	2	29	politician
Goering	German	68	17	42.5	politician
Goethe	German	98	59	78.5	writer
Gotha	German	21	9	15	city; railway industry or museum
Gotterdammenrung	German	0	0	0	city
Gottfried	German	37	24	30.5	first name
Gottingen	German	6	4	5	city
Graz	German	25	19	22	city
Gretchen	German	13	8	10.5	first name
Gretel	German	18	15	16.5	literary character
Grindelwald	German	26	9	17.5	city
Gstaad	German	16	14	15	city
Gunter	German	33	24	28.5	horserider; politician or scientist
Gunther	German	19	16	17.5	first name
Gutenberg	German	24	18	21	printmaker
Habsburg	German	182	35	108.5	royal family
Hamburg	German	373	188	280.5	city
Hamelin	German	23	1	12	city
Handel	German	211	76	143.5	artist
Hanover	German	2	13	7.5	city
Hans	German	62	28	45	first name
Hansa	German	12	6	9	common surname; wash device company or part of the football team name
Hansel	German	24	18	21	literary character
Hapsburg	German	34	19	26.5	royal family
Hartz	German	9	6	7.5	mountain/valley
Haydn	German	191	67	129	artist
Hegel	German	239	49	144	scholar
Heidegger	German	22	12	17	scholar
Heidelberg	German	55	43	49	city
Heilbron	German	18	6	12	city
Heimlich	German	0	0	0	medician or part of the medical term
Heine	German	46	15	30.5	writer or proper name
Heisenberg	German	62	16	39	city

Helmholtz	German	27	12	19.5	physician
Hermann	German	11	74	42.5	first name
Herr	German	467	68	267.5	family name; form of address
Hesse	German	54	35	44.5	province/region; writer
Hilbert	German	1	7	4	scientist
Himmler	German	13	19	16	politician
Hindemith	German	13	8	10.5	artist
Hindenburg	German	38	1	19.5	politician
Hitler	German	1634	357	995.5	politician
Hofmannsthal	German	9	4	6.5	writer
Hofmeister	German	1	1	1	scientist
Hohenzollern	German	21	11	16	royal family
Hohner	German	26	5	15.5	manufacturer
Holbein	German	56	33	44.5	artist
Holstein	German	54	14	34	city
Honecker	German	156	56	106	politician
Humboldt	German	36	21	28.5	scholar
Humperdinck	German	15	8	11.5	artist
Immelmann	German	8	4	6	pilot
Innsbruck	German	53	17	35	city
Interlaken	German	12	7	9.5	province/region
Isolde	German	27	13	20	first name
Joachim	German	83	44	63.5	first name
Jung	German	151	8	79.5	scientist
Jungfrau	German	21	13	17	mountain/valley
Jurgen	German	3	19	11	first name
Kaiserslautern	German	31	5	18	city
Kant	German	236	54	145	scholar
Karajan	German	291	12	151.5	artist
Kekule	German	3	2	2.5	scientist
Kepler	German	97	25	61	scientist
Kiel	German	29	23	26	city
Kirchhoff	German	33	4	18.5	scientist
Kitzbuhel	German	38	5	21.5	city
Klaus	German	242	19	130.5	first name
Klebs-Loffler	German	0	0	0	medicians or name of virus
Klee	German	58	26	42	artist
Klemperer	German	54	9	31.5	artist
Klosters	German	23	9	16	province/region; travel institution

Koblenz	German	25	12	18.5	city
Koch	German	173	68	120.5	medician or nazi activist
Kochel	German	2	2	2	scholar
Kohl	German	66	146	106	politician
Kokoschka	German	19	9	14	artist
Konigsberg	German	4	4	4	city
Kreutzer	German	7	6	6.5	artist
Krupp	German	23	13	18	industrialist
Kultur	German	8	4	6	civilisation
Kurt	German	297	123	210	first name
Langerhans	German	12	6	9	scientist
Lauda	German	137	11	74	sportsman
Leibnitz	German	14	12	13	scholar
Leica	German	36	9	22.5	small screen apparatus
Leipzig	German	176	83	129.5	city
Liebfraumilch	German	3	3	3	wine
Liebig	German	14	8	11	scientist
Liechtenstein	German	86	44	65	country/state
Linz	German	29	17	23	city
Liszt	German	88	23	55.5	artist
Loeb	German	24	12	18	artist or scientist
Lohengrin	German	17	5	11	literary character
Lorelei	German	12	6	9	literary character
Lorenz	German	122	39	80.5	physician
Lowenbrau	German	2	2	2	beer/beer company
Lubeck	German	11	5	8	city
Ludwig	German	168	75	121.5	first name
Lufthansa	German	37	28	32.5	airline company
Luftwaffe	German	146	52	99	air military corps
Luger	German	14	6	10	arms
Luneberg	German	1	1	1	province/region
Luthe	German	0	0	0	city
Mahler	German	119	3	61	artist or writer
Main	German	125	68	96.5	river
Mainz	German	93	37	65	city
Mannheim	German	19	42	30.5	city
Marburg	German	15	12	13.5	city
Marienbad	German	6	4	5	city
Marlene	German	127	58	92.5	first name

Matthaus	German	3	3	3	sportsman
Mauser	German	9	4	6.5	inventor; arms
Mayer	German	153	61	107	physician or scholar
Meclenburg	German	0	0	0	city
Meissen	German	26	8	17	city
Meistersinger	German	21	4	12.5	watch
Melanchton	German	0	0	0	scholar
Mendelssohn	German	8	32	20	artist
Messerschmidt	German	3	3	3	airline company
Metternich	German	32	15	23.5	politician
Mobius	German	8	5	6.5	family name; capital fund
Mohs	German	6	3	4.5	scientist
Mollweide	German	0	0	0	scientist
Monchen-Gladbach	German	0	0	0	city
Mond	German	57	13	35	family name; manufacturer
Mossbauer	German	2	2	2	scientist
Mozart	German	1198	188	693	artist
Muller	German	12	53	32.5	scientist or footbal player
Munchausen	German	15	5	10	scientist; building
Munich	German	621	285	453	city
Munster	German	93	45	69	city
Nassau	German	1	1	1	city
Neanderthal	German	63	35	49	type of man
Neckar	German	6	3	4.5	river; part of the name of stadium
Neisse	German	3	3	3	river
Nessler	German	1	1	1	manufacturer
Neubrandenburg	German	3	3	3	city
Neumann	German	94	24	59	scientist or writer
Nibelung	German	1	1	1	royal family
Nibelungenlied	German	2	2	2	epic poem
Niebuhr	German	1	6	3.5	traveller or historian
Niedersachsen	German	0	0	0	city
Niersteiner	German	1	1	1	member of Nierstein municipality
Nietzsche	German	325	36	180.5	scholar
Nordhein-Westfalen	German	0	0	0	province/region
Nuremberg	German	141	77	109	city
Oberammergau	German	12	6	9	city
Oberland	German	24	7	15.5	city
Oder	German	19	14	16.5	river

Offenbach	German	3	2	2.5	city
Olbers	German	0	0	0	scientist
Oldenburg	German	39	15	27	city
Opel	German	4	2	3	car
Ophuls	German	2	1	1.5	artist
Orff	German	8	4	6	artist
Osnabruck	German	7	5	6	city
Paderborn	German	16	6	11	city
Parsifal	German	28	13	20.5	opera
Parzival	German	1	1	1	epic poem
Perutz	German	6	5	5.5	city; scientist
Petri	German	64	22	43	scientist
Piesporter	German	2	2	2	wine
Pilatus	German	12	7	9.5	mountain/valley; airline company
Pilsen	German	3	3	3	city
Planck	German	58	25	41.5	scientist
Plattdeutsch	German	1	1	1	langauge
Poggenpohl	German	3	3	3	manufacturer
Porsche	German	26	12	19	car
Potsdam	German	63	42	52.5	city
Quellenforschung	German	0	0	0	source research in scholarly studies
Radetzky	German	2	2	2	soldier; hotel
Ratskeller	German	0	0	0	type of building
Reich	German	329	19	174	province/region
Reichstag	German	47	23	35	type of building
Rhine	German	26	18	22	river
Rhineland	German	71	35	53	province/region
Ribbentrop	German	18	11	14.5	politician
Richter	German	189	85	137	American seismologist; painter; artist or chemist
Richthofen	German	9	6	7.5	traveller; general or pilot
Ricoh	German	16	1	8.5	manufacturer
Riefenstahl	German	2	2	2	artist
Riemann	German	5	6	5.5	scientist
Riesling	German	36	21	28.5	family name; wine
Rollei	German	1	1	1	industrialist
Rommel	German	5	2	3.5	soldier
Rostock	German	39	18	28.5	city

Ruhr	German	92	43	67.5	province/region
Saar	German	24	8	16	river; province/region
Saarbrucken	German	5	2	3.5	city
Saarland	German	22	15	18.5	province/region
Sachs	German	16	71	43.5	writer
Salzburg	German	272	52	162	city
Schengen	German	31	14	22.5	city
Schiller	German	97	41	69	writer
Schlegel	German	23	11	17	scholar or ornithologist
Schleicher	German	7	3	5	general or other family name
Schleswig	German	6	6	6	city
Schliemann	German	6	6	6	scientist
Schneider	German	94	57	75.5	writer or artist
Schnitzler	German	3	3	3	writer
Schoenberg	German	9	2	5.5	artist or surname
Schopenhauer	German	9	13	11	scholar
Schroder	German	26	19	22.5	politician or other family name
Schrodinger	German	4	3	3.5	scientist
Schubert	German	139	66	102.5	artist
Schumacher	German	98	45	71.5	sportsman
Schumann	German	9	31	20	artist
Schwartz	German	73	41	57	common surname
Schwartzenegger	German	1	1	1	artist
Schwartzschild	German	0	0	0	scientist
Schwartzwald	German	0	0	0	city
Schwarzkopf	German	83	34	58.5	general or other family name
Schweitzer	German	37	22	29.5	scholar or artist
Seebeck	German	0	0	0	scientist
Siegfried	German	14	45	29.5	literary character or popular first name
Sieglinde	German	1	1	1	literary character
Siemens	German	466	147	306.5	manufacturer
Sigmund	German	63	5	34	first name
Sigurd	German	11	8	9.5	literary character or historical figure
Spandau	German	27	21	24	province/region
Spatlese	German	1	1	1	wine
Spengler	German	25	6	15.5	scholar
Spohr	German	0	0	0	artist

Sprechstimme	German	2	2	2	technique of vocal production
St Gotthard	German	4	3	3.5	mountain/valley; saint
St Moritz	German	28	21	24.5	city
Stalag	German	12	8	10	prisoner of war camp in Nazi Germany
Stein	German	33	91	62	common surname
Steiner	German	177	45	111	scholar or nazi officer
Stern	German	771	442	606.5	scientist or writer; newspaper/magazine
Stich	German	76	24	50	politician; tennis palyer or other family name
Stockhausen	German	21	13	17	artist
Strauss	German	236	87	161.5	artist
Struwwelpeter	German	1	1	1	literary character
Stuka	German	9	5	7	aircraft bomber
Stuttgart	German	297	128	212.5	city
Suppe	German	0	0	0	artist
Tannhauser	German	3	3	3	literary character
Taunus	German	3	2	2.5	mountain/valley
Telemann	German	19	1	10	artist
Thyssen	German	51	14	32.5	industrialist
Tilsit	German	5	4	4.5	city
Tirpitz	German	41	11	26	soldier; battleship
Trabant	German	18	11	14.5	car
Trier	German	58	37	47.5	city
Trubner	German	0	0	0	artist
Tubingen	German	2	2	2	city
Tyrol	German	55	27	41	province/region
Ulrich	German	63	43	53	first name
Ursprache	German	1	1	1	langauge
Vaduz	German	7	6	6.5	city
Volk	German	18	14	16	family name; people
Volkswagen	German	2	11	6.5	car
Vorsprung durch Technik	German	5	3	4	tagline used in advertising campaigns
Wagner	German	486	14	250	artist
Waldemar	German	7	6	6.5	first name
Waldheim	German	46	19	32.5	city; politician

Waldstein	German	1	1	1	mountain/valley; scholar; historian or patron of Beethoven
Walter	German	1763	545	1154	engineer; scientist or first name
Wankel	German	3	2	2.5	engineer; scientist or first name
Wartburg	German	1	1	1	car
Wassermann	German	2	2	2	scientist
Weber	German	487	118	302.5	common surname
Webern	German	38	11	24.5	artist
Weil	German	26	21	23.5	city; common surname
Weimar	German	98	58	78	city
Weismann	German	44	9	26.5	artist
Weiss	German	71	44	57.5	common surname; ice cream brand
Weizmann	German	28	8	18	scientist
Weltanschauung	German	2	13	7.5	philosophy/philosophical term
Weltschmerz	German	3	3	3	philosophy/philosophical term
Werner	German	235	84	159.5	chemist; writer; geologist; religious activist or other family name
Wernicke	German	15	4	9.5	physician
Weser	German	8	4	6	river
Wiesbaden	German	27	14	20.5	city
Wiesenthal	German	4	4	4	engineer
Wildenstein	German	24	9	16.5	art dealer
Winterthur	German	33	6	19.5	city; insurance company
Wittenberg	German	12	8	10	city
Wolf	German	182	96	139	common surname
Wolff	German	177	59	118	scholar or scientist
Wolfgang	German	374	11	192.5	first name
Worms	German	0	0	0	city; surname
Wotan	German	5	5	5	deity; language
Wuppertal	German	5	5	5	city
Wurttenberg	German	0	0	0	city
Wurzburg	German	11	9	10	city
Zeiss	German	27	17	22	industrialist
Zermatt	German	27	13	20	province/region
Ziegler	German	33	18	25.5	common surname
Zugspitze	German	7	4	5.5	mountain/valley
Zurich	German	336	131	233.5	city

Zwingli	German	14	7	10.5	religious activist
Aegina	Greek	36	6	21	province/region
Avgolemono	Greek	1	1	1	dish
Corfu	Greek	85	3	44	province/region
Delphi	Greek	111	25	68	city
Epidaurus	Greek	11	8	9.5	city
Heraklion	Greek	11	8	9.5	city
Katharevousa	Greek	0	0	0	language
Papadopoulos	Greek	3	3	3	sportsman or politician
Papandreou	Greek	81	28	54.5	politician
Paphos	Greek	8	6	7	city
Parnassus	Greek	1	7	4	mountain/valley
Patras	Greek	14	6	10	city
Paxos	Greek	16	2	9	province/region
Piraeus	Greek	36	16	26	city
Salonica	Greek	4	3	3.5	city
Samos	Greek	52	11	31.5	city; province/region
Seferis	Greek	7	3	5	writer
Siros	Greek	0	0	0	province/region
Theodorakis	Greek	5	4	4.5	artist
Thera	Greek	37	7	22	province/region
Thessalonica	Greek	6	6	6	city
Chaim	Hebrew	46	34	40	first name
Chanukah	Hebrew	0	0	0	festival
Eretz	Hebrew	29	5	17	country/state; magazine or part of the name: the land (Eretz) of Israel
Hanukah	Hebrew	0	0	0	festival
Torah	Hebrew	56	22	39	holy/classic scriptures
Agni	Hindi	2	2	2	deity
Allahabad	Hindi	12	8	10	city
Amritsar	Hindi	25	15	20	city
Andhra Pradesh	Hindi	38	19	28.5	province/region
Arjuna	Hindi	15	9	12	literary character
Asoka	Hindi	4	4	4	king
Bhagwan	Hindi	0	0	0	religion
Bihar	Hindi	8	31	19.5	province/region
Brahmaputra	Hindi	13	9	11	river
Buddha	Hindi	175	64	119.5	deity/spirit-inspired being

Deccan	Hindi	12	7	9.5	province/region
Gandhi	Hindi	672	112	392	politician
Ganesh	Hindi	16	1	8.5	deity
Granth	Hindi	4	3	3.5	holy/classic scriptures
Gujarat	Hindi	4	22	13	province/region
Gujarati	Hindi	21	13	17	ethnic group; language or writing system
Gupta	Hindi	41	26	33.5	city; common surname
Gwalior	Hindi	9	3	6	city
Haryana	Hindi	51	17	34	province/region
Himalaya	Hindi	45	16	30.5	mountain/valley
Indore	Hindi	4	3	3.5	city
Jaipur	Hindi	4	14	9	city
Jalalabad	Hindi	12	7	9.5	city
Karma	Hindi	61	33	47	philosophy/philosophical term
Lucknow	Hindi	2	17	9.5	city
Mahabharata	Hindi	19	1	10	epic poem
Maharashtra	Hindi	45	22	33.5	epic poem
Maharishi	Hindi	26	18	22	deity/spirit-inspired being
Mahayana	Hindi	9	8	8.5	religion
Marathi	Hindi	2	2	2	language
Meerut	Hindi	3	3	3	city
Nehru	Hindi	98	29	63.5	politician
Pathan	Hindi	9	4	6.5	ethnic group
Patna	Hindi	16	9	12.5	city
Pradesh	Hindi	199	55	127	province/region
Rabindranath	Hindi	1	5	3	writer
Ramayana	Hindi	3	3	3	epic poem
Shiva	Hindi	69	2	35.5	a form of Ishvara or God in the later Vedic scriptures of Hinduism
Sikh	Hindi	192	68	130	ethnic group
Sind	Hindi	8	26	17	province/region
Sindhi	Hindi	1	7	4	ethnic group; language
Siva	Hindi	23	6	14.5	deity
Srinagar	Hindi	42	2	22	city
Tabla	Hindi	6	4	5	musical instrument
Taj Mahal	Hindi	41	27	34	type of building
Urdu	Hindi	56	34	45	language
Vedanta	Hindi	0	0	0	philosophy/philosophical term

Balaton	Hungarian	1	4	2.5	lake/ocean/sea
Bartok	Hungarian	17	14	15.5	artist
Biro	Hungarian	121	72	96.5	inventor
Budapest	Hungarian	44	198	121	city
Dohnanyi	Hungarian	0	0	0	artist
Esterhazy	Hungarian	1	1	1	nobleman
Gabor	Hungarian	21	15	18	physicist or artist
Kaposi	Hungarian	29	12	20.5	scientist; medical term
Kodaly	Hungarian	3	2	2.5	artist
Lehar	Hungarian	15	7	11	artist
Lukacs	Hungarian	38	6	22	politician or scholar
Magyar	Hungarian	36	14	25	ethnic group; language
Pecs	Hungarian	16	8	12	city
Pest	Hungarian	0	0	0	city; province/region
Petofi	Hungarian	0	0	0	writer
Sandor	Hungarian	31	11	21	first name
Solti	Hungarian	16	5	10.5	artist
Szeged	Hungarian	8	4	6	city
Tokay	Hungarian	5	4	4.5	alcohol/spirits
Ibo	Ibo	14	8	11	island; ethnic group; language
Bjork	Icelandic	4	4	4	artist
Hekla	Icelandic	1	1	1	mountain/valley
Keflavik	Icelandic	11	4	7.5	city
Reykjavik	Icelandic	62	33	47.5	province/region
Surtsey	Icelandic	6	5	5.5	city
Balie Atha Cliath	Irish	0	0	0	city
Cobh	Irish	6	3	4.5	city
Dail	Irish	42	19	30.5	administrative/political term
Deirdre	Irish	133	39	86	legendary character
Dun Laoghaire	Irish	1	7	4	city
Eire	Irish	161	19	90	country/state
Eithne	Irish	13	5	9	first name
Fianna Fail	Irish	5	18	11.5	political party
Gaeltacht	Irish	6	5	5.5	province/region
Laois	Irish	6	5	5.5	province/region
Ni	Irish	631	158	394.5	country
O'Fiaich	Irish	0	0	0	religious activist

Oireachtas	Irish	4	4	4	"national parliament"[2] or legislature of the Republic of Ireland
Padraig	Irish	18	13	15.5	golfer or first name
Saoirse	Irish	0	0	0	newspaper/magazine
Sean	Irish	962	263	612.5	first name
Seanad	Irish	2	2	2	administrative/political term
Sinn Fein	Irish	387	85	236	administrative/political term
Taoiseach	Irish	42	18	30	politician
Tearlach	Irish	0	0	0	first name
Abruzzi	Italian	12	12	12	city; province/region
Agrigento	Italian	4	4	4	city
Aida	Italian	31	19	25	opera
Albinoni	Italian	5	4	4.5	artist
Alitalia	Italian	17	12	14.5	airline company
Amalfi	Italian	3	8	5.5	city; province/region
Amati	Italian	4	4	4	craftsman
Ancona	Italian	2	11	6.5	city; province/region
Andrea	Italian	283	142	212.5	first name
Andrea del Sarto	Italian	4	3	3.5	artist
Angelico	Italian	1	9	5	artist
Annigoni	Italian	1	1	1	first name
Antonioni	Italian	1	9	5	artist
Anzio	Italian	1	7	4	city
Ariosto	Italian	12	6	9	writer
Armani	Italian	95	4	49.5	artist
Arno	Italian	27	23	25	river
Arturo	Italian	36	27	31.5	first name
Ascona	Italian	2	2	2	city
Asiago	Italian	1	1	1	city; province/region
Assisi	Italian	68	4	36	city
Asti	Italian	12	6	9	city; province/region
Avogadro	Italian	0	0	0	last name
Bari	Italian	73	29	51	city
Bartolommeo	Italian	1	7	4	artist
Bel Paese	Italian	4	3	3.5	cheese
Bellini	Italian	43	25	34	artist
Bergamo	Italian	17	13	15	city
Bernini	Italian	56	24	40	artist

Bertolucci	Italian	9	8	8.5	artist
Boccaccio	Italian	15	12	13.5	writer
Bodoni	Italian	5	5	5	printmaker
Bologna	Italian	2	92	47	city
Borlotti	Italian	1	2	1.5	type of vegetable
Boticelli	Italian	1	1	1	artist
Brindisi	Italian	24	13	18.5	city; province/region
Cagliari	Italian	57	2	29.5	city
Cagliostro	Italian	1	1	1	physician
Calabria	Italian	39	26	32.5	province/region
Campari	Italian	24	18	21	alcohol/spirits
Canaletto	Italian	57	26	41.5	artist
Cannelloni	Italian	11	6	8.5	dish
Cannizzaro	Italian	5	4	4.5	scientist
Canossa	Italian	2	2	2	province/region
Capodimonte	Italian	1	5	3	province/region
Capri	Italian	13	56	34.5	province/region
Capua	Italian	6	5	5.5	city
Caravaggio	Italian	0	0	0	artist
Carpaccio	Italian	6	5	5.5	artist
Carrara	Italian	17	11	14	city
Caruso	Italian	26	14	20	artist
Casanova	Italian	44	34	39	traveller
Castel Gandolfo	Italian	3	2	2.5	city
Cavalleria Rusticana	Italian	6	6	6	film
Cavour	Italian	14	6	10	politician
Cellini	Italian	26	11	18.5	artist
Cenci	Italian	7	2	4.5	writer
Cherubini	Italian	1	3	2	artist
Cimabue	Italian	7	5	6	artist
Como	Italian	73	24	48.5	city; lake
Corelli	Italian	3	1	2	artist
Correggio	Italian	1	9	5	artist
Corti	Italian	15	4	9.5	common surname; music band
Cosi Fan Tutte	Italian	8	6	7	opera-related term
Cremona	Italian	18	15	16.5	city; province/region
D'Annunzio	Italian	0	0	0	writer
da Vinci	Italian	55	47	51	artist
Dallapicolla	Italian	0	0	0	artist

Dante	Italian	146	92	119	writer
Don Carlos	Italian	23	16	19.5	politician; opera
Don Giovanni	Italian	46	17	31.5	writer
Don Pasquale	Italian	6	5	5.5	scholar
Donatello	Italian	7	4	5.5	artist
Donizetti	Italian	16	13	14.5	artist
Eco	Italian	2	54	28	writer
Elba	Italian	34	1	17.5	province/region
Enrico	Italian	34	26	30	physician
Este	Italian	21	18	19.5	city; dynasty
Eurydice	Italian	4	4	4	educational institution
Fellini	Italian	12	1	6.5	artist
Fermi	Italian	3	11	7	physician
Ferrara	Italian	38	22	30	city; province/region
Ferrari	Italian	423	97	260	sportsman
Fibonacci	Italian	1	5	3	trader
Fra	Italian	68	4	36	cartographer
Frascati	Italian	15	15	15	city; province/region
Friuli	Italian	11	7	9	province/region
Galileo	Italian	358	8	183	physician
Garda	Italian	51	4	27.5	province/region; lake
Garibaldi	Italian	31	17	24	soldier
Genoa	Italian	187	76	131.5	city
Ghia	Italian	3	15	9	car
Gianni	Italian	82	53	67.5	writer
Gigli	Italian	22	11	16.5	artist
Gioconda	Italian	4	4	4	portrayal
Giotto	Italian	38	22	30	artist
Giovanni	Italian	275	11	143	politician or popular first name
Giulietta	Italian	2	2	2	first name
Giuseppe	Italian	79	45	62	first name
Golgi	Italian	17	6	11.5	physician
Gorgonzola	Italian	11	6	8.5	province/region
Guido	Italian	473	39	256	artist or popular first name
Lamborghini	Italian	43	22	32.5	industrialist
Lambrusco	Italian	1	8	4.5	wine
Lampedusa	Italian	2	1	1.5	writer
Lancia	Italian	6	27	16.5	car
Lazio	Italian	142	29	85.5	city

Lepanto	Italian	6	5	5.5	city; province/region
Locarno	Italian	11	7	9	city; province/region
Lombardo	Italian	15	1	8	sportsman
Lombroso	Italian	25	5	15	scholar in political science
Loren	Italian	54	27	40.5	artist
Lucia	Italian	22	84	53	first name
Lugano	Italian	43	17	30	city
Machiavelli	Italian	41	3	22	writer
Maestro	Italian	145	88	116.5	title
Malpighi	Italian	4	3	3.5	physician
Marconi	Italian	126	59	92.5	inventor
Marengo	Italian	5	4	4.5	city; province/region
Marsala	Italian	9	6	7.5	city
Mascagni	Italian	3	1	2	physician
Mastroianni	Italian	1	1	1	artist
Medici	Italian	66	32	49	banker
Menotti	Italian	22	4	13	artist
Messina	Italian	43	2	22.5	city; province/region
Michelangelo	Italian	123	75	99	artist
Modena	Italian	34	26	30	city
Modigliani	Italian	594	28	311	artist
Monte	Italian	332	144	238	city; mountain; province/region
Montessori	Italian	11	1	6	scholar
Monteverdi	Italian	73	17	45	artist
Monza	Italian	6	15	10.5	city
Mozzarella	Italian	46	21	33.5	cheese
Mussolini	Italian	277	96	186.5	politician
Nessun dorma	Italian	8	7	7.5	artist
Olbia	Italian	15	4	9.5	city
Olivetti	Italian	353	97	225	manufacturer
Orsini	Italian	8	5	6.5	politican
Otranto	Italian	16	8	12	city; province/region
Padua	Italian	56	37	46.5	city
Paganini	Italian	23	15	19	artist
Pagliacci	Italian	2	2	2	artist
Palermo	Italian	91	46	68.5	city
Palestrina	Italian	34	9	21.5	city
Panini	Italian	8	6	7	dish
Pantelleria	Italian	7	3	5	province/region

Parma	Italian	192	4	98	city
Pasolini	Italian	9	6	7.5	writer
Pavarotti	Italian	74	42	58	writer
Peano	Italian	7	3	5	scholar
Pergolesi	Italian	8	5	6.5	artist
Perugia	Italian	59	15	37	city
Perugino	Italian	5	4	4.5	artist
Peruzzi	Italian	2	2	2	sportsman
Pestalozzi	Italian	8	6	7	scholar
Piacenza	Italian	15	11	13	city; province/region
Pinocchio	Italian	19	18	18.5	opera-related term
Pirandello	Italian	7	7	7	writer
Pisa	Italian	163	57	110	city
Po	Italian	0	0	0	river
Portofino	Italian	9	6	7.5	hotel
Prodi	Italian	3	2	2.5	politican
Puccini	Italian	5	2	3.5	artist
Rapallo	Italian	7	3	5	province/region
Ravenna	Italian	92	23	57.5	city
Reggio	Italian	18	12	15	city
Rialto	Italian	44	27	35.5	province/region
Ricci	Italian	64	22	43	artist
Rigoletto	Italian	18	12	15	opera
Rimini	Italian	48	26	37	city
Risorgimento	Italian	13	1	7	war period
Roberto	Italian	216	94	155	first name
Rocco	Italian	39	21	30	first name
Romagna	Italian	1	5	3	province/region
Rome	Italian	3341	772	2056.5	city
Rossini	Italian	82	36	59	artist
Salerno	Italian	25	18	21.5	city
Salieri	Italian	21	8	14.5	artist
San Remo	Italian	26	9	17.5	city
Savonarola	Italian	6	6	6	religious activist
Scarlatti	Italian	18	13	15.5	artist
Schiaparelli	Italian	32	8	20	scientist
Scutari	Italian	9	8	8.5	city; province/region
Sergio	Italian	49	37	43	first name
Siena	Italian	115	32	73.5	city; province/region

Signora	Italian	97	25	61	title
Sorrento	Italian	65	12	38.5	city
Stromboli	Italian	8	4	6	province/region
Taranto	Italian	15	9	12	city; province/region
Tasso	Italian	16	9	12.5	writer
Ticino	Italian	16	6	11	river; province/region
Tiepolo	Italian	17	13	15	artist
Tintoretto	Italian	25	16	20.5	artist
Tivoli	Italian	233	57	145	city
Toscanini	Italian	56	15	35.5	artist
Traviata	Italian	24	15	19.5	opera
Turandot	Italian	6	4	5	opera
Uccello	Italian	1	6	3.5	artist
Uffizi	Italian	29	16	22.5	type of building
Umberto	Italian	73	39	56	first name
Verdi	Italian	124	54	89	artist
Verona	Italian	139	54	96.5	city
Veronese	Italian	49	24	36.5	artist
Verrazano	Italian	1	1	1	sailor
Vespucci	Italian	3	3	3	traveller
Vinci	Italian	6	52	29	artist
Visconti	Italian	54	13	33.5	nobleman
Vivaldi	Italian	53	36	44.5	artist
Zabaglione	Italian	8	7	7.5	dish
Zeffirelli	Italian	1	8	4.5	artist
Bushido	Japanese	1	1	1	Japanese code of conduct
Fuji	Japanese	95	51	73	river; mountain; a part of many complex names. e.g. festivals, sports events, companies, etc.
Fujitsu	Japanese	371	117	244	manufacturer
Hiroshima	Japanese	112	66	89	city
Hokkaido	Japanese	37	22	29.5	province/region
Honshu	Japanese	13	6	9.5	province/region
Iwo Jima	Japanese	6	6	6	province/region
Kawasaki	Japanese	74	46	60	city; politician; artist; sportsman; manufacturer
Kobe	Japanese	21	11	16	city
Kyoto	Japanese	61	37	49	city
Kyushu	Japanese	23	13	18	province/region

Mitsubishi	Japanese	193	86	139.5	manufacturer
Nagasaki	Japanese	63	45	54	city
Nagoya	Japanese	24	18	21	city
Narita	Japanese	6	3	4.5	city
Nikkei	Japanese	97	49	73	stock market index
Nippon	Japanese	137	77	107	country; manufacturer
Nissan	Japanese	329	113	221	car
Noh	Japanese	9	8	8.5	musical drama
Okinawa	Japanese	43	18	30.5	province/region
Orinoco	Japanese	23	18	20.5	river
Osaka	Japanese	124	61	92.5	city
Ryukyu	Japanese	0	0	0	province/region
Sanyo	Japanese	4	26	15	manufacturer
Sapporo	Japanese	7	6	6.5	city
Satsuma	Japanese	19	9	14	city; province/region
Shikoku	Japanese	4	2	3	province/region
Shinto	Japanese	25	12	18.5	religion
Sumitomo	Japanese	71	33	52	manufacturer
Suzuki	Japanese	16	53	34.5	manufacturer
Tokyo	Japanese	184	414	299	city
Toshiba	Japanese	23	1	12	manufacturer
Toyota	Japanese	369	122	245.5	car
Yamaha	Japanese	216	46	131	motor or musical equipment company
Yokohama	Japanese	36	3	19.5	city
Phnom Penh	Khmer	194	5	99.5	city
Chaebol	Korean	7	4	5.5	business conglomerate
Daewoo	Korean	13	1	7	manufacturer
Hyundai	Korean	86	45	65.5	manufacturer
Inchon	Korean	3	3	3	city
Panmunjom	Korean	21	13	17	city
Pusan	Korean	12	1	6.5	city
Pyongyang	Korean	7	28	17.5	city
Samsung	Korean	94	4	49	manufacturer
Seoul	Korean	287	12	149.5	city
Taegu	Korean	7	5	6	city
Caesar	Latin	456	155	305.5	king
Riga	Latvian	98	49	73.5	city
Kaunas	Lithuanian	4	4	4	city

Skopje	Macedonian	15	1	8	city
Perak	Malay	2	2	2	province/region
Perlis	Malay	1	1	1	province/region
Maori	Maori	149	4	76.5	ethnic group; language
Bulawayo	Ndebele	45	16	30.5	city
Nkomo	Ndebele	33	11	22	politician
Amundsen	Norwegian	53	1	27	scientist
Bergen	Norwegian	65	37	51	city
Bokmal	Norwegian	0	0	0	language
Grieg	Norwegian	35	21	28	artist
Haakon	Norwegian	6	5	5.5	king
Ibsen	Norwegian	3	23	13	writer
Lofoten	Norwegian	28	7	17.5	province/region
Longyearbyen	Norwegian	3	1	2	province/region
Munch	Norwegian	11	64	37.5	artist
Olaf	Norwegian	72	18	45	first name
Oslo	Norwegian	182	115	148.5	city
Roald	Norwegian	75	32	53.5	first name
Stavanger	Norwegian	125	11	68	city; province/region
Svalbard	Norwegian	32	8	20	province/region
Trondheim	Norwegian	41	19	30	city; province/region
Utsira	Norwegian	0	0	0	province/region
Bahai	Persian	7	4	5.5	religion
Qom	Persian	14	7	10.5	city
Bialystok	Polish	3	2	2.5	city
Bydgoszcz	Polish	4	4	4	city
Gdansk	Polish	55	33	44	city
Gorecki	Polish	1	1	1	artist
Jan	Polish	1415	46	730.5	first name
Jaruzelski	Polish	51	22	36.5	politician
Katowice	Polish	1	9	5	city
Kosciusko	Polish	1	1	1	mountain; historical figure
Krakow	Polish	49	31	40	city
Lodz	Polish	18	1	9.5	city
Paderewski	Polish	8	7	7.5	artist
Pilsudski	Polish	4	3	3.5	politician
Rzeszow	Polish	2	2	2	city
Strzelecki	Polish	0	0	0	mountain; traveller
Szczecin	Polish	5	3	4	city

Torun	Polish	4	3	3.5	city
Wajda	Polish	4	4	4	artist
Walesa	Polish	264	6	135	politician
Wojtyla	Polish	6	6	6	religious activist
Wroclaw	Polish	14	7	10.5	city
Algarve	Portuguese	82	31	56.5	city
Bahia	Portuguese	21	13	17	city
Beira	Portuguese	23	12	17.5	city; province/region
Belem	Portuguese	1	1	1	city
Belo Horizonte	Portuguese	2	2	2	province/region
Benguela	Portuguese	9	6	7.5	city; province/region
Brasilia	Portuguese	18	17	17.5	city
Cabinda	Portuguese	22	1	11.5	province/region
Camoens	Portuguese	3	3	3	writer
Chagas	Portuguese	5	4	4.5	scientist
Coimbra	Portuguese	16	11	13.5	city; province/region
Copacabana	Portuguese	15	7	11	province/region; beach
Dias	Portuguese	23	16	19.5	politician
Douro	Portuguese	8	6	7	province/region
Estoril	Portuguese	36	21	28.5	racing track
Faro	Portuguese	29	16	22.5	city
Funchal	Portuguese	24	4	14	city
Gomes	Portuguese	47	18	32.5	soldier (politician) or writer
Juninho	Portuguese	0	0	0	sportsman
Lisbon	Portuguese	339	137	238	city
Luanda	Portuguese	5	25	15	city; province/region
Macau	Portuguese	26	17	21.5	province/region
Madera	Portuguese	3	3	3	province/region
Mato Grosso	Portuguese	5	5	5	river; province/region
Oporto	Portuguese	18	13	15.5	city
Para	Portuguese	0	0	0	city
Parana	Portuguese	11	8	9.5	river
Pernambuco	Portuguese	1	1	1	province/region
Porto Alegre	Portuguese	9	5	7	province/region
Rio de Janeiro	Portuguese	139	92	115.5	city
Rodriguez	Portuguese	176	49	112.5	first name
Santos	Portuguese	132	54	93	city; common surname
Sao Paulo	Portuguese	92	52	72	city
Sao Tome	Portuguese	3	3	3	city

Setubal	Portuguese	2	2	2	city
Soares	Portuguese	55	24	39.5	politician
Vasco da Gama	Portuguese	8	6	7	traveller
Verde	Portuguese	14	44	29	wine
Xingu	Portuguese	4	2	3	river
Cluj	Romanian	18	15	16.5	city
Constanta	Romanian	7	5	6	city
Oradea	Romanian	9	7	8	city
Ploesti	Romanian	2	1	1.5	province/region
Timisoara	Romanian	72	25	48.5	city
Vlad	Romanian	14	8	11	dynasty
Aeroflot	Russian	45	23	34	airline company
Amur	Russian	21	1	11	river
Andrei	Russian	188	96	142	first name
Ashkhabad	Russian	1	1	1	city
Astrakhan	Russian	16	14	15	city
Azerbaijan	Russian	376	81	228.5	country/state
Azov	Russian	5	5	5	city
Babi Yar	Russian	4	2	3	city
Baikal	Russian	35	18	26.5	mountain/valley; lake
Bakst	Russian	11	4	7.5	artist
Bakunin	Russian	6	6	6	anarchist
Birobidzhan	Russian	0	0	0	city
Bokhara	Russian	23	8	15.5	city
Bolshoi Ballet	Russian	14	9	11.5	artistic institution
Boris	Russian	651	23	337	first name
Borodin	Russian	14	9	11.5	artist; writer or scientist
Borodino	Russian	2	2	2	city
Brezhnev	Russian	199	54	126.5	politician
Brodsky	Russian	27	11	19	writer
Bug	Russian	1	1	1	river
Bukhara	Russian	2	11	6.5	city
Chekhov	Russian	47	34	40.5	writer
Chernobyl	Russian	367	117	242	city
Dnepropetrovsk	Russian	2	2	2	city
Dnieper	Russian	11	7	9	river
Dniester	Russian	8	5	6.5	river
Eisenstein	Russian	28	17	22.5	common family name; architect or artist

Feodor	Russian	0	0	0	first name
Gagarin	Russian	18	1	9.5	scientist
Glazunof	Russian	0	0	0	artist
Glinka	Russian	5	2	3.5	artist
Godunov	Russian	12	8	10	regent; scientist or ballet artist
Gogol	Russian	17	13	15	writer
Gorbachev	Russian	1723	23	873	politician
Gorki	Russian	8	7	7.5	city; political activist
Gromyko	Russian	19	9	14	politician
Grozny	Russian	8	6	7	city
Igor	Russian	95	64	79.5	first name
Ilyushin	Russian	25	9	17	industrialist
Inkerman	Russian	1	1	1	city
Irkutsk	Russian	42	16	29	city
Kamczatka	Russian	0	0	0	province/region
Kandinsky	Russian	51	19	35	artist
Karenina	Russian	25	14	19.5	literary character
Karpov	Russian	12	24	18	sportsman
Katyn	Russian	14	6	10	city
Kazan	Russian	33	15	24	city
Kerenski	Russian	0	0	0	politician
Khachaturian	Russian	2	2	2	artist
Khrushchev	Russian	263	41	152	politician
Kiev	Russian	244	115	179.5	city
Kirov	Russian	381	22	201.5	city; politician
Komsomol	Russian	27	8	17.5	youth movement
Korsakoff	Russian	4	2	3	artist or scientist
Kyzyl Kum	Russian	0	0	0	province/region
Lena	Russian	7	34	20.5	river
Lenin	Russian	578	144	361	politician
Leningrad	Russian	229	93	161	city
Lermontow	Russian	0	0	0	writer
Lysenko	Russian	1	5	3	scientist or artist
Mendeleyev	Russian	4	2	3	scientist
Mikhail	Russian	39	149	94	first name
Molotov	Russian	36	24	30	politician
Moskva	Russian	3	3	3	city
Moskvich	Russian	6	2	4	car
Moussorgsky	Russian	4	2	3	artist

Murmansk	Russian	31	25	28	city
Mussorgsky	Russian	32	9	20.5	artist
Nagorno-Karabakh	Russian	27	13	20	country/state
Nakhichevan	Russian	37	1	19	country/state
Nekrasov	Russian	2	2	2	writer; politician or filmaker
Nesselrode	Russian	14	2	8	politician
Neva	Russian	34	9	21.5	river
Nijinsky	Russian	44	18	31	artist; horse
Nizhni Novogrod	Russian	0	0	0	city
Novaya Zemlya	Russian	23	15	19	province/region
Novgorod	Russian	4	19	11.5	city
Novosibirsk	Russian	22	16	19	city
Nureyev	Russian	2	14	8	artist
Ob	Russian	119	63	91	city; lake
Odessa	Russian	39	3	21	city
Oistrakh	Russian	3	2	2.5	artist
Okhotsk	Russian	14	1	7.5	city; sea
Oleg	Russian	65	45	55	first name
Olga	Russian	144	51	97.5	first name
Omsk	Russian	17	1	9	city
Onega	Russian	1	1	1	city; river; lake
Ordzhonikidze	Russian	9	3	6	politician
Ouspensky	Russian	5	4	4.5	scholar
Pasternak	Russian	13	6	9.5	writer
Pavlov	Russian	124	5	64.5	common surname
Pavlova	Russian	34	17	25.5	artist
Perm	Russian	269	65	167	city
Petrograf	Russian	0	0	0	inventor
Petropavlovsk	Russian	1	1	1	city
Potemkin	Russian	15	14	14.5	soldier
Prawda	Russian	0	0	0	newspaper/magazine
Primakov	Russian	4	15	9.5	politician
Prokofiew	Russian	0	0	0	artist
Przewalski	Russian	4	3	3.5	scientist
Pushkin	Russian	93	38	65.5	writer
Raisa	Russian	8	5	6.5	first name
Rasputin	Russian	29	16	22.5	charlatan
Rimsky-Korsakov	Russian	13	7	10	artist

Romanov	Russian	28	14	21	royal family
Rostov	Russian	439	13	226	city
Rostropovich	Russian	13	5	9	artist
Sakhalin	Russian	0	0	0	province/region
Sakharov	Russian	52	2	27	scientist
Salyut	Russian	7	5	6	spaceflight programme
Samara	Russian	35	8	21.5	city; river
Samarkand	Russian	26	18	22	city; province/region
Saratov	Russian	37	7	22	city
Scriabin	Russian	23	8	15.5	artist
Sergei	Russian	19	76	47.5	first name
Sevastopol	Russian	0	0	0	city
Shostakovich	Russian	11	21	16	artist
Smolensk	Russian	111	12	61.5	city
Solzhenitsyn	Russian	26	17	21.5	writer
Soyuz	Russian	28	16	22	spaceflight programme
St Petersburg	Russian	442	127	284.5	city
Stalin	Russian	617	185	401	politician
Stanislavski	Russian	4	4	4	artist
Stravinsky	Russian	169	5	87	artist
Sverdlovsk	Russian	17	14	15.5	city
Tashkent	Russian	51	24	37.5	city
Tchaikovsky	Russian	156	24	90	artist
Tolstoy	Russian	16	62	39	writer
Tomsk	Russian	16	7	11.5	city
Trubetzkoy	Russian	2	1	1.5	scholar
Turgenev	Russian	33	17	25	writer
Tuva	Russian	11	6	8.5	country/state
Ural	Russian	28	2	15	river; mountain/valley; province/region
Ustinov	Russian	18	14	16	artist or politician
Vladimir	Russian	25	115	70	first name
Vladivostok	Russian	28	19	23.5	city
Volga	Russian	1	31	16	river
Volgograd	Russian	8	8	8	city
Vostok	Russian	3	3	3	province/region; spacelift programme
Yakutsk	Russian	9	3	6	city
Yeltsin	Russian	1191	131	661	politician

Yenisei	Russian	2	12	7	river
Yerevan	Russian	22	16	19	city
Yevtushenko	Russian	6	4	5	writer
Yuri	Russian	14	5	9.5	first name
Zhirinovsky	Russian	4	4	4	politician
Zhukov	Russian	63	6	34.5	politician
Zinoviev	Russian	19	13	16	writer or politician
Bheinn	Scottish-Gaelic	25	5	15	mountain/valley
Gaidhealtachd	Scottish-Gaelic	0	0	0	province/region
Sassenach	Scottish-Gaelic	9	6	7.5	Englishman or Lowland Scot
Seonaid	Scottish-Gaelic	0	0	0	a given name for a woman; Scottish Executive Online News and Information Distributor.
Seumas	Scottish-Gaelic	17	3	10	first name
Sgurr	Scottish-Gaelic	85	15	50	mountain/valley
Herzegovina	Serb-Croatian	32	1	16.5	province/region
Mohorovicic	Serb-Croatian	2	2	2	scientist; asteroid
Pristina	Serb-Croatian	27	15	21	city
Sarajevo	Serb-Croatian	479	9	244	city
Tuzla	Serb-Croatian	28	9	18.5	city
Vukovar	Serb-Croatian	32	12	22	city
Zagreb	Serb-Croatian	138	65	101.5	city
Karadzic	Serbian	57	26	41.5	politician
Milosevic	Serbian	152	44	98	politician
Pale	Serbian	0	0	0	province/region
Radovan	Serbian	48	27	37.5	first name
Subotica	Serbian	5	2	3.5	city; province/region
Bratislava	Slovak	76	48	62	city
Ljubljana	Slovene	32	17	24.5	city
Aconcagua	Spanish	3	2	2.5	mountain/valley
Albeniz	Spanish	1	1	1	artist
Alcazar	Spanish	13	7	10	type of building
Alfonso	Spanish	189	65	127	Mexican politician or writer
Algeciras	Spanish	16	1	8.5	city
Alhambra	Spanish	38	22	30	type of building
Alicante	Spanish	17	14	15.5	city; province/region
Allende	Spanish	86	34	60	writer
Almeria	Spanish	8	6	7	city; province/region
Alonzo	Spanish	2	1	1.5	first name

Alphonso	Spanish	1	7	4	politician
Altamira	Spanish	9	7	8	city; caves/rocks
Altiplano	Spanish	9	7	8	province/region
Alvarez	Spanish	48	21	34.5	Mexican (Spanish) politician or Cuban artist
Andalusia	Spanish	6	18	12	province/region
Angel	Spanish	55	31	43	writer
Angeleno	Spanish	1	1	1	inhabitant of Los Angeles
Antofagasta	Spanish	6	5	5.5	city
Aragon	Spanish	137	34	85.5	province/region
Arrecife	Spanish	0	0	0	city
Asturias	Spanish	65	18	41.5	writer
Avila	Spanish	38	24	31	city; province/region
Aymara	Spanish	14	4	9	language
Bacardi	Spanish	23	13	18	alcohol/spirits
Baja	Spanish	11	1	6	province/region
Balboa	Spanish	69	11	40	adventurer/conquistador
Balenciaga	Spanish	5	3	4	fashion house
Ballesteros	Spanish	23	46	34.5	sportsman
Barcelona	Spanish	756	238	497	city
Bilbao	Spanish	17	53	35	city
Bogota	Spanish	55	33	44	city
Bolivar	Spanish	1	6	3.5	politician
Borges	Spanish	19	1	10	writer
Buenos Aires	Spanish	197	98	147.5	city
Bunuel	Spanish	5	5	5	artist
Caballe	Spanish	7	3	5	artist
Cadiz	Spanish	98	32	65	city
Cancun	Spanish	21	6	13.5	city
Caracas	Spanish	11	44	27.5	city
Carlos	Spanish	626	194	410	Spanish heir to the throne; Spanish (Argentinian) writer or politician
Carmen	Spanish	244	1	122.5	writer
Cartagena	Spanish	4	22	13	city
Casals	Spanish	3	3	3	artist
Castro	Spanish	47	11	29	politician
Cervantes	Spanish	2	14	8	writer
Ceuta	Spanish	13	8	10.5	province/region

Chavez	Spanish	28	8	18	Mexican artist or Venezuelian politician
Chiapas	Spanish	3	3	3	province/region
Chicana	Spanish	1	1	1	cultural identity
Chicano	Spanish	0	0	0	artistic institution
Ciudad	Spanish	14	13	13.5	city; province/region
Colima	Spanish	1	1	1	city; province/region
Cordoba	Spanish	41	2	21.5	city; province/region
Cordova	Spanish	1	5	3	city
Cortez	Spanish	12	11	11.5	adventurer
Corunna	Spanish	17	5	11	city; province/region
Costa Brava	Spanish	24	18	21	province/region
Costa del Sol	Spanish	54	38	46	province/region
Costa Rica	Spanish	312	113	212.5	country/state
Cotopaxi	Spanish	2	2	2	mountain/valley
Cristobal	Spanish	9	9	9	city; province/region
Cuba	Spanish	745	195	470	country/state
Dali	Spanish	73	45	59	artist
Darien	Spanish	11	8	9.5	city; province/region
De Soto	Spanish	27	15	21	adventurer
Diaz	Spanish	44	24	34	artist
Diego	Spanish	361	184	272.5	Spanish writer; artist or first name
Dolores	Spanish	66	24	45	Spanish politcal activist or first name
Domingo	Spanish	87	58	72.5	Argentinian politician or Spanish singer
Don Quixote	Spanish	4	27	15.5	literary character
Ebro	Spanish	13	6	9.5	river
Ecuador	Spanish	263	12	137.5	country/state
El Monte	Spanish	2	2	2	city
El Paso	Spanish	11	1	6	city
El Salvador	Spanish	44	11	27.5	country/state
Escorial	Spanish	8	6	7	type of building
Evita	Spanish	17	14	15.5	Argentinian trades union activist
Falange	Spanish	49	4	26.5	political party
Falla	Spanish	14	7	10.5	artist
Fernandez	Spanish	115	57	86	Mexican writer or artist
Fernando	Spanish	69	11	40	Spanish writer or writer
Fidel	Spanish	132	68	100	politician

Franco	Spanish	954	134	544	politician
Fray Bentos	Spanish	2	2	2	city
Galapagos	Spanish	99	32	65.5	province/region
Garcia	Spanish	15	8	11.5	common surname
Geraldo	Spanish	8	6	7	first name
Gonzales	Spanish	23	17	20	common surname
Goya	Spanish	75	41	58	artist
Granada	Spanish	47	24	35.5	city; province/region
Granados	Spanish	5	5	5	artist
Guadalajara	Spanish	29	13	21	city
Guadalcanal	Spanish	14	1	7.5	province/region
Guadalquivir	Spanish	14	1	7.5	river
Guantanamo	Spanish	8	5	6.5	city; province/region
Guarani	Spanish	5	2	3.5	currency or language
Guayaquil	Spanish	12	9	10.5	city
Guernica	Spanish	11	27	19	city; river; painting
Guevara	Spanish	37	14	25.5	Spanish writer or South American revolutionist
Havana	Spanish	149	86	117.5	city
Ibiza	Spanish	93	34	63.5	province/region
Inez	Spanish	61	7	34	first name
Jacinta	Spanish	4	4	4	first name
Jerez	Spanish	26	13	19.5	wine
Jorge	Spanish	181	71	126	first name
Jose	Spanish	388	27	207.5	first name
Juan	Spanish	464	181	322.5	first name
Juanita	Spanish	23	9	16	first name
Junipero	Spanish	0	0	0	religious activist
La Nina	Spanish	1	1	1	lake/ocean/sea
Lanzarote	Spanish	6	21	13.5	province/region
Las Palmas	Spanish	25	7	16	city
Leon	Spanish	543	2	272.5	first name
Lima	Spanish	288	95	191.5	city
Lopez	Spanish	8	36	22	province/region; common surname
Lorca	Spanish	46	13	29.5	first name
Loyola	Spanish	17	15	16	religious activist
Macarena	Spanish	2	2	2	dance
Madrid	Spanish	867	278	572.5	city

Majorca	Spanish	139	62	100.5	province/region
Malaga	Spanish	42	25	33.5	city
Managua	Spanish	69	32	50.5	city
Manresa	Spanish	8	5	6.5	city
Martinez	Spanish	94	32	63	common surname
Mendez	Spanish	92	6	49	common surname
Meta	Spanish	23	13	18	city
Mexicali	Spanish	2	2	2	city
Mexico	Spanish	162	488	325	city
Miguel	Spanish	546	8	277	first name
Minorca	Spanish	16	8	12	province/region
Monterey	Spanish	22	19	20.5	city
Murillo	Spanish	13	11	12	artist
Nicaragua	Spanish	474	112	293	country/state
Noriega	Spanish	384	55	219.5	politician
Oaxaca	Spanish	5	4	4.5	city; province/region
Ortega	Spanish	131	45	88	politician
Oviedo	Spanish	19	11	15	city
Pablo	Spanish	14	63	38.5	first name
Pachuco	Spanish	0	0	0	Mexican fashion style pioneer
Palma	Spanish	16	37	26.5	city
Pamplona	Spanish	29	14	21.5	city
Panama	Spanish	66	165	115.5	country/state
Panza	Spanish	9	8	8.5	literary character
Paraguay	Spanish	116	61	88.5	country/state
Pedro	Spanish	219	98	158.5	first name
Pepe	Spanish	82	18	50	artist
Pepita	Spanish	18	12	15	artist
Peron	Spanish	36	12	24	Argentinian politician or trades union activist
Peru	Spanish	775	225	500	country/state
Picasso	Spanish	793	114	453.5	artist
Picchu	Spanish	31	3	17	mountain/valley
Pinochet	Spanish	139	42	90.5	politician
Pizarro	Spanish	32	11	21.5	adventurer
Plata	Spanish	15	11	13	river
Popocatepetl	Spanish	1	1	1	mountain/valley
Prado	Spanish	88	24	56	type of building
Puerto Rican	Spanish	36	23	29.5	inhabitant of Puerto Rico

Puerto Rico	Spanish	18	96	57	country/state
Quito	Spanish	47	29	38	city
Quixote	Spanish	54	33	43.5	literary character
Ramon	Spanish	63	37	50	common surname
Ramos	Spanish	116	44	80	writer
Raquel	Spanish	27	14	20.5	first name
Raul	Spanish	4	34	19	first name
Rio	Spanish	497	223	360	city; province/region
Rio Grande	Spanish	34	25	29.5	river
Rioja	Spanish	4	2	3	province/region; wine
Rivera	Spanish	127	41	84	province/region
Rosario	Spanish	35	21	28	city; province/region
Rosinante	Spanish	0	0	0	horse
Salamanca	Spanish	88	13	50.5	city; province/region
Salvador	Spanish	62	152	107	city
San Jose	Spanish	155	93	124	city
San Miguel	Spanish	25	14	19.5	city; province/region
San Sebastian	Spanish	33	21	27	city
Sancho	Spanish	38	9	23.5	literary character
Sancho Panza	Spanish	5	5	5	literary character
Santa Maria	Spanish	15	35	25	province/region
Santander	Spanish	69	27	48	city
Santeria	Spanish	3	2	2.5	religion
Santo Domingo	Spanish	19	16	17.5	city
Saragossa	Spanish	35	19	27	city
Segovia	Spanish	15	12	13.5	city; Spanish guitarist or town
Senor	Spanish	31	25	28	form of address
Senorita	Spanish	4	4	4	form of address
Serra	Spanish	91	22	56.5	city; religious activist
Seville	Spanish	32	18	25	city
Sierra Leone	Spanish	164	61	112.5	country/state
Sierra Madre	Spanish	5	5	5	mountain/valley
Soledad	Spanish	1	1	1	TV character
Tampico	Spanish	16	4	10	city
Tarragona	Spanish	14	8	11	city
Tenerife	Spanish	97	6	51.5	province/region
Tierra del Fuego	Spanish	27	19	23	province/region
Tijuana	Spanish	1	7	4	city; province/region
Titicaca	Spanish	15	5	10	lake/ocean/sea

Toledo	Spanish	81	42	61.5	city
Torquemada	Spanish	6	6	6	adventurer/conquistador
Torremolinos	Spanish	9	9	9	city; province/region
Trujillo	Spanish	35	26	30.5	city; province/region
Tupamaro	Spanish	0	0	0	poilitical party
Uruguay	Spanish	387	119	253	country/state
Valderrama	Spanish	12	9	10.5	province/region; sportsman
Valencia	Spanish	182	57	119.5	city
Valladolid	Spanish	26	9	17.5	city
Valparaiso	Spanish	9	8	8.5	city
Vega	Spanish	71	38	54.5	writer
Velasquez	Spanish	18	12	15	artist
Vigo	Spanish	53	8	30.5	city
Xavier	Spanish	59	36	47.5	first name
Ximenes	Spanish	2	2	2	common surname
Yorba Linda	Spanish	0	0	0	city
Yucatan	Spanish	39	11	25	province/region
Zapata	Spanish	1	9	5	politician
Zaragoza	Spanish	28	14	21	city
Mombasa	Swahili	52	23	37.5	city
Aland	Swedish	0	0	0	province/region
Berzelius	Swedish	3	1	2	scientist; secret society at Yale University
Birgitta	Swedish	5	5	5	first name
Bjorn	Swedish	67	41	54	first name
Bofors	Swedish	45	18	31.5	manufacturer
Borg	Swedish	85	34	59.5	sportsman
Ericsson	Swedish	13	35	24	manufacturer
Gothenburg	Swedish	68	42	55	city; province/region
Hammarskjold	Swedish	6	2	4	politician; educational institution
Krona	Swedish	39	12	25.5	currency
Lars	Swedish	63	38	50.5	first name
Lund	Swedish	49	37	43	city
Nobel	Swedish	37	177	107	scientist; Nobel prize
Norrkoping	Swedish	7	5	6	city
Orebro	Swedish	2	2	2	city; province/region
Palme	Swedish	39	16	27.5	politician
Sibelius	Swedish	59	23	41	artist
Stockholm	Swedish	335	196	265.5	city

Strindberg	Swedish	9	7	8	writer
Swedenborg	Swedish	8	3	5.5	scientist
Uppsala	Swedish	37	19	28	city
Phuket	Thai	29	11	20	city
Botswana	Tswana	221	81	151	country/state
Ankara	Turkish	117	53	85	city
Ataturk	Turkish	2	15	8.5	politician
Dalaman	Turkish	11	1	6	city
Doner	Turkish	2	2	2	dish
Istanbul	Turkish	329	17	173	city
Mehemet	Turkish	7	3	5	first name
Smyrna	Turkish	32	2	17	city
Twi	Twi	0	0	0	language
Lahore	Urdu	78	44	61	city
Haiphong	Vietnamese	14	5	9.5	city
Hanoi	Vietnamese	159	57	108	city
Aberaeron	Welsh	4	4	4	city
Aberavon	Welsh	20	12	16	city
Aberfan	Welsh	11	9	10	city
Abergele	Welsh	27	11	19	city
Abersoch	Welsh	6	4	5	city
Abersychan	Welsh	3	1	2	province/region
Aberystwyth	Welsh	268	87	177.5	city
Aeronwy	Welsh	0	0	0	Dylan Thomas' son
Afon	Welsh	12	7	9.5	city
Aled	Welsh	26	18	22	estate agency
Alun	Welsh	5	36	20.5	first name
Amlwch	Welsh	15	9	12	city
Aneirin	Welsh	2	2	2	writer
Angharad	Welsh	89	5	47	literary character
Anuerin	Welsh	0	0	0	politician
Arenig	Welsh	2	2	2	caves/rocks
Arfon	Welsh	23	11	17	province/region
Arwel	Welsh	7	6	6.5	scholar
Arwyn	Welsh	7	7	7	literary character
Bangor	Welsh	642	112	377	city
Beddgelert	Welsh	11	8	9.5	city
Bedwellty	Welsh	0	0	0	province/region
Bettws	Welsh	3	2	2.5	administrative/political term

Betws-yn-Rhos	Welsh	1	1	1	administrative/political term
Bews-y-Coed	Welsh	0	0	0	city
Blaenau	Welsh	58	23	40.5	city
Blodwen	Welsh	0	0	0	opera or flower enterprise
Brynmawr	Welsh	3	3	3	city
Buddig	Welsh	0	0	0	industrialist; brand name
Bwlch	Welsh	17	1	9	city
Cader Idris	Welsh	9	6	7.5	mountain/valley
Caerau	Welsh	1	1	1	province/region
Caernarfon	Welsh	17	38	27.5	city
Caerphilly	Welsh	62	28	45	city
Caersws	Welsh	12	1	6.5	city
Cain	Welsh	158	74	116	literary character
Capel	Welsh	113	69	91	city
Caradog	Welsh	4	4	4	legendary character
Ceinwen	Welsh	1	1	1	first name
Ceredig	Welsh	0	0	0	king
Ceredigion	Welsh	25	17	21	dynasty/kingdom
Cleddau	Welsh	5	4	4.5	river
Clwyd	Welsh	36	91	63.5	province/region
Coch	Welsh	16	7	11.5	type of building
Coed	Welsh	32	12	22	city; public institution
Conwy	Welsh	98	31	64.5	city; province/region
Corwen	Welsh	13	6	9.5	city
Creigiau	Welsh	0	0	0	city
Criccieth	Welsh	9	6	7.5	city
Cyfeiliog	Welsh	0	0	0	prince
Cynan	Welsh	9	1	5	king
Cynon	Welsh	24	14	19	province/region
Dafydd	Welsh	3	19	11	first name
Deganwy	Welsh	19	8	13.5	city
Dolgellau	Welsh	2	11	6.5	city
Dovey	Welsh	22	15	18.5	railway station
Dyfed	Welsh	149	73	111	dynasty/kingdom
Dyffryn	Welsh	13	12	12.5	province/region
Dyfrig	Welsh	0	0	0	religious activist
Dylan	Welsh	29	15	22	writer
Dynevor	Welsh	6	2	4	royal family
Ednyfed	Welsh	5	2	3.5	historical figure

Eglwys	Welsh	1	1	1	religion
Eifion	Welsh	3	3	3	sportsman
Eirian	Welsh	1	1	1	first name
Eirlys	Welsh	1	1	1	first name
Emrys	Welsh	4	2	3	first name
Emyr	Welsh	4	22	13	first name
Eurig	Welsh	0	0	0	politician
Eynon	Welsh	12	4	8	city
Fach	Welsh	23	9	16	city; caves/rocks; lake
Fan	Welsh	14	4	9	mountain/valley
Fawr	Welsh	31	15	23	mountain/valley
Felindre	Welsh	4	2	3	city
Festiniog	Welsh	3	1	2	city
Ffestiniog	Welsh	71	23	47	province/region
Ffion	Welsh	0	0	0	first name
Ffynnongroew	Welsh	0	0	0	city
Fron	Welsh	12	9	10.5	city
Froncysyllte	Welsh	0	0	0	city
Gabalfa	Welsh	4	2	3	province/region
Gaenor	Welsh	0	0	0	film character
Gareth	Welsh	712	181	446.5	knight of the round table
Gelligaer	Welsh	0	0	0	city
Geraint	Welsh	45	27	36	literary character
Glanyrafon	Welsh	0	0	0	province/region
Glyder	Welsh	3	2	2.5	mountain/valley
Goronwy	Welsh	6	4	5	literary character
Gorsedd	Welsh	1	1	1	artistic institution
Gorseinon	Welsh	2	2	2	city; administrative/political term
Graig	Welsh	7	6	6.5	province/region administrative/political term
Gregynog	Welsh	8	4	6	type of building
Gruffydd	Welsh	11	8	9.5	prince
Gwaun-cae-Gurwen	Welsh	0	0	0	city; administrative/political term
Gwenllian	Welsh	1	1	1	names of medieval female characters
Gwynedd	Welsh	23	69	46	province/region
Gwynfor	Welsh	3	3	3	artist
Hafod	Welsh	5	5	5	city
Harlech	Welsh	41	2	21.5	city

Heulwen	Welsh	0	0	0	first name
Hirwain	Welsh	0	0	0	city
Huw	Welsh	88	41	64.5	first name
Hyder	Welsh	2	2	2	manufacturer
Hywel	Welsh	116	16	66	first name
Iestyn	Welsh	11	9	10	king
Ifor	Welsh	4	5	4.5	manufacturer
Illtud	Welsh	0	0	0	saint
Ioan	Welsh	7	6	6.5	first name
Iolo	Welsh	2	2	2	literary character
Iorwerth	Welsh	16	6	11	historical figure
Ithon	Welsh	0	0	0	river; mountain
Iwan	Welsh	17	14	15.5	politician
Lago	Welsh	17	13	15	first name
Leuan	Welsh	42	22	32	first name
Llanberis	Welsh	54	24	39	city
Llandeilo	Welsh	15	1	8	city
Llandovery	Welsh	27	15	21	city
Llandrindod	Welsh	41	22	31.5	city
Llandudno	Welsh	185	57	121	city
Llanelli	Welsh	223	49	136	city
Llanfairfechen	Welsh	0	0	0	city
Llanfairpwll	Welsh	6	3	4.5	city
Llangollen	Welsh	72	31	51.5	city
Llanwrst	Welsh	1	1	1	city
Llanwrtyd	Welsh	5	4	4.5	city
Lleyn	Welsh	11	8	9.5	province/region
Llywelyn	Welsh	11	6	8.5	historical figures
Mabinogion	Welsh	3	3	3	collection of prose stories
Machynlleth	Welsh	85	28	56.5	city
Maelor	Welsh	27	13	20	province/region
Maendy	Welsh	1	1	1	city
Maentwrog	Welsh	0	0	0	city
Maesteg	Welsh	29	18	23.5	city
Mair	Welsh	131	4	67.5	common surname
Mawddach	Welsh	1	5	3	river
Meirion	Welsh	29	9	19	prince; educational institution
Menai	Welsh	48	27	37.5	city; province/region
Meredydd	Welsh	1	1	1	first name

Merfyn	Welsh	5	3	4	scholar
Merioneth	Welsh	9	6	7.5	province/region
Merthyr	Welsh	57	39	48	city; province/region
Merthyr Tydfil	Welsh	24	2	13	city; province/region
Meurig	Welsh	8	4	6	city; prince or chemist
Morfa	Welsh	16	9	12.5	city; building
Morfudd	Welsh	0	0	0	legendary character
Myfanwy	Welsh	1	5	3	first name or Welsh artist
Nant	Welsh	28	15	21.5	part of the name of the Welsh language or heritage centre
Nantffrancon	Welsh	0	0	0	mountain/valley
Nantgarw	Welsh	3	2	2.5	city
Ogwr	Welsh	4	4	4	province/region
Olwen	Welsh	18	8	13	legendary character
Owain	Welsh	25	2	13.5	historical figure
Padarn	Welsh	3	3	3	lake
Pantycelyn	Welsh	1	1	1	writer
Penarth	Welsh	37	19	28	city
Penmaemawr	Welsh	0	0	0	province/region
Penrhos	Welsh	7	6	6.5	educational institution
Penrhyndeudraeth	Welsh	2	2	2	city
Pentre	Welsh	1	9	5	city
Penybont	Welsh	0	0	0	city
Pen-y-groes	Welsh	0	0	0	city
Plaid Cymru	Welsh	39	24	31.5	political party
Plynlimon	Welsh	4	3	3.5	mountain/valley
Pontardawe	Welsh	1	1	1	city
Pontardulais	Welsh	9	2	5.5	city; administrative/political term
Pontllan-fraith	Welsh	0	0	0	educational institution
Pontypoll	Welsh	0	0	0	type of building
Pontypridd	Welsh	66	31	48.5	city
Porth Dinilaen	Welsh	0	0	0	province/region
Porthmadog	Welsh	65	17	41	city
Prescelly	Welsh	0	0	0	mountain/valley
Prestatyn	Welsh	75	23	49	city
Pwllheli	Welsh	58	23	40.5	city
Rheidol	Welsh	24	1	12.5	river
Rhian	Welsh	5	4	4.5	first name
Rhiannon	Welsh	19	1	10	legendary character

Rhianydd	Welsh	0	0	0	artist
Rhinog	Welsh	2	1	1.5	mountain/valley
Rhiwbina	Welsh	0	0	0	province/region
Rhodri	Welsh	29	7	18	historical figures
Rhondda	Welsh	5	33	19	river; province/region administrative/ political term
Rhonwen	Welsh	0	0	0	first name
Rhos	Welsh	21	9	15	city
Rhosllanerchrugog	Welsh	0	0	0	city
Rhosneigr	Welsh	1	1	1	city
Rhuddlan	Welsh	42	19	30.5	city
Rhyd-ddu	Welsh	5	3	4	city
Rhydderch	Welsh	2	1	1.5	prince
Rhyl	Welsh	12	33	22.5	city
Rhys	Welsh	129	54	91.5	writer
Ruabon	Welsh	15	9	12	city
Seiriol	Welsh	1	1	1	saint
Sulwen	Welsh	0	0	0	writer
Sulwyn	Welsh	0	0	0	first name
Talacre	Welsh	4	4	4	city
Talfan	Welsh	2	2	2	writer or first name
Taliesin	Welsh	85	6	45.5	writer
Talybont	Welsh	17	8	12.5	city
Tal-y-llyn	Welsh	2	2	2	lake
Tawe	Welsh	0	0	0	river
Teifi	Welsh	4	9	6.5	river
Teleri	Welsh	1	1	1	literary character
Tonfanau	Welsh	1	1	1	city
Tonypandy	Welsh	12	8	10	city
Tonyrefail	Welsh	1	1	1	city
Torfaen	Welsh	36	1	18.5	province/region
Towy	Welsh	12	4	8	river
Towyn	Welsh	3	15	9	city
Trawsfynydd	Welsh	22	11	16.5	city
Trefor	Welsh	6	4	5	city
Tryfan	Welsh	7	6	6.5	mountain/valley
Tudur	Welsh	1	1	1	religious activist
Ty	Welsh	44	29	36.5	manufacturer
Tywyn	Welsh	24	16	20	city

Urdd Gobaith Cymru	Welsh	0	0	0	youth movement
Waunfawr	Welsh	1	1	1	city
Wylfa	Welsh	8	6	7	nuclear power station
Ynys	Welsh	37	17	27	province/region
Ynys-ddu	Welsh	0	0	0	administrative/political term
Ynysybwl	Welsh	2	2	2	city
Ystalyfera	Welsh	1	1	1	city
Ystrad	Welsh	5	5	5	city
Ystradgynlais	Welsh	6	3	4.5	city
Ystwyth	Welsh	2	2	2	title
Mandela	Xhosa	61	126	93.5	politician
Soweto	Xhosa	96	42	69	province/region
Thabo	Xhosa	8	8	8	politician
Xhosa	Xhosa	13	12	12.5	ethnic group; language
Ibadan	Yoruba	15	9	12	city
Buthelezi	Zulu	98	41	69.5	politician
Cetshwayo	Zulu	0	0	0	king
Hluhluwe	Zulu	0	0	0	city; river
Inkatha	Zulu	196	49	122.5	political party

Common words and phrases

Lemma	Language	Solutions	Texts	CRACn1	Meaning
rand	Afrikaans	118	57	87.5	a currency unit
sjambok	Afrikaans	4	3	3.5	the traditional heavy leather whip
verkrampte	Afrikaans	0	0	0	descriptive of rigidly conservative political attitudes of an Affrikaner nationalist
fatwa	Arabic	55	22	38.5	a legal pronouncement in Islam
halal	Arabic	13	10	11.5	permissible or edible
intifada	Arabic	167	51	109	uprising
mecca	Arabic	193	139	166	any important site for any particular group of people
mujaheddin	Arabic	172	31	101.5	those who engage in Islamic holy war – jihad („struggle")
sharia	Arabic	43	25	34	the body of Islamic law
dim sum	Cantonese	9	6	7.5	light Chinese meal

wok	Cantonese	48	21	34.5	a versatile round-bottomed cooking vessel
feng shui	Chinese	14	1	7.5	Chinese art of interior design
guoyu	Chinese	0	0	0	the standardized spoken Chinese language
kung fu	Chinese	150	24	87	chinese martial arts
pinyin	Chinese	0	0	0	a system of romanization for Standard Mandarin
renminbi	Chinese	4	4	4	the official currency in the mainland of the People's Republic of China
shih-tzu	Chinese	0	0	0	a breed of dog originating in Tibet
tai chi	Chinese	31	8	19.5	an internal Chinese martial art
yuan	Chinese	426	48	237	a currency unit
krone	Danish	32	17	24.5	a currency unit
oersted	Danish	5	3	4	a unit of magnetic field strength
Verner's law	Danish	1	1	1	descriptive of sound changes
sauna	Finnish	298	92	195	sauna
a gogo	French	3	3	3	a phrase for 'until you are satisfied'
a la carte	French	89	29	59	a phrase for 'from the menu'
a la Grecque	French	1	1	1	a phrase for 'in the Greek fashion'
a la mode	French	3	3	3	a word for 'fashionable'
abbe	French	56	10	33	a word for 'abbot'
actualite	French	0	0	0	a word for 'reality'
adieux	French	2	2	2	a word for 'farewell'
affaire	French	12	8	10	a word for 'matter'
agent provocateur	French	8	8	8	a word for 'agitator'
aide-de-camp	French	22	16	19	a word for 'adjutant'
aide-memoire	French	8	6	7	a word for 'notebook'
aiguille	French	13	7	10	a word for 'needle'
allemande	French	3	3	3	an adjective 'related to German'
ambience	French	191	133	162	a word for 'atmosphere'
amour-propre	French	9	9	9	a word for 'own love'
ampere	French	2	2	2	a word for a unit of electric current
ancien regime	French	2	2	2	a word for former political ruling powers
anis	French	20	5	12.5	a type of flowering plant
apercu	French	1	1	1	a word for 'outline'
aperitif	French	49	40	44.5	a word for 'appetizer or starter'

appellation controlee	French	0	0	0	a high quality product
apres	French	12	11	11.5	a word for 'after, behind'
arete	French	5	2	3.5	a word for 'bone'
arriere-pensee	French	0	0	0	a word for 'hidden thought'
arriviste	French	11	11	11	a word for 'career-maker'
arrondissement	French	15	14	14.5	a word for 'district'
art deco	French	2	2	2	a kind of art
art nouveau	French	73	48	60.5	a kind of art
artiste	French	32	24	28	a word for 'artist'
assai	French	3	2	2.5	a word for 'a lot'
assez	French	4	4	4	a word for 'sufficiently'
assignat	French	0	0	0	a word for 'voucher'
atelier	French	28	18	23	a word for 'workshop'
au contraire	French	8	8	8	a word for 'opposite'
au courant	French	0	0	0	a word for 'up-to-date'
au fait	French	32	32	32	a word for 'actually'
au fond	French	2	2	2	a phrase for 'as a matter of fact'
au gratin	French	5	3	4	a phrase for 'baked, roasted'
au jus	French	0	0	0	a phrase for 'with juice'
au lait	French	2	2	2	a phrase for 'with milk'
au naturel	French	0	0	0	a phrase for 'without seasoning'
au pair	French	53	33	43	a word for a 'baby-sitter'
au revoir	French	26	25	25.5	a word for 'see you'
au vin	French	14	9	11.5	a word for 'with wine'
aubade	French	6	4	5	a poem or song of or about lovers separating at dawn
auberge	French	36	10	23	a word for 'inn, tavern'
auteur	French	13	10	11.5	a word for 'author'
autoroute	French	17	15	16	a word for 'motorway'
avant-garde	French	313	111	212	a word that refers to people or works that are experimental or innovative, particularly with respect to art, culture and politics
baccarat	French	15	12	13.5	a type of glass
badinage	French	18	18	18	a word for 'trick'
baguette	French	20	15	17.5	a variety of bread
ballade	French	16	8	12	a word for 'ballad'

bandeau	French	5	4	4.5	literally a 'band' worn around a woman's breasts in French
barre	French	6	4	5	a word for 'pole, bar'
barré	French	9	6	7.5	a word for 'divided, crossed out'
battue	French	0	0	0	a word for 'battue'
bearnaise	French	1	1	1	a phrase for descriptive of the province
beau	French	133	83	108	a word for 'elegant, refined'
beau geste	French	11	11	11	a phrase for 'nice gesture'
beau monde	French	5	4	4.5	a phrase for 'grand world'
beaux-arts	French	56	19	37.5	a word for 'fine arts'
bechamel	French	8	5	6.5	a word for 'white sauce'
beignet	French	0	0	0	a type of cookie
belle	French	357	156	256.5	a word for 'beautiful, refined'
belle epoque	French	12	10	11	a phrase for 'nice epoch'
belles-lettres	French	0	0	0	a word for 'fiction'
berceuse	French	9	4	6.5	a word for 'guardian'
bete noire	French	4	4	4	a word that refers to someone/ something unwanted or even hated
betise	French	0	0	0	a word for 'stupidity'
bidet	French	39	22	30.5	a word for 'bidet'
bien-pensant	French	1	1	1	a word for 'optimistic'
bijouterie	French	1	1	1	a word for 'jewellery'
billet-doux	French	4	4	4	a word for 'love letter'
bisque	French	11	8	9.5	a type of cancer soup
bistro	French	69	35	52	a type of bar
blanquette	French	8	3	5.5	a type of grape
blanquette de veau	French	0	0	0	a type of veal dish
blasé	French	18	17	17.5	a word for 'bored'
blouson	French	16	14	15	a word for 'sweatshirt'
boeuf	French	32	28	30	a word for 'ox'
boeuf bourguignon	French	4	3	3.5	a type of beef dish
bombe	French	5	3	4	a word for 'bombe'
bon	French	176	100	138	a word for 'good, suitable'
bon voyage	French	10	10	10	a kind of farewell saying
bonne bouche	French	0	0	0	a word for 'good taste'
bonne femme	French	3	3	3	a word for 'woman or wife'
bonvivant	French	7	7	7	a word descriptive of 'the pleasure seeking person'

bouchée	French	0	0	0	a word for 'nibble'
bougi	French	0	0	0	a word for 'candle'
bouillabaisse	French	12	10	11	a type of fish soup
bouillon	French	14	11	12.5	a word for 'stock or broth'
bouillon cube	French	2	1	1.5	a word for 'stock cube'
boules	French	12	10	11	a type of ball game
bourgeois	French	1009	257	633	a word for 'townsman'
bourgeoisie	French	581	133	357	a word descriptive of the social class
bourrée	French	0	0	0	a bundle of brushwood
brie	French	57	25	41	a type of cheese
brioche	French	15	8	11.5	a type of sweet cake
broderie anglaise	French	19	16	17.5	a type of English embroidery
brut	French	46	26	36	a word for 'crude'
bureau	French	1433	465	949	a word for 'office, desk'
bureau de change	French	7	6	6.5	a point where you can exchange foreign currency
c'est la vie	French	14	11	12.5	a phrase for 'this is life'
cabernet sauvignon	French	59	22	40.5	a type of wine
cabochon	French	56	5	30.5	a word for 'expensive stone'
café	French	574	215	394.5	a word for 'coffee or cafe'
cafetiere	French	133	82	107.5	a type of jug
canard	French	23	21	22	a word for 'duck (also newspaper gossip)'
capote	French	19	9	14	a word for 'overcoat'
carnet	French	3	2	2.5	a word for 'notebook'
carte blanche	French	57	54	55.5	a word for 'power of attorney'
cassis	French	26	16	21	a word for 'blackcurrant'
cassoulet	French	5	4	4.5	a type of dish based on beans and pork
cause célèbre	French	7	7	7	a phrase for 'known reason'
causerie	French	0	0	0	a word for 'conversation'
cayenne	French	43	16	29.5	a type of pepper
centime	French	1	1	1	a type of currency unit
chaconne	French	8	3	5.5	a type of dance
chacun a son gout	French	0	0	0	a phrase for 'everbody has a different taste'
chaise longue	French	18	13	15.5	a type of an upholstered couch in the shape of a chair

chamber	French	27	19	23	a word for 'room, chamber'
champagne	French	1794	495		a type of alcoholic drink
champignon	French	0	0	0	a word for 'mushroom'
chanson	French	47	13	30	a word for 'song'
chanteuse	French	12	11	11.5	a word for 'female singer'
charge d'affaires	French	4	4	4	a word for 'plenipotentiary'
chartreuse	French	19	15	17	a word for 'cloister'
chasseur	French	6	4	5	a word for 'hunter'
chateau	French	220	98	159	a word for 'castle, palace'
chateaux	French	24	20	22	a word for 'palaces, castles' (plural)
chatelaine	French	13	12	12.5	a word for 'castellan'
chaud-froid	French	0	0	0	a word for 'poultry, venison'
chef d'oeuvre	French	0	0	0	a word for a 'masterpiece'
chemindefer	French	2	2	2	a word for 'railway'
cherchez la femme	French	3	3	3	a phrase for 'search for a wife'
chevre	French	1	1	1	a word for 'goat'
chez nous	French	4	4	4	a phrase for 'in our country'
chignon	French	21	17	19	a word for 'bun'
chinoiserie	French	13	12	12.5	a word denoting a complicated procedure
cinema-verite	French	0	0	0	a style of filmmaking
cocotte	French	3	2	2.5	a word for 'cocotte'
cognac	French	133	66	99.5	a type of drink
coiffeur	French	1	1	1	a word for 'hairdresser'
coiffure	French	16	0	8	a word for 'hairstyle'
collage	French	154	77	115.5	a word denoting the idea of combination of different elements or patterns
colon	French	6	1	3.5	a word for 'settler'
colporteur	French	1	1	1	a word for 'distributer'
comme il faut	French	0	0	0	a phrase for 'as much as is required'
concierge	French	30	16	23	a word for 'caretaker'
concours	French	2	2	2	a word for 'competition'
concours d'elegance	French	0	0	0	a phrase for 'beauty contest'
confrere	French	1	1	1	a word for 'countryman'
conge	French	0	0	0	a word for 'a leave'
consomme	French	2	2	2	a word for 'consumed'
conte	French	19	6	12.5	a word for 'told, made up'
conte	French	6	4	5	a word for 'a tale'

contretemps	French	11	11	11	a word for 'an obstacle'
coqauvin	French	14	9	11.5	a phrase for 'chicken in wine'
coquille	French	0	0	0	a word for 'shell'
cordon bleu	French	45	33	39	a phrase for 'excellent cook'
cordon sanitaire	French	9	9	9	a phrase for 'plumbing system'
cortege	French	30	22	26	a word for 'procession'
corvee	French	1	1	1	a word for 'serfdom'
coulis	French	14	10	12	a word for 'puree, paste'
coulomb	French	13	3	8	a unit of electricity
coup d'etat	French	18	17	17.5	a word for 'military coup'
coup de theatre	French	2	2	2	a phrase for 'a twist'
coups	French	93	70	81.5	a word for 'strikes, shots'
court-bouillon	French	0	0	0	a word for 'stock'
couture	French	125	56	90.5	a word for 'tailoring'
creche	French	131	78	104.5	a word for 'infant bed'
'creme	French	79	42	60.5	a word for 'cream'
creme de la creme	French	10	9	9.5	a phrase for 'creamy cream'
creme de menthe	French	3	3	3	a type of mint cream
creme fraiche	French	0	0	0	a phrase for 'fresh cream'
crepe	French	54	35	44.5	a type of very thin cooked pancake
cri de Coeur	French	5	5	5	a phrase for 'voice of heart'
crime passionel	French	0	0	0	a phrase for 'love-driven murder'
croissant	French	28	23	25.5	a type of buttery flaky pastry
crouton	French	0	0	0	a word for 'heel of bread'
cru	French	90	17	53.5	a word for 'crude,' often descriptive of wine brand
crudités	French	2	2	2	a word for 'raw dish'
cuisine	French	481	156	318.5	a specific set of cooking traditions and practices
cuisine minceur	French	1	1	1	a type of modest cuisine
cul-de-sac	French	124	94	109	a word for 'dead end'
culotte	French	4	4	4	a word for 'knee long trousers'
curé	French	19	4	11.5	a word for 'parish-priest'
dance	French	37	19	28	a word for 'dance'
danseur	French	5	3	4	a word for 'male dancer'
danseuse	French	5	5	5	a word for 'female dancer'
daube	French	6	4	5	a kind of stewed meat
dauphin	French	59	21	40	a word for 'dolphin'

débâcle	French	131	112	121.5	a word for 'downfall'
debutante	French	24	22	23	a word for 'novice'
declasse	French	1	1	1	a word for 'outclassed'
decolletage	French	2	2	2	a phrase for 'wear a low-cut dress'
decollete	French	4	4	4	a word for 'cleft created by the partial exposure of a woman's breasts'
decoupage	French	5	4	4.5	a word for 'the art of decorating an object by gluing colored paper cut outs onto it in combination with special paint effects'
deja vu	French	23	22	22.5	a phrase for 'already seen'
deluxe	French	98	46	72	a word for 'luxurious'
demarche	French	4	4	4	a word for 'walk, conduct'
demi-mondaine	French	1	1	1	a word descriptive of 'underworld'
demimonde	French	4	2	3	a phrase for 'underworld'
demitasse	French	1	1	1	a phrase for 'half cup'
demode	French	2	2	2	a word for 'old-fashioned'
demoiselle	French	14	5	9.5	a word for 'young girl'
denier	French	10	5	7.5	a type of unit of currency
dénouement	French	42	36	39	a word for 'a series of events that follow a dramatic or narrative's climax'
derigeur	French	5	5	5	a word for 'indispensable'
dernier cri	French	1	1	1	a phrase for 'latest fashion'
deshabille	French	1	1	1	a word for 'in the nude'
détente	French	42	27	34.5	a word for 'pastime'
detrop	French	6	6	6	a word for 'redundant'
diablerie	French	0	0	0	a word for 'devilish practices'
Dieu est mon droit	French	2	2	2	a phrase for 'God is my law'
digestif	French	3	3	3	a word for 'digestion pill'
dirigisme	French	7	5	6	a word for 'state capitalism'
dirigiste	French	14	12	13	a word for 'state capitialist'
distingue	French	0	0	0	a word for 'distinguished'
distrait	French	3	3	3	a word for 'absent-minded'
divertissement	French	11	6	8.5	a word for 'distraction, entertainment'
dossier	French	123	74	98.5	a collection of papers, or other sources, containing detailed information about a particular person or subject

double entendre	French	4	4	4	a phrase that has two meanings, one 'innocent and literal,' the other 'risque or bawdy;' also a word for 'innuendo'
douceur	French	7	6	6.5	a word for 'sweetness'
doyen	French	58	56	57	a word for 'dean'
doyenne	French	11	10	10.5	a word for 'elder'
droit de seigneur	French	5	5	5	a word for 'feudal law'
duchesse	French	26	15	20.5	a word for 'countess'
eau	French	53	35	44	a word for 'water'
eau de cologne	French	0	0	0	a word for 'Cologne water'
eau de Nil	French	3	2	2.5	a phrase for 'water of the Nile'
eau de vie	French	3	3	3	a word for 'vodka'
ecarte	French	2	1	1.5	a word for 'dismissed'
eclair	French	10	8	9	a word for 'flash of lightning'
eclat	French	3	3	3	a word fro 'sheen, lustre'
ecraseur	French	0	0	0	a word for 'crasher'
ecu	French	279	88	183.5	a word for 'shield'
elan	French	19	15	17	a word for 'elk'
elan vital	French	1	1	1	a phrase for 'enthusiasm for life'
embarras de richesse	French	0	0	0	a phrase for 'too wide a selection of options'
embonpoint	French	5	4	4.5	a word for 'corpulence'
embouchure	French	3	1	2	a word for 'mouthpiece'
embourgeoisement	French	11	6	8.5	a word for 'acquisition of bourgeois habits'
eminence grise	French	7	7	7	a phrase for 'grey eminence'
en bloc	French	40	36	38	a phrase for 'in bulk'
en brochette	French	0	0	0	a phrase for 'on a pointed end'
en brosse	French	0	0	0	a word for 'closely-cropped hair'
en famille	French	7	6	6.5	a phrase for 'in family'
en fete	French	4	4	4	a word for 'cheerful, gay'
en garde	French	0	0	0	a phrase for 'on guard'
en masse	French	115	97	106	a phrase for 'in a crowd'
en passant	French	12	11	11.5	a phrase for 'passing by'
en route	French	524	344	434	a phrase for 'on one's way'
enceinte	French	5	4	4.5	a word for 'pregnant woman'
encroute	French	2	2	2	a word for 'covered in shell'
enfant terrible	French	24	16	20	a phrase for 'prodigy child'

engagé	French	5	3	4	a word for 'engaged'
enjambement	French	1	1	1	the breaking of a syntactic unit (a phrase, clause, or sentence) by the end of a line
ennui	French	27	27	27	a word for 'boredom'
ensuite	French	205	51	128	a word for 'next'
entente	French	53	34	43.5	a word for 'agreement'
entourage	French	208	141	174.5	a word for 'surroundings'
entracte	French	2	2	2	a word for 'interval'
entre nous	French	3	3	3	a phrase for 'between me and you'
entrechat	French	1	1	1	a word for 'leap, bound'
entrecote	French	1	1	1	a word for 'rib steak'
entree	French	12	12	12	a word for 'first course'
entremets	French	6	1	3.5	a word for 'starter'
entrepot	French	7	6	6.5	a word for 'warehouse'
entrepreneur	French	288	192	240	a word for 'businessman'
entresol	French	1	1	1	a word for 'mezzanine'
epee	French	6	4	5	a word for 'sword'
escargot	French	6	5	5.5	a word for 'snail'
esprit d'escalier	French	0	0	0	a phrase for 'plodding sense of humour'
esprit de corps	French	34	29	31.5	a phrase for 'corporality'
estaminet	French	3	3	3	a word for 'eatery'
estragon	French	0	0	0	a word for 'tarragon'
etagere	French	0	0	0	a word for 'etagere'
etude	French	8	6	7	a word for 'etude, study'
etui	French	0	0	0	a word for 'case holder'
explication	French	32	19	25.5	a word for 'explanation'
exposé	French	25	23	24	a word for 'prime minister's inaugural speech'
extraordinaire	French	28	27	27.5	a word for 'extraordinary'
fabliau	French	186	1	93.5	satirical verse tale
faience	French	31	14	22.5	a type of tin-glazed earthenware ceramic
faineant	French	1	1	1	a word for 'layabout'
fait accompli	French	66	62	64	a phrase for 'accomplished deed'
faites vos jeux	French	0	0	0	a phrase for 'put a bet'
farandole	French	0	0	0	a type of dance
farceur	French	3	3	3	a word for 'joker'

farci	French	1	1	1	a word for 'stuffed'
farouche	French	3	3	3	a word for 'ardent, zealous'
faute de mieux	French	0	0	0	a phrase for 'lack of something better'
faux amis	French	1	1	1	a phrase for 'false friends'
faux pas	French	27	27	27	a word for 'blunder'
faux-naif	French	0	0	0	a phrase for 'pretending to be naïve'
femme	French	81	61	71	a word for 'a woman'
femme fatale	French	32	29	30.5	a phrase for 'bad woman'
fête	French	114	61	87.5	a word for 'holiday'
fête champetre	French	2	1	1.5	a phrase for 'folk festivities'
fiancé	French	109	66	87.5	a word for 'engaged' or 'a person engaged'
filet mignon	French	2	2	2	a phrase for 'surloin'
fils	French	19	15	17	a word for 'son(s)'
fin de siecle	French	6	6	6	a phrase for 'the end of the century'
fine	French	1	1	1	a word for 'end'
fines herbes	French	0	0	0	a phrase for 'aromatic herbs'
flambe	French	2	1	1.5	a word for 'lost'
fleur-de-lis	French	12	10	11	a word for 'Bourbon lily'
foie gras	French	2	1	1.5	pharse for 'liver plate'
folie a deux	French	0	0	0	a word for 'psychosis'
folie de grandeur	French	4	4	4	a phrase for 'superiority complex'
fondu	French	1	1	1	a word for 'melted'
force majeure	French	38	20	29	a phrase for 'major force'
formidable	French	1	1	1	a word for 'wonderful'
fouette	French	0	0	0	a word for 'whipped, flogged'
foyer	French	362	187	274.5	a word for 'fire, home'
franc	French	295	101	198	unit of currency
frappe	French	4	4	4	a word for 'imprinted'
frise	French	1	1	1	a word for 'freeze'
frisson	French	95	83	89	a word for 'shiver'
fromage frais	French	65	19	42	a phrase for 'fresh cheese'
frottage	French	4	4	4	a word for 'polishing'
galere	French	0	0	0	a word for 'galley'
gamin	French	3	2	2.5	a word for 'rascal'
gamine	French	10	8	9	a word for 'little girl'
garcon	French	0	0	0	a word for 'boy, waiter'

gateau	French	31	22	26.5	a word for 'cake'
gaucherie	French	8	8	8	a word for 'clumsiness'
gendarme	French	43	16	29.5	a word for 'French policeman'
genre	French	592	224	408	a word for 'type, species'
gigot	French	14	8	11	a word for 'leg'
gigue	French	1	1	1	a word for 'haunch'
gite	French	2	2	2	a word for 'shelter, lair'
gouache	French	103	28	65.5	a thick, opaque watercolour paint
goujon	French	0	0	0	a word for 'bolt'
gourmand	French	5	5	5	a word for 'glutton'
gourmet	French	126	91	108.5	a word for 'taster'
grand	French	102	45	73.5	a word for 'grand, sumptuous'
grand mal	French	0	0	0	a word for 'big disaster'
grand prix	French	767	155	461	a word for 'main prize'
grisaille	French	14	10	12	type of fabric
habitue	French	2	2	2	a word for 'accustomed'
haute	French	143	85	114	a word for 'tall'
hauteur	French	32	30	31	a word for 'height'
honi soit qui mal y pense	French	4	4	4	a phrase for 'shame on the person who thinks ill about it'
hors de combat	French	0	0	0	a phrase for 'unfit to fight'
hors-d'oeuvre	French	13	9	11	a word for 'snack'
idee fixe	French	1	1	1	a phrase for 'obsessive thought'
idee recu	French	0	0	0	a phrase for 'accepted views'
idiot savant	French	3	3	3	a phrase for 'pseudoscholar'
impasse	French	172	135	153.5	a word for 'dead end'
ingenue	French	8	7	7.5	a word for 'innocent'
ingression	French	35	28	31.5	a word for 'raid'
insouciance	French	36	33	34.5	a word for 'carefree (female)'
insouciant	French	21	20	20.5	a word for 'carefree (male)'
jabot	French	9	6	7.5	a word for 'ruffle'
jardiniere	French	2	2	2	a word for 'gardener's wife'
je ne sais quoi	French	6	6	6	a phrase denoting 'something unknown'
jete	French	1	1	1	a word for 'outcast, dismissed'
jeu d'esprit	French	3	3	3	a phrase for 'mental distraction'
jeunesse doree	French	1	1	1	a phrase for 'golden youth'
joie de vivre	French	24	21	22.5	a phrase for 'joy of life'

jongleur	French	17	2	9.5	a word for 'juggler'
julienne	French	13	12	12.5	a word for 'julian'
kepi	French	3	2	2.5	a type of hat
kir	French	10	7	8.5	a type of blackurrant syrup
laisser-faire	French	136	90	113	a word for 'permit'
langouste	French	3	1	2	a word for 'cray fish'
langoustine	French	7	4	5.5	a word for 'chicken kebab'
langue	French	46	22	34	a word for 'language, speech'
langue de chat	French	0	0	0	a type of cookie
lese majeste	French	1	1	1	a phrase for 'injury to majesty'
liaison	French	985	497	741	a word for 'union'
limacon	French	1	1	1	a word for 'snail'
lingerie	French	72	49	60.5	a word for 'personal underwear'
liqueur	French	88	56	72	a type of alcoholic drink
litterateur	French	0	0	0	a word for 'writer, man of letters'
lorgnette	French	3	3	3	a word for 'binoculars'
luxe	French	68	47	57.5	a word for 'luxury'
lycee	French	8	5	6.5	a type of secondaty school
lyonnaise	French	5	5	5	a word descriptive of Lyon
mademoiselle	French	205	43	124	a word for 'young girl'
madrilene	French	0	0	0	a word descriptive of Madrid
maitre d'hotel	French	4	4	4	a phrase for 'manager of the hotel'
mal de mer	French	2	2	2	a phrase for 'sea sickness'
manege	French	1	1	1	a word for 'riding arena'
mangetout	French	10	3	6.5	a word for 'waster'
manque	French	4	4	4	a word for 'failed, illborn'
maquillage	French	2	2	2	a word for 'make-up'
maquis	French	28	15	21.5	a word for 'thickets'
mariniere	French	0	0	0	a word for 'naval'
marron	French	29	9	19	a word for 'bronze'
massif	French	118	35	76.5	a word for 'section of a planet's crust that is demarcated by faults or flexures'
materiel	French	10	7	8.5	a word for 'material'
melange	French	16	16	16	a word for 'combination'
menage	French	17	13	15	a word for 'marriage'
menage a trois	French	3	3	3	a phrase for 'marriage triangle'
merlot	French	11	8	9.5	a type of wine

mesalliance	French	0	0	0	a word for 'misalliance'
métier	French	9	9	9	a word for 'profession'
meuniere	French	0	0	0	a word for 'miller'
midinette	French	1	1	1	a word for 'shop assistant'
mignon	French	13	9	11	a word for 'pretty'
milieu	French	159	111	135	a word for 'middle or environment'
millefeuille	French	2	2	2	a type of French cookie
mirepoix	French	0	0	0	a word for 'pitch checking device'
mise-en-scene	French	4	2	3	a phrase for 'directed by'
mistral	French	37	15	26	a word for 'strong, dry and cold wind'
mitrailleuse	French	0	0	0	a type of machine gun
moi	French	233	85	159	a word for 'me'
mon veneris	French	1	1	1	a phrase for 'my love'
montage	French	55	34	44.5	a technique in film editing
morne	French	11	7	9	a word for 'dejected, gloomy'
mot	French	130	93	111.5	a word for 'A word, say'
mot juste	French	8	8	8	a phrase for 'proper A word'
moué	French	9	9	9	a word for 'sour face'
mouillé	French	1	1	1	a word for 'depth, moisturing of a load'
musette	French	11	3	7	a word for 'bagpipe, sanitary bag'
musique concrete	French	0	0	0	a phrase for 'specific music'
naivete	French	24	18	21	a word for 'simplicity, lack of experience, naivety'
napolitaine	French	0	0	0	a word descriptive of Naples
navarin	French	2	2	2	a word for 'mutton dish'
negligee	French	11	9	10	a word for 'neglected'
neve	French	2	1	1.5	a type of snow
nicoise	French	1	1	1	a word descriptive of Nice
noblesse oblige	French	13	13	13	a phrase for 'knithood obliges'
noisette	French	3	3	3	a word for 'hazel nut'
nom de guerre	French	4	4	4	a phrase for 'assumed name'
nom de plume	French	12	12	12	a phrase for 'literary pseudonym'
nouveau	French	194	93	143.5	a word for 'novel, new'
nouveau riche	French	17	15	16	a word for 'newly rich'
nouvelle cuisine	French	27	21	24	a word for 'a new recipe'
nouvelle vague	French	10	7	8.5	a phrase for 'new wave'

nuance	French	93	69	81	a word for 'subtlety, small difference'
nul point	French	0	0	0	a phrase for 'no matter'
objet d'art	French	8	7	7.5	a word for 'masterpiece'
objet trouve	French	0	0	0	a phrase for 'thing found'
oeil-de-boeuf	French	0	0	0	a word for 'oval or round window'
oeuvre	French	73	47	60	a word for 'work'
opera bouffe	French	0	0	0	a phrase for 'opera comique'
ordinaire	French	3	2	2.5	a word for 'ordinary'
outre	French	3	3	3	a word for 'exaggerated'
paillette	French	0	0	0	a word for 'metal tip'
pain au chocolat	French	0	0	0	a type of croissant
papier-mache	French	3	3	3	a phrase for 'paper mass'
papillon	French	13	5	9	a word for 'butterfly'
papillote	French	3	2	2.5	a word for 'curl-paper'
par excellence	French	94	85	89.5	a phrase for 'in the right meaning of the A word'
parfum	French	11	6	8.5	a word for 'fragrance'
parterre	French	11	9	10	a formal garden construction on a level surface consisting of planting beds
parti pris	French	3	3	3	a phrase for 'occupation taken'
parvenu	French	11	10	10.5	a word for 'newly rich'
pas	French	229	112	170.5	a word for 'step, walk'
pas de deux	French	46	12	29	a phrase for 'double step'
passant	French	13	12	12.5	a word for 'passer-by'
passe	French	16	16	16	a word for 'past'
passe-partout	French	0	0	0	a phrase for 'ornamental frame'
pastis	French	23	11	17	a type of alcohol
pate	French	89	58	73.5	a word for 'dough, pastry'
pate de fois gras	French	9	6	7.5	a word for 'liver plate'
patois	French	135	21	78	a type of dialect
pavé	French	1	1	1	a word for 'cobbled surface'
peignoir	French	6	3	4.5	a long outer garment for women
penchant	French	161	129	145	a word for 'inclination'
penillion	French	0	0	0	a type of Gallic song
pension	French	0	0	0	a word for 'fee, pension, guesthouse'
pere	French	21	15	18	a word for 'father'

perruquier	French	0	0	0	a word for 'hairdresser' (pej.)
persienne	French	0	0	0	a word for 'blind, shutter'
persiflage	French	4	3	3.5	a word for 'sneer'
petanque	French	6	5	5.5	a type of ball game
petillant	French	3	3	3	a word for 'energetic, sparking'
petit bourgeois	French	20	17	18.5	a word descriptive of a member of social class
petit four	French	3	1	2	a type of biscuit
petit mal	French	0	0	0	a phrase for 'small trouble'
petit point	French	5	4	4.5	a phrase for 'petty issue'
petit pois	French	2	2	2	a word for 'green peas'
petite	French	159	102	130.5	a word for 'little, young, faible'
picot	French	1	1	1	a word for 'chip'
piece de resistance	French	13	12	12.5	a phrase for 'piece of resistance', i.e. 'the best part of something'
pied-a-terre	French	1	1	1	a phrase for 'temporary accommodation'
pierrot	French	1	1	1	a word for 'clown, fool'
pince-nez	French	30	18	24	a style of spectacles, popular in the nineteenth century
pique	French	60	54	57	a word for 'pike'
pisaller	French	0	0	0	a word for 'half measure, the last resort'
pissoir	French	2	2	2	a word for 'urinal'
piton	French	15	8	11.5	a type of naked mountain peak
plage	French	26	15	20.5	a word for 'beach'
planchette	French	8	6	7	a word for 'slat'
plat du jour	French	8	6	7	a phrase for 'dish of the day'
plein-air	French	6	6	6	a phrase for 'the open air'
plie	French	0	0	0	a phrase for 'bent.' A smooth and continuous bending of the knees
plus ca change	French	4	4	4	a phrase for 'it changes more'
pointilliste	French	4	4	4	a word for 'punchliner'
pompadour	French	4	3	3.5	a word for 'cherry colour'
pompon	French	2	2	2	a type of a decorative ball of fluff
portiere	French	3	1	2	a word for 'curtain, drape'
portugaise	French	1	1	1	a word descriptive of Portugal
poste restante	French	5	4	4.5	a phrase for 'to be collected at post-office'
pot pourri	French	30	12	21	a phrase for 'burned pot'

potage	French	7	6	6.5	a word for 'vegetable soup'
pot-au-feu	French	8	2	5	a type of broth with meat
potiche	French	0	0	0	a word for 'vase'
pourboire	French	1	1	1	a word for 'tip'
poussin	French	4	3	3.5	a word for 'fuck' (vulg.)
pret-a-porter	French	1	1	1	a phrase for 'ready to wear'
prie-dieu	French	1	1	1	a type of prayer desk primarily intended for private use
prix fixe	French	0	0	0	a phrase for 'fixed price'
protégé	French	43	38	40.5	a word for 'protected'
quarte	French	2	2	2	a word for 'quarter'
quel	French	29	17	23	a word for 'which, who'
quinze	French	3	3	3	a word for 'fifteen'
raconteur	French	37	35	36	a word for 'storyteller'
raison d'être	French	48	43	45.5	a phrase for 'reason for being'
rapport	French	296	208	252	a word for 'good relationship'
rapporteur	French	29	14	21.5	a word for 'speaker'
rapprochement	French	115	75	95	a word for 'closeness'
ratatouille	French	27	16	21.5	a type of regional dish
recherche	French	8	7	7.5	a word for 'wanted'
regime	French	3445	795	2120	a word for 'ruling power'
renaissance	French	1	1	1	a period of cultural movement
rendezvous	French	263	140	201.5	a word for 'meeting'
rentier	French	13	11	12	a word for 'gentleman of leisure'
repechage	French	3	2	2.5	a word for 'elicitation'
repetiteur	French	1	1	1	a word for 'private tutor'
repondez s'il vous plait	French	0	0	0	a phrase for 'answer please'
repousse	French	1	1	1	a word for 'embossment'
restaurateur	French	42	34	38	a word for 'restaurant manager'
retrousse	French	1	1	1	a word for 'hitched up, tucked up'
risque	French	14	13	13.5	a word for 'risky'
rite de passage	French	11	7	9	a phrase for 'temporary habit'
rocaille	French	0	0	0	a word for 'shells, pebbles'
roman	French	57	1	29	a word for 'novel as a literary genre'
roman a clef	French	0	0	0	a type of novel
roman fleuve	French	2	2	2	a type of novel
rouge	French	395	131	263	a word for 'red'

rouge et noir	French	6	1	3.5	a word for 'red and black'
s'il vous plait	French	10	8	9	a word for 'please'
sabot	French	0	0	0	a word for 'clog'
saint	French	18	3	10.5	a type of person worshipped in church
sang froid	French	5	5	5	a phrase for 'cold blood'
sans	French	95	65	80	a word for 'without'
sans-culotte	French	1	1	1	a word descriptive of the members of the social class
sauve qui peut	French	1	1	1	a phrase for 'run away who can'
savant	French	16	12	14	a word for 'learned'
savoir-faire	French	6	6	6	a word for 'agility'
savoir-vivre	French	0	0	0	a word for 'knowledge of social forms'
seance	French	37	22	29.5	a word for 'an attempt to communicate with spirits'
sedan	French	35	20	27.5	a type of delicate fabric
seigneur	French	27	22	24.5	a word for 'God or Lord'
se-tenant	French	0	0	0	a phrase for 'behaving himself/ herself'
soi-disant	French	4	4	4	a word for 'so-called'
soignee	French	5	5	5	a word for 'well-groomed'
soiree	French	26	10	18	a word for 'evening'
soixante-neuf	French	5	4	4.5	a word for 'sixty-nine'
sommelier	French	12	11	11.5	a word for 'cellar'
son et lumiere	French	4	4	4	a phrase for 'sound and light'
soupcon	French	5	5	5	a word for 'suspicion'
succes	French	1	1	1	a word for 'success'
succes d'estime	French	0	0	0	a phrase for 'indisputable success'
succes de scandale	French	0	0	0	a phrase for 'savour of scandal'
succes fou	French	0	0	0	a phrase for 'rocketing success'
table d'hote	French	15	8	11.5	a word for 'table you share with sb in a restaurant'
tachisme	French	2	2	2	a word for 'French style of abstract painting'
tapenade	French	0	0	0	a type of Provence dish
telepherique	French	3	1	2	a word for 'cable line'
tete-a-tete	French	3	2	2.5	a phrase for 'head-to-head,' private meeting
tête-beche	French	0	0	0	a word for 'bonehead'

tic doloureux	French	0	0	0	a phrase for 'painful cramp'
timbale	French	7	2	4.5	a word for 'pot'
timbre	French	58	40	49	a word for 'stamp'
toilette	French	20	17	18.5	a word for 'toilet'
tole	French	1	1	1	a word for 'prison'
ton	French	1	1	1	a word for 'sound, tone'
toupet	French	22	14	18	a word for 'audacity, cheek'
tour de force	French	56	49	52.5	a word for 'feat'
tout court	French	11	11	11	a phrase for 'briefly and concisely'
tout ensemble	French	1	1	1	a word for 'altogether'
trompe l'oeil	French	18	15	16.5	a word for 'optical illusion'
tulle	French	24	18	21	a type of material
vacherin	French	1	1	1	a word for 'bad-tempered'
vichyssoise	French	3	3	3	a word descriptive of the inhabitant of Vichy
vignette	French	30	25	27.5	a word for 'graphic design or a road tax'
vin	French	93	52	72.5	a word for 'wine'
vin blanc	French	1	1	1	a phrase for 'white wine'
vin du pays	French	0	0	0	a phrase for 'regional wine'
vin ordinaire	French	0	0	0	a phrase for 'ordinary wine'
vin rouge	French	3	3	3	a phrase for 'red wine'
vingt-et-un	French	2	2	2	a word for 'twenty one'
virement	French	12	7	9.5	a word for 'turn'
voila	French	16	13	14.5	a word for 'here'
voile	French	23	15	19	a word for 'sail'
voix	French	22	13	17.5	a word for 'voice'
vol-au-vent	French	7	5	6	a phrase for 'wind-in-the-sail'
volte-face	French	30	29	29.5	a phrase for 'sudden change'
voyeur	French	42	35	38.5	a word for 'peeping Tom'
wagon-lit	French	1	1	1	a word for 'sleeping car'
ablaut	German	2	1	1.5	a word descriptive of vowel change
ach	German	7	7	7	a type of phoneme
ach-laut	German	0	0	0	a type of phoneme 'ach'
achtung	German	15	9	12	a word for 'attention'
angst	German	111	66	88.5	a word for 'fear'
anschluss	German	13	11	12	a word for 'access'
apfelstrudel	German	1	1	1	a phrase for 'apple pie'

auf Wiedersehen	German	7	7	7	a phrase for 'see you'
auslese	German	8	3	5.5	a word for 'late harvest wine'
autobahn	German	61	24	42.5	a word for 'motorway'
berg	German	191	63	127	a word for 'mountain'
bergschrund	German	11	6	8.5	a word for 'a deep crevasse'
bierkeller	German	1	1	1	a word for 'beer cellar'
bildungsroman	German	5	5	5	a word for 'novel'
blitzkrieg	German	26	20	23	a word for 'quick war'
bratwurst	German	2	2	2	a word for 'baked sausage'
Brocken spectre	German	1	1	1	a phrase for 'physical phenomenon observed at a high altitude in the mountains'
Bunsen burner	German	15	14	14.5	a word for 'gas burner, used mainly in chemistry laboratories'
dachshund	German	17	9	13	a type of dog
ding an sich	German	1	1	1	a phrase for 'a matter in itself'
doppelganger	German	14	10	12	a word for 'lookalike'
echt	German	6	5	5.5	a word for 'true, real'
edelweiss	German	8	5	6.5	a word for 'white snow'
ersatz	German	30	28	29	a word for 'compensation'
festschrift	German	6	6	6	a word for 'guest book'
foehn	German	0	0	0	a word for 'wind'
fohn	German	5	3	4	a word for 'warm wind'
fraulein	German	23	8	15.5	a word for 'young girl'
fuhrer	German	39	14	26.5	a word for 'guide'
gauleiter	German	12	4	8	type of horse
gemeinschaft	German	8	6	7	a word for 'community'
gesellschaft	German	24	13	18.5	a word for 'society'
gestalt	German	47	24	35.5	a phrase for 'figure,' whole image
gestapo	German	95	52	73.5	a word for 'political police'
gesundheit	German	1	1	1	a word for 'health'
glockenspiel	German	20	9	14.5	a word for 'play of bells'
gluhwein	German	5	3	4	a word for 'mulled wine'
gneiss	German	26	4	15	a word for 'common and widely distributed type of rock'
groschen	German	0	0	0	a phrase for 'pennies'
gymnasium	German	148	99	123.5	a type of school
hausfrau	German	4	4	4	a word for 'housewife'

herrenvolk	German	3	3	3	a word descriptive of a political slogan meaning 'Nation of Lords'
hertz	German	111	34	72.5	a unit of measurement
homburg	German	19	18	18.5	a type of hat
ich dien	German	0	0	0	a phrase for 'I serve'
ich-laut	German	0	0	0	a type of phoneme
junker	German	40	12	26	a word for 'young master'
kaiser	German	127	58	92.5	a word for 'emperor'
kirsch	German	30	18	24	a word for 'cherry'
knackwurst	German	2	2	2	a word for 'sausage'
kummel	German	6	3	4.5	a word for 'caraway'
lebensraum	German	14	11	12.5	a word for 'living space'
lederhosen	German	9	8	8.5	a word for 'leather trousers'
leitmotif	German	29	17	23	a phrase for 'main motif'
lied	German	11	4	7.5	a word for 'song'
loess	German	29	8	18.5	a fine, silty, windblown (eolian) type of unconsolidated deposit
mach	German	2	2	2	a phrase for 'do it'
oflag	German	2	2	2	an abbreviation of the name of a war camp
ohm	German	52	16	34	a unit of electrical resistance
panzer	German	19	16	17.5	a word for 'tank'
pfennig	German	24	15	19.5	a unit of currency
pumpernickel	German	7	5	6	a type of bread
quark	German	61	17	39	a type of cottage cheese
realpolitik	German	31	26	28.5	a word for 'real politics'
roentgen	German	9	5	7	a type of x-ray
Rorschach test	German	18	8	13	a method of psychological evaluation
sauerbraten	German	2	2	2	a word for 'roast'
sauerkraut	German	10	10	10	a type of sour cabbage
schadenfreude	German	13	13	13	a word descriptive of joy derived from sb's disaster
schuss	German	1	1	1	a word for 'shot'
Sieg Heil	German	13	9	11	a word for 'victory'
skat	German	15	2	8.5	a word for 'trick-taking card game'
spitz	German	13	7	10	a word for 'spire, tip'
sprachgefuhl	German	0	0	0	a word for 'language feel'
sprechgesang	German	2	2	2	a word for 'recitation'

stollen	German	11	3	7	a word for 'drift'
stoss	German	3	2	2.5	a word for 'strike'
strudel	German	12	11	11.5	a word for 'apple pie'
sturm und drang	German	5	5	5	a type of warfare activity
torte	German	8	7	7.5	a type of cake
ubermensch	German	0	0	0	a word for 'superman'
verboten	German	7	5	6	a word for 'forbidden'
wanderlust	German	18	14	16	a feel for wandering
Wernicke's area	German	0	0	0	a phrase descriptive of the location of language skill on a human brain
wiener	German	51	31	41	a word descriptive of Vienna
wunderkind	German	10	9	9.5	a word for 'prodigy infant'
wurst	German	2	2	2	a word for 'baked sausage'
zeitgeist	German	29	20	24.5	a word for 'spirit of time'
zeppelin	German	130	63	96.5	a type of airborne vehicle
Zollner's lines	German	3	2	2.5	a phrase for 'optical illusion'
zugzwang	German	0	0	0	a word for the situation in which one player is put at a disadvantage because he or she has to make a move
bouzouki	Greek	6	5	5.5	a type of Greek musical instrument
feta	Greek	18	9	13.5	a type of cheese
filo	Greek	56	16	36	a type of Greek flaky pastry
gyro	Greek	43	5	24	a type of pita sandwich
moussaka	Greek	14	10	12	a type of traditional aubergine (eggplant)-based dish
ouzo	Greek	16	7	11.5	an type of anise-flavored liqueur
retsina	Greek	20	13	16.5	a type of resinated white (or rosé) wine
souvlaki	Greek	1	1	1	a word for 'popular type of fast food'
taramasalata	Greek	26	15	20.5	a word for 'Greek meze'
taverna	Greek	39	14	26.5	a word for 'small restaurant'
tzatziki	Greek	3	3	3	a word for 'meze, or appetizer, also used as a sauce or dip'
tallith	Hebrew	0	0	0	a type of shawl
babu	Hindi	6	3	4.5	a word for 'native Indian clerk'
basmati	Hindi	29	9	19	a variety of long grain rice
bhangra	Hindi	3	3	3	a type of lively dance
bhindi	Hindi	2	2	2	a type of dish

chapatti	Hindi	3	3	3	a type of bread
chela	Hindi	24	6	15	a word for 'disciple'
dak	Hindi	2	2	2	an type of East Indian tree
deva	Hindi	2	1	1.5	a word for 'deity or supernatural being'
dhal	Hindi	24	19	21.5	a word for 'pulses which have been stripped of their outer hulls'
dhansak	Hindi	1	1	1	a word for 'mixture of spices'
dharma	Hindi	25	13	19	a word for 'natural law or reality'
dhobi	Hindi	6	2	4	a word for 'washerman in India'
dhoti	Hindi	6	6	6	a type of original and timeless garment of men's wear in India
garam masala	Hindi	13	6	9.5	a blend of dry-roasted ground spices
guru	Hindi	206	140	173	a teacher in Hinduism, Buddhism, and Sikhism
hanuman	Hindi	6	4	5	one of the most important personalities in the epic
hatha	Hindi	4	3	3.5	a type of yoga
jai	Hindi	11	10	10.5	a word for a popular name
lakh	Hindi	1	1	1	a type of unit in the Indian numbering system
lassi	Hindi	2	2	2	a type of traditional South Asian beverage
maharaja	Hindi	11	6	8.5	a word for 'king'
maharanee	Hindi	0	0	0	a word for 'wife of maharaja'
mahatma	Hindi	23	19	21	a phrase for 'Great Soul'
nirvana	Hindi	204	53	128.5	a word for 'the culmination of the pursuit of liberation'
pandit	Hindi	9	4	6.5	a word for 'Hindu Brahmin'
puggree	Hindi	0	0	0	a type of cloth band or scarf wrapped around the crown of a hat or sun helmet
puja	Hindi	2	2	2	a type of religious ritual
purda	Hindi	0	0	0	the practice of requiring women to cover their bodies
raj	Hindi	94	55	74.5	a phrase for 'prince, royalty'
roti	Hindi	2	2	2	a type of bread
sandhi	Hindi	0	0	0	a word for 'a wide variety of phonological processes'
saree	Hindi	1	1	1	a word for 'traditional garment worn by many women'

sari	Hindi	55	31	43	a word for 'traditional garment worn by many women'
satyagraha	Hindi	11	3	7	a word for 'the philosophy of nonviolent resistance'
sitar	Hindi	20	15	17.5	a type of classical instrument
suttee	Hindi	5	5	5	a word for 'funeral custom'
svarabhakti	Hindi	0	0	0	a type of epenthesis
swami	Hindi	13	4	8.5	a word for 'primarily Hindu honorific, loosely akin to master'
czardasz	Hungarian	0	0	0	a type of dance
forint	Hungarian	26	12	19	a unit of currency
vizsla	Hungarian	0	0	0	a breed of dog
cead mile failte	Irish	0	0	0	a phrase for 'a hundred thousand welcomes'
garda	Irish	130	32	81	a word for 'the national police of the Republic of Ireland'
accelerando	Italian	2	2	2	a word for 'accelerating'
aggiornamento	Italian	6	1	3.5	a word for 'updating'
aioli	Italian	2	2	2	a word for a 'bird net'
al dente	Italian	15	12	13.5	a word for 'not overcooked'
alla	Italian	43	21	32	a type of grammatical process
andante	Italian	44	15	29.5	a word for 'ordinary, current'
andantino	Italian	5	2	3.5	a word for a technical term in music
arietta	Italian	0	0	0	a word for 'wind, melody'
arrivederci	Italian	13	5	9	a word for 'see you'
autostrada	Italian	11	8	9.5	a word for 'motorway'
bambino	Italian	12	7	9.5	a word for 'child'
basso pro fundo	Italian	0	0	0	a phrase for 'deep, low'
bel	Italian	158	62	110	a word for 'pretty, considerable'
bel canto	Italian	7	6	6.5	a word for 'beautiful singing'
ben trovato	Italian	0	0	0	a word for 'found or invented'
cacciatore	Italian	1	1	1	a word for 'hunter'
calamari	Italian	2	2	2	a word for 'squids'
carabinieri	Italian	37	12	24.5	a word for 'Italian policemen'
cassata	Italian	7	2	4.5	a word for 'icecream or cake'
chianti	Italian	62	19	40.5	a type of wine
chiaroscuro	Italian	22	20	21	a word for the term in the art meaning 'a contrast between light and dark'

ciabatta	Italian	7	6	6.5	a type of white bread made with wheat flour and yeast
cicisbeo	Italian	0	0	0	a word for 'ladies' man'
cinquecento	Italian	8	6	7	a word for 'five hundred'
commedia dell'arte	Italian	0	0	0	a type of comedy
condottiere	Italian	4	4	4	a word for 'leader'
contessa	Italian	22	14	18	a word for 'princess'
da capo	Italian	7	4	5.5	a phrase for 'from beginning'
dal	Italian	227	63	145	a type of grammatical process
dal segno	Italian	0	0	0	a phrase for 'from the sign (gesture, signal)'
dolce farniente	Italian	0	0	0	a phrase for 'sweet loafing about'
dolce vita	Italian	15	10	12.5	a phrase for 'life in the lap of luxury and inertia'
doloroso	Italian	1	1	1	a word for 'painful'
donna	Italian	1008	142	575	a word for 'woman'
duce	Italian	21	13	17	a word for 'leader'
espresso	Italian	27	23	25	a type of coffee, a letter or a train
fantoccini	Italian	0	0	0	a word for 'puppets'
farfalle	Italian	0	0	0	a word for 'butterflies' or a type of pasta
fettuccine	Italian	6	3	4.5	a word for 'tagliatelle'
finito	Italian	2	2	2	a word for 'perfect, effected'
focaccia	Italian	4	2	3	a word for 'pancake'
fusilli	Italian	4	3	3.5	a type of helical shaped pasta, is usually about four centimetres long
gnocchi	Italian	10	6	8	a word for 'dumplings'
gran turismo	Italian	2	1	1.5	a phrase for 'a well developed tourism'
grappa	Italian	13	10	11.5	a type of vodka
grave	Italian	1	1	1	a word for 'grave, serious'
intaglio	Italian	15	11	13	a word for 'jewel'
lasagna	Italian	13	9	11	a word for 'square noodle'
linguini	Italian	3	3	3	a type of pasta
lira	Italian	108	53	80.5	a currency unit or a musical instrument
lollo rosso	Italian	4	3	3.5	a type of man
macchiato	Italian	0	0	0	a word for 'spoiled' or a type of coffee
maestoso	Italian	5	1	3	a word for 'majestic'

marinara	Italian	4	3	3.5	a naval clothing for kids
mascarpone	Italian	7	2	4.5	a type of cream, cheese
molto	Italian	12	9	10.5	a word for 'a lot, a multitude'
morbidezza	Italian	1	1	1	a word for 'softness, delicacy'
non troppo	Italian	2	2	2	a word for 'insufficiently'
opera buffa	Italian	8	3	5.5	a type of opera
opera seria	Italian	7	4	5.5	a word for 'serious masterpiece'
osso bucco	Italian	4	1	2.5	a word for a 'tall lanky man'
palladio	Italian	26	14	20	a word for 'protection'
panettone	Italian	6	3	4.5	a type of cake
papabile	Italian	2	2	2	a word descriptive of a cardinal who may become Pope
paparazzo	Italian	3	3	3	a type of sensationalist journalist
parlando	Italian	0	0	0	a word for 'speaking'
penne	Italian	12	7	9.5	a type of pasta
pentimento	Italian	1	1	1	a word for 'remorse'
pesto	Italian	42	17	29.5	a word for 'pulp'
piazza	Italian	297	63	180	a word for 'square'
pieta	Italian	5	5	5	a word for 'mercy, adoration'
poco	Italian	5	3	4	a word for 'a little'
presa	Italian	0	0	0	a word for 'clamp'
prosciutto	Italian	19	11	15	a word for 'ham'
provolone	Italian	2	1	1.5	a type of cheese
quattrocento	Italian	17	10	13.5	a word for 'four hundred'
radicchio	Italian	7	6	6.5	a word for 'chicory'
rallentando	Italian	3	1	2	a word for a technical musical term
ricotta	Italian	27	14	20.5	a type of cheese
rilievo	Italian	0	0	0	a word for 'importance, observation'
ripieno	Italian	1	1	1	a word for 'filling'
risotto	Italian	50	20	35	a word for 'rice dish'
saltimbocca	Italian	2	2	2	a word for 'veal escalopes'
scagliola	Italian	7	2	4.5	a word for 'plaster'
scena	Italian	5	5	5	a word for 'stage, theatre'
scherzando	Italian	4	3	3.5	a word for 'joking'
scherzo	Italian	61	11	36	a word for 'joke, trifle'
scirocco	Italian	2	2	2	a type of wind
secco	Italian	4	3	3.5	a word for 'dry'

segno	Italian	0	0	0	a word for 'sign, gesture, trace'
segue	Italian	2	2	2	a word for 'spy, follow'
semplice	Italian	7	3	5	a word for 'simple, naive'
sempre	Italian	2	2	2	a word for 'always, constantly'
senza	Italian	3	3	3	a word for 'without'
sgrafitto	Italian	0	0	0	a word for 'ornamental technique in painting'
simpatico	Italian	6	5	5.5	a word for 'nice'
sotto voce	Italian	27	24	25.5	a phrase for 'in a whisper'
spiccato	Italian	0	0	0	a word for 'detached'
spumante	Italian	13	9	11	a type of champagne
stretto	Italian	2	2	2	a word for 'tight-fitting, narrow'
terza rima	Italian	2	2	2	a type of poem
tessitura	Italian	8	5	6.5	a word for 'weaving'
toccata	Italian	14	4	9	a word for 'touch'
trattoria	Italian	31	19	25	a word for 'restaurant, inn'
troppo	Italian	3	3	3	a word for 'too much'
tutti	Italian	35	18	26.5	a word for 'everybody'
vaporetto	Italian	5	5	5	a word for 'steam ship'
verismo	Italian	0	0	0	a word for 'realism'
volta	Italian	27	18	22.5	a word for 'a turn, one time'
aikido	Japanese	15	6	10.5	a way of 'harmonizing energies' or a modern Japanese martial art
futon	Japanese	17	11	14	a type of mattress that makes up a Japanese bed
hiragana	Japanese	0	0	0	a type of Japanese syllabary, one of the four Japanese writing systems
ikebana	Japanese	7	3	5	a type of Japanese art of flower arrangement
kabuki	Japanese	9	6	7.5	a form of traditional Japanese theater
kakemono	Japanese	1	1	1	a word for 'scroll painting'
kana	Japanese	8	5	6.5	a word for the syllabic Japanese scripts
kanji	Japanese	20	13	16.5	the Chinese characters that are used in the modern Japanese logographic writing system
karate	Japanese	307	76	191.5	a type of martial art of Okinawan origin
katakana	Japanese	4	3	3.5	a type of Japanese syllabary, one of the four Japanese writing systems

kendo	Japanese	20	8	14	a type of martial art of Japanese fencing
kimono	Japanese	47	29	38	a type of traditional garment of Japan
koan	Japanese	0	0	0	a story, dialog, question, or statement generally containing aspects that are inaccessible to rational understanding, yet that may be accessible to intuition
netsuke	Japanese	4	3	3.5	a word for 'the miniature sculpture'
origami	Japanese	23	15	19	a word for 'the art of paper folding'
shiatsu	Japanese	14	10	12	a word for 'hands-on therapy technique originating in Japan'
shogun	Japanese	68	16	42	a type of military rank and historical title in Japan; the rank is equivalent to 'general,' a high officer in an army
sukiyaki	Japanese	0	0	0	a type of dish
sumo	Japanese	44	29	36.5	a type of competition contact sport
sushi	Japanese	23	18	20.5	a type of food
tanka	Japanese	3	3	3	a genre of poetry (see Waka)
tatami	Japanese	4	2	3	a type of traditional flooring
tempura	Japanese	2	2	2	a classic Japanese deep fried batter-dipped seafood and vegetables
teriyaki	Japanese	13	7	10	a type of cooking sauce such for fish or meat
tofu	Japanese	23	14	18.5	a type of food of Chinese origin
torii	Japanese	0	0	0	a word for 'traditional Japanese gate'
tsunami	Japanese	88	38	63	a word for 'series of water waves caused by a large volume of a body of water'
tsutsugamushi	Japanese	0	0	0	a type of mite-borne typhus, and tropical typhus
hangul	Korean	2	2	2	a word for 'the native alphabet of the Korean language'
taekwondo	Korean	73	4	38.5	a word for 'the most popular of the Korean martial arts'
sambal	Malay	1	1	1	a type of condiment used in Indonesia, Malaysia, Singapore and Sri Lanka, made from a variety of peppers
ngaio	Maori	17	12	14.5	a variety of tree, the Ngaio, also known as the Mousehole tree
grosz	Polish	0	0	0	a currency unit

kielbasa	Polish	0	0	0	a type of food
mazurka	Polish	8	6	7	a type of lively Polish dance in triple time
zloty	Polish	47	21	34	a currency unit
auto-da-fe	Portuguese	4	3	3.5	a type of public trial ceremonial
cruzeiro	Portuguese	10	9	9.5	a type of monetary unit of Brazil
fado	Portuguese	7	2	4.5	a type of Portuguese folk music
jacana	Portuguese	12	4	8	a type of bird
vinho verde	Portuguese	5	5	5	a type of wine
kara	Punjabi	0	0	0	a type of bracelet
kirpan	Punjabi	0	0	0	an article of defence
kuccha	Punjabi	0	0	0	a word for 'kitchen'
babushka	Russian	35	6	20.5	a word for 'grandmother and/or old lady'
balaclava	Russian	69	49	59	a word for 'helmet or ski mask'
balalaika	Russian	6	4	5	a type of stringed instrument of Russian origin
chernozem	Russian	1	1	1	a kind of soil
dacha	Russian	45	22	33.5	a word for 'house in the country occupied part of the year by its owner'
glasnost	Russian	125	76	100.5	a word for 'publicity' or 'openness'
kolkhoz	Russian	6	2	4	a form of collective farming in the Soviet Union
Markov process	Russian	1	1	1	a phrase for 'scholarly axiom'
oblast	Russian	39	24	31.5	a subnational entity of Bulgaria, the Russian Federation, Ukraine, and the now-defunct Soviet Union
perestroika	Russian	198	92	145	a word for 'the economic reforms introduced in June 1987 by the Soviet leader Mikhail Gorbachev'
piroshki	Russian	0	0	0	a type of dish
Przewalski's horse	Russian	0	0	0	a type of horse also known as the Mongolian Wild Horse, or Takhi
samizdat	Russian	10	10	10	a word for 'clandestine copying and distribution of governmentsuppressed literature'
samovar	Russian	10	6	8	a type of heated metal container traditionally used to brew tea in and around Russia, as well as in other Slavic nations and Turkey
taiga	Russian	15	4	9.5	a biome characterized by coniferous forests

om	Sanskrit	6	2	4	a type of sacred syllable intoned in mantras and prayers, etc. that symbolizes the Vedic scriptures
pibroch	Scottish-gaelic	3	2	2.5	an ancient type of music, native to the Scottish Highlands
slainte	Scottish-gaelic	2	2	2	a word for 'cheers'
adios	Spanish	12	7	9.5	a word for 'goodbye'
aficionado	Spanish	23	22	22.5	a word for 'devotee of bullfighting'
aguardiente	Spanish	0	0	0	a type of coarse Spanish brandy
alameda	Spanish	30	18	24	a word for 'public walk shaded with trees'
amontillado	Spanish	6	2	4	a type of medium sherry of a matured type
angostura	Spanish	10	8	9	a word for 'narrowing or crossing on a river'
arroyo	Spanish	5	4	4.5	a word for 'steep gully'
barrio	Spanish	21	12	16.5	a word for 'quarter of a town'
bodega	Spanish	10	7	8.5	a word for 'storehouse'
bola	Spanish	19	10	14.5	a word for 'missile'
burrito	Spanish	7	3	5	a word for 'tortilla'
burro	Spanish	7	5	6	a word for 'small donkey'
caballero	Spanish	30	11	20.5	a word for 'Spanish gentleman'
cabana	Spanish	6	4	5	a word for 'shelter at a beach or swimming pool'
cajon	Spanish	0	0	0	a kind of box drum
camarilla	Spanish	1	1	1	a word for 'cabal, clique'
caudillo	Spanish	22	5	13.5	a word for 'military or political leader'
centavo	Spanish	1	1	1	a type of currency unit
chihuahua	Spanish	30	16	23	a very small dog of a breed originating in Mexico
chorizo	Spanish	7	4	5.5	a pork sausage
cojones	Spanish	3	3	3	a word for 'courage'
conquistador	Spanish	11	10	10.5	a word for 'conqueror'
conquistadores	Spanish	11	10	10.5	a word for 'conquerors'
copita	Spanish	0	0	0	a word for 'glass of sherry'
cordillera	Spanish	64	18	41	a word for 'belt of mountains and valleys'
corrida	Spanish	6	6	6	a word for 'bullfight'
costa	Spanish	588	230	409	a word for 'coast designed as a holiday resort'

fino	Spanish	5	5	5	a type of pale-coloured dry sherry
frijole	Spanish	0	0	0	a word for 'bean'
gazpacho	Spanish	8	7	7.5	a type of soup
guacamole	Spanish	12	7	9.5	a type of Mexican dish
hasta la vista	Spanish	5	4	4.5	a word for 'goodbye'
hasta manana	Spanish	0	0	0	a fixed phrase for 'see you tommorow'
infanta	Spanish	8	7	7.5	a word for 'heir to the thorne' (the eldest daughter)
infante	Spanish	23	6	14.5	a word for 'heir to the throne' (the eldest son)
jalapeno	Spanish	3	3	3	a word for 'chilli pepper'
jojoba	Spanish	9	3	6	a type of desert shrub whose oil is used in cosmetics
junta	Spanish	185	66	125.5	a type of administrative council or committee
machismo	Spanish	68	44	56	a word for 'masculine pride'
mantilla	Spanish	15	12	13.5	a type of light scarf
manzanilla	Spanish	8	7	7.5	a type of pale, very dry sherry
mariachi	Spanish	2	2	2	a type of itinerant Mexican folk band
paella	Spanish	28	16	22	a type of Spanish dish
pampa	Spanish	4	4	4	a word for 'treeles plains'
parador	Spanish	8	3	5.5	a word for 'hotel owned by the government'
pelota	Spanish	14	3	8.5	a type of ball game
peseta	Spanish	68	34	51	a type of currency unit
peso	Spanish	61	20	40.5	a type of currency unit
peyote	Spanish	8	6	7	a type of desert cactus used as hallucinogenic drug
pina colada	Spanish	2	2	2	a type of drink
piñata	Spanish	0	0	0	a word for 'decorated container filled with sweets or small gifts'
pinta	Spanish	20	14	17	a word for 'colored spot, painted mark'
plaza	Spanish	173	86	129.5	a word for 'market place'
poncho	Spanish	23	17	20	a word for 'cloak'
presidio	Spanish	3	2	2.5	a word for 'garrison, fort'
pronunciamento	Spanish	0	0	0	a political manifesto
pulque	Spanish	1	1	1	a type of drink
quesadilla	Spanish	0	0	0	a word for 'tortilla'

real	Spanish	14	14	14	a word for 'kingdom'
salsa	Spanish	15	11	13	a type of sauce or dance
sangria	Spanish	33	19	26	a type of drink
sierra	Spanish	595	217	406	a word for 'mountain range'
solera	Spanish	4	1	2.5	a blend of sherry
tapas	Spanish	26	16	21	a word for 'savoury snack'
tortilla	Spanish	33	17	25	a type of cake
angstrom	Swedish	2	2	2	a unit of length
fartlek	Swedish	1	1	1	a type of sports training
ore	Swedish	0	0	0	a word for 'the one-hundredth subdivision of the Swedish krona'
smorgasbord	Swedish	9	9	9	a word for 'any buffet with a variety of dishes'
dolma	Turkish	2	2	2	a word for 'family of stuffed vegetable dishes originating in Turkey'
ach-y-fi	Welsh	0	0	0	a word for ' expression of disgust'
bach	Welsh	1	1	1	a word for 'hook, clasp or productive morpheme'
cwm	Welsh	33	25	29	a word for 'valley'
Cymru am byth	Welsh	0	0	0	a fixed phrase for 'Wales forever'
cynghanedd	Welsh	2	2	2	a type of music or a word for 'harmony'
eisteddfod	Welsh	44	22	33	a type of cultural festival
iechyd da	Welsh	0	0	0	a phrase for 'good health'
impala	Zulu	15	14	14.5	a type of antelope
impi	Zulu	0	0	0	a word for 'any armed body of men'
indaba	Zulu	1	1	1	a word for 'important conference or discussion between members of different tribes'

INDEX OF NAMES

N

O

P

Q

R

S

Mail orders for the publications of
Wydawnictwo Uniwersytetu Wrocławskiego
should be directed to

Dział Handlowy
Wydawnictwa Uniwersytetu Wrocławskiego Sp. z o.o.
50-137 Wrocław, pl. Uniwersytecki 15
tel./fax +48 (0)71 3752507, tel. +48 (0)71 3752885
e-mail: marketing@wuwr.com.pl
www.wuwr.com.pl

Wydawnictwo Uniwersytetu Wrocławskiego
invites to the bookshops:

- www.wuwr.com.pl
- Księgarnia Uniwersytecka
  50-138 Wrocław, ul. Kuźnicza 49/55, tel. +48 (0)71 3752923
- Księgarnia Naukowa im. Mikołaja Kopernika
  50-138 Wrocław, ul. Kuźnicza 30/33, tel. +48 (0)71 3432977

All books sold in the bookshop can be ordered by mail.

Foreign words and phrases in English, 2009