

Ida Stria

Inventing languages, inventing worlds
Towards a linguistic worldview for artificial languages



Inventing languages, inventing worlds
Towards a linguistic worldview for artificial languages

Ida Stria

Inventing languages, inventing worlds
Towards a linguistic worldview for artificial languages

Wydział Neofilologii UAM w Poznaniu
Poznań 2016

Projekt okładki:
Ida Stria

Recenzja:
prof. Federico Gobbo

Copyright by:
Ida Stria

Wydanie I, Poznań 2016

ISBN 978-83-947609-1-5

DOI: 10.14746/9788394760915

Wydanie:
Wydział Neofilologii UAM w Poznaniu
al. Niepodległości 4, 61-874 Poznań
e-mail: dziekneo@amu.edu.pl
www.wn.amu.edu.pl

Acknowledgements

My thanks go to Prof. Jerzy Bartmiński, Prof. Sabine Fiedler, Prof. Federico Gobbo, Prof. Ilona Koutny, Prof. Jerzy Pogonowski, dr Sławomir Sikora and dr Agnieszka Schönhof-Wilkans for all their invaluable comments.

All errors are my own

Q: How many Lojbanists does it take to change a broken light bulb?

A: Two: one to decide what to change it into and one to figure out what kind of bulb emits broken light.

Table of Contents

Acknowledgements	5
List of figures	9
List of tables	10
Introduction	11
1. Linguistic worldview	15
1.1 Short history	15
1.2 Definitions and crucial concepts	18
1.3 The language – culture – thought relation	26
1.4 Methodology	29
2. Artificial languages	40
2.1 Definitions	40
2.2 History and motives of creation	47
2.2.1 Universal languages	50
2.2.2 Formal languages	63
2.2.3 International auxiliary languages	66
2.2.4 Artistic languages	76
2.2.5 Modern glossopoeia	85
2.3 Typology and classifications	89
2.3.1 Traditional classifications: structure and source of material	90
2.3.2 Traditional classifications: purpose	93
2.3.3 Blanke’s functional classification	100
3. Natural vs. artificial	104
3.1 Properties of human language	104
3.2 The scale of naturalness	114
3.2.1 The continuum of deliberate influence	114
3.2.2 Lyons’ classification	121
3.3 Borderline cases	129
3.3.1 Pidgins and creoles	129
3.3.2 Language revitalisation and revival	133
3.3.3 Other cases	137
4. Esperanto as a transitional case	145
4.1 Naturalness of Esperanto	145
4.2 Sociolinguistic situation of Esperanto speakers	148
5. Linguistic worldview for artificial languages	160
5.1 The subject: author or community?	160
5.2 Possible objects of study	169
5.3 Linguistic worldview in non-native languages	180
5.4 Linguistic worldview for Esperanto	184
6. Research problems on the example of a pilot study	195
6.1 The respondents	197
6.2 The questions	204
6.3 Conclusions	225

Final remarks.....	229
The boundaries of artificiality	229
A linguistic worldview for artificial languages.....	230
References.....	233
Appendix: The questionnaire	249
Summary in Esperanto / Resumo	261

List of figures

Figure 1 The (simplified) relationship language – culture – perception	28
Figure 2 Relationships between artificial and international languages	46
Figure 3 Federico Gobbo’s coordinate system (2014).....	94
Figure 4 Classification of invented languages by Albani & Buonarotti (1994).....	96
Figure 5 The Gnoli triangle (around 1997).....	97
Figure 6 Coloured version of the Gnoli triangle modified by R. A. Brown	97
Figure 7 Jan van Steenbergen’s hexagon (2008).....	98
Figure 8 Baron’s natural language spectrum (1994).....	116
Figure 9 Baron’s computer language spectrum (1994).....	118
Figure 10 Scale of artificiality/naturalness	120
Figure 11 Age distribution among the participants	198
Figure 12 Native languages of the respondents	199
Figure 13 Usage of Esperanto among the respondents.....	200
Figure 14 Usage of four most popular languages according to activities	201
Figure 15 Usage of Esperanto vs. native languages according to activities	202
Figure 16 In which language would you read a book if the original language was not known to you?	202
Figure 17 Usage of Esperanto vs. native languages according to activities	203

List of tables

Table 1 Classificatory matrix (simplified with examples).....	99
Table 2 Hockett's design features applied to artificial languages.....	108
Table 3 Hockett's design features in formal languages	114
Table 4 Languages and their naturalness.....	124
Table 5 Languages and their naturalness II.....	127
Table 6 A comparison of pidgins and creoles with ethnic languages and IALs.....	130
Table 7 A comparison of revitalised and minority languages with ethnic languages and IALs	136
Table 8 A comparison of dormant classical languages with ethnic languages and IALs.....	140
Table 9 Types of artificial languages and borderline cases according to usability	169
Table 10 Approaches to artificial languages and some borderline cases ...	179
Table 11 Samples according to language	188
Table 12 The features of a "true" Esperantist arranged by aspect	222
Table 13 The features presupposed in the formula 'John is an ardent Esperantist but...'	223
Table 14 Symbols of an Esperantist	225

Introduction¹

Artificial languages are not a common subject of today's linguistics. They are deemed rigid abstract systems suitable for performing the role of interlanguages in translation, serving as a basis for developing programming languages or describing human thinking. Such an understanding of 'artificial languages' is, however, very limiting as it does not include universal schemes, international auxiliary languages or artistic languages. The latter types, to the minds of linguists, are mere curiosities not deserving of any serious study. Natural languages are assumed to constitute the main subject of linguistics. Yet the author of this book believes that such a binary division into 'natural' and 'artificial' not only can be disproved but also is detrimental to the development of linguistics as a science trying to uncover the mechanisms lying behind human linguistic categorisations. Artificial languages are of considerable interest – since they are designed freely by humans, it may mean they obey the same rules as other human languages (so-called “natural” languages).

It has to be remembered that artificial languages constitute a heterogeneous group and that each subgroup is characterised by distinct properties. These properties determine the scope and limitations of studying particular languages. Therefore, this work aims at revealing the internal diversity of artificial languages and placing them on a scale between the two aforementioned poles. Moreover, as artificiality is a gradable property, so is the possibility of applying diverse linguistic frameworks to various artificial systems.

Until now, little importance has been given to artificial languages. They have been the central topic of several historical linguistic works. A considerable amount of literature has also been published on artificial languages as the subject of

¹ Fragments of this book were previously published as chapters or papers in a slightly changed form as Stria 2013; Stria 2015a; Stria 2015b; and Stria 2015c.

interlinguistics, especially focusing on international auxiliary languages such as Esperanto.

This book aims to assess the possibility of studying artificial languages in the framework of cultural linguistics and, more precisely, studying the concept of the linguistic worldview in the framework of the theory and practice developed in the Ethnolinguistic School of Lublin. Current research on the linguistic worldview is focused on natural ethnic languages and recently comparative studies have been postulated. The present work is the first attempt to show possible applications of the theory to various artificial languages, as well as some limitations resulting from such an approach.

The concept of the linguistic worldview is not a new one, albeit a relatively unpopular one. It has a prominent place in Polish ethnolinguistics, where it is extensively discussed; however, it is little known outside Poland. It has been widely investigated by Polish scholars from various research centres (mainly in the *Etnolingwistyka* journal, the Wrocław series of *Język a Kultura* and the so-called “Red Series” – a series of proceedings published in Lublin, and recently in Gład et al. 2013). The initiator of the idea and the *de facto* founder of the Ethnolinguistic School of Lublin is Jerzy Bartmiński, whose work is of central importance (Bartmiński 2012a; Bartmiński 2012b).

The study of the linguistic worldview for artificial languages is a new idea. It has customarily been studied for ethnic languages. The only research article known to the author of this book which discussed the problem of the linguistic worldview for one of the international auxiliary languages, i.e. Esperanto, is the article by Koutny (2010). However, to the author’s best knowledge, the application of this concept to a wider range of artificial languages has been scarcely investigated from the theoretical point of view. The book also raises the problem of the linguistic worldview of multilinguals, which has not been explored previously.

The remainder of the book is organised into six chapters.

Chapter 1 discusses the concept of the linguistic worldview and its historical development. The chapter is devoted to the theory and methodology established in the Ethnolinguistic School of Lublin as well as the variety of proposed definitions and some practical consequences of adopting one of them.

In **Chapter 2**, artificial languages are presented from a historical and a typological perspective. Their definitions and classifications are examined.

The division of languages into natural and artificial is investigated in **Chapter 3**. Artificial languages are analysed according to various sets of properties. The class of borderline cases is described.

Chapter 4 focuses on Esperanto as a transitional case between an artificially created language and a fully developed natural language with its own speech community.

Chapter 5 is an attempt to apply theoretically the paradigm of the Ethnolinguistic School of Lublin to artificial languages and some borderline cases. It outlines possible limitations to such research based on crucial concepts of the framework applied to each type of languages. The chapter also presents a proposition for potential future study of the linguistic worldview in Esperanto.

Chapter 6 is a practical application of the earlier propositions to Esperanto. It presents a pilot study conducted in 2015 in the form of a questionnaire, and compares it to some of the results presented by Koutny (2010).

The book concludes with final remarks.

1. Linguistic worldview

1.1 Short history

The central concept of this book, namely the *linguistic worldview*² featured in the title, has a long history. According to Jerzy Bartmiński (2012a: 22), the initiator and propagator of the idea in Polish linguistics and the founder of the Ethnolinguistic School of Lublin (henceforth ESL), it can be traced as far back as to Aristotle's *topoi* (*loci communes*), that is, common, generally recognised judgements which are part of the argumentation. However, the first obvious reference to dissimilar characteristics of different languages can be found in Martin Luther's *Sendbrif vom Dollmetschen* (1530, and even more clearly worded in *Tischreden* 5, 5521: "ein ittliche sprag hatt ir eigen art"³). In the 16th c., scholars were interested not only in the difficulties of translating the Bible into national languages but also in the newly discovered languages of the Far East and the ways in which they mirrored the reality (see section 2.2.1). Studying languages was for many a way to arrive at truth or the ideal god-made order of things. This is visible in the thought of Leibniz. He claimed that language mirrors the internal structure of intellect, which is common to all men, all being created by the same god. Thus, there is a kind of "universal grammar" or common traits of the mind. However, aspects of reality are differently realised in particular languages because of diverse circumstances and varied experience, which have made man abandon the order of things instituted by god (Święczkowska 1998: 36f., 58f.).

² Otherwise known as the *linguistic picture of the world*. For the discussion of the term see Bartmiński (2012a: chap. 7) and Tabakowska (2013).

³ The quote is very often wrongly given as coming from *Sendbrif vom Dollmetschen*. The passage of *Tischreden* reads as follows: "Man kan nicht ublich in Hebreo die wort geben, wie sie stehhen, der sprach nach; denn ein ittliche sprag hatt ir eigen art, und ist ein wort so latum, das man im nicht kan genug thun".

It was only in the 18th century that the idea of the influence of cultural and sociological factors on language and thinking gained a strong foothold in philosophy. J. G. Hamann and J. G. Herder saw language as shaping thought and as a device reflecting the spirit of nations (Andrzejewski 1989: 163, 174f.; Anusiewicz 1999: 263). They opposed the Leibnizian idea of language merely reflecting thought. Hamann went so far as to declare that “reason is language (*logos*)” (Andrzejewski 1989: 149).

The idea of the linguistic worldview (*Weltansicht*) explicitly appeared for the first time in the works of W. von Humboldt in the first half of the 19th century (see Głaz et al. 2013: 11–24; Underhill 2009; Żuk 2010; cf. Allwood 1983). Humboldt believed that language is “a work of the subject and his *a priori* spiritual activity” (Andrzejewski 1989: 150). He stressed the importance of individuals, who, although being “cultivated with and within language, they simultaneously cultivated language by leaving their own personal impressions upon it” (Underhill 2009: 122). Language is a perpetual *energeia*, through its structure shaping thinking and “a nation’s spirit”, i.e. culture. However, the constraints language enforces on thinking can be overcome by creative use. Changes in culture and human activity also influence language and each specific culture is contained in the nation’s vernacular (Andrzejewski 1989: 154). What follows is that different cultures produce different languages and vice versa.

The concept presented by Humboldt, namely, that language allows users to form a worldview (*Weltansicht*) through its inherent specific structure has come to be mistaken with the concept of *Weltanschauung*, that is, socially constructed worldviews which may be different in the same language (e.g. socialist and Christian views of German-speaking people) but the same across languages (e.g. liberal views of English and German speakers). *Weltansichten* are language-dependent and, at the same time, culture-dependent (Underhill 2009: 55ff.)

These relativist Humboldtian views were developed in two diverging yet similar ways: by German researchers such as L. Weisgerber and by American anthropologists and linguists (Boas, Sapir, and Whorf).

The German Neo-Humboldtians tried to uncover the worldview by studying the division of languages into semantic categories. Weisgerber claimed that the worldview (again *Weltansicht* rather than *Weltanschauung*) is a mental construct lying between the world and the community (*Zwischenwelt*). It is not a reflection of the world but its dynamically changing conceptual interpretation (Andrzejewski 1989: 181f.). H. Gipper went as far as calling language the “key to the world”, as it is a tool that helps conceptualise perceived reality. For him, however, semantic analysis would not suffice; morphology and syntax were just as vital (Anusiewicz 1999). An important contribution of German researchers was the general term *Weltbild*, later on translated into Polish as *obraz świata*.

The American anthropologists were occupied with the grammatical distinctions made by languages. As an immigrant, F. Boas believed in differences between languages stemming from their diversely shaped environments. He also saw direct links between language and culture, not claiming, however, that one necessarily conditions the other. E. Sapir and B. Whorf, in turn, postulated that language shapes an individual’s way of thinking by imposing its categories onto perceived reality (Lucy 1992b; Underhill 2009). As Lucy (1992b: 258) writes: “[in Whorf’s approach] individual thought was inferred from the language analysis and empirically verified by reference to related cultural patterns of belief and behavior. This approach allowed Whorf to emphasize the general significance of language patterns for behavior [...]”. The comparative approach was for him the basis of worldview studies.

Both groups believed that cultural traces can be found in language and that people can only access reality through lan-

guage, which categorises things according to its internal structure (Gład et al. 2013: 11–24; Pajdzińska 2013).

In Poland, the linguistic worldview (henceforth LWV) has been explored in the Ethnolinguistic School of Lublin (the main representative being J. Bartmiński), Wrocław (e.g. J. Anusiewicz), and Warsaw (R. Grzegorzczkova, axiolinguists with J. Puzynina). The understanding of the worldview as traces of culture in language, and not language in culture, positions the ESL as stemming from the Polish ethnographic tradition and at the same time from (neo)humboldtism. The contemporary discussion owes much to cognitive linguistics as well. Bartmiński often quotes the accomplishments of A. Wierzbicka on the one hand and Russian linguists such as J. Apresyan and N. Tolstoy on the other (cf. Chlebda 2013). It is clearly visible that the crucial concepts (discussed in more detail in the following section) of the linguistic worldview theory are taken from many sources, and thus the theory presents an interesting example of postmodern agreement between different branches of the humanities.

1.2 Definitions and crucial concepts

Humboldt is most known from his relativist approach to languages. Although claiming the capacity for language is the same in all humans, he also stressed the differences between languages being “organs of the peculiar ways of thinking and feeling of nations⁴” (Andrzejewski 1989: 162f.). These differences lie primarily in the worldviews. The classical definition by Bartmiński & Tokarski (1986: 72, quoted in Anusiewicz et al. 2000: 28), modified and expanded first in 2006⁵ by Bartmiński, defines the LWV as “[...] a certain set of judgements

⁴ All quotations translated by the author.

⁵ 1st ed. of *Językowe podstawy obrazu świata*; a book translated and published in a modified form as *Aspects of Cognitive Linguistics* in 2009 (1st ed.).

more or less entrenched in the language, contained in or implied by the meanings of words, which reveals the characteristics and manners of existence of objects from the non-linguistic world”⁶. According to the latest definition by Jerzy Bartmiński (2012a: 23) the LWV is

a language-entrenched interpretation of reality, which can be expressed in the form of judgments about the world, people, things or events. It is an interpretation, not a reflection [...]. The interpretation is a result of subjective perception and conceptualization of reality performed by the speakers of a given language; thus, it is clearly subjective and anthropocentric but also intersubjective (social). [...] It influences [...] the perception and understanding of the social situation by a member of the community.

The focal point of this definition is the dynamic character of the worldview contained in language. The LWV is not only an interpretation but also a social and a changing one.

This approach follows closely the idea presented by Humboldt, who is “famous for having rejected the idea of language as a *product*, claiming that it was an activity, a *producing*” (Underhill 2009: 30). The individual is in the centre of this activity – produces language as a reaction to the perceived world, and these perceptions shape language. However, the inherent form of language may also guide the user through the world. The categories of the language are, admittedly, imposed on thinking, but rather as useful patterns than impassable boundaries.

These traits (dynamic, interpretational nature of language) are not always found in other definitions. R. Grzegorzczkova

⁶ Orig. “[...] pewien zespół sądów mniej lub bardziej utrwalonych w języku, zawartych w znaczeniach wyrazów lub przez te znaczenia implikowanych, który orzeka o cechach i sposobach istnienia obiektów świata pozajęzykowego.”

sees the JOS⁷ as a fixed structure rather than a fluctuating one; “I would like to understand the JOS”, she writes (1999: 41), “as a conceptual structure established (fossilised) in the system of a given language, hence in its grammatical and lexical properties (word meanings and their connectivity), realised, as everything in language, in texts (utterances)”⁸. However, language does not simply reflect the world but interprets it (Grzegorzyczkowa 1999: 42, cf. *ibid.*: 45).

W. Pisarek (1978, quoted in Żuk 2010), being the first in Poland to use the name ‘językowy obraz świata’, writes that the JOS is “reflected in a given national language” (“odbity w danym języku narodowym”); thus the Polish ‘obraz’ should be translated as ‘picture’ in both; ‘picture’ being a fossilised, unchanging structure captured in one precise moment in time and mirrored in language). A short discussion of all the above can be found in Bartmiński (2012a: 24).

Bartmiński quotes also a definition by R. Tokarski (2001:366), who wants the JOS to be a set of regularities (to which Bartmiński opposes, saying that ‘regularity’ is an abstract concept; 2012a: 23):

zawartych w kategoryalnych związkach gramatycznych (fleksyjnych, słowotwórczych, składniowych) oraz w semantycznych strukturach leksyki, pokazujących swoiste dla danego języka sposoby widzenia poszczególnych składników świata oraz ogólniejsze rozumienie organizacji świata, panujących w nim hierarchii i akceptowanych przez społeczność językową wartości.

⁷ The acronym *JOS* [językowy obraz świata; ‘linguistic picture/image/view of the world’] is used here to avoid confusion when quoting Polish definitions, where the full name can be translated in several ways into English. Later on in this chapter the LWV acronym is used again.

⁸ Orig. “Językowy obraz świata chciałabym rozumieć jako strukturą pojęciową utrwaloną (zakrzepłą) w systemie danego języka, a więc w jego właściwościach gramatycznych i leksykalnych (znaczeniach wyrazów i ich łączliwości), realizującą się, jak wszystko w języku, za pomocą tekstów (wypowiedzi).”

contained in the categorial grammatical relationships (inflection, word formation, syntax) and the semantic structures of vocabulary, showing language specific ways of seeing the individual components of the world and the broader understanding of the arrangement of the world, the hierarchy in it, and the values accepted by the language community.

The JOS can then be found at all language levels. Still, Tokarski acknowledges the primacy of the lexicon, in line with other Polish scholars. The definition does not, however, reveal anything more on the nature of the JOS – is it an interpretation or a reflection? Is it dynamic or static? These dilemmas are not present in two other Polish descriptions of the idea.

The first is a proposition by Anusiewicz (1994: 113, quoted in Anusiewicz et al. 2000: 29):

określony sposób ujmowania przez język rzeczywistości (zarówno pozajęzykowej, jak i językowej), istniejący w semantycznych, gramatycznych, syntaktycznych i pragmatycznych kategoriach danego języka naturalnego [...] to określony sposób odwzorowania świata dany w pojęciowym rozczłonkowaniu zawartym w języku ujmującym ten świat.

a certain way of describing reality (both non-linguistic and linguistic) by language, existing in the semantic, grammatical, syntactic and pragmatic categories of a natural language [...] aspecific way of mapping the conceptual world contained in the conceptual segmentation in the language describing that world.

It is clearly visible that Anusiewicz does not add anything new to the previous definition by Tokarski except for the pragmatic element⁹. However, he uses the term ‘mapping’, which might suggest that the JOS here is a reflection of reality, a picture.

⁹ Tokarski’s article appeared for the first time in a book under the same title in 1993, published by Wiedza o Kulturze.

The second definition of this type is that of J. Maćkiewicz (1999: 8), who considers the JOS to be a part of a bigger, conceptual picture of the world:

Obraz świata to [...] odbicie doświadczenia poznawczego jakiejś społeczności, [...] określony sposób odwzorowania otaczającego świata, [...] takie modelowanie rzeczywistości, które umożliwia człowiekowi poruszanie się w niej.

The picture of the world is [...] a reflection of the cognitive experience of a community, [...] a certain way of mapping the surrounding world, [...] a modelling of the reality that enables man to navigate in it.

This picture can be present in different semiotic systems, one of which is language. Therefore, the JOS is the picture found in a language. It is an abstract model of the reality, being a part of the linguistic competence of the user (Maćkiewicz 1999: 10). Clearly, the term ‘picture’ is justified here, as the concept presented is a static¹⁰ mirroring of reality, albeit dependent on the subject(s), their cognition, experience and their perspective. Interestingly, though, she supports both Bartmiński and Grzegorzczkova in claiming that language is an interpreter of reality (Maćkiewicz 1999: 12) and a reflection of the speaker’s mentality as well (cf. Bartmiński 2012a: 24).

An interesting example is provided by Allwood (2004). He is one of few modern Western scholars giving a definition of a concept otherwise not often described. He defines the JOS (Swedish *den språkliga världsbilden*, a direct counterpart of German *sprachliches Weltbild*) as a compound of judgments about the world (similar to Bartmiński, who wants to see each notion in a language to be a bundle of traits as described by various subjects): “ett holistiskt system av tankar som används för att identifiera, karakterisera, kategorisera, förklara och förstå många (potentiellt alla) fenomen i omvärlden” (“a holis-

¹⁰ This word is also used by Maćkiewicz herself (1999: 12).

tic system of thoughts used to identify, characterise, classify, explain and understand many (potentially all) phenomena in the world around”). Allwood writes about a system of thoughts, which might call to mind the claim Bartmiński makes about the objects of the linguistic worldview (2012a: 67): they are mental objects, not a reflection of reality (especially if one considers such objects as unicorns or fairies).

Thus, it is clear that Bartmiński admits that some objects are created by language. The linguistic worldview not only cannot be a simple mapping of the world (as there are no unicorns) but also cannot create reality (the existence of the word ‘unicorn’ does not create one). The only entities language produces are mental ones – either images of real world objects or of culture (Bartmiński 2012a: 13). This view is supported by Maćkiewicz (1999: 11f.), who states that words merely name the elements of culture and do not create reality. The LWV is thus a reflection of mentality and not of reality (Bartmiński 2012a: 24). The idea of an intermediate plane (i.e. language) between the world and the individual’s mind is noticeably derived, through Weisgerber’s *Zwischenwelt*, from Humboldt’s concept of language not merely as a mental structure portraying reality but as a complement to mind, interpreting the nature of objects and of our own thinking in order to clarify ideas (Andrzejewski 1989: 159).

R. Grzegorzczkova (1999: 42, 45), however, does not agree with this, claiming that language cannot create the objects of the LWV and that the entities humans talk about are real world objects (even though differently portrayed). The question then is what to do with culturally created notions, such as the ‘unicorn’ (on the details of the language-culture relationship in the LWV paradigm see section 1.3). Nevertheless, all three scholars believe that language is a tool by which the world is interpreted.

The LWV in fact is one of a set of seven interrelated concepts, namely: the LWV itself, stereotypes as components

of the LWV, cognitive definition as a tool for describing linguistic stereotypes, profiling, the values of the subject (both individual and collective), their point of view and perspective, and finally the subject. In this chapter, I am only cursorily describing some of the concepts to give the reader a general view of the theory. I am focusing on the importance of the subject, which will play a crucial role in ch. 3 and 5. Cognitive definition and the stereotype as useful tools will be discussed in section 1.4. The reader should of course refer to the canonical *Aspects of cognitive ethno-linguistics* (Bartmiński 2012a). In close detail the concept of the LWV is examined in M. Guz's book (in Polish) *Językowy obraz świata u wybranych przedstawicieli lingwistyki niemieckiej, amerykańskiej i polskiej* (2012).

The reality as perceived by the individual is shaped in diverse ways across languages and cultures, and the choices made in a speech community are then stored in language. Such an operation of shaping is called 'profiling' in the ESL framework, that is, describing a perceived object in terms of facets (its characteristics). The experiential frame is the base, to which "various cultural codes" contribute (Bartmiński 2012a: 89).

The interpretation is done by *homo loquens*, and it is the position of the subject and not the position of the object talked about that plays the most important role in the LWV theory of Bartmiński. The very name of the central notion shows that the theory in this variant is subject-related: the view (or vision) must be someone's view. Not surprisingly, Bartmiński (2012a: 76) recognises the second variant, the object-related one ('picture of the world', a direct translation of the German *sprachliches Weltbild*). Also here the position of the subject and their point of view is important, but the focus lies with the static portrait of the entity.

Viewpoint as a parameter of perceiving plays a significant role in the theory of the LWV. It is a cultural factor, "a set of directives shaping the content of words", while the

perspective is a bundle of semantic properties of the object resulting from the point of view/viewpoint¹¹ (Bartmiński 2012a: 77f.). A perspective can comprise several different viewpoints, and therefore is a more comprehensive notion. The structure of a cognitive definition and the perspective described in it depends largely on the point of view assumed by the subject (for more details see Bartmiński 2012a: 81ff.). Therefore, the subject and their viewpoint can be reconstructed by a careful and detailed profiling of a notion.

The profiling is a categorial arrangement of facets, i.e. bundles of judgements made by the subject from a particular viewpoint, which form the cognitive definition. The object in the definition is described in terms of stereotypical, recurring features. Bartmiński (2012a: 63–65) understands stereotypes as “stable connections of meanings”, stored in collective memory and reproduced as *topoi* (utterances purely semantic in form, not fixed), *formulae* (established combinations transparent semantically to some extent) and idioms (formally fixed combinations, “without clear semantic motivation”). The cognitive definition is a contextual one, that is, one containing connotations of the object described. Such a definition corresponds to the ‘definition through postulates’ of Ajdukiewicz, being “an arrangement of sentences (‘postulates’) that contains the word being defined in various contexts and that meets two conditions: of non-contradiction (i.e. it has a solution) and of non-ambiguity (i.e. it has no more than one solution)” (Bartmiński 2013b).

All these basic concepts in the ESL theory clearly indicate that a worldview cannot exist without its most important element – not only a subject but also a community with a common cultural background, without which the facets cannot be reproduced and reconstructed.

¹¹ ‘Point of view’ and ‘viewpoint’ are treated synonymously by Bartmiński, although the preferred term is the latter one.

1.3 The *language – culture – thought* relation

The problem of the relations mentioned in the title of this section is a complicated one when it comes to the ESL theory. A whole paper by Łozowski (2013) is devoted to the influence of language on culture and vice versa. The linguistic worldview notion suggests that the core element is language. Does this, however, influence culture, or is it influenced by it? Where is the site of cognition and thinking?

It is generally agreed that the programme proposed by Bartmiński treats culture as an inseparable part of the worldview and that co-linguistic¹² data (i.e. culturally determined behaviour) are important material for studies (Maćkiewicz 1999; Żuk 2010). Although Anusiewicz, Dąbrowska and Fleischer (2000) propose that the status of culture be elevated in their theory of the cultural worldview, this, in fact, adds nothing new to the LWV programme. The only difference is the suggested material for research: it is based on the standard variety of Polish, rather than on folk varieties, and includes facial expressions, gestures etc., which in any case the ESL does not exclude.

Bartmiński admits that the linguistic worldview can have two interpretations: subject- and object-oriented. The first one (the preferred one in the ESL) focusses on the performance of speakers, pragmatics and social considerations. The second one is language-oriented; the object is contained in language itself (Bartmiński 2012a: 76). According to Łozowski the subject-oriented interpretation can be drawn directly from the term ‘linguistic worldview’, which syntagmatically corresponds to ‘[językowy] [obraz świata]’ and suggests that conceptualisations entrenched in language derive from cultural experience (2013: 352). This view might be supported by Allwood’s claim (2004) that a worldview is a system of thoughts, and therefore a crucial part of culture. The object-oriented interpretation re-

¹² In Bartmiński (2012b) the term is ‘ad-linguistic’.

quires language to influence culture, which is a view coming from the works of Sapir and later Whorf (at least partly).

Resolving this issue is not made easier by Bartmiński himself. On the one hand, he states that culture is one of many components of language (2012a: 9), on the other, he claims that culture includes language (2012a: 12). Łozowski (2013: 364-366) rightly points out that Bartmiński, calling his programme ‘cognitive’, suggests subject-oriented interpretation, whereas writing that “language conditions culture, for without it one cannot participate in culture or in social life” (Bartmiński 2012a: 12), he assumes the views of Sapir and his followers.

The very definition of the LWV (a socially intersubjective interpretation of reality entrenched in language) together with the claim that the ESL investigates stereotypes implies the primacy of culture over language. However, the continuous references to Sapir (although Bartmiński explicitly rejects the strong version of the relativism hypothesis) could testify to the opposite.

Łozowski (2013: 367), quoting Bartmiński, claims the solution to the problem is to acknowledge the bridging position of values. However, even though Bartmiński writes that values lie at the foundation of both language and culture, he does not indicate the source of them, other than the human subject. Does this mean that values arise in mentality/cognition? Or are they produced within culture (which would suggest that culture takes primacy)?

Let us look at how other scholars approach this issue. Anusiewicz (1994: 28, quoted in Anusiewicz et al. 2000: 26) believes that the observations of the world (the cognitive act) chosen by a community create linguistic categories forming an interpretation of the observed world, which in turn influence the cognitive categories. It is visible here that cognition and language are mutually related, although the categories are culturally dependent (the social factor is decisive). Grzegorzyczkowa (1999: 40f.) presents the Ogden-Richards’ triangle – words are connected to real-world objects through the mentality of the speakers. Only in the last paragraph of her

paper does Grzegorzczkova mention the place of culture, which considerably affects communication (1999: 46). Maćkiewicz (1999: 11f.) approaches the problem similarly to Anusiewicz: a human subject perceives objective reality and categorises and interprets it by means of language, which is a reflection of their mentality. The influence of culture is decisive in the process of creating new objects in the worldview. The LWV is a common cognitive basis to which every member of a community must refer. Therefore, Maćkiewicz also adopts the weak version of the Sapir-Whorf hypothesis by saying that language “facilitates and suggests such and no other interpretations” (1999: 18).

It seems that this model of the cycle of influence (see Figure 1) is not only widespread but also accepted by Bartmiński. He argues that both language and culture are “an artefact of human cognitive activity” (2012a: 102) and that cultural patterns should be included in the experiential frame along with conceptualisations (2012a: 89).

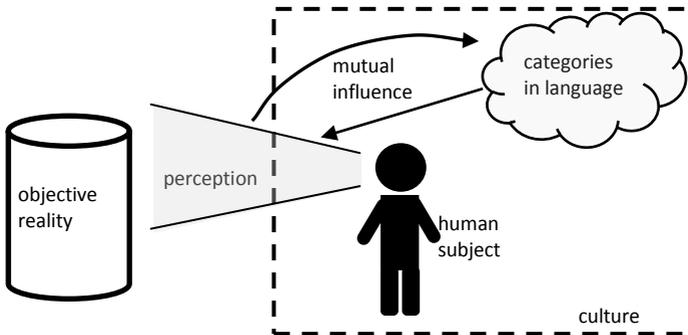


Figure 1 The (simplified) language – culture – perception relation

This not only follows closely the assumptions of Humboldt, who claimed that language helps shape/structure thoughts, and at the same time that world-perceiving shapes the language in use, but is also compatible with the conception

of Underhill (2009: 134f.), who differentiates between world-perceiving and world-conceiving (the latter one visible in texts):

- *world-perceiving*, for the changing and developing perception we have of the world,
- *world-conceiving*, for the changing and developing manner in which we draw that world into the realm of thought and form concepts and frameworks to represent things and our experience of the world.

According to Underhill (2009: 135) the *Weltanschauung* might also be split into three, this time, different notions: ‘cultural mindset’ (i.e. a general, relatively rigid social conception of the world), ‘personal world’ (i.e. the individual’s mindset) and finally ‘perspective’ (i.e. interactive, shifting viewpoint). It is the interactions of individuals that shape the worldview and therefore the language.

It can be concluded that thinking, language and culture are inextricably intertwined in Bartmiński’s approach. This should not be surprising given the fact that the idea of the LWV comes from Humboldt, for whom thinking and language were complementary activities of the human spirit, cultivating culture and being cultivated by it (Andrzejewski 1989: 153f.; Underhill 2009: 65f.).

1.4 Methodology

Scholars working in the ESL framework propose taking into account different levels of language, from vocabulary to grammar (Anusiewicz et al. 2000; Bartmiński 2012a; Grzegorzczkova 1999). Bartmiński (2012a: 71) explicitly expresses his conviction that in order to obtain “content adequacy” (i.e. the inclusion of folk knowledge) in the cognitive definition of an X, a researcher may refer to several different types of data: the language system, texts, interviews with native speakers and sociological and ethnographic data (cultural

use of an object and speakers' behaviour in relation to it). However, the cognitive definition may well be based on only one of these types. Only recently, the EUROJOS project postulated that the examination is to be based on a wide array of data to account for a language in its dynamic entirety.

Such an approach to language as a changing structure is based on the belief of Humboldt, who considered it as an activity, an *energeia*. Any definition of language should therefore be based on individual instances of *parole* (see Andrzejewski 1989: 157), that is real-life data.

In his paper on the role of etymology for the LWV reconstruction Bartmiński (2013a: 235) elaborates on the material and its features:

rekonstrukcja [językowego obrazu świata] opiera się na szerokiej bazie materiałowej, odwołuje się do semantycznej analizy kategorii gramatycznych i słownictwa, do łączliwości leksykalnej (stałej i okazjonalnej) i metafor, do struktury logiczno-semantycznej i treści tekstów zarówno kliszowanych (tj. różnych gatunków folkloru: przysłów, zagadek, pieśni, bajek itd.), jak też kreowanych (w ich warstwie presuponowanej, implikowanej). Przydatne dla rekonstrukcji JOS są teksty wywołane, tj. odpowiedzi na pytania o rozumienie słów przez ich użytkowników oraz opisy rytualnych zachowań i wierzeń dotyczących nazywanych przedmiotów. W rekonstruowaniu JOS znaczący udział ma też analiza etymologiczna [...].

[the LWV's] reconstruction is based on a broad material basis; it refers to the semantic analysis of grammatical categories and vocabulary, lexical connectivity (permanent and occasional) and metaphors, the logical-semantic structure and content of texts both clichéd (i.e. various folklore genres: proverbs, riddles, songs, fables etc.) as well as created (in their presupposed, implied layer). For the reconstruction of the LWV, elicited texts are useful, that is, answers to questions about a user's understanding of words and descriptions of ritual behaviours and beliefs about the objects referred to. In reconstructing the LWV, etymological analysis also plays a significant part [...].

Word etymology not only shows how language users conceptualise the world and reveals the choices lying behind the words, but also helps the researcher to decide what the internal hierarchy of meanings of a notion should look like and to prioritise primary meanings (Bartmiński 2013a: 236; Underhill 2009: 108).

These assumptions have been developed to the fullest in the EUROJOS project (Abramowicz et al. 2009; EUROJOS 2008). The methodological instruction¹³ distinguishes several sources of material within system and “real-life” data. The data should be extracted from dictionaries (with the reservation that they do not reflect the typical worldview of the everyday language, and that only general entries should be considered while specialised meanings are left out), texts, corpora and questionnaires. The system data should include the “whole network of lexico-semantic relations” both paradigmatic and syntagmatic:

- hypernyms and hyponyms
- opposites
- synonyms
- derivatives (word-formative and semantical)
- complexes and collections
- collocations/phrasemes
- proverbs (treated as minimal clichéd texts; with the proviso that original and not borrowed ones should primarily be studied)

Great modern monolingual **dictionaries** are the source of basic lexeme definitions with their collocations and relevant quotes. Bartmiński (2012a: 67, 71) remarks that the cognitive definition (both the material analysed and the metalanguage)

¹³ A methodological instruction: the principles and stages of proceeding in the development of the entries within the EUROJOS research programme (available at: http://ispan.waw.pl/default/images/eurojos/eurojos_instrukcja_ma_ibg_wch_jb_26_ii_2011_.doc)

should be primarily occupied with the colloquial variant of a language. However, the boundary between the scientific and the colloquial is blurred. Stylistically neutral usage examples should be extracted from modern **texts** and **corpora** (about 200-300 contexts) balanced in respect of the style (popular scientific and journalistic) and political orientation (both left- and right-wing journals; this follows the assertion that *Weltansicht* is not the same as *Weltanschauung*; see sections 1.1 and 1.3). However, the researcher has to bear in mind that stereotypical judgements are not introduced as simple assertions but rather hidden in presuppositions, and therefore require thorough analysis. It is important to focus on those judgements that are statistically reproduced. Instances of individual judgments can be included, provided they are culturally and ethnolinguistically relevant. The last step in the procedure is a **questionnaire** with only one obligatory question, namely “In your opinion, what is a true X like?” Other than this, questionnaires can contain open-ended questions as well. These include (cf. Bartmiński 2012a: 132–148, 178–198) supplying only one word best describing the X, naming objects characteristic of X, supplying a noun to an adjective, providing synonymous expressions or filling in blanks (collocation test). The study sample should be a balanced one, with at least 100 respondents. Interestingly, Bartmiński (2012a: 132–148, 178–198) differentiates between ‘ideal’, ‘typical’ and ‘true’ features (‘true’ corresponding to Lakoff’s ‘real’). It seems that ‘ideal’ represents an exemplary X (prescriptive view), ‘typical’ represents an average X (descriptive view), whilst ‘true’ combines both. Therefore, a questionnaire with a question about the ‘true’ features of an X is sufficient, although the other two types may also be included.

Stereotyped judgements about notions may further be studied through Osgood’s semantic differential. Bartmiński (2007) proposes a three-step procedure in which relevant attributes should be selected. Step (1) is the said questionnaire, in which the respondents are prompted to provide one feature

of an X. Based on the responses, the researcher (2) creates a set of antonymic pairs as a base for (3) the subsequent semantic differential. It is proposed that the antonyms should be placed on a scale graded from 3 to 0 and back to 3 (3 being 'extremely', 2 – 'as usual', 1 – 'a little, some' and 0 – 'neither A nor B, neutrally'). This approach is well advised. It does not impose on respondents biased, mechanically created antonymic pairs and allows for studying languages non-native to the researcher. The semantic differential is not only useful in assessing the strength of single notions but can be even employed in analyses of synonymous or semantically close notions. Bartmiński (2007: 79f.) shows that the results differ for such synonymous notions as *chłop* 'peasant' and *rolnik* 'farmer'.

The answers are treated as texts (of a specific genre; see Bartmiński 2012a: 179) and can be divided into denotative associations about the object coming from the user's extralinguistic knowledge and connotative associations about the object's name coming from linguistic knowledge. Their relevance should be tested through the 'but' test: ready-made sentences should be given to the respondents to assess their acceptability (Bartmiński 2007: 82f.). Alternatively, the respondents could finish such sentences themselves, from which their presuppositions can be extracted (as in Bartmiński 2012a: 134, 180).

The answers in the EUROJOS project are later grouped, coded and statistically analysed¹⁴. The material is divided into domains (e.g. social aspect, ideological aspect, physical aspect, etc.) in which specific descriptors find their place. For example, if the respondents give names of non-basic colours (e.g. coral, crimson and burgundy), the answers should be grouped under one basic descriptor/keyword

¹⁴ According to Komunikat po międzynarodowej konferencji pt. „Teoria językowego obrazu świata i metody jego rekonstrukcji. Problem eksplikacji wartości” połączonej ze spotkaniem warsztatowym (EUROJOS-VII) (obtained in personal correspondence).

(here: red) and allocated to a domain (here: physical aspect). Features within the domains (i.e. facets of a cognitive definition) are to be presented as “minimal diagnostic contexts”, that is

in the form of sentences communicating stereotypical judgments of the object. In other words, these are not abstract names of features but sentences or their equivalents: “A horse pulls wagons,” “A horse is a saddle animal,” “A horse is a healthy animal (as a rule),” “A horse can sense a person’s death,” etc. These sentences function in a pragmatic-modal frame that one can express as “the speakers think that...” and relate it to “a stereotypical horse.” (Bartmiński 2013b: 170)

AN EXAMPLE OF ANALYSIS

To better understand this concept an example is useful. M. Grzeszczak in her 2009 paper reconstructs the cognitive definition of DEMOCRACY. She uses three complementary sources of data: system data, questionnaires and texts from two daily newspapers. In the paper

a political understanding of democracy is accepted as a political system in which a specific form of government is practised [out of six dictionary definitions]. On the basis of three types of data, the author identifies the defining (base) features of democracy, correlated with specific aspects of the concepts. (2009: 83)

Grzeszczak recognises three senses of the concept: economic, social and ethical. Her cognitive definition is broken down into facets of both “system” and semantic type. After each facet’s name, a brief explication is presented, containing an explanation, as follows¹⁵:

1. [Name and its etymology]

Democracy is a word borrowed from Greek through Latin and it is a Europeanism.

¹⁵ The version presented here is a much-shortened one.

2. [Collections]
According to the consulted dictionaries, *democracy* co-occurs with *Rzeczpospolita* (Pol ‘Republic’). In the analysed texts *democracy* co-occurs with concepts-values, e.g. *democracy and justice*
3. [Collocations]
stable democracy, liberal democracy (in dictionaries), *modern democracy, true democracy* (in texts)
4. [Synonyms]
the rule of people
5. [Opposites]
dictatorship; in texts *democracy* is opposed to *communism*
6. [Who is the source/subject of (superior) authority?]
nation or *people* in dictionaries, *people* or *most citizens* in questionnaires
7. [What is democracy based on? / What is the foundation of democracy?]
According to respondents *democracy* is most closely linked with ethical values such as *freedom/liberty*
8. [What guarantees that democracy works?]
the rule of law understood as *properly functioning law* (questionnaires) and *tolerance*
9. [What does democracy give to / guarantee to people]
freedom and political rights to all citizens (a dictionary), *free elections* (dictionary, questionnaires)
10. [Results of democracy / What democracy brings:]
optimal conditions for development of every man (questionnaires)
11. [Democracy as an object, goal, and sth desired / What does democracy require?]
According to a dictionary, *democracy* needs *fighting for*. In light of text data, it needs to be *learnt, supported* and *built*.
12. [What is democracy NOT? / What is a contradiction to democracy?]
As examples compounds with *-kratía* are given.

13. [What are the diseases of (modern) democracies?]
corruption (in both questionnaires and texts)
14. [What restricts democracy? / What are threats to democracy?]
globalisation (texts)

Such an approach is presented in and recommended by the *Słownik stereotypów i symboli ludowych* ('Dictionary of folk stereotypes and symbols'), the flagship publication of the ESL.

The paradigm presented here describes in much detail the types of material and ways of collecting it. Bartmiński specifies also how the relevant attributes of an X should be chosen, through Osgood's semantic differential and the contradiction test. A shortcoming of this framework is the scant use of corpora. The methodological instruction recommends that only 200-300 contexts should be taken into consideration. Why two hundred, and which ones these should be is not indicated. Indeed, corpora may serve as a major source of automatically extracted collocations (a more reliable one than even big monolingual dictionaries) and concordances with their frequencies. The fact that the data can be automatically generated significantly shortens the analyses, as well as increases their accuracy and reproducibility (Kamasa 2014).

Although Bartmiński postulates taking into account all possible connotations of an X (in which a corpus would be helpful, preventing omissions), he also notes that some of them may "turn out to be occasional or coincidental" (2012a: 68). The problem of extraction of the criterial features is still an unsolved one. Bartmiński comes closest to a solution in his "Kryteria ilościowe w badaniu stereotypów językowych" ['Quantitative criteria in in studying linguistic stereotypes'; 2007, first published in 1988] where he states that qualitative and quantitative measures should be collated (i.e. the results from contradiction tests and from frequency and/or rank lists). He remarks that the semantic acceptability of contradictory statements correlates with a feature's rank. For example, it is acceptable to say

“He’s a doctor, but a poor one” because the feature presupposed here (wealth) is a stereotyped high-ranked characteristic of doctors. The correlation is, however, unidirectional (Bartmiński 2007: 83):

obejmuje cechy o wysokich wskaźnikach wyborów, natomiast nie pozwala na wykluczenie cech ze środka czy z końca listy; np. cecha *postępowy* ma przy rolniku zaledwie 21% wyborów, ale zdanie o kimś, że jest rolnikiem, ale nie jest postępowy — jest zdaniem „normalnym” [...]. Znaczy to, że o ile wysoka ranga cechy implikuje akceptowalność zdania z presupozycją tej cechy, o tyle niska ranga cechy nie implikuje jego nieakceptowalności.

[it] includes high-rank characteristics but does not allow to exclude characteristics from the middle or the end of the list; e.g. the feature *progressive* of a farmer is in only 21% of choices, but a sentence that someone is a farmer, but is not progressive — is a “normal” sentence [...]. This means that while the high rank of a feature implies the acceptability of a sentence with the presupposed feature, a low rank does not imply its unacceptability.

This problem is not easily solved. Even though corpora facilitate collection and preliminary preparation of linguistic material, automatically generated word lists are merely a starting point. The data are subject to subsequent subjective decisions by the researcher.

A key problem with this framework, though, is that there are no indications as to how the results should be grouped and processed, how the domains and descriptors are chosen and the facets distinguished. For example, Prorok and Gład (2013: 187f.) write that

[the defining] sentences will be arranged in special semantic categories (facets), that is, groups treated as homogenous from a certain point of view: names, categorisation, complexes and collections, oppositions and gradation, origin, transformation, appearance and properties, or actions directed at iron.

How these “special semantic categories” are selected does not follow. The authors seem to suggest that the researcher should rely on their native competence or knowledge of competent judges.

Niebrzegowska-Bartmińska (2015: 32) describes a facet as

wiązka cech, składająca się na eksplikację, odkrywana drogą analizy materiału, a nie narzucana z zewnątrz, a wtórnie traktowana też jako siatka porządkująca materiał i ułatwiająca porównywanie opisów.

a bundle of features, constituting an explication, discovered through analysis of the material and not imposed from the outside; secondarily it is treated as a grid ordering the material and facilitating comparison of descriptions.

The material is considered here as “revealing its own structure”, an assumption which may be challenged. After all, even if the material shows some patterns, it is the researcher who decides subjectively which domains “reveal themselves” and imposes such and not another grid on the data. It is also very important to note that facets may serve as the *tertium comparationis* (TC) in multilingual research. As said, there is some degree of arbitrariness in establishing a full list for a concept. Besides, such a procedure requires native or near-native knowledge of the language in question, making comparative research by one researcher difficult or even impossible.

The problem of the TC is discussed in detail by Bartmiński (2012a: 214–218), who begins by noting that the choice of an object of comparison can be made from two perspectives: onomasiological and semasiological. Both are problematic: in the onomasiological perspective “the comparative procedure is relatively straightforward in the case of unambiguously identifiable objects” (Bartmiński 2012a: 216), but abstract concepts differ greatly across languages and cultures; in the semasiological perspective establishing clear boundaries of meanings (division of the world into lexemes) is challeng-

ing. He proposes two solutions: working on a concept in a specific cultural sphere, where in unclear cases a reference can be made to their common source (e.g. Latin, in the case of the Mediterranean culture) or applying a universal semantic metalanguage. Again, both seem insufficiently non-arbitrary. Firstly, there is no guarantee that a concept in all analysed languages comes from the same source. Furthermore, the concept in question has to be first well scrutinised. This means, in fact, that a concept in any language can serve as the TC only if first described in detail. Secondly, the Natural Semantic Metalanguage (NSM), which Bartmiński references explicitly, can also be criticised as arbitrarily devised or, at least, as being imposed in advance, instead of coming from the data at hand, the latter plainly advocated by the ESL¹⁶.

To conclude, I would like to quote G. Lazard (2001: 365), who declares that the solution to the problem of the TC is

to form hypotheses and elaborate sets of concepts for each grammatical domain for use as a *tertium comparationis*. Such concepts are **logically arbitrary and necessarily based on intuition** [my emphasis, I.S.]; however, the intuitions behind them are better for being inspired by a large body of experience with linguistic structures. Only via empirical research can these hypothesis [*sic*] be validated. If they are not validated, they have to be replaced by others that better account for the data. I am afraid there is no other possible procedure for typological comparison.

¹⁶ For a discussion of the NSM see *Theoretical Linguistics* (2004). 29 (3).

2. Artificial languages

2.1 Definitions¹⁷

A comprehensive definition of an artificial language is hard to find, partly because this branch of linguistics has been neglected or looked down upon (a famous quote by Chomsky goes “Esperanto is not a language. It’s just parasitic on other languages”). In fact, the very name is controversial, and many synonyms can be found in the literature. The term in itself comprises a good deal of various, only remotely related codes and systems. It may refer to both *langue* and *langage* (see Lyons 1991). Therefore, a clear definition is needed to draw a boundary between what is considered natural and the rest of the systems. Under the name ‘artificial’ the following languages are often placed (cf. Albani & Buonarotti 1994: 9; Carlevaro 1989: 177; Eco 1997: 2f.; Sakaguchi 1998: 26–28):

- programming or machine languages (COBOL, Assembler, C#)
- formal languages (propositional calculus)
- experimental languages:
 - philosophical (Toki Pona)
 - logical (Lojban)
 - pasygraphies (Pictopen) and pasylalies
- international auxiliary languages or planned languages (hereafter IALs; Esperanto, Novial)
- artistic languages (Klingon, Quenya)
- normative languages:
 - superdialectal (Rumantsch Grishun, Standard Arabic)
 - standard literary languages (*Ausbausprachen*)
 - revived (Cornish)
 - classical languages (Sanskrit, Latin)
- controlled languages (Caterpillar Fundamental English)
- reconstructions (Proto-Indo-European)
- pidgins and creoles
- oneiric languages and glossolalias.

¹⁷ This section is partly based on Stria 2013.

Here the term ‘artificial’ is used as a broad name for all languages that are the result of deliberate and conscious (creation and) planning. Therefore, the last type will not be examined in detail, because it does not meet the requirement of being deliberately created, although it is discussed in the classification of Albani & Buonarotti (1994) (Figure 4 in this chapter). It is worth mentioning that in their dictionary the term ‘imaginary’ is preferred to ‘artificial’ or ‘invented’, because it stresses the role of fantasy and imagination in the process of creation of fictitious languages. Their understanding of the term is very broad, and the creators of the languages described in the book range from professional linguists to outcasts of society (Albani & Buonarotti 1994: 8):

[...] con il termine «lingua immaginaria» intendiamo semplicemente una «lingua non naturale», dove l’attributo «naturale» sta ad indicare una lingua il cui apprendimento avviene per trasmissione orale dai genitori e dall’ambiente circostante. Sotto questo profilo «immaginaria» è ogni lingua di tipo artificiale, frutto dell’elaborazione a tavolino di una o più persone [...].

[...] by the term ‘imaginary’ we mean simply a ‘non-natural language’, where the attribute ‘natural’ denotes a language which is learnt through oral transmission from parents and from the surrounding environment. In this respect, ‘imaginary’ is any type of artificial language, the result of work at the desk of one or more people.

A shorter list of systems called ‘artificial’ is given by Blanke (1997: 3):

1. Regularized and standardized literary language, as distinguished from dialects [...].
2. Ethnic languages, highly regularized to maintain them at a particular stage of development (Sanskrit, Church Latin) or to modernize them (Modern Hebrew, Bahasa Indonesia, Landsmål).

3. Consciously created languages to facilitate international communication [...], that is, planned languages.
4. Nonredundant, formulaic, or symbolic languages to facilitate scientific thought [...].
5. Programming languages for computers [...].
6. Machine languages for automatic translation.

The existing definitions of what an artificial language is are very general and often vague. Fettes (2005) barely enumerates several systems which can be classified as artificial. The entry in Malmkjær (2002) highlights the need of a speech community to call a language natural and the fact that artificial languages are deliberately created:

An artificial language is one that has been created for some specific purpose or reason, as opposed to a natural language, such as those spoken by most speech communities around the world, which is normally thought of as having evolved along with its speech community, and for which it is not possible to find some ultimate source of creation. The machine codes and various programming languages we use with computers (see artificial intelligence) and the languages of logic (see formal logic and modal logic) are all artificial languages [...]

Likewise, Bausani (1970: 7) states that the artificiality of a language stems from it being created non-spontaneously and non-naturally (“nicht-naturbedingt”). A similar definition is given by the *Columbia Encyclopedia* online (2014; under ‘international languages’):

An artificial language is an idiom that has not developed in a speech community like a natural tongue but has been constructed by human agents from various materials, such as devised signs, elements or modified elements taken from existing natural languages, and invented forms.

It does not follow exactly which languages are artificial. A pidgin has also been “constructed by human agents” from

diverse components “taken from existing natural languages”, but its naturalness is never questioned (probably because an unconscious character of creation). In the same way, revived languages such as Modern Hebrew or Cornish can be treated as artificial but typically are not (further details in section 3.3).

This problem is mentioned by Bartlett (2009). At first, the intention is emphasised:

An artificial language is a language that has been deliberately designed for a purpose by one person or a small group of people over a relatively short period of time. (Adapted with permission from a definition by Richard K. Harrison, personal communication, 2004.)

Later on, however, the author of the entry briefly mentions the difficulties that arise when this particular definition is used. Should pidgins be considered artificial? Are reduced languages like Basic English natural? The author also indicates that although the definition includes programming languages, these are not discussed, as they do not serve the purpose of human-to-human communication.

Both Large (1994) and Kennaway (2010) (even though the entry in the 3rd edition of the *Routledge Linguistics Encyclopedia* has been revised and seriously modified) offer definitions analogous to the previous ones. Some other names for artificial languages are given, such as ‘conlangs/constructed languages’ and ‘planned languages’.

Some (similar) constraints on the categories ‘natural’ and ‘artificial’ are put by Duličenko (2001). He lists three structures: ‘planned languages’ (or ‘international artificial languages’), ‘language constructs’ and ‘ethnic languages’.

Planlingvo (PL), aŭ alie internacia artefarita lingvo (IAL), estas objekto, kiu troviĝas inter lingvokonstruo (LK) kaj etna lingvo (EL). LK estas plene produkto de racia kreado ellaborita de unu homo aŭ, plimalofte, de malgranda kolektivo. EL estas spontanaracia kreado ellaborita dum longa tempo de unu konkreta etna

kolektivo. LK estas pankronia – nur de momento de ĝia ensociiĝo komencas kreiĝi ĝia diakronio: EL, male, estas elkomence socia kreado kaj tio ĉi estas bazo de ĝia diakronio. LK estas refoja, ĝi kreiĝas laŭ similo al EL; EL estas tiamaniere elkomenca laŭ sia genezo kaj uziĝas kiel bazo por konstruo de PL.

Planned language (PL), or in other words international artificial language (IAL), is an object that lies between a language construct (LC) and an ethnic language (EL). A LC is fully a product of rational creation of one person, or, rarely, of a small collective. An EL is spontaneous-rational creation elaborated during a long time of one distinct ethnic collective. A LC is panchronic – only from the moment of its socialisation its diachrony begins. An EL, on the contrary, is from the beginning a social creation, and this is the basis of its diachrony. A LC is secondary; it is created in resemblance to ELs. An EL thus is primary by its origin and is used as a basis for construction of PLs.

An interesting case is presented by Gobbo (2012). He discusses the taxonomy given by Lyons (1991), while pointing out some problems connected with it. According to Lyons, there is a four-class division of naturalness (see detailed description in section 3.2.2). Gobbo studies this classification in detail, coming to a conclusion that the only types of systems which can without doubt be called artificial are programming languages (2012: 190). However, Schubert (1989: 9) considers such systems (along with machine and formulaic languages) to be “highly restricted subsets of a language” and does not account for them in his article.

All those definitions are in fact alike and leave room for many borderline cases. It does not help to look at definitions of language in general, as there are probably hundreds of them, producing cases that are even more problematic. Just to give an example: *Britannica* (2014; under *language*) quotes three linguists who explicitly write that language is “vocal” (Bloch & Trager) and a combination of “speech-sounds” (Sweet). That would rule out sign languages. Gobbo (2012) requires language to undergo “graphisation” and “socialisa-

tion". While the latter condition is not surprising in light of the definitions given above, the first one seems odd because it would mean denying the name 'language' to a large number of small non-codified ethnic languages. Gobbo explains, of course, that he merely means a higher degree of abstractness and not the necessity of an alphabet, while ruling out the possibility that language would be simply equated with *parole*.

In the interlinguistic literature the term 'artificial' as opposed to 'natural' is regarded as "crudely misleading" (Schubert 1989) because it suggests that languages created to facilitate international communication are in fact identical to machine or formulaic languages. As has already been mentioned, the name 'artificial' encompasses various systems and it will be used throughout this book as a general name for all languages that underwent any kind of planning. Other names have also been used throughout history: 'universal', 'international', 'auxiliary', 'constructed', 'planned', and 'invented'. Okrent (2010) calls those languages 'invented', whereas Bausani uses both 'artificial' and 'invented' "als Sammel- und Oberbegriff für die Universal- und Geheimsprachen"; ("as a collective and generic term for the universal and secret languages" – note that the latter are not cryptic codes but concealed, secretive languages; 1970: 7). The titles of Eco's (1997) and Large's (1985) books already hint at which terms the authors prefer: 'perfect' (alongside 'universal' in the text) and 'artificial' respectively. Blanke (1989) distinguishes between two groups of terms that sometimes overlap, i.e. the term 'constructed/artificial', which points towards the creation of the language and 'planned/universal', which describes the language's function. This book uses the term 'artificial' rather than the currently popular 'constructed' to avoid confusion between all artificial language-systems and modern day hobby constructed languages also abbreviated as *conlangs*.

A simple and clear explanation of how to distinguish between 'universal' and 'auxiliary' languages is given by Liu (2006: 44). For him (and many interlinguists, i.a. Blanke)

the term ‘universal’ denotes languages created from the 17th c. onwards meant to be common for all and forever, and to replace ethnic languages. The term ‘(international) auxiliary’ and synonymous ‘(international) planned’ is used to represent languages created to facilitate human communication but not necessarily to replace existing languages. This is shown in Figure 2.

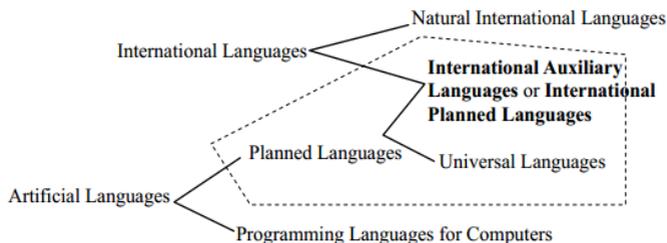


Figure 2 Relationships between artificial and international languages. Reprinted from Liu 2006: 44

All the definitions given above describe artificial language as a consciously/deliberately invented language with a known creator. In fact, this is a “genetic” trait, and does not say anything about the development of artificial languages and the internal variation of the group. If such a binary division is assumed, the only languages not clearly belonging to either natural or artificial are controlled languages (e.g. Basic English), created by deliberately limiting a natural language. Such definitions leave a lot of room for borderline cases, these being controlled languages, revived and revitalised languages (where historical continuity is broken, the language has no or few native speakers and the revived version is a common effort of a group of enthusiasts, e.g. Modern Hebrew), linguistic reconstructions (hypothetical languages recreated by linguistics), pidgins (simplified vernaculars based on several languages created on purpose but linguistically naïvely) and finally various types of signed languages. A more detailed discussion of

these cases, taking into account various characteristics other than origin, is presented in section 3.3, while the typology of artificial languages will be taken up in section 2.3.

2.2 History and motives of creation

Most definitions of artificial languages call attention to the fact that such languages are created for a specific purpose. There are ten motives listed by Bartlett (2009). Some languages are intended to “replace an entire family of languages” (Tutonish), others to be used as auxiliary systems for international communication (Esperanto). There are of course languages designed for artistic use (Quenya, Klingon) or for personal enjoyment (hobby languages on the Internet). Along with secret languages, one might find mystic ones, although some of them are not conscious creations. A few languages belong to the category of philosophical and experimental systems created to test some hypothesis (Loglan, Láadan). Some cater for “special communication needs”, e.g. Blissymbolics for people with disabilities. The last two types are languages “allegedly usable in psychoanalysis” (aUI) and those for communication with extra-terrestrial life forms (Lincos).

Bausani (1970: 16–43) gives six types of what he calls “Spracherfindertum” (“language invention”):

- linguistic taboo (language engineering)
- borrowings from prestigious varieties and creation of compounds
- relexification in argots and pidgins
- child languages
- oneiric languages and glossolalias
- poetic creations and riddles

A similar, very detailed list of language creation cases is offered by Meyer (1901, quoted in Sakaguchi 1998: 28–31):

- spontaneous or intentional disruption of natural development (lexical creations, language games etc.)

- spontaneous or intentional changes (children's secret languages, nicknames etc.)
- remodelling of the usual linguistic material (baby talk, jargons, skaldic poetry, purism etc.)
- artificial languages in narrow sense (Volapük, Esperanto etc.)
- sound-symbolism (glossolalias, riddles etc.)
- abstract language creation (*a priori* languages)
- arbitrary creations (taboo languages)
- sign systems (sign languages, gestural languages, flag semaphores, symbolic languages and formal languages)

Meyer concludes by stating that only the last five examples can be regarded as artificial languages, because they are wholly artificial, while the first three only partly. It has to be indicated, though, that sign languages have been proved natural and that this list is unclear in many places (on sign languages see sections 2.2.1 and 3.3.3; see also other classifications by Bausani in sections 2.3.1 and 2.3.2). All of these types can give rise to new artificial constructs; however, they mostly only amount to a small part of the future language.

Nevertheless, the main reason for creating a language is probably one of many kinds of striving for perfection. As Eco (1997: 2f.) notes, it can be perfection in terms of function/structure, resulting in *a priori* systems, in terms of universality, resulting in *a posteriori* languages, and in terms of practicality, resulting in polygraphies (i.e. universal scripts). Nevertheless, the motivations go even further, leading to attempts to rediscover the original language of Adam or to reconstruct the primordial language.

Historically, Bausani (1970: 15) divides artificial languages into three categories:

- primitive languages and those of the Middle Ages
- Renaissance and the 17th c.
- auxiliary languages of the 19th and 20th centuries

This, obviously, is a very simple division and the timeline of artificial languages is much more complicated. Even the beginnings are very difficult to determine. As one of the oldest examples Albani & Buonarotti (1994: 15) give the language of the Frogs of Aristophanes (405 BC). As for a work referring to a possibility of an artificial language, the first one is probably *Cratylus* by Plato (4th c. BC). Some fragmentary information is available on the language of Galen (2nd c.) and later the gestural language of Beda Venerabilis (7th/8th c.; see Duličenko 1990).

However, any sensible timeline or list of artificial languages is in fact impossible to create, because there exist hundreds if not thousands of “languages” created for artistic purposes only, in order to make a book or a poem more colourful and interesting and, more recently, myriads of schemes invented “for the sheer joy of it” and published on the internet. The most comprehensive lists have been assembled by Albani & Buonarotti (1994; no less than 1,100 pre-internet era languages), Duličenko (1990; exactly 917 IALs from a language project of Galen of the 2nd c. until the 1970s) and rent (2010; precisely 500 from *Lingua Ignota* of the 12th c. up to Proto-Central Mountain of 2007). As for the newer languages, the website www.langmaker.com (closed in 2009, although some mirror websites still maintain the copy) was the best source of knowledge with over 1,000 entries. Because of the limited scope of this study, the works of authors “who poetically distort natural language such as Lewis Carrol” (Higley 2007: 9) will not be discussed here. Only two short sections (2.2.4 on artistic languages and 2.2.5 on “modern glossopoeia”) are devoted, respectively, to some major literary creations as well as to the newer schemes found on the Internet. Thus, the languages excluded are those created consciously but “for fun” (“bewusste, fast spielerische”), such as children’s language games or poetic creations, as well as unconscious glossolalias (“unbewusste Zwangsreden”; both quotations from Bausani 1970: 79).

Bausani (1970: 38ff.) argues that most pre-Renaissance artificial languages were mystic or ceremonial attempts at substitutions and additions into ethnic languages (“primitive languages”). He lists five types: taboo, magic, secret, glossolalias and ceremonial. Thus, although mentioning two languages before the Renaissance and briefly describing the Renaissance itself, Bausani devotes most attention to languages from the 17th century onwards. Having in mind the purposes of artistic creations (and sometimes lack of material for analysis), many assume that the earliest artificial language is *Lingua Ignota* from the 12th c., created by St. Hildegard of Bingen (see Bartlett 2009; Large 1985: 3; Okrent 2010: 10). Higley (2007) seems to believe *Lingua Ignota* to be the first **consciously created** and **systematic** invented language. Bausani (1970: 76f.) also claims that the language cannot be classified as a simple glossolalia. As the second more developed artificial language both Bausani (1970: 89f.) and Higley (2007: 66) indicate *Bālaibalan* from around the turn of the 15th and 16th c. It is worth mentioning that Bausani actually calls *Bālaibalan* “the first real language invented in an educated milieu” (1970: 83).

Those two languages are the only widely known languages that have been consciously created before the 17th century. Other examples from the time include glossolalias, ceremonial formulas and gibberish of mystic works (Bausani 1970: 44–71; Higley 2007; Large 1985: 3).

2.2.1 Universal languages

The seventeenth century brought the advent of universal schemes. These were not only spoken languages but also pasygraphies, gestural as well as musical languages. The need arose for various reasons, some very old ones and some newly discovered.

The religious conflicts of the 16th c. led to the slow rise of the vernaculars. Later on, local languages grew in impor-

tance in printing and in education, which resulted in the secondary babelisation of the scholarly world in the second half of the 17th c. (in 1687 Newton published *Principia Mathematica* in Latin, but by 1704 the language of *Opticks* was English). Latin, which was the sole language of education and the only language taught, experienced a decline, and there was no other language to replace it anytime soon (French rose to importance in the last decades of the century).

The failure of Latin seems surprising, but at the same time, it can be easily explained. It took too long to learn properly and the results were not satisfactory: scholars could read and write it, but weren't able to communicate fluently face to face because the differences in pronunciation rendered any conversation almost impossible (Large 1985: 7). A common view was that the language had been somehow "corrupted" through the centuries of use. And not in spite of, but because of the many attempts to "resurrect" the classical version, Latin remained dead. As Knowlson (1975: 28) noticed, "[...] the Renaissance insistence upon a return to pure classical Latin meant that this language tended to appear to many to be increasingly unsuited to modern needs". Interestingly, such problems never occurred in the Islamic world, because there was always only one standard Literary Arabic, which made any universal language dispensable (Bausani 1970: 92f.). Nevertheless, there was a need of communication between the East and the West. More and more merchants travelled to the far reaches of Asia. They did not speak Latin, which anyway was of no use to them in China or India (Large 1985: 6–8).

The discoveries made in the Far East and in Egypt aroused interest in pasygraphies, i.e. universal scripts. The belief that one symbol meant exactly one idea or represented precisely one object seemed very attractive. "Analogies were found in musical notes, Arabic numerals, Chinese ideograms and Egyptian hieroglyphics, all of which it was believed provided a representation of concepts which could

be understood regardless of language” (Large 1985: 11; see also Slaughter 1982: 85). Many believed quite wrongly that Egyptian hieroglyphics and Chinese characters were not representations of words or phrases but mirrored reality as it is; the Egyptian symbols bore, however, some resemblance to the objects they signified, whereas Chinese ideograms were a matter of convention (see Eco 1997: 160, 212f.; Knowlson 1975: 15f.; Large 1985: 12f.).

The confusion of languages brought to attention the biblical myth of Babel and the beginnings of language itself. The Bible described how the first man named all the things in the world according to what they truly were, which would mean that being an *onomatopete* ‘name giver’, he was also a *nomothete* ‘law giver’ (Higley 2007: 6; also Eco 1997: 8). The scholars of the 17th century tried to handle the disarray by attempting to find the original language of Adam. Claims were made that this language was Hebrew (Eco 1997: 80–85), and some even argued that it was some Germanic language, e.g. Flemish as claimed by Goropius Becanus already in 1569 or Swedish by Georg Stiernhielm in 1671; Olaus Rudbeck maintained that Swedish was the source of all Germanic languages in 1675 (Eco 1997: 96–98; Świączkowska 1998: 43–45).

The interest in the mystic powers of language realised itself in occult and secret works connecting the kabbalah (especially the art of *gematria*; stemming from the fact that each of the Hebrew letters had a numerical and symbolic value) with the art of cryptography and shorthand (particularly two works of Trithemius: *Steganographia* written ca. 1499 and *Polygraphia* from 1508, both published posthumously, respectively 1606 and 1518) as well as with Egyptian hieroglyphs.

At the same time fewer and fewer scholars believed it was Hebrew that was the first language in the world and decided – through comparing living languages – to find the true, original “language of Adam”. Before the comparative

method came into use, such attempts were visible in the works of Leibniz, who tried to compare different languages in order to find their common root and, through that, the oldest possible form of language (Świączkowska 1998: 49).

THE BEGINNINGS: PASYGRAPHIES AND NUMERICAL LANGUAGES

All these reasons brought about the mass creation of all kinds of schemes. The oldest ones drew on the works of Ramon Lull (13th/14th c.) and Trithemius. Many were inspired by Johannes Becher, who designed one of the best known pasygraphies published in 1661 – *Character pro notitia universali linguarum* (Bausani 1970: 99f.). He was also the inspiration for his contemporary Athanasius Kircher, who wrote *Polygraphia nova et universalis* (1663, the title being a clear reference to Trithemius). Both these works, as well as the earlier *Universal Character* (1657) by Cave Beck utilised the idea of creating a universal system based on Arabic numerals (Bausani 1970: 99f.; Knowlson 1975: 21).

Beck compiled a dictionary of Latin words and assigned Arabic numerals to every one of them. He added letters indicating grammatical forms (e.g. *p* meant the personal pronoun and *pf* its feminine form) and later syllables for specific numbers, which would make it possible to pronounce the words. However, this would mean that the user would have to remember every entry of the dictionary and the number given to it (Eco 1997: 201).

Becher's system was even more complicated. It employed numbers for words (10,000 entries in the dictionary) as well as for grammatical endings. For fear that the Arabic numbers were not known to everyone he also devised a system of graphical representations for them – a hardly legible one (Eco 1997: 201f.).

However, the earliest attempts to devise a pasygraphy of such kind, of which there is not much detail known, are the following four (Knowlson 1975: 44–48):

- 1627, a prospectus for “*escriture universelle*” by Jean Douet (the only one of the four to be published);

- 1628, in a letter from Kinder to Beveridge, the scheme of a Monsieur Champagnolle is mentioned; little is known about this scheme – the one presented in the letter is most likely Kinder’s own scheme;
- 1629, a project of an unknown author criticised in a letter from Descartes to Mersenne.

An important addition to the list was made by Francis Lodwick (spellings vary; Lodowyck, Lodowick) in 1647. His pamphlet *A Common Writing* contained a universal character (a kind of shorthand), where a great number of words could be derived from a limited number of roots (radicals) by using a set of grammatical marks in the form of strokes and hooks. Undoubtedly, this was an original idea and a great contribution to semantics. Moreover, he based his classification not on nouns (as it was usually done at that time), but on verbs (Eco 1997: 260f.). The scheme was never successful, though, as it took English as its base and chose the radicals rather subjectively, not really being universal. Memorising the list of radicals and the set of additional symbols seemed to be a challenge. Lodwick was well aware of the weaknesses of his scheme and therefore settled on devising another project, namely *The Ground-Work, Or Foundation Laid, (or So Intended) For the Framing of a New Perfect Language: And an Universall or Common Writing* (1652). Consequently, the project was a revised and much simpler version of the previous one. Its main topic is lexical derivation; Lodwick proposes that a language should be composed of monosyllabic radicals. Several years later he discussed “the natural order of things”, i.e. possible ways of rational classifications of concepts in his *Of an universall reall character* (preserved in manuscript). The work again followed the Aristotelian tradition (Salmon 1972: 106).

1652 was also the year of publishing a collection of miscellanea by Sir Thomas Urquhart entitled *Ekskubalauron* (supposedly Greek for “gold from a dung hill”; Large 1985:19;

Okrent 2010: 28). The collection included a proposal for a universal language, which Urquhart never completed. 1653 brought another creation of Urquhart's, namely *Logopandectision*. However, none of his works included a full description of the languages he praised in his first work.

PHILOSOPHICAL LANGUAGES

The following years provided a whole collection of other *a priori* projects: *Lingua Philosophica* from *Ars Signorum* by G. Dalgarno (1661), an unnamed project from *De Arte Combinatoria* by G. W. Leibniz (1666), *Panglottie* by Comenius (1665-1666) and the Philosophical Language from *An essay towards a real character* by J. Wilkins (1668).

Dalgarno's scheme was actually a dictionary of allegedly logically organised ideas. He broke down more complicated concepts into basic elements for which he developed a polygraphy (Okrent 2010: 47-49). There were 17 classes of concepts, each marked with a letter (Bausani 1970: 99). Dalgarno reduced both nouns and verbs to a single class of radicals. At first, his idea was to attach grammatical particles to lexical radicals (as was common at that time) but later he decided to eliminate this distinction and derive particles from root radicals. Thus, a verb would consist of a root radical and a modified radical denoting an activity (Cram 1992: 197). Some of the classification schemes for the project were to be prepared by Wilkins. Dalgarno was not happy with the result of Wilkins's work and decided not to include the tables in his system. Unlike Wilkins, he did not wish to develop a character for every possible notion, but rather a neat and concise system of radicals and a simple grammar allowing for combinatory expressions. This disagreement, and the apparent inspiration coming from the works of Comenius, pushed Wilkins to write his *Real Character* in 1668.

Comenius, known as a teacher and educator, wrote his work touching this topic, *Via lucis*, around 1641/1642. In it, he merely outlined the principles of a new universal language – it

was to be based on nouns and to mirror the natural order of things. The work circulated in manuscript until published in 1668 (dates differ¹⁸). A more detailed proposal called *Panglottie* was described in the fifth volume of his *De rerum humanarum emendatione consultatio catholica* (probably 1665-1666¹⁹). Still the noun was the basis for derivation. However, in the same collection of manuscripts (unfinished and only printed in 1966) can be also found the *Novae linguae harmonicae tentamen primum* in which Comenius turned his attention to verbs rather than nouns (Salmon 1972: 40, 124).

At the same time when Comenius began working on *Via lucis*, another scholar, the aforementioned John Wilkins, published *Mercury, or the Secret and Swift Messenger* (1641). The work dealt more with cryptography than with universal language and contained comprehensive discussion of various modes of communication. In one chapter Wilkins considered musical notes, Arabic numerals and Chinese script, among others, as systems which could contribute to creating a basic language for all humanity consisting of ideas and notions rather than words (Large 1985: 28f.; Salmon 1972: 15).

Wilkins's scheme of 1668 seems to be the most known and quoted (and probably the most comprehensive), although he would not have published his project but for the earlier attempts by Dalgarno and Comenius. Wilkins was largely inspired by Chinese characters and their apparent direct connection to objects without mediation through words. However, he believed that there were too many of them to be easily memorised. Therefore, he wanted to create a language which could be presented in the form of such characters (but much simpler) on the one hand and as a philosophical system that would classify all the existing things on the other. Wilkins

¹⁸ Thus, Eco (1997: 215), Knowlson (1975: 88) and Salmon (1972: 26); Large (1985: 8) gives 1642.

¹⁹ These dates given by Duličenko (1990: 37) and Skalická (2005: 4f.). In Okrent (2010: 299) the date is 1665.

believed that a word would directly reveal its meaning if it was built from letters corresponding to categories. The notions were organised in a binary tree of genera and species (taking origin in the Porphyrian tree). Each class was marked with (a) specific (Greek or Latin) letter(s). The word for 'dog', for example, would be *zita*, that is, *zi* standing for category XVIII (beasts), *t* for subcategory V (oblong-headed) and *a* for sub-subcategory 1 (bigger kind); furthermore it could be written as a character composed of strokes assigned to specific categories (Okrent 2010: 51f.). It is worth noting here that both Dalgarno and Wilkins made use of "the art of memory" found in the works of Ramon Lull.

Another important contribution to the history of universal languages was made by G. Leibniz. The contribution was two-fold: firstly, Leibniz wanted to uncover the oldest human language; secondly, he tried to find the way to mathematise human thought. His ideas were strongly connected to the, popular at the time, search for the language of Adam (Święczkowska 1998: 9). It seems that Leibniz never believed that it was possible to find such a language, but he was a supporter of the idea that all languages in the world have indeed some common roots. He presented a plan for classifying all languages according to grammatical and lexical similarities in order to show their mutual relations and thus find the common denominator. He was clearly aware that such a task would prove almost impossible and therefore proposed that several expert groups should be appointed to classify and compare groups of languages (Święczkowska 1998: 44–46).

Leibniz also believed that human thought could be rationalised. This could only be achieved by providing the people of the whole world with a logical language with which everyone could operate, as with numbers in mathematics. For this reason, in 1666, Leibniz published *De Arte Combinatoria*. He assumed that all notions could be broken down into primitive ideas or atoms (in accordance with the Aristotelian tradition) which could be combined and calculated (Salmon 1972: 40f.;

Święczkowska 1998: 131ff.). He later substantially developed his ideas from a simple classification of thoughts into a project of a universal language in 1678, where the digits were represented by subsequent consonants and the decimal units by five vowels and diphthongs (Large 1985: 40; Święczkowska 1998: 138). The idea of representing numbers as sounds is apparently adopted from Lodwick's *Ground-work* (Salmon 1972: 41). Finally, he developed his *calculus ratiocinator*. He also considered creating a musical language consisting of notes and tones, and simplifying Latin (more in Święczkowska 1998: 138–145).

1666, the year of publishing of young Leibniz's *De Arte Combinatoria*, was also the year of issuing the first truly deliberately created *a posteriori* language, that is, Ruski Jezik by J. Križanić (Duličenko 1990: 38ff.; Okrent 2010: 299).

GESTURE AS A UNIVERSAL LANGUAGE

The idea that gestures could be a universal language is not a new one. Gestural communication has been mentioned in many works, i.a. by Plato (in the *Cratylus*), Saint Augustine, Rabelais, Descartes, Francis Bacon and John Wilkins. (Knowlson 1975: 215–217). However, it was only in the 17th and 18th c. that scholars took a serious interest in gesture both as an educational tool and a universal language, even though the prevailing conviction was that spoken, as well as sign languages are arbitrary and require learning, with only basic signs being common to all men. The first successful English teacher to the deaf, John Wallis, recognised the importance of gestures for bridging the initial gap between teacher and deaf pupils. The idea that gestures might form a complex universal language appeared for the first time at the beginning of the 17th c. in Giovanni Bonifacio's *L'Arte de' cenni*, in which he showed a wide array of oratory gestures and the ideas which they expressed. The inspiration most probably came from the gestures of Renaissance rhetoric. In *Chirologia, or the Naturall Language of the Hand* (1644) John Bulwer maintained that gesture was the language of Adam and that it was better suited for

international communication than spoken languages, being “more striking, speedier to use and more natural; [it] could therefore be universally understood without being learned or translated” (Large 1985: 54f.).

Sign language as a universal language and teaching the deaf were the topics of *Institution des sourds et muets, par la voie des signes méthodiques; ouvrage qui contient le projet d'une langue universelle, par l'entremise des signes naturels assujettis à une méthode* by the Abbé de l'Épée (1776). He believed that the signs by which the deaf communicated had to be perfected by introducing linguistic rules. The system invented by de l'Épée was based on signs naturally evolved among the deaf of Paris and invented ones he agreed upon with his pupils. The signs were systematised and a set of rules and grammatical markers mimicking spoken French were introduced. The Abbé de l'Épée believed that abstract ideas should be analysed into natural iconic signs – thus they would retain naturalness and yet be complex and systematic (Knowlson 1975: 218). His ideas were similar to those of Leibniz – the universal language was to be expressed through combinations of signs, which would be simple like arithmetic symbols, rather than complex like the arbitrarily developed Chinese characters. Those latter ones, according to de l'Épée, do not bear a natural resemblance to what they signify.

After de l'Épée's death in 1789, the institution founded by him was run by his former student, the Abbé de Sicard. Sicard put the emphasis more on the practical side of the endeavour and published a thematic dictionary of signs and their grammatical relations. The system found little support as an attempt at a universal language. It was too complex and tedious to be adopted universally; complex signs were often abbreviated, thus losing the natural analogies with real-world objects. Another accusation it faced was that it lacked script. Thus, practical reasons prevailing, de l'Épée's system continued as a means to communicate with the deaf. It gave rise to what is now known as Signed French, i.e. manually coded spoken

French (later followed by other signed versions of oral languages). As there was no institution to perpetuate naturally evolved sign language, almost up until this day the deaf were forced to learn an artificially created sign language.

WANING INTEREST

The *a priori* languages continued until modern times, although they were created less often. After the outbreak of such schemes in the 17th century, the next century brought only half a dozen. As Knowlson (1975: 139) put it: “interest in the construction of an artificial universal scheme [...] appears to have flagged considerably, and only the occasional, unexciting language scheme was produced at that time”. The eighteenth century brought forth French as the language of international communication. This was probably the main reason for the decline of interest in universal schemes. However, in the second half of the century some scholars attempted to revive the ideas of Leibniz by applying symbolic notation to logic (see Knowlson 1975: 141). The two best known schemes – *Langue Universelle* by Delormel and *Pasigraphie* by de Maimieux were a product of the last decade of the 18th century.

The scheme by Delormel (1795) was based – similarly to those of Dalgarno and Wilkins – on a classificatory system, where the basic elements were syllables. His script was a modified Latin alphabet with ten vowels and twenty consonants, reflecting the decimal system (although no pronunciation was provided), and some ambiguous letters for consonants were eliminated (LARGE 1985: 46).

The scheme by Delormel (1797) – vapid and unoriginal – was much less impressive than that of de Maimieux, which gained much attention and is reported not only to have been taught in schools in Germany and France, but was also admired by Napoleon himself (Knowlson 1975: 155; Okrent 2010: 80). The term ‘pasygraphy’ was used for the first time in the title of de Maimieux’s book (Duličenko 1990: 21). This indicates that the scheme was first planned to be used

only in writing. Only later did de Maimieux create the oral version, which he called ‘pasilalie’. The system consisted of 12 characters and 12 regular rules, along with some additional dots for grammatical relations (Knowlson 1975: 155–156; Large 1985: 47). The symbols were ordered in three tables (of 3-, 4- and 5-character words) and the characters themselves served to indicate the exact position of the word sought.

The nineteenth century brought several more *a priori* systems, i.a. *Lengua Universal y Filosofica* (B. Sotos Ochando, 1852), *Solresol* (J. F. Sudre, 1866), *Lingua Lumina* (J. W. Dyer, 1875), *Blaia Zimondal* (C. Meriggi, 1884), *Chabé-Aban* (E. Maldant, 1886), and *Lingue Universelle* (C. L. A. Letellier, dates differ²⁰). Only one of those systems, *Solresol*, was successful. The others were rather peculiar schemes not suited to the needs of human communication. The language of Sotos Ochando was a typical “universal” scheme with letters assigned to specific notions and a set of affixes. *Lingua Lumina* had 33 diphthongs, *Blaia Zimondal* was based on an (alleged) onomatopoeic similarity to things, *Chabé-Aban* included inflected articles and *Lingue Universelle* comprised a division, typical of this kind of scheme, into 10 categories with subdivisions, finally giving 100,000 species, and imitated French grammar with double negation (Rónai 1969: 26–38).

In its time, *Solresol* attracted great attention and won several awards (1855 at the Universal Exhibition in Paris and 1862 at the London Exhibition). J. F. Sudre invented the language in 1817 and worked on it until his death in 1862. The project was published posthumously in 1866 as *Lingue Musicale Universelle*. It was based on the idea that the *Solfeggio* (or *Solfège*) consisted of seven universally recognised notes

²⁰ 1852-1886 according to Duličenko (1990: 108); published in four volumes between 1852 and 1855 according to Large (1985: 60) and published in 1886 according to Okrent (2010: 301); Eco (1997: 306) gives 1832-1855.

with seven corresponding syllables. The language could also be represented using seven Arabic numbers, colours, specially devised stenographic characters, sung, played or shown with fingers. There was no philosophical or logical classification behind the project, only the belief that music is universal. The system was a combination of musical notes and each note signified the class to which each four- or five-note combination belonged. Single notes expressed words such as ‘yes’ and ‘no’ (respectively *si* and *do*) and two-note words formed pronouns. Yet, the system proved rather tedious – grammatical categories were to be distinguished by positioning an accent over a syllable (*sirelasi* ‘to constitute’, *sîrelasi* ‘constitution’) and the opposites were made by inversion (*misol* ‘good’, *solmi* ‘evil’), which counteracted the idea of the first note representing the encyclopaedic class. The number system was hexadecimal and the language included some portion of Gallicisms (Eco 1997: 305–306; Large 1985: 61–63; Rónai 1969: 29–30). It would also be very difficult to distinguish between such phrases as *famisi domido* ‘to carry the universe’ and *fami sidomido* ‘this place’ (Bausani 1970: 112).

Some more or less known schemes sprung up even in the 20th c. In 1921, Thiemer created his numerical language Timerio, where, for example, the phrase ‘I love you’ is written as 1-80-17. Ro, by E. P. Foster, published in 1908,²¹ strongly resembles the previous 17th c. schemes based on the classification of ideas. Both in Ro and in Universel (this name in Okrent 2010 and Duličenko 1990, while Rónai (1969) gives Universal) of A. J. Decormis (1948), as well as in Babm of F. Okamoto (1962) a single letter denoted a class of ideas. All things were “logically” categorised and similar objects were given similar names, e.g. in Babm *bomb* signified cattle in general, *bomd* the pig, *bomf* the sheep, *bomg* the horse etc. Therefore, as Rónai (1969: 40) wrote:

²¹ This date in Okrent (2010: 304), 1906 in Bausani (1970: 136), 1906-1908 in Duličenko (1990: 177).

[I]ogisch mag das sein, aber es erschwert die Einprägung, ganz davon abgesehen, dass die Beschreibung eines Zoobesuchs fürchterlich monoton sein müsste.

it might be logical, but it complicates the memorisation, quite apart from the fact that the description of a visit to the zoo would be terribly monotonous.

It is interesting that *a priori* languages continued to appear even after their constant failures, and the relative successes of *a posteriori* languages in the 19th and particularly in the 20th century. It seems that their authors in modern times were very often inspired not only by formal languages, but also by various attempts to formalise the description of natural languages and to find the universals in them.

2.2.2 Formal languages

The history of formal languages is closely connected to that of universal languages. Many of the ideas found in Aristotle's works are not only found in universal languages (the Porphyrian tree) but also constitute the foundations of predicate logic. Another contributor to both universal and formal languages was Ramon Lull. He was a major influence on Kircher, Wilkins and Dalgarno (the art of memory) as well as on Giordano Bruno and even Leibniz (Dutilh Novaes 2012: 68; Knowlson 1975: 84f.). However, it was not until the 17th c. that the "notational explosion" took place (on Descartes and Leibniz see more in Dutilh Novaes 2012: 76–81). Both the search for the perfect language and the increased interest in mathematics and logic stemmed from the intellectual climate of the time – the need to push the human mind onto new tracks and to mathematise it to answer the needs of the new era of reason.

The modern era of formal languages begins with Boolean algebra (the first symbolic system with strict rules of formation), Hilbert's programme (formalisation of all mathematics)

and Gödel's incompleteness theorems (1. in no consistent system it is possible to prove its completeness, 2. in no consistent system it is possible to prove its consistency). I am not going to discuss the languages of mathematics and logic further – the reader may refer to Dutilh Novaes 2012 for a short survey of formal languages or Murawski 1995 for the philosophy of mathematics. The important reason for naming the latter two scholars, though, is that they heavily inspired the development of programming languages, which will be presented in a short while. Firstly, however, it needs to be stated that the opposition artificial – natural has its counterpart, i.e. formalised – non-formalised.

Not all artificial languages are formal, obviously (e.g. IALs). Moreover, not all systems called formal truly are so. Marcus (2004: 15) asserts that

if the propositional calculus and the predicate calculus are purely formal languages, we cannot make a similar claim for the language of mathematics (which has a mixed structure, with a natural and an artificial component, the latter being only partially formalized) or for the computer programming languages, sharing features with both natural and artificial languages and with both formal and non-formal languages.

'Formal language' is usually defined as a finite set of operations over a finite non-empty alphabet. However, the predicate calculus and programming languages are infinite "in their general competence" (Marcus 2004: 16). The features of programming languages which they share with natural languages are also, among others, idiomaticity (although more iconic or indexical in nature), ambiguity and being their own metalanguage. Having this in mind, we may proceed to the history of programming languages.

The first and second generation were low-level machine languages (sequences of binary numbers; also known as 1GLs) and assembly languages (shorthand codes; 2GLs) not requiring compilers and interpreters, that is, similar to notation systems,

sets of instructions issued directly to the hardware. From the third generation onwards (the fifth one began in the '80s and continues) low-level languages have been replaced by high-level languages.

O'Regan (2012: 124) cites Plankalkül by Konrad Zuse as the earliest high-level language. Developed in the 1940s and published in 1948, it relied heavily on arithmetical and algebraic notation. However, much better known are FORTRAN (1957, IBM) and COBOL (1959, CODASYL based on Grace Hopper's work). These two languages remain in usage until this day.

The beginning of a new era within 3GLs is marked by the development of ALGOL, the first structured programming language (as opposed to so-called "spaghetti coding", i.e. programming lacking internal structure; Baron 1994). Its first version was published in 1958, but it is the 1960 version that is of most interest. John Backus and Peter Naur developed a metalanguage for describing ALGOL 58's syntax and first implemented it in ALGOL 60. This metalanguage corresponds to Chomskian context-free grammars (type 2). Chomsky developed his ideas to mathematise the description of natural languages; however, both his grammar and the hierarchy of grammars are also applicable to programming languages, as shown by Ginsburg and Rice in 1961 (Marcus 2004: 19). ALGOL 60 was the first language to be formalised this way. It gave rise to such programming languages as Pascal and C, which have developed into their more modern offspring of today.

The next revolution was the shift from procedural programming to object-oriented programming. Sets of procedures were substituted by classes with properties, of which instances were created. This allowed for optimising the code and facilitating writing and debugging of programs. Significant languages in this paradigm include C++, C#, Java, Perl and Ruby.

To the formalised languages also belong some auxiliary languages. In 1960, Hans Freudenthal published Lincos, a language created to communicate with extra-terrestrial intel-

ligence. He proposed using radio wavelengths to send out messages based on arithmetic and logic (Okrent 2010: 212f.; Sakaguchi 1998: 143f.). This system more closely recalls an *a priori* universal scheme than an actual language.

Visibly, formal languages do not constitute a homogenous group. The degree of artificiality and formalisation varies strongly. In fact, the only truly formalised language may be the language of logic, whereas those of mathematics and programming share a great deal with natural languages.

2.2.3 International auxiliary languages

The difference between universal languages and international auxiliary languages lies not only in their philosophical assumptions but also in the structure itself. The universal languages created for practical, and at the same time for philosophical reasons to facilitate communication, but most of all logical thinking, were *a priori* schemes constructed of invented and subjectively chosen elements. Auxiliary languages, as the name indicates, were all languages built to help people communicate, sometimes locally, sometimes globally. They were *a posteriori* systems, which means they were constructed on the basis of already existing languages.

Albani & Buonarotti note in the entry ‘Ausiliaria internazionale, lingua’ (1994: 49) that there are three requirements that have to be met by an IAL to be successful, all declared by the Délégation pour l'Adoption d'une Langue Auxiliaire Internationale, an association created by L. Couturat and L. Leau in 1901:

Una lingua A[usiliaria] deve soddisfare le seguenti condizioni:
1) essere capace di servire alle relazioni abituali della vita sociale, agli scambi commerciali e ai rapporti scientifici e filosofici; 2) essere di facile acquisizione per tutte le persone d'istruzione elementare media e in particolare per le persone di civilizzazione europea; 3) non essere una delle lingue nazionali.

An auxiliary language must satisfy the following conditions: 1) be capable of being used for the ordinary relations of social life, for commercial and for scientific and philosophical dealings; 2) be easily acquired by every person of average elementary education and in particular by people of European civilization; 3) not be one of the national languages.

Such conditions, although stated only in 1901, were implicit in most of the schemes. Only a few *a posteriori* languages were not meant to satisfy the European (or generally Western) needs and not all of them were created to serve as international helping systems.

The first *a posteriori* language is most likely Bālaibalan, created in the 16th c. Its lexicon and grammar are of Arabic, Turkish and Persian origin. However, this language was not meant for auxiliary purposes, but rather for secret communication within the movement in which it came to being (the Hurufism, see Bausani 1970: 82ff.). Some other minor *a posteriori* schemes are accounted for in Duličenko 1990.

1666 was an important year 1666, not only for the history of universal and formal languages, being the year of the publishing of *De Arte Combinatoria* by Leibniz, but also as the year of the first true *a posteriori* auxiliary scheme – Ruski Jezik by J. Križanić (Duličenko 1990: 38ff.; Okrent 2010: 299). A pan-slavic system, written both in Cyrillic and in Latin script, was the language of several books and treatises by Križanić (Duličenko 2006: 6).

The 18th c. brought only two major *a posteriori* projects. One of them is Scriptura Oecumenica. Written in 1732, it was created by an unknown German scholar using the pseudonym Carpophrophilus. The scheme was in fact regularised Latin with no synonymous word forms and the declination replaced with four articles (Bausani 1970: 109).

The second famous 18th c. scheme is one of the most interesting attempts to sketch the principles of creation of an IAL – a four page long conclusion to the entry ‘Langue’ (1765) by J. Faiguet in the famous *Encyclopédie*. He modelled

his Langue Nouvelle on French, using French word roots and greatly simplifying the grammar (Eco 1997: 294). The scheme included some *a priori* elements, such as numerals and tense endings. Some have noted that these endings are identical to those of Esperanto (Bausani 1970: 109; Large 1985: 51f.).

But the true epoch of IALs began in the second half of the 19th c. In 1890,²² A. Liptay, the creator of Langue Catholique/Lengua Catolica stated that a language for all should be discovered and not invented (Rónai 1969: 70).

VOLAPÜK AND OTHER MIXED SYSTEMS

The most prominent of all the early IALs is definitely Volapük (1879). Created by a German priest J. M. Schleyer, it soon gained popularity and became spoken by hundreds of enthusiasts. The system, however, was a complicated one. Not only did it include umlauts, but it also incorporated a multitude of endings giving a total of 505,440 verbal forms (Bausani 1970: 114; Large 1985: 68). On the one hand, Schleyer wanted his language to be easy to learn, even for the Chinese, and therefore excluded the ‘r’ sound, and reduced word roots to be only one syllable long. On the other hand, he imagined it being able to express every nuance possible and thus devised a very complex grammatical system. Very soon, the language became subject to numerous amendments and improvements which made Schleyer furious. He resisted any changes, which led to further animosities and finally to people turning away from Volapük and joining Esperanto. In defiance of Schleyer, there appeared reformed versions of the language or its outright opposites (Eco 1997: 319–320; Okrent 2010: 106–107), such as Nal Bino (S. Verheggen, 1886), Balta (E. Dormoy, 1887), Bopal (S. de Max, 1887), Spelin, (J. Bauer, 1888), Dil (J. Fieweger, 1893), Veltparl (W. von Arnim, 1896), the famous Idiom Neutral (W. K. Rosenberger, 1902) and Spokil

²² Rónai (1969: 67) quotes the 1892 edition, while Duličenko (1990: 123) gives as the date of creation 1890.

(A. Nicolas, 1904). Another reason for abandoning Volapük was probably the language's *a priori* system of affixes and unintelligibly transformed words. Because 'r' and multiple syllable roots were to be avoided, the German word *Berg* 'mountain' became *bel*, and the English *brother* – *blod*. The most striking example is the word for iron, *lel*. It is said to be derived from the Latin *ferro*, a word that contains the forbidden 'r'. If it is omitted, the Volapük transcription gives *fel*. However, this word is already taken (for 'field') and so are *fil*, *fol*, *ful* with the vowel changed, as well as *gel*, *hel*, *jel* and *kel*, with the onset consonant changed. Therefore, the nearest solution is the disguised *lel* (Rónai 1969: 43; cf. *Schere* 'scissors' and *jil* in Large 1985: 68). Although the vocabulary of Volapük was of European origin (mostly modelled on English) and thus *a posteriori*, the grammar contained many *a priori* elements. Hence, the language is usually classified as mixed, and not purely *a posteriori*.

Until now, only about a dozen mixed languages have been designed. Some of them, like Bolak/Langue Bleue (L. Bolack, 1899) or Qôsmianî (W. M. L. Beatty, 1922), are typical projects with an *a posteriori* lexicon changed beyond recognition and *a priori* grammar and some peculiarities. For example, Qôsmianî violates Universal 481, requiring the mood suffix to stand before the progressive suffix (Libert 2013), while Langue Bleue (the name coming from the colour of the sky, which covers the whole world) uses *a*, *o*, *e*, *i* as prefixes showing gradation (respectively *not at all*, *a little*, *much*, *very much*) and the whole concept is illustrated in the form of a thermometer (Rónai 1969: 57, 79). Some projects are, however, notable: Loglan (J. Cooke Brown, 1955) and its offshoot Lojban (Logical Language Group, 1989). Although Loglan and Lojban are primarily classified both by Sakaguchi (1998: 104, 129) and Libert (2000) as *a priori*, it is important to note that their vocabulary is statistically derived from natural languages. Libert remarks that

while, as noted, the term ‘a priori’ is often taken to apply mainly or only to vocabulary, in Loglan/Lojban the vocabulary and sound system are largely a posteriori and it is the grammar which is of the a priori sort (2000: 7).

Loglan began as a project based on propositional logic intended to test the Whorfian hypothesis (detailed description in Sakaguchi 1998: 132–137). In the late 80s, a split in the movement occurred (based on similar premises as the disputes in the Volapük movement) and gave rise to Lojban. On one hand, Lojban is based on the same principles as Loglan (i.e. logicity) but on the other, it incorporates syntactical and morphological features of various natural (and non-natural, e.g. Láadan) languages to form a culturally neutral language capable of maximal expression. Okrent (2010: 249) rightly indicates that the validity of emotional markers cannot be evaluated within the scope of formal logic. Thus, uniting the principles of formal languages with properties present in ethnic languages, both Loglan and Lojban may be assessed as mixed.

A *POSTERIORI* LANGUAGES: EARLY CREATIONS AND ESPERANTO

Apart from several mixed systems, most IALs were of the *a posteriori* type. Okrent (2010) lists close to 150 such languages from Ruski Jezik (1666) to Guosa (1981). At least fifteen of them are reforms of Esperanto and a number of Ido, itself being a reform of Esperanto. Several projects are based on Latin as the first auxiliary language still in use among scientists (most notably Latino sine Flexione by G. Peano, 1903). Some derive their vocabulary and/or grammatical systems from a specific language group, e.g. from Germanic – Tutonish (E. Molee, 1902), Alteutonik (E. Molee, 1915) and Slavic – Slavski Jezik (B. Holý, 1920), Mezhduslavjanski Jezik (L. Podmele, 1958). There was a surge of projects claiming to be universal, with names suggesting that the language could be used by the whole world, e.g. Universalglot (J. Pirro, 1868), Mundolingue (J. Lott, 1890), Mondi Lingua (A. Lavagnini, 1955),

Mondial (H. Heimer, 1943), Lingua Komun (F. Kürschner, 1900) or Komun (F. Musil, 1946). In spite of the pompous slogans, few of the authors based their systems on languages other than European ones (e.g. El-Afrihili written by a Ghana historian K. Kumi Attobrah in 1970 with vocabulary coming mostly from Swahili, or Guosa by A. Igbinewka from 1981 with the origins of its lexicon in Niger-Congo and Chadic language families). Some authors were surprisingly prolific: R. de Saussure (brother of the famous linguist) is known to have constructed at least 8 languages and P. Stojan wrote 9 schemes in 15 years (Okrent 2010: 144, 297).

Nevertheless, the biggest rival of Volapük and the only truly internationally employed IAL turned out to be a language whose name now lays no claims to universality: Esperanto ('hopeful'). L. L. Zamenhof, the father of Esperanto, wrote in his *Unua Libro (International Language: Introduction and Complete Grammar)* called by Esperantists 'the First Book' that for a language to be global it is not enough to call it so (Zamenhof 1887). A Polish Jew born in Bialystok (then in the Russian Partition of Poland) brought up in a multilingual, multinational and multicultural environment, Zamenhof since childhood was aware of the problems that mushroomed in such circumstances. His reasons for the creation of a new auxiliary language were purely humanitarian. He spoke and understood several languages, which – although he was no linguist by profession – helped him conceive his first project when he was about 15. Yet, because of his father's lack of enthusiasm (to say the least, cf. Rónai 1969: 54) towards the project, young Ludwik had to abandon his work on Lingvo Internacia during his studies in Moscow. When he returned to Warsaw (where his family had lived for several years), he started over. At last, using his wife's dowry, Zamenhof – under the pseudonym Dr Esperanto – published the first book on his international language in 1887 (Duličenko 2006: 98–100; Large 1985: 71–72; Okrent 2010: 94–98).

Esperanto spread rapidly, despite the popularity of Volapük. In fact, Zamenhof himself for a short time wanted to give up on his project, knowing that Schleyer's language was held in so high esteem. However, he soon discovered that Volapük was hard to learn, and that its author was averse to any petitions for changes. Zamenhof in turn thought that any auxiliary language should be easy to learn and use (and therefore he regularised the grammar) and that after publication it belonged not to the author but to the users. *La Unua Libro* first published in Russian, soon after translated into Polish, French and German, and later into English, contained an introduction, the 16 rules of grammar, ca. 900 word roots and some texts in Esperanto. In 1889, Zamenhof compiled the so-called *Adresaro* – a list of 1,000 Esperantists who had translated into Esperanto. This enabled enthusiasts to communicate. The same year brought the first Esperanto journal *La Esperantisto* issued by the first Esperanto Society (Nuremberg, 1888) as well as the third and at the same time the last Volapük congress (Large 1985: 73; Okrent 2010: 106–107). The first Esperanto congress took place in 1905. In the same year Zamenhof published the *Fundamento de Esperanto* with the 16 rules from 1887, a mini-dictionary from 1894 and example sentences, also from 1894. The *Fundamento* is now called “*netuŝebla*” (‘untouchable’; declared as such in 1905) among Esperantists as they have agreed never to change the original 16 rules. Zamenhof himself declared at the first congress that he renounced all rights to the language and wanted it to develop spontaneously by being used, but also that the *Fundamento* should be the basis for everyone (Large 1985: 76–77). This, along with the statement that Esperanto should always be neutral, was the content of the Declaration on Esperanto, a text that is almost always read at Esperantic meetings as the fundamental document of the movement.

Since then the language has suffered some defeats, but is still used in an almost unchanged form as the most popular IAL with 1-2 million users (numbers differ, see Corsetti 2012: 69

and Wandel 2015; Piron (1989b) quotes 3.5 million). However, a dozen Esperanto offshoots were constructed at the beginning of the 20th c. as an attempt to improve its alleged flaws, among which were the five circumflexed letters (ĉ, ĝ, ĥ, ĵ, ŝ) and ŭ with a breve, the accusative ending -n (the only existing case), illogical affixes, the vocabulary being “randomly” chosen (sometimes accused of being too Romance, sometimes not Romance enough) and many others (Large 1985: 118–123; Rónai 1969: 62–65).

Esperantists sought the support of official institutions for their case. Many of them welcomed the creation of a Délégation pour l'Adoption d'une Langue Auxiliaire Internationale by L. Couturat and L. Leau in 1901. The Delegation seriously considered two projects: Esperanto and Idiom Neutral. The latter, published in 1902, was the product of W. K. Rosenberger, the President of the former International Academy of Volapük. Rosenberger, with the approval of the Academy, designed an *a posteriori* language, formally based on its predecessor, but more Romance (French) than Germanic (or straightforwardly German) and with easier grammar rules. The Delegation decided that the projects should not be presented by their creators. Therefore, O. Jespersen pleaded the case of Idiom Neutral, while L. de Beaufront that of Esperanto (Large 1985: 80–82). None of the projects seemed right to adopt. Zamenhof's shock was all the greater, when he discovered that his former supporter de Beaufront, together with L. Couturat, chose to reform Esperanto and in 1907 introduced Ido – the name being a suffix, standing for ‘offspring’ (Rónai 1969: 58). Ido eliminated the circumflexed letters, the concord between adjectives and nouns, changed some verbal endings and word roots to be more “natural,” as well as used the accusative optionally in sentences with inverse word order (Bausani 1970: 126; Rónai 1969: 82). However, Ido not only never gained substantial success, as Esperanto had fervent supporters and extensive literature, but also failed to promote the changes as the ultimate reform. Scores of new languages

emerged as a result of the so-called Schism (Duličenko 2006: 112f.; Rónai 1969: 82) – A. Libert (2008) examines 30 of the “daughters of Esperanto”, but lists also more than 10 other schemes somehow connected to Esperanto.

A POSTERIORI LANGUAGES: LATER RIVALS OF ESPERANTO

Ido itself suffered from a schism in the 1920s. In 1922 E. de Wahl (or von Wahl, a former Volapükist, Esperantist and Idist) presented Occidental (also called Interlingue) and six years later O. Jespersen, dissatisfied with the schemes he knew, published his very own Novial (literally ‘new IAL’). Both schemes showed naturalistic tendencies; they comprised European international vocabulary and somewhat simplified grammars. Naturalism was not a new idea. In 1903, the famous mathematician G. Peano wrote an article with a proposal for Latino sine flexione (also known as Interlingua). He reduced the grammar of Latin and in the course of time selected 14,000 word roots for his dictionary from 1915 (Large 1985: 142–145).

The problem of such a large number of popular schemes called for a solution. Thus, in 1924 the International Auxiliary Language Association (IALA) was founded in New York. It appointed a team of renowned linguists (E. Sapir, M. Swadesh, R. Jakobson and A. Martinet) to compare six existing languages (Esperanto, Esperanto II, Ido, Occidental, Latino sine flexione and Novial²³) with each other and with natural languages. Later on, the IALA’s goal became to promote the scheme of its choosing. However, none of the six seemed perfect, and therefore the team decided to standardise the international vocabulary and to find the common denominator of the projects. The outcome was a new language called – not surprisingly – Interlingua. Its first dictionary was published in 1951 under the supervision of A. Gode (often referred to as the main creator of Interlingua). He thought the language to be

²³ These six given by Large (1985: 146), while Rónai (1969: 83) lists only five (without Novial).

planned and not constructed, because it was “pan-Occidental” and as naturalistic as possible (Large 1985: 146ff.). Only a word found in three out of five European languages (Italian, Spanish, Portuguese, French or English, and if need be Russian and German were consulted) was to be included in the vocabulary. The language kept a complicated European system of tenses and modes (Bausani 1970: 132–135).

Another example of an IAL is Interglossa from 1943 (L. Hogben), subsequently reformed and renamed Glosa (1981, W. Ashby and R. Clark). Hogben realised that any scheme intended to be used internationally by scientists should include widely known vocabulary (here Greek, as it is often used in science) and a simple isolating grammar (similar to that of Chinese). He removed inflection and grammar categories so that a word could be a noun, a verb, an adjective etc. depending on the context. The syntax is then clearly delimited (see Rónai 1969: 18). A similar system was made public in 1957 by the Vietnamese linguist Phạm Xuân Thái. The *Lingua sistemfrater* (or simply *Frater*) comprised Greek and Latin word roots and a grammar not unlike that of Interglossa (see Rónai 1969: 22).

REDUCED NATURAL LANGUAGES

A different kind of IAL is simplified (also minimal or controlled) natural languages, although it is hard to say where the borderline between naturalistic *a posteriori* and simplified natural languages is. In fact, Janton (1993: 9–12) writing about minimal languages includes also projects which are classified by other authors as naturalistic *a posteriori*, for example *Latino sine flexione* by Peano, *Ruski Jezik* by Križanić and several projects of Molee, *Wede* by Baumann (1915, “Welt Deutsch”) and *Anglic* by Zachrisson (1930). He states that the European language creators were mostly oriented on Latin and Romance languages as the basis of vocabulary and structure, and therefore “it is often difficult to tell whether a given naturalistic language is a form of simplified Latin or is modelled on Romance lan-

guages” (Janton 1993: 10). The last project Janton mentions is Basic English, a project by C. Ogden (1935).

Ogden’s project was an original attempt to simplify the vocabulary of English without changing the grammar or the orthography. He limited the scope of the lexicon to just 850 words and substituted “unnecessarily complicated” verbs with particle verbs. Thus, with those 850 words and their combinations Ogden claimed to have covered 20,000 usually used words (Large 1985: 163). However, Basic (British American Scientific International Commercial) English required from the user constant circumlocutions and a great dose of inventiveness. At the same time, it also required a fairly good command of English as the grammar was not simpler and the use of periphrasis and compounding made the language much more idiomatic and irregular in meaning. The project’s capacity for expressing nuances was likewise limited (Janton 1993: 12; Large 1985: 169f.).

Until this day, Esperanto seems to be the only successful IAL. It has attracted the most supporters and speakers, and it continues to be used during congresses of the Universal Esperanto Association and national Esperanto associations, youth congresses, conferences and even at schools and universities. The interest in creating an IAL however has waned recently, to make way for the modern inventions.

2.2.4 Artistic languages

Languages created for artistic purposes comprise a wide range of types – from a few words or sentences to several passages or mere sketches, up to whole systems elaborated over many years. In those works where the text is only sprinkled with strangely looking creations, an invented language is often used as a stylistic device to enliven the piece²⁴. This seems to be the

²⁴ As Albani & Buonarotti (1994: 12–13) indicate in their chart, constructed languages appear not only in literary works (both prose and poetry)

case in Aristophanes' *Frogs* (405 BC), where the author uses onomatopoeic, sentence-like structures. The play is possibly the oldest example of the use of an artistic artificial language.

Other motives for creation of an artlang are comparison of cultures and/or languages, introduction of a new socially important idea or an attempt to make the civilisation presented more authentic. Some authors simply enjoyed using their linguistic knowledge to amuse themselves and the reader (Knowlson 1975: 112ff.).

ARTISTIC LANGUAGES OF THE 16TH – 18TH CENTURIES

One of the earliest works where the reader encounters a developed artificial language is T. More's *Utopia* (1516). Although the author wrote, and translated into Latin, a quatrain in the Utopian language, only three editions included the text. Utopian is an *a posteriori* system following the syntax of Latin and mimicking its case endings, with words invented and borrowed from Greek and Persian. The language also has its own alphabet modelled on Greek or Glagolitic letters (Higley 2007: 64f.; Knowlson 1975: 115).

Three other well-known language specimens are presented by F. Rabelais in *Pantagruel* (1532). The fragments are modelled in phonetic resemblance to existing languages to add humour and vitality to the text. A similar technique is adopted by J. Swift in his *Gulliver's travels* (1726), although here the author uses a plethora of methods: anagram, omission, transposition of letters etc. (Fiedler 2011: 15; Knowlson 1975: 116, 125).

Imaginary voyages became a very popular theme in literature in the 17th and 18th c. The authors made it a tradition to lend credibility to the journeys to faraway lands by providing their societies with invented languages. Such is the case of Cyrano de Bergerac's *Histoire comique des États et Empires de*

but also in the visual arts (e.g. as invented alphabets), cinema, theatre, music and comics. Works of all these types are treated here as texts; however the examples provided come mostly from prose and film works.

la Lune (published posthumously in 1657). The inhabitants of the moon speak a language consisting of musical notes. The idea came to Bergerac probably from an earlier work by J. Wilkins, who, before setting out to create a universal language (see section 2.2.1), wrote *Mercury, or the Secret and Swift Messenger* (1641) in which a whole chapter “was devoted to a consideration of a possible musical language, Domingo Gonsales and *The Man in the Moone* [a novel by F. Godwin from 1620s, published posthumously in 1638] being referred specifically by name” (Knowlson 1975: 121). A musical language also features in L. Holberg’s *Nicolai Klimii Iter Subterraneum* (1741), in which the hero encounters instrument-like creatures who do not speak but play their own bodies (Knowlson 1975: 122).

It seems that inventing languages was particularly popular in French literature. Four examples are quoted by Knowlson (1975): *La Terre australe connue* by G. de Foigny (1676), *Histoire des Sevarambes* by D. Vairasse (or Veiras) from 1677-79, *Voyages et Aventures de Jacques Massé* by S. Tyssot de Patot (1710) and *La Découverte Australe* by Restif de la Bretonne (1781). De la Bretonne uses French written backwards and describes the language of “man-lions” as consisting of no more than thirty words supported by a number of gestures (Knowlson 1975: 114,124). This scheme appears to be no more than a light-heartedly created prop. This is not the case when it comes to the other three creations. Foigny believes that a language can provide the key to knowledge, not only by words directly representing things but also by letters denoting different qualities and the nature of those things. For example, the five vowels of the language stand for five elements. This procedure largely recalls the English universal language schemes (Knowlson 1975: 130f.). Similarly, Veiras wants specific combinations of sounds to represent specific qualities or objects. However, his language is based on onomatopoeia and euphony, while in the language of de Foigny many combinations unpronounceable for Westerners occur

(Knowlson 1975: 130, 133). There are 10 vowels and 40 consonants, along with 30 diphthongs and triphthongs, as well as marks for tones and the quality of the vowel. Veiras represented the system with an invented alphabet bearing a close similarity to the polygraphy of Lodwick (Knowlson 1975: 135f. ; cf. section 2.2.1). More than twenty years later, *Voyages et Aventures* appeared, in which Tyssot de Patot presented a much simpler language with only 7 vowels and 13 consonants not serving the purpose of direct representation of things, as well as a regular grammar with only three tenses and several suffixes designating grammatical categories (Knowlson 1975: 137).

An unusual example of an invented language is the fake Formosan presented in *A Historical and Geographical Description of Formosa* (1704). It was created by one George Psalmanaazaar, a Frenchman²⁵ whose real name remains unknown. He claimed to be an expert in Formosan (being a native from Formosa, now Taiwan), on which he even wrote a course book and for which he invented an alphabet. He taught the language to several people at Oxford, claiming its similarity to Japanese. The language was meant to be simple and regular, having only three genders in the singular represented by three articles and only one in the plural. The tenses were indicated by a rising or a falling intonation (Knowlson 1975: 125–129).

MODERN CREATIONS

Subsequent literary works (often fantasy and sci-fi) frequently feature short examples of languages of different kinds – both *a priori* and *a posteriori* – such as incantations to Cthulhu in H. P. Lovecraft's novels (*a priori*), Nadsat in A. Burgess's *A Clockwork Orange* (slang largely based on Russian and cockney rhyming slang), Fremem in F. Herbert's

²⁵ So writes Knowlson (2007: 125); Higley (2007: 72) reveals that Psalmanaazaar himself admitted to being the son of German parents.

Dune series (*a posteriori* based on Arabic), Kesh in U. K. Le Guin's *Always Coming Home* or the language of reptilians in H. Harrison's *Return to Eden*. However, two types of language inventions are especially widely commented: the dystopian jargon in G. Orwell's *Nineteen Eighty-Four* (1949) and the tongues of fictional races in J. R. R. Tolkien's *Lord of the Rings* (1955).

Orwell clearly believed that one can manipulate thinking with the help of language. Although the reader of *Nineteen Eighty-Four* is left with only about 50 in-text examples of the use of Newspeak, more detailed information can be found in an appendix (Jackson 2011: 50f.). The dystopian Newspeak (itself being a Newspeak word) is intended to brainwash the citizens of Oceania, a totalitarian super-state. The work on a new definitive edition of the Newspeak dictionary is in progress so that by 2050 each word would correspond to exactly one concept. Abbreviations, affixing and compounding are used continuously in order to prevent any "engagement of the brain" and to develop rhythmic, monotonous prosodic patterns (Jackson 2011: 53f.). Newspeak can be compared to Ogden's Basic English in that it is based on English and its vocabulary is very limited. However, while Ogden never intended to regularise the grammar, Orwell limits the scope of expression of the language by cutting out any possible irregularity.

Tolkien's adventure with invented languages started early. He is claimed to have known Latin, Ancient Greek, Old English and Gothic already as a schoolboy (Weiner & Marshall 2011: 95). His first attempts to create an artificial language were a contribution to a cousin's Nevbosh and a systematic extension of limited extant Gothic vocabulary fashioned on the basis of Germanic sound laws and cognate words (Weiner & Marshall 2011). Tolkien also had some knowledge of Old Icelandic, the Germanic, Celtic and Romance families of languages. At some point before 1915, he began to be interested in Finnish after reading the *Kalevala*. Around that time, the first sketches of the Elvish languages were made. Tolkien

first created the lexicons of Qenya and Gnomish, later renaming them Quenya and Noldorin, and finally renaming Noldorin once again to Sindarin and making it one of the dialects of Primitive Qendian (these being Quenya, Sindarin, Noldorin, Telerin, Ilkorin and Danian). Through the years, the systems of the predecessors to the languages known from *The Lord of the Rings* were transformed noticeably. Although Tolkien's taste for particular language configurations changed since then, the forms found in *The Lord of the Rings* (first published 1954-55) could not be altered after being printed. Therefore, the novel might be treated similarly to the *Fundamento de Esperanto* – what has been written once cannot be changed later.

Tolkien's languages were not intended for anything other than the pleasure of creation. Not infrequently did he mention that the appearance of the Elvish tongues is a matter of personal taste – the elegance of the solutions and a pleasurable pronunciation. However, another important reason for devising such complicated linguistic relations internally and externally was lending credibility to the stories and legends of Tolkien's literary world (Okrent 2010: 282ff.; Weiner & Marshall 2011). The early versions of the narratives made direct links to the real world. Tolkien wanted the realm of Elves to be the predecessor of ancient Europe. Therefore, there were also direct links between Elvish and human languages. For Tolkien, *mythopoeia*²⁶ was inevitably entwined with *glossopoeia*. As Weiner and Marshall (2011: 98) remark, the resemblances between vocabulary forms in Elvish and European languages are the result of contacts between the folks and the retention of some Elvish names in real-world legends. Such “backwards” etymology is visible even in *The Lord of the Rings*, in which several names and roots are traceable to Old Icelandic (*Gandalf*) or Old English (the language of Rohan). The grammatical characteristics of Quenya are culled from

²⁶ This word used by Tolkien for the first time in his paper *A Secret Vice* – apparently a lecture delivered in 1931 (in Tolkien 1983).

those of Latin, Greek, Finnish and Germanic (i.e. high agglutination, the dual both in Quenya and Sindarin), Sindarin exhibits initial consonantal mutations similar to those of Welsh and regular vocalic alternations as in Germanic, whilst Khuzdul (Dwarvish) has got template morphology inspired by Semitic languages (Weiner & Marshall 2011).

Tolkien's example is atypical: his literary worlds were created as secondary to the languages. His activity, a "secret vice" (Adams 2011: 13f.; Okrent 2010: 282f.), might be regarded as one of the early examples of modern *conlanging* (i.e. creating languages for pleasure). This would suggest that languages of fiction are a subset of languages designed "for fun". Nevertheless, languages created for films and games are not only public but mostly make paying jobs. The practice of hiring linguists to produce a credible language for a TV series or a film starts with Klingon in the *Star Trek* productions. Klingon is a language of a fictional warrior alien race in *Star Trek* films, TV series and connected video games and novels, along with various products. The first lines of the language are from the 1979 film *Star Trek: The Motion Picture*. The dialogues were created by one of the actors in the film, James Doohan. Five years later, the executive producer decided to hire linguist Marc Okrand to devise an actual language for the third *Star Trek* film (Okrand et al. 2011: 112–113). In 1982, Okrand had created some pieces of the Vulcan language for the second *Star Trek* film, lip-syncing Vulcan dialogues with English speaking actors. His new task included creating a Klingon phonology pronounceable for the actors, a grammar and some additional vocabulary, all based on earlier pieces of dialogue. Klingon was not yet to be a fully developed language.

During the work, Okrand invented the "real" name of the language, namely "Tl'ingan" /tʰi.ŋan/, which, due to its unusual pronunciation (the voiceless alveolar lateral affricate hardly occurs in languages of the world, e.g. Cherokee, Nahuatl, Tlingit and Tswana) was misheard and rendered "Klingon" by humans. To make Klingon more unfamiliar, Okrand

decided to employ rarely occurring features all found in natural human languages but not in this specific combination. These include an unusual sound inventory (e.g. voiced retroflex stop, glottal stop and no voiced/voiceless velar stops), rare OVS word order (less than 1% of world languages), and agglutination of up to 9 consecutive verb suffixes (Okrent 2010: 268–270; Okrand et al. 2011: 116–119, 122f.). Sabine Fiedler (2011: 11) points out that this “strangeness” is a planned effect of an aesthetic pursuit rather than a practical one:

Plansprachen haben im Sinne Wüsters (1931) die Aufgabe, die internationale sprachliche Kommunikation zu erleichtern. Um dieser gerecht zu werden, spielen Kriterien wie leichte Erlernbarkeit und Universalität eine Rolle, welche für die vor allem auf ästhetische Wirkungen ausgerichteten fiktionalen Sprachen nicht relevant sind oder sogar kontraproduktiv sein können. So war es bekanntlich Mark Okrands Intention, mit dem Klingonischen eine besonders fremdartige Sprache zu schaffen, die außerdem den Charakter der kriegerischen Klingonen widerspiegelt.

Planned languages, as defined by Wüster (1931), have the task of facilitating international linguistic communication. In order to meet this, criteria such as ease of learning and universality play a role; to the fictional languages created mainly for aesthetic effects these are not relevant, or may even be counterproductive. It is generally known it was Mark Okrand’s intention to create Klingon as a particularly strange language that also reflects the character of the belligerent Klingons.

Subsequent *Star Trek* productions featured also some lip-synced Klingon and passages created by authors other than Okrand. This did not prevent the codification of Klingon. Followed by the appearance of a journal devoted to Klingon linguistics published by the Klingon Language Institute (*Hol-Qed*, 1992–) was the first edition of *The Klingon Dictionary* (1985). A collection of Klingon proverbs with commentary appeared in 1996 and a year later a book on sociolinguistic variations of Klingon – both by Okrand (Okrand et al. 2011: 125). To

this day his is the ultimate linguistic authority among Klingon speakers (who, according to some research, do not constitute a speech community, but rather consist of two separate groups: a sub-culture of Trekkies, i.e. die-hard fans of *Star Trek*, and some dedicated linguists²⁷).

However, this community of fans is immensely important capital for the producers of films or games. The idea of hiring a linguist even before the film began to be shot to craft a language suitable to the tastes of future fans spread widely and resulted in such creations as Na'vi from *Avatar* (2009) and Dothraki and High Valyrian from the *Game of Thrones* series (2011–, an adaptation of George R. R. Martin's series of novels). Dothraki is a language that made its first appearance in the books as names, individual words and short phrases. It has been developed further by the linguist David J. Peterson, member and co-founder of the Language Creation Society, who has also authored the book *Living Language Dothraki* (2014). The language is head-initial and inflectional with SVO word order and now has more than 3,000 vocabulary items (Peterson 2011; Peterson n.d.). Na'vi, in turn, has been created from scratch by the linguist Paul Frommer. The language has tripartite case alignment (absolute intransitive subject, ergative transitive subject and accusative direct object; genitive, dative and a topic marker are also used), a dual, a trial and a plural number and a vocabulary of more than 1,500 items (Frommer n.d.).

The games industry has also come up with the idea of creating artificial languages – sometimes as part of the gameplay (so the user has to “learn” whichever words and sentences they may encounter) or as an artistic addition to the game, which Portnow (2011: 140) calls incomplete “flavour languages”, that is, unsystematically operating sentences being merely gameplay embellishments. However, two languages from the

²⁷ For a summary see Okrand et al. 2011; full record in Hendriks-Hermans (1999) and Wahlgren (2004).

beginning of computer games' history are worth noting: Gargish from the 1990 game *Ultima VI* and D'ni from the *Myst* series by Cyan Worlds (1993). Both difficult for the English speaker and having their own scripts, they are inextricably tied to the gameplay. Gargish is a free word order, pro-drop language with an unusual feature: when spoken, tenses and parts of speech are indicated by gestures and intonation. In writing, those are marked by a series of suffixes. Word order in D'ni, in turn, is rather fixed (SVO) but the order of modifiers can be changed to highlight the importance. Its most uncommon feature is a base 25 number system used to create mathematical puzzles in the gameplay (Portnow 2011: 141–146).

Many artistic creations are merely “flavour languages”. Authors sprinkle their works with invented words and phrases to enliven the stories or make them more realistic. However, recent years show that languages created for artistic purposes may also become a hobby – a well-paid one for the inventors and a lifelong love for the fans.

2.2.5 Modern glossopoeia²⁸

Modern language creations are not easily classified. On the one hand, they are typically built in resemblance to diverse ethnic languages and therefore are partly *a posteriori*, but on the other hand, often the elements are invented and *a priori*. Some systems are created to exercise the limits of human language and deliberately violate the hypothetical universals. Many of them are created merely for pleasure or linguistic interest and some are meant to replace a whole family of languages. They could be classified as artistic, auxiliary or philosophical. Therefore, as there are hundreds of projects, semi-languages and elaborate creations published on the Internet,

²⁸ Tolkien in 1931 used the term ‘mythopoeia’ in his lecture entitled “A Secret Vice”. In resemblance to this word, a new one has been coined: “glossopoeia”. Its authorship and time of creation are unknown.

there will be presented only several relatively well-known conlangs in chronological order.

Talossan is one of the earliest conlangs still used. Invented in 1980 by Robert Ben Madison, it is a language of the micronation of Talossa, a kingdom founded by Madison in 1979 in his bedroom (Barandovská-Frank 2011: 33; Rogers 2011: 217). In 1995,²⁹ the Seneschal Geoffrey Thomas offered Talossan citizenship on the Internet to anyone interested. The kingdom split in 2004 into the Kingdom of Talossa and the Talossan Republic. Talossan is based on Romance languages and currently has more than 28,000 words. There are over a hundred citizens of Talossa, but most are English speaking; however, the language features as one of the most (if not the most) important nation-building components, not only in the description of the kingdom on its website but also in the national anthem (Barandovská-Frank 2011: 42, 47).

In 1996 Andrew Smith created Brithenig (Rogers 2011: 40–42), a Latin-descended language which might have displaced Celtic languages in the British Isles if Latin had been more influential. Just as in the case of Tolkien, Smith decided to authenticate his language by inventing Ill Bethisad, an alternate world with an alternate history. Ill Bethisad has become a collaborative project with several languages, detailed maps and histories (see <http://www.bethisad.com/>). One of the most prominent languages developed within the project is Wenedyk invented in 2002 by Jan van Steenberg (Rogers 2011: 243–244). Wenedyk is Romanised Polish used in the Republic of Two Crowns (roughly the Polish-Lithuanian Commonwealth).

Slovio was created in 1999 and made public in 2001 by Mark Hučko. It is a Panslavic auxiliary language (Mannewitz 2009). Partly in response to its artificiality, Slovianski was invented in 2006. Various “dialects” (i.e. simplified and

²⁹ This is the year which Okrent (2010: 313) and Rogers (2011: 217) give as the date of publishing the language.

schematic) of Slovianski were abandoned in favour of the naturalistic version in 2009. It has since been the corporate endeavour of Ondrej Rečnik, Gabriel Svoboda, Jan van Steenberg and Igor Polyakov. Efforts to bridge the gap between Slovio and Slovianski (i.a. by Steeven Radzikowski) led to a merger of several projects into one, known, as of 2011, as Interslavic. It can be written in both Latin and Cyrillic and is highly naturalistic, drawing on all major Slavic languages. Slovio remains a separate language (van Steenberg 2013).

Toki Pona is a philosophical language created by Sonja Ellen Kisa in 2001 (Blahuš 2011). Its purpose is not to serve as an auxiliary language but to express as much as possible in a minimal language. The language avoids overly complex “euphemisms” such as “collateral damage” instead of “killing of civilians”. It is also said to have been inspired by the Dao philosophy. It is a pidgin-like language based on English, Tok Pisin, Chinese and several other languages, with only fourteen phonemes and a very simple grammar, and only a little more than 120 root words (Blahuš 2011; Rogers 2011: 226f.)

In 2004, the first version of Ithkuil was made public. A language invented by John Quijada, it was in the making from 1978. In 2011, the final version was published on the Internet. It incorporates the “consonantal phonology and verbal morphology of Ubykh and Abkhaz, certain Amerindian verbal moods, Niger-Kordofanian aspectual systems, Basque and Dagestania nominal case systems, Wakashan enclitic systems, the Tzelal and Guugu Yimidhirr positional orientation systems, the Semitic trilateral root morphology, the evidential and possessive categories of Suzette Elgin’s Láadan, and the schematic word-formation principles of Wilkins’ Analytical Language and Sudre’s Solresol” (Okrent 2010: 290). It has both a very difficult pronunciation and grammar, so that even its creator has never learnt to speak it (Quijada 2011).

One of the newest conlangs is Lingwa de Planeta (Lidpla/LdP), created in 2006 and published in 2010 by Dmitri Ivanov et al. from the University of Sankt Petersburg (Kiril-

lov 2012). The language is meant as a world auxiliary language, being based on Arabic, Chinese, English, French, German, Hindi, Persian, Portuguese, Russian and Spanish and using most widespread word-forms (see also Libert 2013: 128). Its grammar is analytical or pidgin-like – tenses, genders and numbers are expressed with (non-obligatory) unbound grammatical markers (Ivanov 2014).

It might be claimed that Lidepla should be classified as an IAL and not as a modern conlang precisely because of its purpose. However, the 1990s mark the beginning of the Internet era, where the boundaries between diverse classes of artificial languages become blurry. It is also important that the relative success of a language is not anymore directly connected to its values (usability or aesthetics) or the socio-political situation, as it was in the cases of the universal languages of the 17th and 18th c. or the 19th c. IALs. Now more than ever, it is a matter of publicity and coincidence.

As the last type of contemporary artificial language, I should mention the so-called esoteric programming languages. They are designed to test the limits of programming, or as mere entertainment. They are hardly useable – Whitespace may serve as an example, a language using exclusively white characters, the difference between which is usually ignored by other programming languages, or Befunge-93, in which the commands may be written either horizontally or vertically in both directions. One of the most prolific creators is Chris Pressey, the creator of, among others, Befunge with its offshoots – his personal website lists 78 of languages (Pressey 2015)³⁰. One of his most interesting inventions is what he himself calls an “abstract artlang (i.e., a conlang designed independently from any conception of society.) The sole design principle was to entirely eliminate word order”. Opus-2 expresses actions through colours, noun meanings through sounds

³⁰ Although the copyright date is given as 2014, the newest changes are from 2015; I assumed that the latest version of the website is 2015.

and adjectives through smells. A sample sentence “The man quickly flees the dangerous child” takes the following form (Mannewitz 2003: 47):

- + pale green
- + Eb, trombone, forte
- + leaning 40 degrees left (sudden)
- + C, tubular bells, piano
- + mothballs (gentle whiff)

As can be seen, modern Internet creations range from hobby languages of invented worlds, hypothetical languages of alternative history worlds to philosophical and auxiliary languages.

2.3 Typology and classifications³¹

Barandovská-Frank (2004: 134; see also Back 1996) rightly points out that generally artificial languages are named according to the purpose of their creation (universal languages, world languages, international languages etc.) and their construction (planned languages, fictional languages etc.). This, however, does not provide the reader with any classification. In his article Blanke (2001: 51f., see also Blanke 1985: 99-110) gives six different types of classifications of artificial languages according to:

- realisation level (graphic/phonetic, i.e. pasygraphies or universal scripts vs. pasyalies or universal languages)
- material and structure, i.e. invention vs. imitation
- linguistic concept, i.e. how well the project is formulated linguistically
- level of details/development
- availability
- role in real communication

³¹ This section is partly based on Stria 2013.

While the last four do not appear in the literature, the first two are common enough.

2.3.1 Traditional classifications: structure and source of material

The second classification mentioned by Blanke is regarded as traditional and is most widespread. It was proposed by L. Couturat and L. Leau (1903; cf. Schubert 2011: 50) as a crude division into three main types:

- *a priori* (based on invented elements)
- mixed
- *a posteriori* (based on elements already given)

This proposal has been quoted by several linguists. A more detailed version can be found in Janton (1993: 6f.) with the following categories:

- *a priori*, (i.e. “philosophical” languages)
- *a posteriori*
 - simplified ethnic (i.e. minimal)
 - mixed
 - schematically derived (Volapük)
 - with partly schematic and partly natural derivation (Esperanto)
 - naturalistic
 - with some schematic traits (Novial)
 - with natural derivation (Occidental, Interlingua)

It can be seen that these categories present a scale of artificiality. The poles determine whether a project is derived from (an) ethnic language(s) or deliberately designed. It is immediately noticeable that the examples given include only the universal and international auxiliary languages (that is to say, ‘planned languages’ according to Blanke 1985: 11). Of course, other types could be incorporated as well, however uncommon the idea. Many similar representations are pro-

posed. One of them is a scheme by Marcel Monnerot-Dumaine (1960, as quoted in Barandovská-Frank 1995), who divides artificial languages into five groups:

- artificial word roots, schematic derivation (philosophical and non-philosophical languages; the latter group includes musical and symbolic languages)
- half-artificial word roots, schematic derivation (e.g. Spelin)
- deformed word roots, schematic derivation (e.g. Volapük)
- non-deformed word roots, further subdivision into languages with half-natural or schematic derivation (e.g. Esperanto) and Romance-derived irregular roots
- non-classifiable

Another elaboration of the scale is presented by Sergej Nikolajevič Kuznecov (1984, in Barandovská-Frank 1995), who divides artificial languages into *a priori*, mixed, *a posteriori* with *a priori* affixes, *a posteriori* without *a priori* affixes and lastly naturalistic *a posteriori*.

Curiously enough, Janton speaks about “simplified ethnic languages”. This is what is understood as “controlled” languages on Blanke’s list of artificial constructs (Blanke 1997: 3). Klaus Schubert (2011: 52f.) points out that those might be divided into two groups: one before 1960, that is, ethnic languages strongly simplified for the sake of international communication, such as Peano’s Latino sine flexione and Ogden’s Basic English, and the other one after 1960, which comprises languages created to facilitate industrial communication, e.g. Caterpillar Fundamental English or ScaniaSwedish (see also Schubert 2001). A. Large (1994) observes that “modified natural languages are closely related to naturalistic *a posteriori* languages, having many similarities in approach and objective. In practice the distinction between a naturalistic artificial language like Latino sine flexione and a modified natural language such as Basic English is relatively minor”. This raises the question of why otherwise

natural languages would then be described as artificial. In accordance with our previous definitions, these languages result from a process of deliberate and conscious planning; furthermore, they are shaped by a particular individual or an identifiable group of individuals. By the same token, this understanding of the term ‘artificial languages’ includes several “natural” types, such as trans-ethnic languages (those that have lost their ethnic ties by becoming internationally spread, such as Medieval Latin), standard forms and *Dachsprachen* (written German, Standard Arabic), revitalised languages (Modern Hebrew, Cornish) as well as pidgins and creoles.

An interesting division based on “freedom of creation” is developed by Bausani (1970: 11f.). Quoting R. Jakobson, Bausani states that there are four degrees of creative freedom in language: on the level of phonemes, words, sentences and utterances. The fourth degree is possible in all languages and therefore artificial languages can be classified into three groups, according to the levels on which the changes are introduced (Bausani 1970: 12):

1. with non-natural syntax but maintaining the morphological and phonetic inventory of the natural language intact (e.g. poetic and ceremonial languages)
2. with invented vocabulary but the morphology of the natural language more or less preserved (e.g. jargon, slang, poetic languages)
3. with new morphology and vocabulary but the sound inventory unchanged (many universal languages)

The author mentions one last possibility (however uncommon): artificial languages where the sound inventory is altered. The previous section (2.2) shows that changes on all levels can be introduced in artificial languages.

2.3.2 Traditional classifications: purpose

Another type of classification categorises artificial languages according to the purpose of creation. There is no consensus, though, among the researchers on how to classify some systems, because they satisfy several goals and the categories in which they are put might overlap. Kennaway (2010) provides their division in the form of section headers (on the division by Bartlett (2009) see section 2.2):

- the perfect language (on universal languages)
- an international language (e.g. Esperanto, Latino sine flexione)
- fiction (e.g. artistic languages of Tolkien)
- recreation (conlangs)
- programming languages

Commonly, three main “waves” of artificial languages are discerned: universal languages, IALs, and conlangs (cf. Okrent 2010).

A detailed typology is proposed by Albani and Buonarotti (1994), where a division is made into sacred and non-sacred languages. Sacred languages are further divided into structured (Bālaibalan) and non-structured with six subdivisions. Non-sacred languages split into languages with communicative and expressive goals (Figure 4) both with further detailed subdivisions. The chart arranges the languages according to the purpose as well as the construction. It also takes into account the form (realisation level) and the source (material), when it comes to the classification of IALs. Nevertheless, this will not be discussed further, because there is no clear explanation of the principles that guided the creation of this classification. The authors mention (Albani & Buonarotti 1994: 8) that the source of the basic division into sacred and non-sacred was Bausani (1974 extended edition in Italian; here the 1970 German edition is used) and the only criterion was the language’s functional aim. Additionally, some of the languages will not be

examined, as they are not consciously created, which is a prerequisite resulting from the definitions.

Bausani (1970: 13), on whose concept the chart of Albani & Buonarotti is based, provides a simple classification:

A. Sacral

1. Genuine and actual religious artificial languages
2. Partly sacral pseudo-languages (glossolalia, magic formulas)

B. Profane

3. Secular (profane) languages serving purely as expressive plays (languages invented by children)
4. Artificial languages with communicative purpose (example: Esperanto)

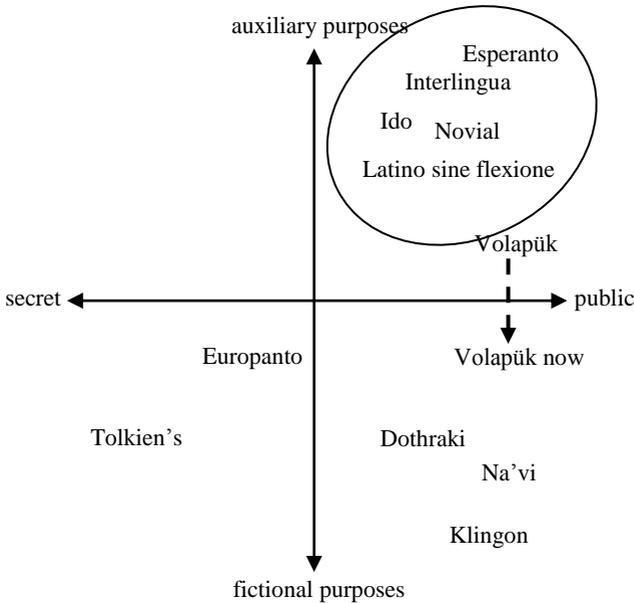


Figure 3 Federico Gobbo's coordinate system (2014)

This is further elaborated by Gobbo (2014, modified version of his 2009 proposal) in the form of a Cartesian coordi-

nate system (Figure 3). Gobbo's claim is that secret languages are those without a key to their grammar and vocabulary. Therefore, Bālaibalan and Tolkien's languages are classified as secret, although the latter ones have been published.

Two questions are worth asking: what is the difference between Quenya of Tolkien and e.g. the early forms of Na'vi (film release 2009, grammar and dictionary published in 2011) and how to treat such languages as Bālaibalan. The answer to the first one is probably that a language can shift its position as in the case of Volapük, which stopped being used as an IAL and features occasionally in the literature. The answer to the second one is problematic. Gobbo quite logically states that it would be nonsense to launch a secret language for auxiliary purposes. However, Bālaibalan might paradoxically be such an example. Evidently, it is a secret language. Its purpose is to be a sacral language for Hurufi cultists and serve their in-group communication. Therefore, the language might be treated as an auxiliary language but only to a very limited extent.

One of the newest propositions widely spread on the Internet is the so-called Gnoli triangle (Figure 5). Claudio Gnoli, dissatisfied with the fact that his constructed language Liva was not easily classified, came up with an idea of a triangle whose vertices were labelled ‘artlang’ (artistic language), ‘auxlang’ (auxiliary language) and ‘loglang’ (logical language; the term ‘engelang’ was proposed later by And Rosta, apparently in 2001).

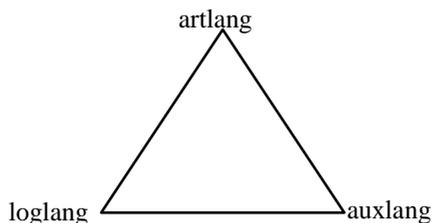


Figure 5 The Gnoli triangle (around 1997)

The triangle was later modified by Raymond A. Brown (Figure 6), who adopted the Maxwell colour triangle and the labels proposed by Rosta (Brown 2014).

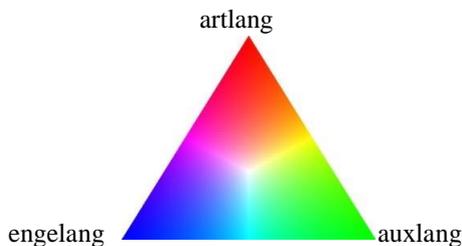


Figure 6 Coloured version of the Gnoli triangle modified by R. A. Brown

Another proposition was made by Jan van Steenberg in 2008. Van Steenberg drew a hexagon, in which he included not only “pure” constructed languages but also language reconstructions and reforms of natural languages. It is one of a very few attempts to incorporate borderline cases (i.e. special cases of

natural languages) into a classification of artificial languages (Figure 7). This, however, being opposed to Gnoli's previous proposition, evoked some criticism. Firstly, it was pointed out that the hexagon does not allow placing a mixed type language somewhere in the middle of the figure (as the Gnoli triangle did), thus excluding fuzzy categories. An example of the problem would be a language designed as both artistic and reconstructed – the hexagon does not allow for zero membership in other categories. Secondly, it was argued that the fields “reconstructed languages” and “reform projects” do not belong to constructed languages. In a discussion on a Yahoo mailing list³² van Steenbergen rebutted both: the Gnoli triangle is used for qualifying constructed languages and not classifying them, and the hexagon for the opposite. The two categories mentioned before, he further argued, are a result of conscious operations on linguistic material (cf. sections 2.1, 3.1 and 3.3).

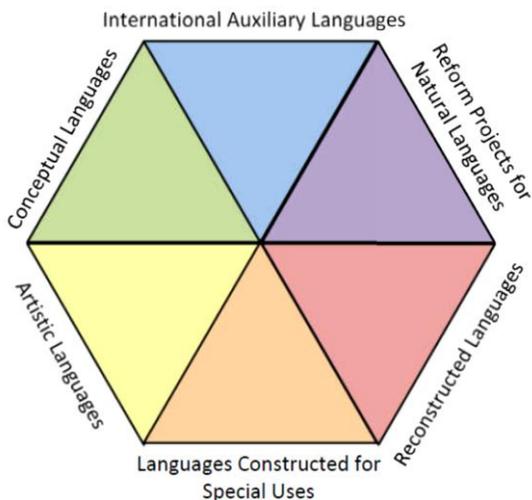


Figure 7 Jan van Steenbergen's hexagon (2008)

³² See <https://groups.yahoo.com/neo/groups/conlang/conversations/topics/179981> (last accessed 07 Jan 2014).

Nevertheless, he developed his classification into a more detailed matrix (see van Steenbergen 2008; a simplified version: Table 1), as the characteristics combining the purpose of creation and the source of material.

		purpose					
		IALs	con- cep- tual	artistic	spe- cial pur- pose	recon- struc- ted	re- forms
sour- ce of ma- terial	<i>a poste- riori</i>	Basic Eng- lish, Espe- ranto	Ce- qli	New- speak, Brithe- nig	Lin- gua Ignot a, Bālai balan	recon- struc- ted Proto- Ger- manic	Ru- mants ch Gris- chun
	<i>a priori</i>	-----	To- ki Po- na, Lo- jban	Klin- gon, Quenya	-----	-----	-----
	not speak able	Ges- tuno	Sol- re- sol	-----	Yerk- ish ³³	-----	-----

Table 1 Classificatory matrix (simplified with examples). Based on van Steenbergen 2008

³³ An artificial language developed for non-human primates, consisting of almost 400 lexigrams (i.e. pictorial symbols for particular lexemes).

2.3.3 Blanke's functional classification

The basic functional classification of artificial languages can be summed up in a binary division into *quasi-langue*, i.e. (invented) projects as well as *parole*, i.e. full-fledged languages (Blanke 2001: 52). Ammon and Hübler (1985, quoted in Sakaguchi 1998: 34) divide artificial languages into:

- full languages (IALs)
- partial languages
 - formal languages
 - jargons
 - specialised terminology of air traffic controllers

Blanke's categorisation brings about a further classification according to the purpose of the creation:

- IALs
- artistic languages (*artlangs*) created for aesthetic reasons
- constructed languages (*conlangs*) invented to exercise the limits of language
- experimental languages to exercise a philosophical idea

Of course, a language classified in one place might be easily moved elsewhere. For example, Toki Pona (created by Sonja Elen Kisa and first presented in 2001) is usually said to be an experimental system created to help express "positive ideas", as it is based on *dao* philosophy. Although Kisa herself never planned the language as an IAL (Blahuš 2011: 51), it might indeed be considered as such, because its first goal was that people communicate. Láadan (by Suzette Haden Elgin, 1982, published in *Native Tongue*, 1984; Okrent 2010: 241-249) was firstly created as part of a fictional world in a book, where it was intended to be used initially by women to express the perceptions of women not lexicalised in existing languages.

Unfortunately conlangs and artlangs are not treated very seriously by linguists – they are seen more as games than real languages. Among these groups, Klingon and Tolkien's crea-

tions have been, nevertheless, described from a somewhat linguistic perspective in Okrent 2010 and Adams 2011 (Klingon is also the subject of two theses available on the Internet: Hendriks-Hermans 1999 and Wahlgren 2004). What makes them seem so interesting is that they are growing in popularity with an ever-expanding group of supporters. Therefore, the birth of a Klingon communicative community might be taking place. Such a community is a requirement to change a mere project into a functioning language. It is also step 12 in an actual functional classification by Blanke (2001: 53-57). He described 28 levels (the first version in Blanke 1985: 107, table 2 had only 18 with the existence of native speakers being the last one; there were 19 steps in his later article 1989: 68f. with step 18 “independent cultural elements” added) through which a language (project) must go to attain the final stage of development (Blanke 2001: 53-57):

1. manuscript
2. publication
3. educational materials
4. advertising
5. journals
6. correspondence
7. translations and original texts
8. oral conversations
9. organisations
10. increase in text production
11. courses
12. small speech community
13. discussion on linguistic issues
14. professional communication
15. events
16. differentiation of the speech community’s structure
17. formation, stabilisation and codification of the norm
18. large events
19. worldwide dissemination
20. interlinguistics

21. heuristic actions
22. external use
23. schools and universities
24. electronic media
25. social differentiation
26. family language
27. original culture
28. language development

According to these criteria, three types of planned languages can be distinguished:

- projects (*Plansprachenprojekte*)
- planned semi-languages (*Semiplansprachen*)
- planned languages (*Plansprachen*)

Blanke applied his criteria only to IALs, but this list can also be used with regard to other languages. If one of them went through, for example, 14 of the stages, it might be considered a semi-language. This categorisation based on usage is sociolinguistically useful, although it must be remembered that it can only be applied to languages designed to facilitate international communication. As Liu (2001: 131) rightly points out “[...] the goals of creoles and planned languages are not the same, so if we evaluate pidgin development according to Blanke’s criteria, a creole language would not be a developed language”. The same is true for quite a great number of ethnic languages. Conversely, Sebba (1997, quoted in Liu 2001: 130) writes that a pidgin normally has to go through the following process in order to become a creole: “increasing stability → increasing vocabulary → increasing expressiveness → increasing functions”. In this respect Esperanto – thought to be a fully realised language with native speakers, the only one among IALs that went through all the stages – could be considered as creolised, although in terms of native speaker interaction it cannot be deemed so. The *denaskaj esperantistoj* (who are always at least bilingual) do not communicate with each other

on a regular basis and have no actual influence on the language (Corsetti 2012; Fiedler 2012; Lindstedt 2006). Therefore, in his revised version Blanke (2001) stated that the last step should be the development of a unique “cultural” phraseology and the evidence of language change. Interestingly though, this situation is in fact regarded as an advantage for learners as they can become fully accepted language community members, and those who master the language will be treated with high esteem (see Liu 2001: 127; Schubert 1989: 13).

3. Natural vs. artificial³⁴

The binary division into ‘natural’ and ‘artificial’ is considered by some linguists – e.g. Lyons (1991), but most notably interlinguists such as Blanke (1985: 26ff.), Koutny (2009), Sakaguchi (2003) and Schubert (1989) – outdated and misleading. This tradition dates back to Aristotle, but was most visible in the 19th-century linguistics strongly influenced by Darwin.

According to Sakaguchi (2003: 238), the idea that some languages are “natural” stems from the likening of languages to organisms that form, develop and die “on their own”, without human intervention (see also Blanke 1985: 19ff.). She points out that artificiality is, in this respect, a question of dependence on, or independence of the will of man. In her article (2003: 236f.) she follows the proposal by R. Keller, who divided the objects of scientific examination into three “worlds”: ‘natural phenomena’, ‘artefacts’ and ‘phenomena of the third kind’, the latter two being products of human activity. Artefacts are deliberately constructed (e.g. buildings, works of art, and also artificial languages), while phenomena of the third kind are created unconsciously (“natural” languages belong here). Thus, the distinction between ‘artificial’ and ‘natural’ would only be pertaining to the manner of their formation.

In this chapter, some possible interpretations of the terms ‘natural’ and ‘artificial’ will be presented along with borderline cases.

3.1 Properties of human language

The number of properties characterising language varies from five/six (Yule 2010: 11–15)³⁵ to sixteen (Hockett and

³⁴ Parts of this chapter pertaining to Esperanto were published as Stria 2015a.

Altman 1968, in Nöth 1995). The features have been repeatedly shown to occur also in animal languages, albeit in various combinations. I am going to discuss their occurrence in artificial languages, natural sign languages and some borderline cases (for more details on these special cases see the following sections).

Hockett (1960) listed thirteen (initially seven) “design features” of language, grouped according to classes of chor-dates (the higher the class, the more features its language displays). Humans are characterised by duality of patterning (double articulation), displacement (ability to talk about things remote in time and place) and productivity/openness; hominoids by discreteness and traditional/cultural transmission and primates by specialisation (communicative function), arbitrariness and semanticity. A further three were added later by Hockett and Altman (1968, in Coleman 2006): prevarication (ability to produce false or meaningless statements), reflectiveness/reflexivity (ability to consciously reflect on language) and learnability. The sixteen features grouped in five dimensions are (Nöth 1995: 235f.):

- relating to the channel
 - vocal-auditory channel
 - broadcast transmission and directional reception
 - rapid fading/transitoriness
- pragmatic
 - interchangeability (speakers can be receivers and *vice versa*)
 - complete feedback (the speaker can monitor their own message by immediately hearing it)
 - specialisation
- semantic
 - semanticity
 - arbitrariness

³⁵ The five main features are displacement, arbitrariness, productivity, duality and cultural transmission, and the sixth is reflexivity (or reflexiveness).

- displacement
- prevarication
- reflectiveness
- semiogenetic
 - traditional transmission
 - learnability (ability to learn other languages)
- characteristics of the code
 - discreteness
 - productivity/openness
 - duality/double articulation (larger entities built from minimal units, not necessarily phonemes)

Sometimes added to the list are structure dependence (as opposed to linearity) and recursion (Coleman 2006). These features are said to characterise every human language. However, they assume that the primary channel of communication for humans is the vocal-auditory one (written language being secondary). Unquestionably, sign languages, as having naturally evolved in human communities, also have to be taken into consideration, although they rather employ the visual-gestural channel. They have the same complexity as oral languages (structure dependence being especially essential, as they are far less linear than oral languages, cf. Wilbur 2011). Arbitrariness may be somewhat lower than in spoken languages, but still makes up a large proportion of signs. The only feature not present in sign languages other than the visual-auditory channel is total feedback (it is not possible for the signing person to see the signs produced, especially if we consider the fact that signing employs also head movements and facial expressions). There is therefore no great difference between sign and oral languages. As Coleman remarks, “children of two deaf parents who communicate using sign language acquire and use sign language according to the usual developmental sequence, and sign language has been found to use the same brain regions as spoken language” (2006: 474). This view is also supported by Tomaszewski (2004), who com-

ments on the seven features presented by Hockett in *A Course in Modern Linguistics* (1958, Polish translation in 1968):

Z perspektywy teorii Hocketta (1968) w odniesieniu do cech, jakimi dysponuje każdy język naturalny, PJM stanowi niewątpliwie kompletny system językowy. Ma wszystkie cechy języka naturalnego: (1) dwoistość, czyli podwójną artykulację, (2) produktywność, (3) arbitralność, (4) zdolność do wzajemnej wymiany polegającą na przemienności ról nadawcy i odbiorcy, (5) specjalizację, (6) przemieszczanie oraz (7) transmisję kulturową. Ostatnia z wymienionych cech PJM występuje w procesie socjalizacji, nauczania i wychowania; za pomocą PJM głusi przekazują z pokolenia na pokolenie wartości kulturowe, zwyczaje, wzorce zachowań, jakimi dysponują.

From the perspective of Hockett's theory (1968) with respect to the features available to any natural language, PJM [*Polski Język Migowy*, 'Polish Sign Language'] is undoubtedly a complete language system. It has all the characteristics of natural language: (1) duality, that is, double articulation, (2) productivity, (3) arbitrariness, (4) the ability to interchange roles of the sender and the recipient, (5) specialisation, (6) displacement and (7) cultural transmission. This last characteristic of PJM occurs in the process of socialisation, education and upbringing; using PJM the deaf transmit from generation to generation cultural values, customs, and behaviour patterns at their disposal.

Visibly, Hockett's design features were meant to describe human spoken languages as contrasted to animal languages. Yet the proposal has met with criticism, on the grounds that some animal species are capable of communicating in a similar fashion to humans (Coleman 2006) and that it does not take into consideration sign languages. Interestingly, as we shall see, some artificial languages do not differ significantly from ethnic spoken languages, if assessed according to Hockett's criteria.

	univer- sal	IALs	Espe- ranto	conlangs/ artlangs	revived	re- constr.	con- trolled	diverse formal	sign
vocal- auditory	?	yes	yes	yes	yes	yes	yes	no	no
broadcast and direc- tional reception	yes	yes	yes	yes	yes	yes	yes	no	yes
transitori- ness	?	yes	yes	yes	yes	yes	yes	no	yes
inter- change- ability	yes	yes	yes	yes	yes	yes	yes	?	yes
complete feedback	yes	yes	yes	yes	yes	yes	yes	?	no
specialisa- tion	yes	yes	yes	yes	yes	yes	yes	yes	yes
semanti- city	yes	yes	yes	yes	yes	yes	yes	yes	yes
arbitrari- ness	?	yes	yes	yes	yes	yes	yes	yes	yes
displace- ment	yes	yes	yes	yes	yes	yes	yes	?	yes
prevarica- tion	?	yes	yes	yes	yes	yes	yes	?	yes
reflective- ness	yes	yes	yes	yes	yes	yes	yes	yes	yes
tradition	no	?	yes	no	yes	no	no	no	no
learnability	no	yes	yes	yes	yes	yes	yes	?	yes
discrete- ness	yes	yes	yes	yes	yes	yes	yes	yes	yes
producti- vity	yes	yes	yes	yes	yes	yes	?	no	yes
duality	yes	yes	yes	yes	yes	yes	yes	yes	yes

Table 2 Hockett's design features applied to artificial languages (? marks problematic areas)

Before the table is discussed, we should take a closer look at some of the design features and their interpretation. Several of Hockett's design features of language do not in fact pertain to language but to species. If *learnability* is understood strictly along the lines of Hockett, that is to say, as human faculty to learn languages, then probably all languages on the list are learnable – meaning that humans have the ability to learn any other language than their own, the language's complexity notwithstanding. However, if it is to be understood as a learnable language, then the only language type on the list which is not learnable but rather potentially memorisable is universal languages (although their rules might be learnable). The same could be said about *transmission*. As Coleman (2006) explains, “the meaningful details of the communication system are not instinctive, but are learned from other members of the species”. All human languages are learned in culture. However, if the transmission is to be understood sociolinguistically as actual generation-to-generation transmission, only some languages should be treated as meeting the requirement (e.g. Esperanto but not formal languages, which have to be taught to children at school; cf. Lyons' environmental acquisition and class Nat₃L in section 3.2.2). *Semanticity* is, according to Lyons, a vaguely defined category (1977: 79f.), too general to be useful. As such, it applies to all human languages. As inherent features of human language, I shall also consider *specialisation*, *discreteness* and *duality*.

All the features apply to two (types of) languages: Esperanto (being transmitted culturally and learnable to the same degree as ethnic languages) and revived languages, which, although they stopped being used at some point of time, are now back in use. The languages fulfilling the least criteria are formal and universal languages. I shall discuss Table 2, beginning with languages having the most features marked with a “yes” (Esperanto and revived languages being treated as natural, and therefore excluded).

International auxiliary languages seem to fulfil all the requirements of human language. The only controversial feature is cultural transmission. If understood as “the ability to speak a particular language [...] passed on from one generation to the next by teaching and learning, rather than by instinct” (Lyons 1977: 82), it applies to IALs to the same extent as it does to natural human ethnic languages. However, if it is understood as transmission to children in the natural process of acquisition, it does not apply, mainly because of the scarce use of IALs. It may be assumed that if more people spoke e.g. Interlingua, the language could be nativised just as Esperanto has become.

Conlangs and artlangs are grouped together as being languages created mainly “for fun”. Their situation is very similar to IALs, that is, they are spoken to a limited degree. Some of them are useable (Klingon, Dothraki etc.), which is testified to by various webpages and meetings. However, they are not transmitted but learned, and the only attempt to create a native speaker of Klingon failed (possibly because of the very limited vocabulary of Klingon at that time and the need to excessively paraphrase; the father explains that the child eventually stopped responding to Klingon as “he didn’t enjoy it”³⁶).

Likewise, if language reconstructions are treated as approximations of a language, that is, if it is assumed that they were spoken at some point in the past, they fulfil all of the criteria set by Hockett apart from transmission as understood by Lyons (1991). If they are treated as abstract constructs not used in everyday communication, some of the features do not apply (vocal-auditory channel, rapid fading and to some extent learnability).

Cultural transmission poses a problem also when applied to controlled languages (manually coded languages in-

³⁶ <http://www.todayifoundout.com/index.php/2012/08/a-man-once-tried-to-raise-his-son-as-a-native-speaker-in-klingon/>

cluded). Evidence shows that teaching children a signed³⁷ language results in creolisation, and that the language is never acquired in full. As Farris (1994: 16) writes, signed languages are “inherently unstable” and “even when they form the only sign input given to Deaf children, the children rapidly modify them in the direction of primary sign languages”. This may be because of the rigidity of structures resulting in low language efficiency. Similarly, it may be assumed that teaching a child controlled English will result in the child abandoning the variety for the sake of the full vernacular. Their productivity is also debatable. Clearly, like all human languages, they must be open to accommodate new meanings. However, they have strict rules, not allowing for full accommodation; for example, Basic English has only 850 words and new meanings are first made to fit the system (‘ornament for ear’ instead of ‘earring’) and only if it is not possible to convert them into Basic, they might be added to the list of international scientific vocabulary or left as they are in square brackets with their Basic English equivalent supplied (Large 1985: 169–172). Words for which there is no sign in signed languages are commonly finger-spelled. To what extent they are then productive is a matter of discussion. Moreover, manually coded languages, just like sign languages, are not vocal-auditory and lack total feedback.

The most challenging types are universal and formal languages. Their heterogeneity makes them difficult to classify. Universal languages may be divided into pasygraphies and pasylalies. The former are obviously transmitted through the visual channel. The latter may be assumed transmittable through the vocal-auditory channel; nonetheless, they do not have a speech community, and therefore this feature does not apply to them. The controversial status of transitoriness is a direct consequence of the vague status of the previous fea-

³⁷ A **signed** language, i.e. a manually coded artificial version of an oral natural language is not to be confused with a **sign** language, i.e. a naturally developed gestural language (see also section 3.3.3).

ture. Arbitrariness and prevarication depend on whether the philosophical assumptions of particular languages and the beliefs of their authors are deemed valid. If it is assumed that the language mirrors the reality and the true order of things, then it is not only iconic to a large extent, but also it is impossible to make false or meaningless statements in the language. However, if we assume – according to our present knowledge – the impossibility of such a language, both features are valid. Tradition and learnability, as interconnected, may both be excluded from the list of features applicable (see above).

Formal languages include two main classes: formal languages of logic and mathematics and programming languages. These classes, too, are heterogeneous. The first three features clearly do not apply to computer languages (directional reception does but broadcast transmission does not). Interchangeability is, however, a matter of philosophical assumptions. If the active subjects of a programming language are considered to be a human being and a computer, then a computer may take the role of sender only if pre-programmed as such. If, on the other hand, two machines are taken as the subjects (assuming they are both pre-programmed to be active in communication, i.e. understand the communication protocol), their species-specific traits have to be taken in consideration; namely, if they are both of the same architecture, they may freely interchange roles. However, if one machine is of specialised architecture and can only perform specific tasks, and the other is of the versatile type, then their communication may only be interchangeable within the scope of the former's tasks. This again confirms the view that some of the features are rather species-specific than pertaining to language itself. The subsequent generations of machines (here: a generic name for both programmable machines and newest-generation computers) are comparable to different species, and their evolution to the development from primates through hominoids to humans. Complete feedback is once more a question of the subject. If the subject is a human, total feedback applies as in any other

language. If, however, the subject is a computer, feedback only applies insofar as the computer is taught to monitor its “utterances”. Prevarication is not possible for computers; however, humans can produce meaningless statements in programming languages. Reflectiveness is obviously not possible for computers, as they are not conscious, living beings. Only humans can reflect on what they have produced. This feature may apply to programming languages only if it is accepted that reflexivity pertains to speaking about language in a language, and not to conscious reflection. In several programming languages, it is possible to metaprogram, that is, use programming code as data (a property known as ‘reflection’). Learnability is a faculty of humans. However, versatile architecture computers are able to “learn” languages if pre-programmed as such. Again, the comparison to species comes to mind. Productivity in general does not apply to formal languages because essentially they are closed languages.

Let us now discuss in some detail other formal languages, namely the language of mathematics³⁸ and logic (chiefly zero- and first-order calculus). They are not vocal-auditory but rather written, and therefore don’t fade rapidly. Because the active subjects of the languages are humans, the features of interchangeability and feedback apply to those formal languages. Displacement in both programming and formal languages is a matter of philosophical discussion, although in mathematics it is possible to describe formally such abstract objects as a Calabi-Yau manifold. Prevarication is pointless in formal languages, but, as a feature inherent to semiosis, might possibly take place in them. Traditional transmission and learnability are only possible if understood as cultural acquisition, not from-birth acquisition.

The design features for formal languages are summed up below in Table 3. It is visible that the type that is the farthest removed from what is usually understood as natural human language is

³⁸ The discussion concerns formalised parts of mathematics.

computer/programming languages. The tables compiled in this section seem to confirm the assertions of Baron (1994), Lyons (1991) and Schubert (1989) (see also other sections in this chapter).

	com- puter	formal
vocal-auditory	NO	NO
broadcast and directional reception	NO	yes
transitoriness	NO	NO
interchangeability	?	yes
complete feedback	?	yes
specialisation	yes	yes
semanticity	yes	yes
arbitrariness	yes	yes
displacement	?	?
prevarication	?	?
reflectiveness	?	yes
traditional transmission	NO	NO
learnability	?	yes
discreteness	yes	yes
productivity	NO	NO
duality	yes	yes

Table 3 Hockett's design features in formal languages (? marks problematic areas)

3.2 The scale of naturalness

3.2.1 The continuum of deliberate influence

As shown in section 2.3, planning can range from changes in one aspect only (graphisation, standardisation of orthography) through standardisation and modernisation (introduction of a standard form, corpus planning) up to revitalisation and planning of a whole language system (cf. Barandovská-Frank 2003; Ferguson 1968; Koutny 2009). It might therefore

be assumed that there is no binary opposition of natural vs. artificial, but instead there is a continuum of “deliberate influence” as presented in Schubert (1989: 22) and Koutny (2009: 118), that is, from an ideal “untouched” ethnic language to an artificial *a priori* system (Figure 10).

The idea of a scale of artificiality is not, in fact, so new. The proposal comes from E. Svadost (1968: 6.2), who divided languages into five classes, according to level of artificiality (i.e. the amount of deliberate influences):

1. [...] языки бесписьменные или речевая стихия бесписьменных говоров национального языка, исторические языки и наречия до или вне их нормализации, до или вне литературных норм;
 2. [...] языки нормализованные – национальные литературные;
 3. [...] проекты международного, интернационального языка, пока не нашедшие применения на практике или нашедшие его экспериментально, созданные на основе языкового опыта человечества, на материале исторических языков (апостериорные языки);
 4. [...] проекты международного языка, созданные в отрыве языкового опыта человечества на основе философской классификации понятий и буквенной символики (априорные языки).
 5. Языки кибернетические, математические языки-коды, а в простейшем виде – всякие коды [...] можно было бы назвать языками пятой степени искусственности (LA-5), но слово *язык* здесь употребляется уже в ином, чем обычное, значении. Такой язык разговорным, устным стать не может даже экспериментально.
-
1. [...] unwritten languages or spoken element of unwritten dialects of the national language, historical languages and dialects before or beyond their normalisation, before or beyond their literary norms;
 2. [...] normalised languages – national literary ones;
 3. [...] international language projects until applied in practice or applied experimentally, created on the basis of lin

- guistic experience of mankind, on the material of historical languages (*a posteriori*);
4. [...] international language projects created in isolation from the linguistic experience of humankind on the basis of a philosophical classification of concepts and on letter symbolism (*a priori*).
 5. Cybernetic languages, mathematical language-codes, and more simply – all sorts of codes [...] could be called languages of the fifth degree of artificiality (LA-5), but the word ‘language’ is used here in a meaning different from usual. Such language cannot be spoken even experimen- tally.

As can be seen, the scale is based on the amount of planning each of the languages undergoes. A similar view that types of language can be shown as a spectrum (Figure 8) is presented by Baron (1994). He remarks that functionally computer and formal languages are in fact sublanguages “since they are designed to operate over highly restricted syntactic and semantic domains”. Notation schemes (e.g. calculus) can be assumed to be languages only because they contain symbols and rules for combination.

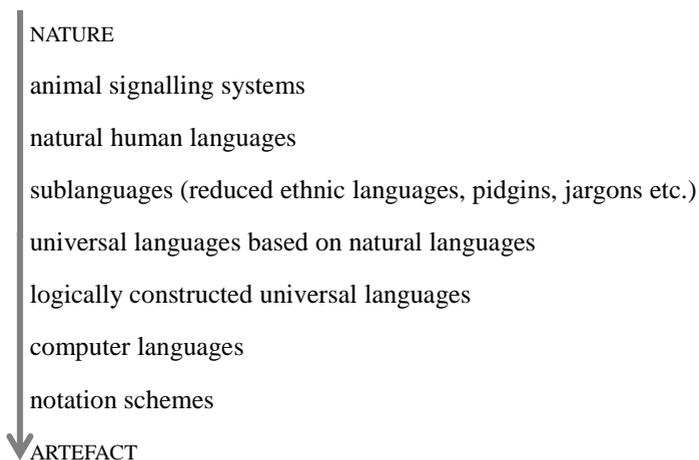


Figure 8 Baron's natural language spectrum (1994)

In the subsequent examination, the concept of *Ausbauisation* discussed by Gobbo (2012: 186f.) might be useful – although his general classification is erroneous (see section 3.2.2). He states that if a language is to be understood as *langue* and not as mere *parole*, it has to have a graphic representation and a speech community identifying itself by means of the language. It is obvious that languages nearer the ‘natural’ pole are classified as such rather due to the existence of the community (because many small ethnic languages do not have standard orthography); while those in the opposite position do not have a community and planning factors play the most important role. Thus, *Ausbauisation* understood as language being a vehicle of identity distinguishes human (near-)natural languages from formal and formalised languages.

Interestingly, the feature often connected with naturalness of languages, namely the existence of native speakers, has only been taken up once previously in section 2.3.3. If one were to evaluate languages according to this feature, the scale would look different; pidgins and *Dachsprachen* would move further down towards the ‘artificial’ pole of the scale, whilst Esperanto would move up as it has about two thousand (bi- and multilingual) native speakers.

Artificial languages themselves can be classified according to their assumed “naturalness”, i.e. imitation of ethnic languages. Janton (1993: 6) observes that “the classification of planned languages takes as its starting point the distinction between *a priori* and *a posteriori* languages – that is, between the tendency to schematize and the tendency to imitate or refer to natural languages” (see also section 2.3.1).

Similarly, Baron (1994) discerns different types of computer languages according to their schematicity (or degree of artificiality):

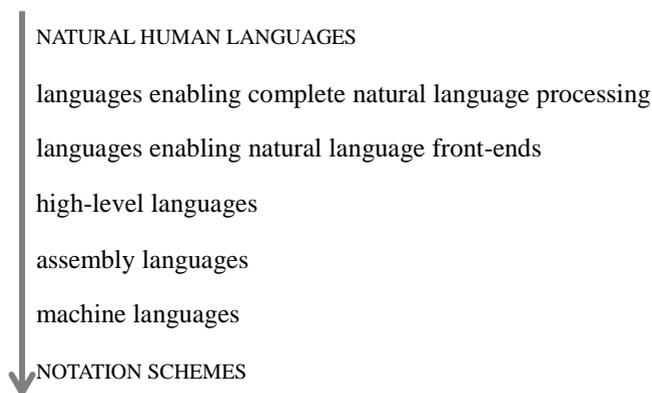


Figure 9 Baron's computer language spectrum (1994)

Computer languages have evolved from notation schemes, through high-level languages (from structured programming languages such as FORTRAN to object-oriented programming languages) to programs using natural language processing, thus mirroring in reverse the history of human language evolution.

Figure 10 presents a scale³⁹ for artificial and natural languages devised partly on the basis of Koutny 2009: 118, Sakaguchi 1998: 26–28, Schubert 1989: 22 and Svadost 1968: 6.2. Ethnic non-standardised languages are considered the most natural, serving as a vehicle for identity and all naturally evolved sign languages (only some being partly standardised). Languages move further to the artificiality pole depending on how many changes and what degree of planning is involved in their development. And although pidgins are unconsciously, or rather linguistically naïvely created, the changes that have brought them into being (i.e. merger of two languages or relexification of one of them) are so vast that they cannot be ignored. Thus, pidgins cannot be placed together with creoles or non-standardised languages.

³⁹ The first version was presented in Stria 2013.

At the same time, their development is not regularised, in contrast to such languages as Nynorsk or Standard Arabic. The most artificial languages are given as fully formalised languages of logics.

The scale may cause controversy because it contains both individual languages and language groups in various stages of development. It is able to take into account their historical development. However, many of these languages (in particular revitalised languages) would have to change their position depending on the stage in their development. It would be difficult to determine what point in time should be taken as the starting point – whether distinct centuries, the beginnings of a language’s existence etc. Temporal changes have been therefore excluded, to simplify the assessment of the overall development of a language.

Interestingly, similar debates are ongoing in the fields of biology and ecology. As Machado (2004) remarks, naturalness is opposed to artificiality, i.e. human influence on ecosystems. Thus, this binary division could rather be presented as naturalness vs. anthropisation. It is, however, difficult to determine if human influence is, in fact, artificial. For some, the starting point of artificiality is modern technology, while aboriginal influence in ecosystems is accepted as natural; for others it is the beginning of agriculture, while yet others say that humans are part of nature and as such all their activities are also natural (Machado 2004: 95f.).

Naturalness has been used as a conservation value and a state descriptor. In this book, I would like to avoid the – rather philosophical – discussion of whether natural languages are “better” than artificial languages, thus not taking naturalness as a value. However, it will be treated as a qualitative descriptor of the human influence on languages.

IDEAL NATURAL LANGUAGE	
Nanai, early creoles, ASL	small non-standardised ethnic languages and creoles, partly standardised sign languages
Bislama	standardised creoles
Standard German	standard literary languages, <i>Ausbausprachen</i>
Latin	dormant classical language
Korean, Hungarian	small changes (often in one aspect only)
pidgins	extensive semi-deliberate changes
Nynorsk, Indonesian	extensive changes in many aspects; far-reaching planning
Literary Arabic, Rumantsch Grishun	highly regularised <i>Dachsprachen</i>
Sanskrit	highly regularised classical language
Modern Hebrew, Cornish	revival ('reinvention')
Basic English	reduced ethnic languages
Proto-Indo-European	linguistic reconstructions
Occidental, Interlingua	naturalistic <i>a posteriori</i>
Esperanto	schematic <i>a posteriori</i>
Volapük, SJM	mixed systems based on ethnic languages
Loglan	mixed system statistically derived from ethnic languages
Solresol	<i>a priori</i>
programming languages	formalised systems based on ethnic languages
predicate calculus	formal languages
ARTIFICIAL LANGUAGE	

Figure 10 Scale of artificiality/naturalness

3.2.2 Lyons' classification

The problem with a precise definition of 'artificial language' is that the notion of 'language' itself is not precise enough. In Romance languages, there is a distinction between a language (*langue, lingua*) and a language-system (*langage, linguaggio*). Therefore, the tacit assumption that ethnic languages are natural languages and formal languages are not stems probably from the fact that the English 'language' is often understood as 'langue' and not 'langage' (cf. Lyons 1991: 49–52). The difference between the two terms is explained in Albani & Buonarotti (1994: 19):

Per i termini «lingua» e «linguaggio» ci siamo attenuti alla distinzione classica – rintracciabile anche nel Nuovo Zingarelli (1993) – che indica nel primo «un sistema grammaticale e lessicale per mezzo del quale gli appartenenti ad una comunità comunicano tra loro», mentre nel secondo 1) «la facoltà di esprimersi mediante l'uso di determinati segni, gesti, oggetti, simboli e simili cui l'uomo attribuisce particolari significati» (linguaggio del corpo, linguaggio dei fiori, ecc.) ed insieme 2) «un particolare modo di parlare di determinati individui e ambienti» (linguaggio dei bambini, linguaggio dei sordomuti, ecc.) e 3) «un sistema di segnali per mezzo dei quali gli animali comunicano tra loro».

As for the terms “lingua” and “linguaggio” we followed the classical distinction – also to be found in the New Zingarelli (1993) – that denotes the first one as “a grammatical and lexical system by which members of a community communicate with each other”, whereas the second one as 1) “the faculty to express oneself through the use of certain signs, gestures, objects, symbols and the like which one attaches special meanings to” (body language, the language of flowers, etc.) along with 2) “a particular way of talking of certain individuals and environments” (children’s language, sign language, etc.) and 3) “a system of signals by which animals communicate with each other”.

Throughout this book, ‘language’ is used in the widest sense, namely, it comprises all language-systems created by humans.

A classification of language-systems is given by Lyons (1991). According to him, there is a four-class division of naturalness:

- Nat₁L – systems conforming to nature
- Nat₂L – ‘species-specific’; *species*, i.e. biological classes and artefacts
- Nat₃L – acquirable in the normal process of maturation/socialisation (e.g. sign languages)
- Nat₄L – “conforming to the linguist’s expectations about what is normal or typical” (Lyons 1991: 61).

It is unfortunately unclear what the specific difference between the Nat₁L and Nat₂L classes is, especially when, under naturalness₁, Lyons (1991: 58) writes:

For present purposes, I am going to select another interpretation of ‘conformity with nature’: I am going to interpret it as meaning “being constrained by the laws of nature”. [...] I am going to interpret this gloss in a rather particular sense: a natural₁ language, I will say, is one that could be used by organisms or devices that are subject to the laws of nature.

Later on, under naturalness₂, he defines Nat₂L as “constrained by nature – i.e. constrained by the physical, or psychophysical, make-up – of the organism or device using it” (1991: 59). What is important though, is that the Nat₂L class is a subclass of Nat₁L, i.e. Nat₁L \supset Nat₂L. The issue of the inclusion of Nat₃L in Nat₂L or vice versa remains unresolved. The last class, i.e. Nat₄L does not seem to be contained in any of the previous three classes.

Before the membership of particular languages to each class is discussed, the notions ‘non-natural’ and ‘unnatural’ should be reviewed. For Lyons ‘non-natural’ means “constructed” or “conventional”, while ‘unnatural’ represents a more pejorative sense

of artificiality and awkwardness (see Gobbo 2012: 185; Lyons 1991: 54, 68). In this interpretation, any language-system created (wholly or partly) by humans is non-natural, which in fact supports the view presented in the previous section. A language can be non-natural without being unnatural, as well as unnatural in one class (e.g. in Nat_3L), while being natural in a different class at the same time (e.g. in Nat_2L).

From the definitions given by Lyons, it follows that any language possible on Earth is natural₁. It is difficult though to conclude what kind of languages belong to class Nat_2L . He states that formal languages are clearly natural₁, but not necessarily natural₂, although some semanticists of the 70s would argue that, on the contrary, “a formal language structurally comparable with the propositional calculus and combined with the first-order predicate calculus [...] is indeed natural₂” (Lyons 1991: 69).

A table with diverse languages distributed according to the class is presented by Gobbo (2012: 188; Table 4), where \top means natural, \perp means unnatural and $[\perp]$ means non-natural. As can be seen, Gobbo claims that sign languages, child speech, pidgins and formal languages are unnatural₁. This results clearly from a mistake⁴⁰ made on pp. 183-184, where he wrongly writes that $\text{Nat}_2\text{L} \supset \text{Nat}_1\text{L}$. The reasoning following from this mistake is that Nat_1L would be a class containing languages commonly referred to as natural, i.e. *langues/lingue*. To account for small ethnic languages in this class, Gobbo discusses the concept of *Ausbauisat*ion (2012: 186f.). He states that for a language to be part of the Nat_1L class a graphisation and a distinct speech community (“where the language itself is a vehicle of identity”, Gobbo 2012: 187) are necessary conditions. Therefore, languages that cannot be considered as *Ausbausprachen* are marked \perp in class Nat_1L . Of course, in

⁴⁰ F. Gobbo acknowledges his error as resulting from a typing mistake (private conversation).

light of what has been said previously about the inclusion of the classes, this categorisation does not hold.

in English	Esperante	Nat ₁ L	Nat ₂ L	Nat ₃ L	Nat ₄ L
Urdu, Chinese, etc.	la urdua, la ĉina, ktp	T	T	T	T
Neapolitan, Cornish, etc.	la napola, la kornvala, ktp	T	T	T	T
Latin, Sanskrit, etc.	Latino, Sanskrito, ktp	T	T	⊥	T
Esperanto, Ido, Interlingua	Esperanto, Ido, Interlingua	T	T	T	⊥
Volapük and similar	Volapuko kaj similaj	T	T	⊥	⊥
sign languages	signolingvoj	⊥	T	T	T
child speech	infanaj variaĵoj	⊥	T	T	[⊥]
pidgins	piĝinoj	⊥	T	T	[⊥]
creoles	kreoloj	T	T	T	T
predicate calculus, etc.	predikata kalkulo, ktp	⊥	T	⊥	[⊥]
BASIC, Python, etc.	BASIC, Pitono, ktp	⊥	⊥	⊥	[⊥]

Table 4 Languages and their naturalness. Reprinted from Gobbo 2012: 188

Furthermore, Gobbo does not recognise computing artefacts as being able to use language, “without a specific action by human beings” (2012: 184). However, he claims that formal languages should be split into two classes, namely that of computational languages and that of non-computational⁴¹ ones, and that at least one of them could be natural₂. He argues that computational languages such as programming languages, being equivalent to the Universal Turing Machine, are clearly unnatural_{1,3} as well as unnatural₂ because computers are not species. Bearing in mind the mistake made by Gobbo earlier, one has to argue against this view and grant naturalness₁ also to these kinds of language-systems. The second class mentioned contains predicate calculus as a generic name for mathematical abstract models of artificial intelligence as described by Turing in 1950 (Gobbo 2012: 188-190). This class, according to Gobbo, could be granted naturalness₂.

Let us discuss Gobbo’s table (Table 4) in view of what has been established thus far. Class Nat₁L should actually contain only T, because it is the class of all languages conforming to nature (although it is not certain what kind of languages could be unnatural here). Class Nat₂L seems to include one exception, namely programming languages; they alone could be marked as unnatural. However, if Lyons’ suggestion that computers are artefacts or pseudo-species is accepted, those languages too are natural₂. Of more interest are classes Nat₃L and Nat₄L.

Gobbo writes that dead languages are not natural₃ anymore. This is obviously true; nevertheless, it should be noted that such languages as Latin belonged to class Nat₃L once and that the continuity of revived languages was broken, and so their naturalness₃ must be considered taking into account a specific period. Another problem is the naturalness₃ of pidgins. If naturalness₃ (understood by Lyons as environmental acquisition) is to be granted to pidgins, then in extreme cases

⁴¹ It is however unclear what exactly is meant by those names.

any language learnt at school could be deemed natural₃, including predicate calculus and programming languages. Pidgins are not and cannot be native languages (unless they are creolised) and they are very often acquired semi-consciously. However, as naturalness₃ is a gradable property, pidgins should be described neither as fully natural₃ nor as fully unnatural₃. However, if naturalness₃ is granted to sign languages, pidgins might then be judged as natural₃ to a lesser degree.

A notable difference between Nat₄L in Gobbo's table and Table 5 in this section is that here this class is divided in two. One reason is that, as it was said before, non-naturalness does not exclude the unnaturalness or naturalness of a language. Another reason is suggested by Lyons himself, who claims that all languages that are the product of human construction are non-natural (1991: 71). This implies that naturalness₄ not only is a gradable property and that it depends on the views of researchers, but also that it should be assessed according to the "genetic" source of the language. Therefore, the first Nat₄L column of Table 5 gets a ⊥ if the language type conforms to the definitions of 'artificial language' given in 2.1 and a T if it has evolved in a speech community and has no particular creator (Malmkjær 2002). This column corresponds to a certain degree to what has been described in 3.2.1. The second column, i.e. Nat₄L (2) contains exclusively the "expectations" of linguists. The number sign (#) marks controversy and the asterisk (*) abstract grammar descriptions.

name	Nat ₁ L	Nat ₂ L	Nat ₃ L	Nat ₄ L	Nat ₄ L (2)
Urdu, Chinese, etc.	T	T	T	T	[⊥]*
Neapolitan, Cornish, etc.	T	T	T	T	*

Latin, Sanskrit, etc.	T	T	T/⊥	T	*
Esperanto, Ido, Interlingua	T	T	T	⊥	[⊥]
Volapük and similar	T	T	⊥	⊥	[⊥]
sign languages	T	T	T	T/⊥	[⊥]#
child speech	T	T	T	T	[⊥]
pidgins	T	T	T	T	[⊥]#
creoles	T	T	T	T	*
predicate calculus, etc.	T	T	⊥	T/⊥	[⊥]#
Basic, Python, etc.	T	T/⊥	⊥	⊥	[⊥]

Table 5 Languages and their naturalness II. Modified Gobbo 2012: 188

Two instances are marked T/⊥ in Nat₄L, namely sign languages and formal languages. Formal languages may seem natural₄ or unnatural₄ depending on the philosophical beliefs of researchers and the specific context of their creation (cf. the discussion about Nat₂L class). If they are assumed to mirror the structure of *langues* or the structure of human thinking (whatever the interpretation), then they are clearly natural₄. But if the view that they are constructed and not discovered is taken, those languages should be marked as unnatural₄. Sign languages present even more trouble. Lyons indicates that sign languages are probably natural to the same degree as *langues*; they are surely natural_{1,3} and perhaps a little less natural₂. A problem occurs when such systems as ‘signed languages’ are considered. For example, in Poland there are two sign/signed languages: Polish Sign Language (PSL), i.e. the natural language

of deaf people and, simultaneously, Signed Polish, i.e. the gestural mixture of PSL and written Polish. Such languages as Polish, American or British Sign Language, not to mention Nicaraguan Sign Language, which has come into existence relatively recently, are natural₄, while Signed Polish or German are constructed, unnatural₄ languages (on the difference between sign and signed languages see below and section 3.3.3).

Another important point of discussion here should be the non-natural class of so-called quasi-N-languages (QNLs). Lyons (1991: 69) wants it to include “all those (more or less unnatural₄) languages which may be constructed from attested N-languages [langues] by deliberately changing one or more of their structural properties”. This class contains, among others, child languages. As can be seen, they are marked T both in Table 4 and in Table 5. This stems from the fact that it is hard to agree with Lyons that child languages are conceived on purpose. As Gobbo (2012: 186) notices, they are “notably unplanned and creative” and therefore should be counted as natural₄ but still non-natural₄. A similar situation concerns pidgins, whose naturalness is discussed in section 3.3.1.

The last column has a [⊥], where a language type is considered non-natural. The first row is marked with an asterisk because, as Lyons writes, abstract grammar descriptions and standardised language-systems satisfy the definition of non-naturalness. According to this line of reasoning, other systems could be marked as non-natural as well, namely, all those that are the product of human creation; therefore there is an asterisk (*) beside all the systems that otherwise would be considered natural. The number sign (#) is placed beside sign languages and pidgins because of the controversy around them. Researchers who study those languages would surely call them natural. Formal languages are less obviously natural₄ though. This, of course, poses no further problems as it has been said that naturalness₄ is gradable and largely depends on the linguist’s expectations and philosophical beliefs.

It is now visible that the proposal of Lyons includes all those types of languages which Blanke enumerated (1997), some of which normally are not considered artificial (cf. the definitions in 2.1 and the scale of naturalness in the previous section), that is, standard languages, highly regularised ethnic languages such as Sanskrit or Israeli/Modern Hebrew, planned languages, scientific notation, programming and machine languages.

3.3 Borderline cases⁴²

3.3.1 Pidgins and creoles

Among the languages that stand somewhere between the poles of artificiality and naturalness (Figure 10) are pidgins and creoles, collectively called ‘contact languages’. To defend the view that they are borderline cases between artificial and natural, the languages should be first compared with ethnic languages. On the basis of Duličenko (1989) and Liu (2001) an ethnic language can be described as a basic conventionalised system that expands over time to new domains, is linked to a certain, mostly monolingual population in a given area and requires constant interaction between speakers to be passed on to the following generations. Out of these criteria of naturalness, pidgins, creoles and planned languages fulfil only some, each one matching a different number of those criteria. As can be seen in Table 6, pidgins share several features with planned languages and several with natural ones, although in terms of artificiality they might be regarded as “almost natural” because their coming into existence is through an unconscious, linguistically naïve process with no particular creator. The latter can also be said about creoles. In fact creoles in terms of Lyons’ classification are not distinguishable from “natural” ethnic languages, but pidgins could be so in terms of naturalness₄ (see 3.2.2).

⁴² This section is partly based on Stria 2013.

feature	ethnic	creole	pidgin	planned/ IAL
development to new domains	yes	yes	no (unless it changes into expanded pidgin)	yes if socialised
linked to a specific population	yes	yes	yes, to several	no
linked to a territory	yes	yes	yes	no
first language and monolingual native speakers	yes, generally	partly	no	no (Esperanto has native speakers, but none are monolingual)
purpose: international communication	no (may be used as such)	no (may be used as such)	yes	yes
conscious planning	no (yes in later stages of development)	no (possibly in later stages of development)	no	yes (not so strict in later stages of development)
known creator	no	no	yes, identifiable population	yes
developed over a relatively short period of time	no	yes	yes	yes

Table 6 A comparison of pidgins and creoles with ethnic languages and IALs

However, it is important to notice that even languages traditionally regarded as natural do not always meet all the criteria. There are ethnic languages on the verge of extinction which do not expand to new domains. Of course, one can argue that historically they did expand and stopped doing so under unfavourable circumstances. Still, many small languages are used only in the basic domains (i.e. day-to-day conversations, religious ceremonies) and never have the chance to develop to be used in fields such as science (in such cases, English is often used instead). When it comes to a specific, more or less delimited territory, diaspora languages need to be mentioned. Clearly, they are natural, although not tied to any particular place. But also here, there could be the following reservation: most diaspora languages have their “homeland”, a territory where the language is constantly spoken. Likewise, monolingual native speakers are not a very good criterion – many communities are at least bilingual.

In general, in Figure 10 (section 3.2.1) contact languages are placed separately, namely, creoles – depending on the degree of development – in the same place as maximally natural languages (because the only essential difference between the two is their source), and pidgins between languages where conscious planning in one major aspect has been involved (Hungarian), and those where extensive vital changes have been introduced (Nynorsk). The modifications of the languages serving as the basis for pidgins are vast, and therefore pidgins should be placed on the scale after such languages as Hungarian, but at the same time those changes were not consciously introduced by any particular individual or institution, as in the case of Nynorsk. Yet, the alterations are described as “semi-deliberate” (Figure 10), because pidgins arise as an answer to the need of a common communication language.

Lindstedt (2006) lists three characteristics of natural languages: the existence of non-codified norms, spontaneous changes (e.g. bound morphemes become free lexical morphemes; more synonymous and polysemous forms) and native

speakers (at least bilingual; similar to e.g. Finnish Romani). According to these criteria, pidgins are not fully natural – their norms are not codified and therefore allow for changes; these changes, though, may form an expanded pidgin or even a creole, if native speakers emerge. However, a pidgin is never a native language, unless it changes into a creole. Esperanto, in turn, is a natural language if Lindstedt's criteria are applied – there are norms possible to learn only in the community, the language develops spontaneously and has a small number of native speakers.

Another similarity between IALs and pidgins is that both groups are created on the basis of several natural languages (IALs with some *a priori* elements). This is observed by Jurkowski (1986: 122):

Język pidżyn, mimo że powstał jako zmieszanie się co najmniej dwóch języków naturalnych, nie jest językiem rodzimym dla nikogo. W pewnym sensie podobny jest on do esperanta. I to nie tylko ze względu na międzynarodową funkcję, jaką spełnia, ale także w sensie genezy – wyrósł on, podobnie jak esperanto, na bazie języków żywych, naturalnych.

The pidgin, although it has come into existence as a mixture of at least two natural languages, is not a native language to anyone. In some sense it is similar to Esperanto⁴³. And this is not only because of its international role but also in regard of its origins; it has developed, similarly to Esperanto, on the basis of living, natural languages.

Corsetti (2010: 374) writes that pidgins, creoles and mixed languages undergo relexification (which can be regarded as the main factor in the creation of pidgins). The same phenomenon concerns IALs, i.e. to an already complete grammatical system, a completely new lexicon is introduced from (a) superstrate language(s).

⁴³ On the native speakers of Esperanto, see chapter 4.

The simplicity of pidgins may also be compared to that of reduced ethnic languages, e.g. Basic English (cf. Sakaguchi 1998: 210–214). Pidgins are characterised by a limited scope of vocabulary and style, simplified phoneme inventory and simplified grammatical structures (loss of inflection and verb conjugation, reduced tense system etc.). Reduced ethnic languages are controlled in terms of vocabulary and style, maintaining the same phoneme inventory and grammatical structures as in the source languages. However, some grammatical constructions may not be recommended or may be restricted in use.

It is worth noting that a more or less precise moment of their birth is known. Therefore, according to some of the definitions from 2.1, pidgins could be considered as (at least to some extent) artificial.

3.3.2 Language revitalisation and revival⁴⁴

In the case of modern standard varieties of ethnic languages, as Duličenko (1989: 53) puts it, a “continuous a-posteriorisation of heterodialectal, but nevertheless monolingual material” can be found (see also Schubert 1989: 9). A similar situation applies to revitalised languages. The material undergoes modernisation and later standardisation as a result of the deliberate efforts of identifiable (groups of) people. Those languages undergo such vast changes that the historical continuity often assumed as required to deem a language as “natural” is broken and the older as well as the newer form of the language might be considered as two separate, although related languages, as in the case of Biblical and Modern Hebrew (Romaine 2011: 186f.). Thus, Romaine (2011) proposes a new label for such languages: ‘reinvented’.

⁴⁴ “Revived” will here refer to those vernaculars whose historical continuity is broken and has to be rebuilt anew (Hebrew, Cornish). “Revitalised”, in turn, are those vernaculars that have been continuously spoken, although only by a handful of speakers (Māori, Welsh).

Languages undergoing revitalisation and revival (minority languages especially) show many sociolinguistic similarities to international auxiliary languages, in particular Esperanto (Kimura 2010; Kimura 2012; Krägeloh & Neha 2014; Romaine 2011; Wood 1979; see also section 4.2). They are used in bilingual or multilingual communities and their use is mostly restricted to everyday informal situations, in contrast to majority languages regarded as a prestigious variety. Furthermore, they might be considered as non-territorial since there are many scattered language islands on the territory of the majority language. There is a strong group identity and language loyalty among the speakers. The sense of belonging to the community develops because of prejudices against the language. Artificial languages have been looked down on by linguists, but their usage has also been persecuted as in the case of minority languages (e.g. Lins 1988; Rónai 1969).

Other similarities have been noticed by Kimura (2009; 2010; 2012) and Wood (1979). Being a member of such a community is often, surprisingly, voluntary. Of course, most people are born in the community, but abandon the language, as in the case of Cornish, where 50% do not speak the language. Those that speak it are not necessarily born Cornish or otherwise (Kimura 2010, Romaine 2011). In the case of Modern Hebrew/Israeli the voluntariness stems from the fact that people speaking different mother tongues willingly abandoned their L1s for the sake of the new community. Therefore, those communities are non-ethnic. It is worth noting that most of these so-called “natural” languages do not in fact have any native speakers. The speakers of rural varieties of Breton use a different version of the language than do the speakers of the revitalised and standardised dialect of Neo-Breton, who have learnt it as a second or third language (Romaine 2011: 217f.). The Cornish language is a great example of revival without native speakers at all. Such an absence of native speakers allows for freer language creation and a sense of influence for users (Fiedler 2006: 77). On the other hand strong language

loyalty and group identity through language leads to clashes and schisms, due to disagreements on which language forms should be adopted. Linguistic discussions of this type create splits between conflicting parties (in revived languages see Romaine 2011: 187f.; on “the Schism” in Esperanto, which led to the creation of another auxiliary language, Ido, see Large 1985: 133–160; Okrent 2010: 99f., 109; on purism in Esperanto see Fiedler 2006: 80). The arguing factions manifest a variety of reasons for the schism. In the case of Cornish, the disagreement is concerned with the fundamentals of the revived language (pronunciation, spelling etc.), because there is no common source for the revival. As mentioned before, the fathers of the revival movement based their versions on two different periods of development. Similarly, Esperantists have agreed on maintaining the principles laid out in the *Fundamento de Esperanto* in 1905 and making the *Fundamento* “untouchable”. The split occurred because of the dissatisfaction of some with the general outline of the language, and resulted in the creation not only of Ido, but later also of several dozen offshoots of both Esperanto and Ido.

A noteworthy fact is that the fathers of the modern varieties can easily be pointed out. Modern Hebrew is the creation of Eliezer Ben Yehuda; the first modern handbook of Cornish was written by Henry Jenner in 1904 and the father of the revival movement was Robert Nance (Kimura 2010; Wood 1979). Similarly, artificial languages in general are the creations of a single person, rarely an institution.

The multilingualism of the speakers might be the reason behind the need for a single unified language (again, the case of Hebrew comes to mind). However, as Kimura (2009; 2010) writes, there is no direct communicative necessity in the minority language, because Cornish speakers all use English, Sorbs speak German etc. The same pertains to Esperanto users, who are mostly well-educated older people speaking several languages. Thus, their language loyalty seems to be a matter of traditions and identity.

feature	ethnic	revitalisa- tions	minority	planned/ IAL
develop- ment to new domains	yes	yes, strictly controlled/ forced	yes, mostly controlled/ forced	yes if so- cialised
linked to a specific population	yes	partly	yes	no
linked to a territory	yes	partly	partly	no
first lan- guage and monolin- gual native speakers	yes	partly	partly	no (Espe- ranto has native speakers, but none are mono- lingual)
purpose: interna- tional communi- cation	no (may be used as such)	no	no	yes
conscious planning	no (yes in later stages of devel- opment)	yes, strict	no (possi- bly in later stages of develop- ment)	yes (not so strict in later stages of develo- pment)
known creator	no	yes, of the revitalised version	no	yes
developed over a rela- tively short period of time	no	no, gener- ally	no	yes

Table 7 A comparison of revitalised and minority languages with ethnic languages and IALs

Corsetti (2010: 377) notes that Esperanto and Hebrew were, from the beginning, complete languages in their grammatical form. The rules of the grammar of Esperanto were laid down in the so-called *Unua Libro* (First Book) in 1887 and later in the *Fundamento de Esperanto* in 1905. Hebrew as a language of religious rituals was also a complete language in this respect. Although it ceased to be spoken around the second century CE, its usage as a written medium continued. The only element that had to be introduced to revive it was modern vocabulary as an addition to about 30,000 extant original Hebrew roots. Esperanto in turn only had 900 vocabulary roots in its first form but over the years, the lexicon has developed in the speech community. The same process is visible in the revival of Cornish. The language has been expanded from a “book language” (Jenner’s variety was based on Late Cornish, Nance’s on Middle Cornish; Kimura 2010: 175; Romaine 2011: 195) to a vernacular of about 20 native speakers (McKinnon 2000; however this number is uncertain) and up to 500 L2 speakers (Romaine 2011: 195; 2011 census).

3.3.3 Other cases

As shown in section 3.1, all languages can be placed on a scale of artificiality/naturalness. The intermediate positions are occupied not only by pidgins or revived ethnic languages but also by classical languages such as Latin, controlled languages such as Basic English and linguistic reconstructions (Proto-Indo-European). As the final point in this section, I shall also discuss in some detail various sign(ed) languages, which, similarly to spoken languages, vary in their degree of naturalness.

CLASSICAL LANGUAGES: SANSKRIT, LATIN

Sanskrit or Latin as classical languages are very often said to be dead. Latin no more functions as an ethnic language with native speakers, while Sanskrit is a highly regularised language mainly used in religious contexts. Nonetheless, their

high status as a cultural and scientific medium has lasted over centuries, making Latin dormant and maintaining Sanskrit in constant usage.

Old Latin is a language first attested in the 3rd c. BCE. In the beginning, it was an ethnic language, consisting of diverse dialects. With the spread of the Roman Empire, Latin became an interlanguage. It functioned in two forms: literary Latin and as an informal vernacular (Vulgar Latin). Contacts with conquered peoples and merchants from outside the Empire led to the creation of dialects, which after the fall of the Empire transformed into Romance languages. This does not mean that Latin stopped existing. It was still used in liturgy and church documents, and as an interlanguage. Mediaeval Latin was the language of administration and the church between the 5th and the 15th c. CE. It shows influence from various sources: Romance languages and literary Classical Latin. In the Renaissance, Latin served as the language of science. It was based mainly on the works of classical authors taught in schools. Scholars using Latin tried to purge it of external influences and vulgarisation. However, in the late 17th c. it became clear that the language could no longer maintain its status. It was abandoned as the chief language of the educated world. Interest in speaking Latin renewed in the 20th c. In 1923 in Paris, the *Société des études latines* was founded. Later decades saw the rise of similar societies with congresses and newspapers published in Latin (Barandovská-Frank 1995).

Sanskrit arose as a codified counterpart of a Vedic vernacular around 400-300 BCE. The subject of the *Aṣṭādhyāyī* by Pāṇini and subsequent commentaries, it came to be taught in schools, while the vernacular from which it stemmed transformed into Prakrits. Although highly formalised and strictly codified, Sanskrit continued as the language of Mahayana Buddhism and Hinduism. These influences spread Sanskrit as a religious interlanguage from the Philippines to Central Asia. Nowadays as many as 14,000 people claim to be native speak-

ers of Sanskrit (2001 census quoted in Ethnologue⁴⁵). As Maurus (2014) observes, “the example demonstrates that a planned language [i.e. Sanskrit] with invariable rules can [...] remain in use indefinitely, and that it can serve as an interlanguage and medium of cultural expression even when principally learned as a second language”.

Visibly, both Sanskrit and Latin have never fallen out of use and, although enriched with new vocabulary, they have remained largely unchanged, due to – as Maurus (2014) writes – a rigorous education and planning processes (the latter true only for Sanskrit). For centuries, Latin had no corresponding vernacular and did not serve as a vehicle for identity, which makes it similar to such revived languages as Hebrew. If assessed according to the criteria presented in sections 3.3.1 and 3.3.2, their modern varieties seem to have more in common with planned languages than with their own initial forms (Table 8).

feature	ethnic	classical languages	planned/IAL
development to new domains	yes	yes	yes if socialised
linked to a specific population	yes	no (not in modern times)	no
linked to a territory	yes	no (not in modern times)	no
first language and monolingual native speakers	yes, generally	Latin no (not in modern times)	no (Esperanto has native speakers, but none are monolingual)

⁴⁵ <http://www.ethnologue.com/language/san>

purpose: international communication	no (may be used as such)	yes (in modern times)	yes
conscious planning	no (yes in later stages of development)	yes	yes (not so strict in later stages of development)
known creator	no	no	yes
developed over a relatively short period of time	no	no	yes

Table 8 A comparison of dormant classical languages with ethnic languages and IALs

LINGUISTIC RECONSTRUCTIONS

The tendency to revive long-gone languages manifests itself also in the linguistic search for a proto-language. A hypothetical ancestor to modern vernaculars is re-constructed through meticulous comparison of equivalent forms in existing languages. The keyword here is “hypothetical” – while the existence of Proto-Indo-European (PIE) is commonly acknowledged, the precise nature of it is not debated, thus making any reconstruction artificial. Moreover, although it may well be argued that a reconstructed language itself was at one point in time a living natural language, certainly texts written in it are artificial. One such example is A. Schleicher’s 1868 Proto-Indo-European fable *Avis akvāsas ka*, repeatedly revised according to the latest linguistic findings.

Interestingly, reconstructed PIE has its supporters as an international auxiliary language. The Dnghu Association promotes a syntactically modernised reconstruction of PIE as an interlanguage for Europe. Its main goals are (Dnghu Association 2007):

- The development of the Modern Indo-European grammatical system, to bring the reconstructed Proto-Indo-European language to its full potential as a living language.
- Teaching it as a second language for all European citizens.
- The adoption of Modern Indo-European by the European Union as its main official language
- The use of Indo-European (its three main dialects) as the main international auxiliary language, to reduce present-day communication and cultural barriers.

The Association regards PIE as the most international and neutral language for Europe. It is claimed that PIE could easily serve as a common language for the EU, as all Indo-Europeans speak some form of dialect of it.

The reconstruction methods may be compared to those of language planning. Similar principles are employed to find a common root for lexemes in IALs. Therefore, as Barandovská-Frank suggests (1995: 36), linguistic reconstruction may be considered as the creation of an autonomous language.

CONTROLLED LANGUAGES

Languages often classified as planned, with the purpose of facilitating international communication, are controlled languages (also called minimal, simplified or regulated; cf. Schubert 2011: 53). Schubert divides them into controlled languages, created before 1960, whose development corresponds to that of IALs (e.g. Ogden's Basic English or Peano's *Latino sine flexione*⁴⁶; see also 2.2.3) and those created after 1960 for industrial purposes (among them Caterpillar English or ScaniaSwedish). The latter ones are frequently the work of linguists

⁴⁶ Traditionally *Latino sine flexione* is included among IALs, while Basic English among controlled languages. The difference stems mainly from the treatment of grammar. Peano simplified Latin's structures to maximally simplify understanding. Ogden, in turn, saw the potential of facilitating communication in restricting the lexicon, leaving grammar structures intact.

by profession, made to order for a company. Such languages are quite similar to naturalistic *a posteriori* languages. What is different, however, is that controlled languages are based on **only one** language and do not change its natural properties, only restricting them to some extent (allowing only one from a set of synonyms, preferring one tense over another etc.). Their artificiality is much higher than that of natural standardised languages, although lower than that of universal systems (see also sections 2.1 and 2.3.1).

SIGN AND GESTURAL LANGUAGES

As the last example of borderline cases, I would like to briefly describe various gestural systems. They too may be placed on a scale of naturalness (Figure 10) as they range from fully natural, non-standardised sign languages of deaf communities through gestural auxiliary languages to artificially created manually coded spoken languages.

The most basic form of signed languages are gestural auxiliary languages (called also Alternate Sign Languages; Faris 1994: 16), such as Monastic sign languages, ritual Australian Aboriginal sign languages or Plains Indian sign languages. The latter ones, used mostly in the 19th-century North America, consisted predominantly of iconic signs and served as contact languages between various Indian tribes (under ‘sign language’ in *Britannica* 2014).

Natural sign languages used for communication among deaf people developed only in permanent places of residence of large numbers of the deaf. Their sources may perhaps lie in so-called ‘home sign’, that is, *ad hoc* basic iconic communication in families where one of the members is deaf. However, the transient nature of such systems makes them impossible to trace. Sign languages are therefore spontaneous reactions to the communication needs of deaf communities; “they effectively fulfil all the social and mental functions of spoken languages; and they are acquired without instruction by children, given normal exposure and interaction” (Sandler 2009). This remains in agreement not

only with Lyons' theoretical discussion but also with Hockett's design features of human language (on the latter, see section 3.1). It is, though, important to notice that language and speech are not to be confused (Lyons 1991: 51f.).

The social conditions of sign languages are very special. Very few deaf are native signers and over 90% of deaf children have hearing parents. This means they have to learn signing at school. Sandler (2009), referring to Fischer 1978, writes that "these social conditions taken together with certain structural properties of sign languages have prompted some linguists to compare them to spoken creoles". It needs to be remembered that deaf people always represent a minority, and that their languages will necessarily be different from spoken languages because of the modality employed. In addition, although linguistic evidence shows they are clearly natural, sign languages and the needs of deaf communities have long been neglected, leading to the creation of artificial manually coded signed languages (hereafter referred to as **signed** languages if the context is sufficiently unambiguous).

Manually coded languages (also: exact signing) are artificial mixed systems, usually taking vocabulary from sign languages and patterning their grammar after local spoken ethnic languages. These systems, dating back to the 18th c. and the teachings of the Abbé de l'Épée (see section 2.2.1), are very often used in the education of hearing-impaired persons. However, as they are much slower in communication than either sign or spoken languages, they tend to be pidginised both by the deaf and the hearing (Farris 1994; Tomaszewski 2004).

Sign languages and manually coded languages are not infrequently treated as one and the same. In Poland, for example, handbooks having in the title the name "Polish Sign Language" are in fact devoted to *SJM*, that is, Polish Signed Language (*Polski Migany, System Językowo-Migowy*). Therefore, in chapter 5 I will discuss both sign and signed languages, to avoid misunderstandings and to show that a modality different from vocal-auditory is no hindrance in studying the LWV.

However, it needs to be remembered that manually coded languages are to natural sign languages as writing systems are to spoken languages: merely a mode of expression. They are looked at here only to clearly show the difference.

4. Esperanto as a transitional case

4.1 Naturalness of Esperanto⁴⁷

Esperanto is a constructed international auxiliary language created by L. L. Zamenhof and first made public in 1887. The language is called ‘artificial’ or ‘constructed’, since its creation is attributed to one man who deliberately built it using elements from chosen natural languages. However, according to what has been said in chapter 3, Esperanto is easily classified as natural. Although on the scale of deliberate changes (see Figure 10) Esperanto is placed a little further towards the artificiality pole, it is the only one of all artificial languages that meets all of Hockett’s criteria of human languages. The naturalness of the language is acknowledged also in three classes of Lyons’ categorisation, namely in Nat₁L (systems conforming to nature), Nat₂L (‘species-specific’) and in Nat₃L (acquirable in the normal process of maturation/socialisation). The first two classes do not need further attention. Class Nat₃L is, in turn, worth discussing.

Contrary to popular belief, Esperanto has a dynamically interacting and growing community. It is used among many thousands of speakers – some sources say up to 3.5 million (see Piron 1989b), while a reasonable estimate would be ca. 2 million according to Corsetti (2012: 69) and Wandel (2015). Gledhill (1998: 10) talks about at least 40,000 fluent speakers, some of whom are even third generation native speakers (on native usage of Esperanto see e.g. Fiedler 2012; Lindstedt 2006). The problem of native speakers will be taken up again later on in this section.

Typologically, Esperanto is built in resemblance to agglutinative languages with a developed system of over 40 prefixes and suffixes, although its fundamental vocabulary is based mostly on Romance and Germanic languages. The word order is usually SVO, but free word order is available thanks to the

⁴⁷ Parts of this section were previously published as Stria 2015a.

accusative marker -n. Manaris et. al (2006) have demonstrated that Esperanto exhibits statistical proportions similar to other European languages (English, French, German, Italian and Spanish were controlled in the experiment) and is generally indistinguishable from them by means of Zipf's law analysis. Parkvall (2010) shows that Esperanto shares most typological features with Indo-European languages of Europe (around 75% and less than 70% for the most related Russian). However, its features in comparison with languages of the world as a whole are similar to 54%.

Several scholars have discussed the changes in Esperanto and shown that in its current form it might be regarded as a naturally changing language. Joshua Herring (2005) observes that of all the predictions about historical changes in Esperanto (frequent in natural languages) only one does not hold, namely the less frequent occurrence of adverbial forms. For example, a higher rate of loanwords with specific, narrower meaning, simultaneous with a more widespread genericity of "native" words and relative clauses replaced by adjectival modifier phrases were confirmed in the study. However, the prediction about disappearing adverbials, disproved not only in this study but also by Piron (1989a) and Gledhill (1998: 69), seems to attest to another very common natural feature, that is, simplification of syntactic patterns. Piron (1989a) and Jansen (2010) present more examples of spontaneous changes in the language: semantic shifts, the tendency of bound morphemes to become autonomous lexemes (i.e. the suffix *-et-* 'small, (a) little, some' becomes an adverb: *ete* 'a little' or an adjective: *eta* 'small'; the suffix *-ebl-* 'possible, able to be done' becomes a verb: *eblas* 'it is possible'), non-verbal forms become verbal (see previous example; predicative adjective constructions such as *estas blua* 'is blue' become verbs: *bluas*) and the obsolescence of some forms.

A comparison of IALs (Esperanto as the most developed example) with borderline cases of natural ethnic languages has been presented in section 3.3. It has been stated that the feature

most often required of a language to be considered natural is the existence of native speakers (not necessarily monolingual; cf. Lindstedt 2006). Yet, it has to be noted that pidgins – although considered natural – are not nativised⁴⁸ and in many languages native speakers are few, multilingual or abandon the language when grown up. A very specific situation is also present in sign languages: over 90% of deaf children have hearing parents and therefore learn signing at school. Esperanto is the only artificial language which has gone through all the stages of Blanke’s functional classification (see section 2.3.3) and which has about a thousand *denaskuloj* (literally in Esperanto ‘from-birth-people’). Teaching Esperanto to one’s children is an extreme form of language loyalty which expands the language into new domains – both linguistic and literary. It provides the users with baby talk, onomatopoeia, and euphemisms, as well as nursery rhymes, songs, riddles and fairy tales. However, as Fiedler (2012) remarks, native speakers of Esperanto are not norm providers, that is, other users do not usually consult them when striving for correctness. Firstly, there are too few of them and the community is principally composed of L2 speakers. Secondly, they might repeat idiosyncratic or erroneous patterns picked up from their parents, who, more often than not, are L2 speakers. Thirdly, about 50% of *denaskuloj* abandon the language at some point in their lives and use their other native language(s) or a foreign language on a day-to-day basis. Therefore, they cannot be treated as the sole determinant of the naturalness of Esperanto or even the correctness.

Accordingly, it can be concluded that, genetically, Esperanto is an artificial language designed to be a natural means of communication, but its sociolinguistic status does not depend on its native speakers. Thus, other factors must be taken

⁴⁸ Native speakers are said to be the most crucial difference between a pidgin and a creole. A pidgin is “native to no one”, whereas “a creole is a nativized pidgin expanded in form and function” (Romaine 2009).

into consideration to substantiate the status of a speech community of this group of speakers. The following section compares the speech community of Esperanto speakers to several other communities from various perspectives: not only sociolinguistic but also purely linguistic, historical or political, both diachronically and synchronically.

4.2 Sociolinguistic situation of Esperanto speakers⁴⁹

Sociolinguistically, Esperanto presents a special case. It shows some similarity to pidgins, creoles, minority and revitalised languages as well as diaspora languages, yet remains a unique type of speech community. This section focuses on Esperanto and recapitulates the sections 3.3.1 and 3.3.2, showing that members of the Esperanto movement (as opposed to speakers of other artificial languages) constitute a speech community.

The characteristics which will serve as a common base for comparisons between ethnic languages, minority languages, revitalised varieties, contact languages and Esperanto, are: (i) development into new domains (i.e. if the language can spontaneously expand), (ii) connection to a distinct ethnic group or a particular population and (iii) connection to a delimited territory, (iv) the existence of monolingual native speakers, (v) international communication as the main purpose of the language, (vi) conscious, strict planning (i.e. forced and directed regulation) and finally (vii) known creator and (viii) known date of formation (based on Table 6 and Table 7).

Choosing a definition of such a community is challenging because many are not applicable to Esperanto. Chomsky and the generativists, for example, emphasise the importance of native speakers. As has been demonstrated in 3.3.1, 3.3.2 and 4.1, native speakers cannot be the decisive factor in the discussion of naturalness and, consequently, of whether this particu-

⁴⁹ This section is partly based on Stria 2015b.

lar group is a speech community. Thus, a better definition would be that by Gumperz (1968, in Patrick 2002), who rightly observes that “speech varieties [...] form a system because they are related to a shared set of social norms”. Kerswill (1994, in Patrick 2002) extends this by claiming that a speech community reaches an “agreement on the social meaning of various linguistic parameters”. The parameters can only be fully understood by the community’s members. This is an important remark, which describes at least part of the Esperantist community, whose language usage is connected to a shared culture. However, Gumperz (*ibid.*) requires also that a group interact regularly and frequently – a condition that cannot be met by Esperanto users, who keep in touch primarily through written messages or meet at occasional congresses or meetings (although lately internet communicators allow for more spoken contact).

A second difficulty is to determine who exactly an Esperantist is, or who constitutes the Esperanto speech community. As Galor (2001) observes:

E-komunumo estas pli mallarĝa ol E-a socia aro; tio devenas de la sekvaj faktoj:

- ekzistas homoj, kiuj konas Esperanton (lingvokonantoj), sed ne uzas ĝin,
- ekzistas lingvouzantoj, kiuj uzas Esperanton (lingvouzantoj) aŭ por celoj ligitaj kun tiuj de la E-komunumoj (e-movado) aŭ por aliaj celoj,
- ekzistas lingvouzantoj, kiuj agas individue aŭ kolektive por grupaj celoj de E-komunumo (lingvoagantoj); tamen ne ĉiuj membroj de la komunumo estas lingvoagantoj

The E[speranto] community is smaller than E[speranto] social collectivity; this stems from the following facts:

- there are people who know Esperanto (*lingvokonantoj*, ‘language knowers’) but do not use it,
- there are *lingvouzantoj*, who use Esperanto (*lingvouzantoj*, ‘language users’) either for purposes connected with those of

the E[speranto] communities (E[speranto] movement) or for other purposes,

- there are *lingvouzantoj*, who act individually or collectively for group purposes of the E[speranto]-community (*lingvoagantoj*, ‘language activists’); however, not all members of the community are *lingvoagantoj*

A related observation is expressed by Wood (1979), who writes that the Esperanto movement consists of:

- *apogantoj* (‘supporters’ not speaking Esperanto),
- *uzantoj* (non-member ‘users’), and
- “mainstream” Esperantists.

Only the latter group seems to be what Gumperz (1968, in Patrick 2002) defines as a speech community. In the core of the Esperanto movement (member speakers are not necessarily affiliated to *Universala Esperanto-Asocio*, *Universal Esperanto Association*, or other official organisations; they can be individual activists) two main phenomena are observable: shared values and identity through language (Galor 2001; Wood 1979) and shared language norms (Fiedler 2006; Fiedler 2012). As previously stated, there are several more specific features as well (not only purely sociolinguistic) that make the community very similar to some natural languages, namely pidgins and creoles, minority and revived languages, and diaspora languages.

COMPARISON WITH REVITALISED/REVIVED LANGUAGES

Revived languages are an extraordinary case of natural languages. They can in fact be considered artificial. Not only is their historic continuity broken when they stop being used but also vast changes and modernisations imposed later to revive the vernacular make the new variety far removed from

the original⁵⁰. The resulting varieties are separate, although related. Revitalisations are the outcome of the deliberate efforts of a person or an institution, and later groups of enthusiasts. Therefore, revitalisations can be dubbed “reinvented” (Romaine 2011). No doubt then, the fathers of the modern varieties can easily be pointed out.

In fact, revived languages and Esperanto were all “book languages” at the dawn of their modern history. Modern Hebrew was a language of religion, used only during ceremonies and in religious texts such as the Mishnah⁵¹ (Kimura 2010; Wood 1979); Esperanto was first and foremost used as a translation medium, until the first Universal Congress in 1905; Cornish became a mature spoken language in the 1970s (Kimura 2010: 172). Revitalised vernaculars, that is, those that are not dead but merely “dormant”, have the advantage of still having a few speakers, although it is surviving documents that offer the basis for revitalisation.

Esperanto proved successful mainly because Zamenhof renounced the rights to his creation and handed it over to the community. He kept a record of enthusiasts (names and addresses) who had translated passages and works into Esperanto and actively used the language, thus allowing the users (not yet speakers for the most part) to communicate and intensify their contacts. Hebrew had to be popularised in a more politically forceful manner: other languages in Israel were actively discouraged. Yiddish especially was to be eradicated, being used by almost a half of Israelis. The purists encouraged Sephardic pronunciation and words of Semitic origin (Romaine 2011: 188, 193). All Israelis were to

⁵⁰ Cf. the case of Hebrew, where elements from other Semitic languages and Yiddish have been incorporated. Zuckermann (2006) argues that Israeli Hebrew is a semi-engineered hybrid of Semitic and Indo-European elements.

⁵¹ Zuckermann and Walsh (2011: 113) point out that the other part of the Babylonian Talmud, the Gemara, written about 300 years later, is largely in Aramaic.

be native speakers of Hebrew. Modern Hebrew and Esperanto have one more characteristic in common: they were both needed as a common language among people of different origins speaking different languages; however, not as a communicative necessity but as a voluntary choice (although in the case of the State of Israel there was a political need, while Esperanto was to be spread because of its “internal idea”, that is, the hope to propagate peace on the basis of a culturally neutral language). Both languages had no native speakers in the beginning. At present, Hebrew can boast great success in nativisation, whilst Esperanto is claimed to have about 2,000 native speakers (Corsetti 2012: 70).

Those speaking revitalised languages need not be of ethnic descent (Kimura 2010; Romaine 2011). They are frequently educated city dwellers who chose to learn the language without having any ethnic connection to it; as stated before, the majority of Neo-Breton speakers are non-Breton, while the rural varieties are spoken mostly by people of Breton ethnicity. The activists play a crucial role in the advancement of the language, constituting very often more than half the speakers. Therefore, being a member of such a community is most often an ideological choice shaping the identity of the speakers. The preference for this particular language over any other is here more important than “birthright membership” (Wood 1979: 433). Language loyalty seems to be a matter of traditions and identity.

COMPARISON WITH MINORITY AND DIASPORA LANGUAGES

Esperanto was first called “a self-elected diasporic linguistic minority” in the article of Wood (1979). This term was spread through the Manifesto of Rauma⁵² of 1980. The language’s situation is comparable to that of diaspora languages’: despite the dispersal of the speakers, they are actively con-

⁵² *Manifesto de Raŭmo*, written in Rauma, Finland, is a document criticising the ideology behind the traditional Esperanto movement.

nected not only through modern channels of communication such as the Internet but also through various associations. Numerous publications, radio programmes and, most importantly, Esperanto-only meetings at local, national and international levels are available to them. Nevertheless, Esperanto speakers do not have a shared homeland from which they have emigrated. What they share, though, is the history of the movement, its beginnings, persecution and the struggle for international recognition (Fiedler 2006: 74–76).

Esperanto shows many sociolinguistic similarities to minority languages. In the case of Esperanto, the distribution of the speakers is clearly non-territorial. Likewise, although the speakers of minority languages inhabit a distinct *Sprachraum*, the territory they live in consists of scattered language islands in an area occupied by the majority language (see the distribution maps of e.g. Rumantsch Grischun or Irish Gaelic⁵³). Because of this non-territoriality and lack of a native country, Esperanto is always a minority language.

Such languages are spoken in bilingual or multilingual communities and their use is mostly restricted to everyday informal situations, in contrast to majority languages, regarded as a prestigious variety. Diglossia is common not only among minority language speakers, but also among the *denaskuloj* of Esperanto – it is rarely possible to use the language at work, to use it for study, or when communicating with the authorities (although there are several groups working for the adoption of Esperanto in the EU or the UN). Yet, minority languages are generally used at home, whereas Esperanto is used mainly at conferences and professional meetings, and its domestic use is limited to a very small number of Esperanto families. Esperanto is used in several professional associations devoted to fields such as law, medicine, astronomy and science in general.

⁵³ Even more so in the case of revived languages, such as Cornish. There is no delimited speech community. The speakers live in and outside the county of Cornwall.

Prejudice against the language often characterises attitudes not only towards minority languages but also towards planned ones. Their usage was banned or persecuted (the legislation in New Zealand imposing English only education practically eradicated Māori). Artificial languages have been looked down on by linguists. Moreover, Esperanto was regarded “a dangerous language” under Hitler and Stalin (e.g. Lins 1988; Rónai 1969). Active Esperantists were suspect and accused of having international contacts (which was known for a fact) and therefore of subversive actions. Nonetheless, at all times the striving for recognition strengthens the sense of belonging, language loyalty, and adds to the identity of the speakers (Kimura 2010; Kimura 2012; Krägeloh & Neha 2014).

Interestingly, as the speakers of minority languages always know the majority language, they lack the communicative necessity to speak the minority language. The community may instead voluntarily use the minority language rather than the “high” variety. The same holds true for revived languages, being for the most part minority languages.

COMPARISON WITH PIDGINS AND CREOLES

Esperanto was created as an auxiliary language, to facilitate international communication. A similar role is played by pidgins – while not created at a writing desk in a meticulously planned way, they arise due to regular contact between two or more groups of speakers of separate, mutually unintelligible languages. The process is spontaneous, unconscious and non-directed (Liu 2001). Both pidgins and Esperanto have their origins in several ethnic languages, although at the beginning they are native to no one. A pidgin can be passed on to nascent generations (nativised) and become a fully “natural” language, i.e. a creole. Such a situation among planned languages takes place only in Esperanto.

The lack of pidgin native speakers and the small number of them in Esperanto gives equal status to all their users – L2 speakers are valued as much as native speakers in deciding on

language norms. The fact that those languages are from the beginning non-ethnic and intercultural adds to their neutrality as intermediary languages. In addition to being multicultural, both the populace speaking a pidgin and the one speaking Esperanto are never monolingual⁵⁴. Also, later native speakers of Esperanto and the subsequent creole are mostly multilingual. Some creoles, of course, become the sole language of some people, whereas Esperanto is always spoken in combination with another language.

Nativisation is connected with another process, namely expansion to new domains. If a pidgin develops over time, it may change into an expanded pidgin and later become a creole; although, as Liu (2001) remarks, there are creoles which have developed from simple pidgins, such as Torres Straits Broken; those developed directly from jargons, such as Hawaiian Pidgin English, and pidgins which have never expanded. Such an expansion is most often spontaneous and unconscious. In Esperanto, the development is twofold: the changes are guided by specialised institutions such as *Akademio de Esperanto* or *Terminologia Esperanto-Centro* or occur naturally through continuous usage in families and in the speech community in general. The speakers do not always follow the guidelines set down by the *Akademio*, as in the case of the words *koruso* and *ĥoro* (both mean ‘choir’) – the first one is the preferred choice of speakers who want to avoid the uncommon ‘h’ with a circumflex, while the latter is recommended by the Academy (or in fact the only one allowed, see *Akademio de Esperanto* 1975).

Corsetti (2010: 374) remarks that pidgins and mixed languages are relexified and that the same process happens in international auxiliary languages, i.e. a new lexicon from (a) superstrate language(s) is introduced to a complete grammati-

⁵⁴ Multilingualism is understood here both in the sense of speaking more than one language in any combination, be it more than one native language or one native language and any number of L2 languages.

cal system. It is worth noting that a more or less exact time of the process is known. In the case of Esperanto, the date of its creation is 1887, while the birth of pidgins or even creoles can be narrowed down to several decades (e.g. Tok Pisin between 1865 and 1890). The creator might be identified as well: Esperanto is an invention of L. L. Zamenhof, and pidgins are created by particular people in a particular place and time.

However, one great difference is visible: contact languages are linked to a specific territory on which they come into existence and further develop, whilst Esperanto in this respect is rather a diaspora language. Even though its beginnings are connected to the territory of what is now Poland, and the majority of speakers to this day are Europeans, Esperanto users are spread across five continents.

Considering all of the above, it can be said that although the Esperanto speech community is similar to other types, it differs from all of them in some aspects, creating a separate category. The tables presented in 3.3.1 and 3.3.2 will serve as a recapitulation of the characteristics that provide the basis for the comparisons.

The development into new domains and conscious planning are, in truth, interdependent. Only pidgins seem to be free from planning and their development is tightly connected to their socio-political situation (i.e. social needs cause a pidgin to expand, creolise or die out). In Esperanto and revived languages extensive planning takes place at the beginning of the history of those languages, while varieties being revitalised and minority languages undergo this process later in their development to further their progress and expand the usage. Creoles and small non-codified languages require standardisation to attain a unified, stable state. Purism and controlled development are a part of the process of achieving and maintaining a certain status. Planning institutions play an important role in the process. In the Esperanto community, the standards are not always dependent on the Academy of Esperanto but rather established in cooperation with the speakers (Fiedler 2006: 80). Their contribu-

tion to the development of language (see especially the role of translators in Fiedler 2006: 79) creates a sense of importance and belonging, and strong language loyalty. The influence of the speakers might be so strong that the planning institutions concede and accept words and structures in widespread use. Therefore, Esperanto as a creation “freed” from planning actions may develop as any other natural language, while revitalised varieties are reliant on strict linguistic control.

Due to the fact that there are no monolingual speakers regularly interacting with each other, Esperanto cannot be considered creolised. There are a small number of native speakers, constituting much less than 1% of the community, who do not set the standards, and therefore the Esperantists cannot rely on the status of the *denaskuloj* in the community (at least 10% is needed to consider a language on its way to being creolised; Liu 2006: 57). The shared norms are rather negotiated within the core of the movement, which comprises an overwhelming majority of L2 speakers (similar to the international usage of English). The nativisation of Esperanto and its continuous usage in families contributes to the lexical and stylistic expansion to new domains. Actually, some small ethnic languages are limited to several basic domains, while in other domains the majority language, or English are rather used.

Esperanto speakers maintain their identity through an outright rejection of English and efforts to introduce vocabulary built in accordance with the rules laid down by Zamenhof, that is rather according to §11 of the Fundamento by word formation, rather than §15 by borrowing, especially directly from English. It is notable that lexical loans in general do not stand in opposition to the Fundamento. The need to resist English influences might stem from different sources. On the one hand, small languages fight against the domination of a larger, internationally used language to survive locally. On the other hand, Esperanto, as a language designed to facilitate international communication, must face competition globally. Purism is therefore not a mere linguistic practice but also a means to

create a sense of unity in the speech communities of Esperanto and low-prestige languages.

The identity of Esperanto speakers is also based on voluntariness – speaking Esperanto and teaching it to their children is their conscious choice as the community is non-territorial and non-ethnic (Wood 1979). When it comes to a specific, more or less delimited territory, diaspora languages need to be mentioned. Clearly, they are not tied to any particular place, although they have their “homeland”, a territory where the language is constantly spoken. Esperanto is thus a “virtual” diaspora language without any native country (it is worth observing that Esperanto users often speak of *Esperantujo/Esperantio*, ‘Esperantoland’, i.e. everything related to Esperanto, its speakers and their activities). Territoriality is, however, a factor in the formation of pidgins and creoles. A pidgin arises in a particular area of contact between two populations of different linguistic and often cultural backgrounds. In the same place a creole is later born. This situation is not easily translated to revitalisations and minority languages. Even though their lifecycle is connected to a defined region, the speakers are scattered across the territory and outside of it, and not necessarily in constant contact.

Monolingual speakers do not seem to exist in such communities. A pidgin is obviously spoken by people already knowing one or more languages. Speakers of Esperanto and other types subject to analysis in this study know several languages, perhaps one of the other languages being the strongest, or most often used (e.g. Welsh and Breton are quoted as having no monolingual native speakers at all). Only Israeli Hebrew might serve as a success story.

However, it should be mentioned once more that native speakers are not necessarily the major driving force of a language. Pidgins are the result of continuous intense contacts between at least two groups of people. Revived languages are very often propelled by non-ethnic enthusiasts. Esperanto is predominantly spoken by non-native speakers of varying eth-

nic and cultural backgrounds. These three types have one more thing in common: their creator is always known. In the case of pidgins, it is identifiable peoples, while Esperanto and revived varieties have “fathers”. Revitalisations are a borderline case – while mostly ethnic and historically uninterrupted, their development depends on the efforts of particular individuals and institutions. In terms of having an identifiable population and creator, the only ones to parallel natural ethnic languages are creoles and minority languages. The latter type also has no distinct beginning, whilst all the others can be traced back to a short period of time during which they came into existence.

As a final point, it should be stated once more that Esperantists constitute a varied group of both speakers and supporters. The sociolinguistic categories often overlap, and thus the movement consists of diverse examples: non-movement speakers, supporters not speaking Esperanto and finally speakers actively participating in the movement and identifying themselves through the language with the values associated with the idea of an international auxiliary language, world peace, tolerance, liberty and equality. Therefore, the movement’s core constitutes a dynamically interacting speech community, although of a special type.

In conclusion, Esperanto satisfies the definitional condition of the LWV, i.e. it has developed a dynamically interacting community in the form of mainstream Esperantists negotiating the social meanings of the language. As Machado (2004: 101) writes: “A secondary forest that develops freely without human interference, will end as a natural system, despite it may have been felled in a more or less remote past.” Such is the case of Esperanto: although designed artificially based on various natural languages, it has evolved to become close to natural.

5. Linguistic worldview for artificial languages

The languages described throughout the previous chapters are not only languages typically considered artificial (universal systems, international auxiliary languages, artlangs, conlangs and formal languages) but also borderline cases (controlled languages, pidgins and creoles, revitalisations and revived languages, linguistic reconstructions as well as sign and signed languages). The present chapter deals with most of them, trying to answer the question of whether and how are they available for analysis in the ESL framework. Languages omitted here will be pidgins and creoles (treated as natural but developing in a multilingual and multicultural environment) and revitalisations (for the same reason). The motivation for including sign languages, which have been identified as natural, is that they are often misleadingly treated as artificial because of their modality and confused with signed languages, which are gestures artificially assigned to oral languages. Revived languages are also included because of the discontinuity and wide-ranging arbitrarily introduced changes in their usage. Furthermore, in section 5.4 I evaluate in more detail the questionnaire method for Esperanto, the only IAL that might be treated as a transitional case between naturalness and artificiality.

5.1 The subject: author or community?

Sections 1.2 and 1.3 discussed the importance of the subject-oriented version of the LWV. The viewpoint and perspective of the perceiving subject are crucial in determining the “ingredients” (i.e. facets and their structuring) of the object perceived. Examined language examples must first reveal the subject who produced the text, for the researcher to decide which approach they should take and which facets are central and which peripheral. Bartmiński (2012a: 29) explicitly states that connotations (together with peripheral meanings) cannot be excluded from the cognitive definition and constitute a key

component of it. Moreover, Bartmiński and Chlebda (2008: 14) claim that core meanings are generally stable components, whilst “weak” and peripheral connotations are subject to contextual and situational changes.

In the analyses of the LWV in artificial languages, the subject is not easily extracted. If the system is to be treated as a language, there must exist a community speaking it and negotiating the meanings to avoid a simple transfer of the author’s ideas. As shown in the previous sections, Esperanto is undoubtedly the only IAL with a speech community large enough to allow development. However, an artificial system can be treated as a product (an artistic creation, a text). The researcher might then expect the idiosyncratic LWV of the author to show through, even if the real author of the analysed text is another person. To decide which case it is, different types of artificial languages have to be discussed.

First, however, it needs to be determined what characterises a “product”. By “product” shall be understood any text produced by the language-system’s author (including grammar rules laid out in script or a vocabulary item list) or any other user, but not language which can be or is used in spontaneous communication in a speech community. Similarly, Bartmiński (2012a: 179) treats questionnaire answers as texts of a specific genre. Generally speaking a text is a closed, logically and coherently arranged set of meanings. It is intentional and can be described in terms of style (cf. Bartmiński & Niebrzegowska-Bartmińska 2012: 36). A text is a concrete realisation of the higher-level abstract text model, that is, a *texteme* (2012: 53).

A helpful description of a text is provided by Beaugrande and Dressler (1981: I), who propose the following seven criteria of textuality:

- cohesion (organisation of the text relying on grammatical dependencies),
- coherence (internal organisation of the concepts and relations),

- intentionality (producer-oriented quality; for the purposes of simplicity it can be assumed here that the intentionality in artificial language texts is never violated);
- acceptability (“concerning the text receiver’s attitude that the set of occurrences should constitute a cohesive and coherent text having some use or relevance for the receiver”),
- informativity (“the extent to which the occurrences of the presented text are expected vs. unexpected or known vs. unknown/certain”),
- situationality (relevance in a certain situation),
- and intertextuality (dependence upon other texts).

Acceptability seems to be important in evaluating the status of an artificial language. As Beaugrande and Dressler write (1981: I.14):

This attitude is responsive to such factors as text type, social or cultural setting, and the desirability of goals. Here also, we could view the maintenance of cohesion and coherence by the text receiver as a goal of its own, such that material would be supplied or disturbances tolerated as required. The operation of inferencing [...] strikingly illustrates how receivers support coherence by making their own contributions to the sense of the text.

Whether a text (here: an artificial language as a product) is acceptable, is not only dependable on the organisation of the text in itself but also on external circumstances such as the social and cultural situation, as well as the state of knowledge of the recipient.

Bartmiński and Niebrzegowska-Bartmińska (2012: 49) present a more detailed – although very similar – list of characteristics of a text:

- has a subject (i.e. sender/author)
- directed to a receiver (Beaugrande and Dressler’s *acceptable*)

- intentional
- specific style and genre
- linear
- structurally integral (Beaugrande and Dressler's *cohesive*)
- logically and semantically consistent (Beaugrande and Dressler's *coherent*)
- informative
- situational
- open to processing (Beaugrande and Dressler's *intertextual*)

The main difference is the lack of the subject among the criteria given by Beaugrande and Dressler, who, according to Bartmiński and Niebrzegowska-Bartmińska, marginalise its role. The Polish authors view this element as central to the concept of textuality. In my view, a subject is a necessary condition for the existence of language. Language, and consequently a text cannot have no subject at all. Therefore, it is a self-evident feature not essential on the list. Textual linearity, which even the authors themselves mention as obviously stemming from the linearity of spoken and written language, is, however, wrongly stated as essential. Evidence from sign languages shows that sequentiality is not a design feature of natural language (Wilbur 2011). Sign languages employ simultaneity as a means to maximise efficiency. A signed message is undeniably a text, even though it is not fully sequential (still, groups of signs might be linear). Style and genre, in turn, are strictly connected with intentionality and acceptability and as such constitute an important part of textology. In this study, however, stylistic considerations will be omitted.

This point of view allows for further analysis of artificial languages. If they are in use and develop in a speech community, they could be treated as language-systems. In the opposite situation, they should be treated as texts (“products”).

Universal languages such as Wilkins's Philosophical Language do not have a speech community and most likely are not

usable at all (no redundancy, strict semantic and grammatical rules etc.). Their structure and form are the result of conforming to the mind-set of the time. They were cohesive and coherent in relation to the contemporary state of knowledge, and therefore informative and acceptable as situationally relevant. At present, those systems are informative of the time, unconvincingly coherent (although probably some are still cohesive in light of the latest linguistic theories) and non-situational, therefore unacceptable (close to being non-texts, i.e. breaking the rules of textuality). Universal languages remain cohesive because of the fact that, although not suitable for communication, they generally do not break the rules of linguistic universals and adhere to the common-sense understanding of how a language-system works. Their coherence is built on the basis of meronomic relations (Bartmiński & Niebrzegowska-Bartmińska 2012: 181), namely, a universal language-system consists usually of a grammar and a lexicon (phonology or script may be additionally included). However, nowadays the meanings and functions ascribed to particular units seem out-of-date, therefore unconvincing, and less acceptable than at the time they were constructed.

Some of the problems resulting simply from an immense change in knowledge resources could be obliterated through intertextuality, although understanding and accepting the systems would require a great deal of mediation as a result of their being non-efficient (see Beaugrande & Dressler 1981: IX.1). It seems that the amount of effort that has to be put into understanding the complexities of a universal language is precisely the reason why such languages are unusable. In this case, the systems should clearly be treated as products, not as languages. As artistic creations, they might be subjected to a LWV analysis according to the propositions made i.a. by Anna Pajdzińska (2013), that is, as instances of the individual worldview of the author embedded in a specific context. Bartmiński and Niebrzegowska-Bartmińska (2012: 117) point out that the artistic style “unveils a personal viewpoint” and

the intention of artistic texts “is not a practical purpose [...] but rather a creation of a certain arbitrary reality”. If universal languages are treated as poetic texts, they become more acceptable as an “interactive, negotiable [...] discourse about the ‘real world’” (Beaugrande & Dressler 1981: IX.9).

The issue becomes more complicated with IALs. The group is not unified and should be divided into Esperanto, planned semi-languages (languages in limited use such as Interlingua) and projects⁵⁵ (see section 2.3.3). Esperanto is the only IAL that has a speech community and develops naturally (cf. sections 4.1 and 4.2). Thus, Esperanto should clearly be considered a living language. Its vocabulary far exceeds the number invented (primarily) by Zamenhof (about 900 in the *Unua Libro*, over 2,500 in the *Fundamento*, thousands in translations of major literary works; the biggest Esperanto monolingual dictionary (Waringhien & Duc Goninaz 2002) records precisely 16,780 entries with 46,890 lexical items). The growing number of words and structures, as well as the multilingualism of the users of Esperanto suggest that meanings are negotiated in the course of communication and not straightforwardly transferred from the ones intended by Zamenhof. The naturalness of Esperanto enables the analysis of its LWV, although the negotiable meanings make it difficult to decide which facets are central and which peripheral. Ken Miner (2011) claims that it is precisely this vagueness is the factor which makes Esperanto linguistics impossible. I would rather insist on applying the LWV theory at the same time, while bearing in mind the most probable outcome: that the prototypical features will be sparse, and the peripheral ones will be far more ample than in ethnic languages.

The same problem of fuzziness applies to other IALs (those still in use seem to be IALA’s Interlingua and Ido, and perhaps Latino sine Flexione and Volapük), especially due to

⁵⁵ This category shall not be further discussed. To great extent, it may be treated just as conlangs.

the very small numbers of speakers. Undoubtedly, they should be treated as languages because they are functional and useable. However, as there is not enough material for analysis, it could be assumed that the LWV would be difficult to obtain in general and the results fuzzy. Additionally, their hypothetical LWVs could not be treated as coming from their authors, because the vocabulary has been extracted from a particular language, or a whole family of languages (see section 2.2.3). Therefore, it is proposed that individual texts (written in the language but not language-systems treated as texts) be analysed as typical literary passages studied against the intended usage as described in textbooks and against the systems of the parent languages. The systems are cohesive (as regular yet based on ethnic languages) and coherent (as generally based on the ethnic organisation of concepts), acceptable as reasonably informative⁵⁶, fully situational and intertextual as based on their source languages. This applies also to revived languages such as Latin or Cornish. They are unambiguously languages but the number of speakers is not sufficient to conduct verifiable analyses of modern usage. However, individual texts – both ancient and modern – may be employed as the basis for analyses.

Similarly, modern arlangs and conlangs may be treated as languages with some limitations. Generally, they are useable (which is shown by fan literature and poetry, and even discussions written in those languages) but they have no stable community in which the meanings and usages can be developed (cf. Hendriks-Hermans 1999; Wahlgren 2004). Vocabulary is very often under development (Frommer n.d.; Peterson n.d.) and the researcher cannot be sure if and how the users

⁵⁶ To my mind, language as a **tool** cannot be informative in the sense proposed by de Beaugrande and Dressler (1981), i.e. providing new, unexpected information and therefore “interesting” (i.e. effective), because it has to be transparent for the speaker in order to be useable. If a system is to be treated as a **product**, it is informative as far as the differences between this particular system and the other scrutinised languages are concerned.

understand it. Therefore, it may be assumed that such languages present the author's view on how the language should be used and what the world in which the language is used should look like. There is, however, some difference between artistic and constructed languages. The former can be treated as typical poetic creations and their LWV may be checked against the projected (fictional) environment in which they are used. The latter present a diversity of solutions. The one trait they have in common is the struggle of the creators not to repeat the patterns of their own mother tongues. For the same reasons as arlangs and partially developed IALs, they have to be treated as semi-languages. Because both arlangs and conlangs mean to be plausible and imitate living languages (i.e. be functional and fully developed) it may be expected they are cohesive and largely intertextual. Their coherence, situationality and acceptability are a matter of individual opinions about how a language should work. On the other hand, it may be assumed that, within the vision of the author, they fulfil the requirements.

Reconstructions of proto-languages (e.g. Proto-Indo-European) present a special case. They are not living languages but hypothetical constructs⁵⁷. They demonstrate the features from Beaugrande and Dressler's list to such a degree as to comply with the newest available historical evidence. Only this compliance makes them acceptable.

A greater degree of abstractness is encountered in the case of formal languages. They seem unusable in the sense of "speakable" languages yet by many are considered perfectly natural (see Lyons 1991: 69 and section 3.2.2) as *textemes*⁵⁸ to ethnic languages ('formal' being here Chomskyan-type formal

⁵⁷ In theory, a reconstruction may perhaps be treated as *textemes* for present-day languages.

⁵⁸ By analogy, an abstract model for language, such as Chomskyan grammars, could probably be called *lingueme*. Unfortunately, this term is already in use (as 'a unit of linguistic structure taking part in replication').

constructs parallel to propositional calculus). Their semantics is contextual, conventional, and therefore strongly situation-dependent. Texts in such languages are cohesive (to the extreme) and informative; they can be, however, unacceptable to the average reader as involving too much mediation.

On the opposite pole are located sign and controlled languages. **Sign** languages are undoubtedly natural. Some degree of artificiality is present in **signed** systems, as they are full signed versions of natural oral languages or auxiliary signed systems for natural oral languages. They cannot be used independently, being grammatically based on spoken languages, and therefore the worldview might be deemed almost identical to that of the language whose structure they follow, albeit simpler in form/controlled (e.g. without idioms or phrasal verbs). It must be noted that for a deaf person not knowing the oral language on which the signed system is based the latter is unacceptable, because of involving too much intertextuality and being non-cohesive in comparison to a naturally acquired sign language.

Not far away from natural languages lie controlled languages. Modelled on ethnic languages but strictly regulated, they should be considered languages, albeit with some peculiarities when compared with their ethnic counterparts. They may encounter limited acceptability because of their very simple structure and limited vocabulary, making texts barely informative. Large (1985: 169f.) remarks that the simplicity is obtained through constant paraphrasing, which retains idiomatity and at the same time requires a great deal of effort, not only from the author of a text but also from the reader.

What follows is that the lack of a speech community does not condemn an artificial language to be treated as a product of auctorial views. Generally, the decisive factors are usability and potential for developing such a community. The degree of abstractness is also of some importance as it resolves the question of formal and reconstructed languages. As a result, artificial languages can be divided into four types: full-fledged

languages suitable for ESL analysis, semi-languages, textemes and finally artistic texts. This is synthetically presented in Table 9 below.

language	type
sign	language
revived	language
controlled	language
Esperanto	language
other IALs	semi-language
conlangs	semi-language
artlangs	semi-language
formal	texteme
reconstructed	texteme
universal/philosophical	text

Table 9 Types of artificial languages and borderline cases according to usability

5.2 Possible objects of study

In the previous section, it has been shown that the LWV of artificial languages can be studied in the ESL paradigm under certain specific conditions. Apart from Esperanto, which has evolved naturally and provides ample material for analysis, artificial languages offer only restricted and controlled texts.

It is now worth casting a closer look at the material at hand. The Ethnolinguistic School of Lublin (EUROJOS 2008, cf. Bartmiński 2012a: chap. 3) postulates that the following elements should be taken into account to create a cognitive definition in this paradigm (more details in section 1.4):

- system data (paradigmatic and syntagmatic relationships) including proverbs treated as minimal recurrent texts but excluding calqued ones;
- lexeme definitions, collocations and quotations from monolingual dictionaries;
- modern style-neutral texts and corpora;
- balanced sample questionnaires.

As stated, questionnaires and larger dictionaries and corpora (modern as well as older) are available only for Esperanto. Some IALs also offer minor corpora and dictionaries; the number of speakers is, however, too small for the questionnaire method. Similarly, revived languages certainly have dictionaries, corpora and texts, but they have to be approached carefully because most of the material comes from the older versions of the language rather than the modern revived vernacular (cf. Latin or Cornish). The number of speakers is not sufficient for questionnaires and the results obtained in such a method may be fuzzy (see Miner 2011). Artlangs and conlangs provide small dictionaries and short texts (mainly dialogues and poetic passages, idioms and proverbs included). The same applies to universal languages, in which the material focuses very often on scientific vocabulary. Linguistic reconstructions offer dictionaries and even hypothesised texts. However, as in the case of controlled languages, the material is rather the set of rules and regulations on how to proceed. The survey method is out of the question for most cases⁵⁹. Sign languages are problematic in the respect that texts and corpora have to be gathered as video recordings (there, however, exist several corpora, such as The British Sign Language (BSL) Corpus or The National Center for Sign Language and Gesture Re-

⁵⁹ The impossibility of the questionnaire method might also stem from the fact that there are no native speakers of most of these languages (this being a requirement of Bartmiński). However, as in the case of Esperanto, advanced speakers may conditionally be asked to participate.

sources (NCSLGR) Corpus in Boston). The same applies to surveys – they cannot be conducted in script, but as recorded signed interviews, to prevent transfer from the language of writing. The most complicated task concerns formal languages. The analysis would have to be based on the grammatical relations and theoretical assumptions of particular texts. The questionnaires would have to relate to technical, formal concepts and their understanding in the framework of particular theories.

What is visible in the proposed choice of material is that lexical analysis is the most vital and intricate part of the research. An important remark here is that the EUROJOS project focuses on the contemporary state of language. Typically in this paradigm a notion is also analysed diachronically and the researcher additionally has to consider the etymology (Bartmiński 2012a: 29; 2013a). Grammar has also been repeatedly proposed as a part of the LWV research (see Bartmiński 2012a: 33; Grzegorzczkova 1999).

Having discussed the material postulated as the basis for analysis, it is time to examine in detail the problems tied to particular aspects of particular artificial languages. The aspects considered are (i) lexicon, (ii) grammar, (iii) etymology, and (iv) phraseology and idiomatic expressions (formulaic language).

Doubtless, the most uncomplicated examples are arlangs. As literary creations they cannot be, obviously, fully analysed in the ESL paradigm. In such languages, the context – the imaginary culture, historical and social circumstances and even geography – drives the appearance of the language. The effectiveness (as understood by Beaugrande & Dressler 1981: IX.9) of such a language depends also on the effectiveness of the story itself. Both the lexicon and the formulaic parts stem directly from the presupposed image of the culture of the speakers. An etymology might also be hinted at, however it does not have to be present at all. It is auctorial in nature but might be somehow logically connected to the ideas of what the imaginary language looks like semantically. There

are also artlangs, whose roots reach to real human languages, for example, the Fremen language in the *Dune* series of F. Herbert (allegedly descended from Arabic and in fact loosely based on it in the books). The grammar of such a language is, of course, a matter of the personal taste of its creator. Any analysis of the LWV of an artistic language – whether carried out in the ESL paradigm or not – would therefore be only a light-hearted experiment, or potentially a study of the coherence between the language and the assumed culture within literary studies. For want of a speech community, the questionnaire method seems pointless. Yet, there is a handful of die-hard fans who speak an artlang (cf. Okrent 2010: 273f.; Wahlgren 2004) and may serve as a cross-check; that is, their questionnaire results could be compared with the text analysis results.

A similar situation concerns conlangs as well as universal languages. Conlangs are hobby creations often devised by linguists by profession. The purpose of the authors is to exercise the limits of human languages, suit their personal tastes or devise a truly “neutral” language, namely one that can express all the possible nuances present in human languages (e.g. Ithkuil or Lojban). Such languages can be, of course, treated as small ethnic languages with little material available and consequently studied in the ESL paradigm. However, such an exploration cannot have a solid material base as there are no speech communities or real-life text instances of those languages (and therefore in Table 10 the ESL mark is qualified with an asterisk) and any such study may only appear as an interesting case showing the possible boundaries of language.

Even more auctorial views are presented in universal languages. The difference between these and conlangs is that the latter are often based on the existing traits of human languages, or, at least, if violating the assumed universals, are described as experiments not pretending to be true language-systems, whereas universal languages are meant to be neutral world languages based on common human abilities (or “the

universal grammar” – parallel to Chomsky intended; cf. section 2.2.1). As stated in the previous section, for want of material, inefficiency and ineffectiveness, universal languages have to be treated as texts, not as fully-fledged languages. The researcher has to bear in mind that the views presented in these texts are entirely auctorial and present a specific vision of the world. The vision is rooted in the mind frame of the period and as such should be analysed as a literary work set in the intellectual climate of the time. This way the grammar and the lexicon could be understood. However, the study of the etymology seems senseless in such cases as the languages of Dalgarno or Wilkins, where a letter represents a specific class of notions. A careful analysis of the organisation of letters and classes may reveal the hierarchy of concepts and the division of the world in the mind of the author but will not uncover the sources of the choice. Evidently, there are no intentional proverbs or idioms in universal languages but close examination of texts written in them might possibly show unconscious calques from their native languages or those most used.

Calques are also common in IALs. Languages based on a particular language family tend to transfer their meanings in a narrower sense into their lexicon. Therefore, the analysis could be a comparison of the semantic fields of an IAL and of its source languages. Idioms, proverbs, phrasemes and the like are almost exclusively calques from source languages (Fiedler 2007). The grammar of IALs follows the original structures, albeit in a simplified fashion. This is probably the only aspect that could be examined in the ESL or neo-Whorfian paradigm. The origins of particular words are very clear in auxiliary languages and therefore do not contribute much to future studies; yet the difference between the contemporary IAL usage and the established meanings in source languages could be treated as a contribution to modern etymology in IALs. Thus, the etymology would not be treated as tracing back the origins of a word in an IAL through its mother languages to the ultimate source language, but a study of seman-

tic change from source language meanings to an IAL established form (as given in dictionaries; if need be additional analysis of texts produced by its users might be taken into account).

Nonetheless, one IAL can be clearly analysed in the ESL paradigm, namely Esperanto. As shown in chapter 4, Esperanto is a naturally changing language, which has developed its own specific culture and provides ample material for study. The vocabulary of Esperanto is descended from (Whorfian) SAE languages (mostly Romance including Latin, around 20% from German and a small percentage from Slavic languages) and therefore, as in the case of other IALs, its usage might be compared to that of original languages (semantic shifts in a developing international community are not unexpected). However, through constant dynamic usage the lexicon may be considered independent. The fact that some lexical units are specific to Esperanto adds to the conviction that its vocabulary should be first and foremost studied in the ESL framework. Nevertheless, as the community is international, multicultural and multilingual, the analysis must include tasks designed to sieve out elements transferred from native languages and cultures (see section 5.3). The particularities of Esperanto, such as proverbs and idioms, are worthy of closer analysis. As in the case of ethnic languages, here to, original and calqued formulaic language exists. Fiedler (2015) explains there are three types of Esperanto proverbs and idioms:

many phraseological units have entered the language through various other languages. This group includes classical loan translations especially from the Bible as well as ad-hoc loans introduced by speakers from their mother tongues more or less spontaneously. Secondly, there is a group of planned, i.e. consciously created, phraseological units. They mainly go back to Zamenhof, the initiator of the language, who published an Esperanto Proverb Collection (*Proverbaro Esperanta*) in 1910. Thirdly, there are phraseological units which have their origin in the language and the cultural life of the speech community.

In the EUROJOS project, the copied structures are left out. It would be, however, interesting to see whether the proverbial forms are used in an unchanged form, and in the opposite case, to what extent the Esperanto ones differ semantically from the source ones. The same is valid for the etymology: it should be studied in retrograde, that is, show if and how the units have changed semantically (with the help of the five-volume *Etimologia vortaro de Esperanto*, which also compares Esperanto words with those of four other IALs). Diverse grammatical constructions are of great importance. Esperanto is an agglutinative/isolating language with about 40 lexical affixes, which, when connected with grammatical markers of class (-o for nouns, -i for infinitives etc.), may serve as independent lexical roots to build lexemes, for example:

<i>skribilo</i> ‘a pen’ (lit. ‘a tool for writing’)	vs.	<i>ilo</i> ‘a tool’
<i>skrib-</i> <i>-il-</i> <i>-o</i>		<i>-il-</i> <i>-o</i>
write tool noun ending		tool noun ending

The structure of the language encourages extensive compounding as in (made entirely of affixes):

<i>malindulino</i> ‘a woman not worthy of respect’				
<i>mal-</i>	<i>-ind-</i>	<i>-ul-</i>	<i>-in-</i>	<i>-o</i>
antonym	worthy	person	female	noun ending

Such compounds cannot be merely treated in terms of morphological transformations but rather of semantic compositionality. Here a questionnaire is invaluable (cf. the results of Koutny (2010)).

Another type of near-natural languages is revived languages. Even though their modern versions are based on undoubtedly natural languages, they require a great deal of plan-

ning and very often cannot boast a large speech community. The only successful example is Israeli, a Hebrew-based hybrid language (Zuckermann & Walsh 2011, cf. sections 3.3.2 and 4.2). As an example of language revival, one might also give modern societies of living Latin. In those cases, the researcher is in fact placed in front of two distinct varieties: one being the old living language with fully developed literature and a speech community, and the other being the revived version, strictly controlled and non-standardised. Therefore, the analysis should also be split in two: a study of remaining texts and modern dictionaries of the particular time, and a study of the modern variety using modern texts, dictionaries and questionnaires. Both should be later collated and compared, to control for semantic fuzziness, shifts in meaning and transfer. Etymology is here just as important as in ethnic languages – not only does it serve to show inner categorisations of words but also explains the choices made in creating neologisms for the modern variety (i.e. accounts for the auctorial decisions of the regulatory bodies). It is also worth seeing whether the speakers borrow any proverbs or idioms from other languages known to them.

Similarly, controlled languages prove to be natural, although with some peculiarities. Their grammatical structures and vocabulary are restricted to the very basic, but the expressive potential is probably near equal to that of everyday varieties of ethnic languages. It seems therefore that the LWV entrenched in them would present only a limited fragment of the reality, and the semantic fields would be restricted to core properties. However, the manuals of such languages are rather a set of dos and don'ts, and not dictionaries or classical handbooks. Texts and corpora are available (almost 90,000 words for ScaniaSwedish in the 1996 study; Almqvist & Sågval Hein 1996). The researcher cannot resort to questionnaires, though. Necessarily, they would only test users' knowledge of the language, and both the questions as well as the answers would pertain to the "full" language, not the controlled one. It is however worth seeing whether controlled circumstances

produce repeated worldviews regardless of the language (i.e. if, for example, ScaniaSwedish and Caterpillar English showed any regular similarities within their worldviews due to controlled usage).

Formal languages present a very unusual case. Their semantics is conventional and dependent on theory and particular usage. However, an investigation of the development of basic concepts (for example the infinity and how the concept has changed mathematics; see Murawski 2012 and Pogonowski 2012) may serve as a surrogate of etymological and contextual studies. Some mathematical problems were unsolvable for a long time as a result of rejection or the specific understanding of a notion. In evaluating the worldview of a specific text (equation, generative grammar tree), the researcher has to account for the contextual knowledge of the text's creator. Due to the absence of traditionally understood meaning, formal languages may only be treated as matrices for natural languages. Such an approach would require studying their grammatical structures according to neo-Whorfian assumptions (Lucy 1992a; 1992b) bearing in mind the strictness of rules governing the syntax of formal languages.

A great degree of abstractness is encountered in linguistic reconstructions. They too may be treated as textemes and as such studied in the neo-Whorfian paradigm. Lexical semantics poses a greater problem. The meanings ascribed to particular lexical items cannot be duly attested. Any analysis of the LWV may only be based on the existence of a lexical unit or the lack thereof, as wider contexts are not present. Yet, in the reconstructions, plausible cultural factors are taken into account, and this fact makes at least a sketchy study of the LWV possible. The researcher may reconstruct a semantic field or a wordnet and compare it with the existing archaeological evidence of cultural importance.

On the opposite pole of artificiality scale are sign languages. Being fully natural, they are analysable in the ESL framework. Until now, metaphor, metonymy and iconicity

have been investigated (Wilcox 2015). However, little attention has been given to the LWV of various sign languages. Here too, tracing back the source of signs could help determine to which language family a sign language belongs. Contrary to popular belief, sign languages are not related in the same way as oral ones (e.g. American Sign Language is descended from French Sign Language, while British Sign Language and its dialects such as Auslan are unrelated to it). Again, it has to be stated that such languages as the ASL, BSL and similar are natural, while Signed English or Signed Polish (*język migany* or *System Językowo-Migowy*) are manually coded forms of oral languages, that is, artificial systems. It is doubtful if the worldview entrenched in signed languages would differ in any respect from that of their oral counterparts.

In conclusion, it is postulated that all artificial languages be studied in the ESL framework, with the following being the most unproblematic cases:

- sign languages (as natural),
- revived languages (against the extinct version; fuzziness may occur due to small number of speakers),
- Esperanto (possible fuzziness due to multiculturalism and multilingualism),
- other IALs (against the source languages; small number of speakers),
- conlangs (experimentally),
- artlangs (experimentally).

	lexicon	grammar	etymology	idioms
artistic	auctorial/ culture first	auctorial	auctorial/ culture first	auctorial/ culture first
universal	auctorial	auctorial	auctorial	-----

IALs	IAL usage vs. source language semantic field	ESL	IAL usage vs. source language semantic field	calqued from source languages
Esperanto	ESL + influence of L1 or other	ESL	E-o usage vs. source language semantic field	ESL + (if calqued) E-o vs. source language usage
sign	ESL	ESL	ESL + family membership	ESL
conlangs	ESL*	ESL*	ESL*	ESL*
revival	ESL/modern vs. old usage	ESL/modern vs. old usage	ESL + auctorial	ESL
controlled	ESL/controlled	ESL/controlled	ESL	-----
reconstructed	retrograde	retrograde	retrograde	-----
formal	conventional	ESL + controlled	history of concepts	-----

Table 10 Approaches to artificial languages and some borderline cases [ESL – the Ethnolinguistic School of Lublin; * marks experimental research]

Controlled languages and signed systems are dubious cases. Because of the fact that they closely follow ethnic languages (controlled languages being minimal versions of ethnic languages and signed systems manual subcodes) their WVs may not differ from those of the languages they represent. What follows is that they stand as possible objects of study,

however unoriginal the findings may be. Formal languages and reconstructions may serve as textemes, that is, background cases against which the findings from natural ethnic languages may be studied. The only instance of a system treated as a literary work would be universal languages, in which the views of the creator could be reconstructed and compared with the contemporary mind-set.

5.3 Linguistic worldview in non-native languages

Artificial languages are never used as the sole languages of their users. Even if Esperanto has been nativised as the only IAL, its speakers are at least bilingual. The bi- and multilingualism of the speakers leads to an assumption that the worldview in this language will contain elements transferred from their native languages. Moreover, such an assumption is supported by the fact that philosophical languages as well as IALs contain multiple calques from either the languages known to the creator or the languages on which the artificial system has been based. Therefore, one of the research stages must contain various tasks designed to detect these elements and determine the impact of source languages on the analysed artificial language.

Aneta Pavlenko (2011b) in her overview of the research in bilingualism presents seven types of relationship⁶⁰ between the L1 and the L2, stating that the second one is probably the most frequent:

- co-existence (maintaining separate frames of reference in both languages)
- the influence of the L1 on the L2 (especially in bilinguals with beginning and intermediate L2 proficiency; so-called linguistic transfer or linguistic interference)
- convergence (so-called “in-between” performance)

⁶⁰ For simplicity's sake, I am going to call any type of cross-language influence *interference* (not to be confused with transfer).

- restructuring (divergence from the L1 patterns and convergence with the L2 ones, especially in advanced L2 speakers)
- internalisation (use of patterns absent in the L1 but present in the L2)
- the influence of the L2 on the L1 (in cases of prolonged exposure to the L2)
- attrition⁶¹

In the cases of languages labelled as texts and semi-languages in the previous sections, the second process will undeniably be the strongest one. A very tentative hypothesis would be that the most avid fans of some artlangs might internalise some elements of the language they speak (e.g. Klingon). This would have to be verified in a separate long-term study. Another separate study would have to concern speakers of Esperanto, in which almost all processes could be present, depending on the proficiency levels and the frequency of usage. However, the better the L2, the bigger the possibility that the speaker will come to prefer some L2 solutions to the L1 ones. This might be the situation of some advanced speakers, who actively participate in the movement and use Esperanto in their work, as well as in their families on a day-to-day basis (restructuring, internalisation and the influence of the L2 on the L1).

Generally the effects of bi- and multilingualism on semantics have been considered from the perspective of cognitive linguistics (e.g. feelings and colours in Athanasopoulos 2009; Athanasopoulos 2011; Pavlenko 2006; a general overview in Pavlenko 2011a). The study of the LWV for bi- and of multilinguals as proposed here is a new idea. Customarily, the LWV has been studied within a single language. Comparative studies have also been postulated (Bartmiński 2012a: 213–221); conducted within the EUROJOS project) but a study of conceptuali-

⁶¹ Attrition and L2 influence on L1 (and in some cases incomplete acquisition of the L1) are at present not very well distinguishable.

sations in bilinguals (in combination native plus non-native, or with two native languages) most likely has never been carried out in the framework of this theory. So far, the LWV has been described almost exclusively for natural languages. A questionnaire method for bilingual speakers of a planned language (specifically Esperanto) has been employed only in the research of Koutny (2010). Thus, it is proposed that additional experimental research on bilingualism be included in any study of the LWV for artificial languages. For Esperanto, the study can be conducted in the form of a two-part questionnaire: the first part would pertain to bilingualism (questions can be based on e.g. Li et al. 2006 or Marian et al. 2007) and the second specifically to the LWV (more on the proposed research for Esperanto in section 5.4). The additional part should take into consideration common factors which “affect the intra- and inter-speaker variation in lexical selection” (Pavlenko 2011b: 204f.), these being (i) linguistic, (ii) referent-specific, (iii) individual and (iv) text- and context-specific properties. In bilingual speakers, the factors influencing the choice of lexical units depends also on (i) the type of bilingualism, (ii) the level of proficiency and (iii) the length of exposure to the L2.

Such a questionnaire for smaller auxiliary languages would be extremely difficult to conduct for want of (advanced) speakers, and the results would not enable the researcher to draw conclusions about the whole language based on a few speakers. It is, of course, possible to carry out a translation test or any other task designed in accordance with the principles described in section 5.4; however, as these languages lack their own culture, it is almost certain the worldviews will be transferred not only from the source languages but also from the source cultures (see the problem of “fuzziness” in the previous section).

Similar problems apply to revived languages such as Latin or Cornish. Not only are they characterised by a small number of speakers but also the revived versions must obviously be

different from the original ones. Additionally the users are multilingual and therefore the results would be prone to show a great degree of fuzziness and transfer from the L1s of the speakers and other varieties upon which the revival movement draws its inspiration.

In artistic and modern constructed languages, interference might reveal itself in the form of errors made by the users. A careful study of texts produced by fans (e.g. poems written for annual contests of Dothraki, Na'vi or Klingon) will most probably show which linguistic categories are difficult for the users. It has to be noted that although the Internet allows for communication of people with different L1s, the majority speaks some Standard Average European language (most often American English, cf. Hendriks-Hermans 1999; Wahlgren 2004 for Klingon; of 135 members of the Language Creation Society 94 are living in the USA and only five in the Far East). This means that some categories appear unusual and difficult only because of the provenance of the users.

Contrary to conlangs and arlangs, which have been created for a specific group/society in a specific environment, philosophical languages were meant to be worldwide neutral languages replacing existing vernaculars or serving as an additional rational system facilitating logical thinking. Their presumed neutrality should be the focal point of the analysis. The transfer of cultural and linguistic elements from the creator's source domains could be made visible by comparing the structure of (a) the creator's native language and culture (the LWV) with (b) the shape of the particular universal language and, finally, with (c) the hypothesised linguistic universals. The study of (b) and (c) could assess the level of neutrality presented by the scheme, whereas that of (a) and (b) would clearly show the assumptions of its creator and the influence of their background on the shape of the language-system.

Nevertheless, the problem of linguistic interference seems to be most important for Esperanto, which is the only artificial language with a sufficient number of speakers and, at the same

time, the only one for which the ESL frame of analysis can be safely proposed. Although the language has developed naturally over the course of many years and can boast ca. 1,000 native speakers, the study by Koutny (2010) suggests that some linguistic categories might be transferred directly from the native languages of the speakers (see also section 5.4). Therefore, it would be advisable that further studies contain tasks designed to reveal those elements. Alternatively, the LWV obtained in further analyses would have to be compared with the existing LWVs in the languages indicated in the questionnaires to control for any possible interference. In this approach, the researcher would be able to separate the original Esperanto worldview from the transferred one. The existence of a homogenous worldview in certain domains would not only confirm the strength of Esperanto culture, but – if a consistent worldview without transfer appeared in the results of native speakers – it could also confirm the weak version of the Sapir-Whorf hypothesis. It has to be noted that the concepts of transfer in Esperanto have been considered in relation to its native users (e.g. Bergen 2001; Lindstedt 2006) but only as a study of nativisation.

5.4 Linguistic worldview for Esperanto⁶²

As stated in the previous chapters, Esperanto is a fully functioning language, which has passed through every one of the 28 stages proposed in Blanke's functional classification (2001; see section 2.3.3; developed specifically for planned languages, but possible to apply to other languages). In fact, some ethnic languages and creoles have not attained all of the levels. The existence of original literature, newspapers, web-pages etc. in Esperanto as well as monolingual dictionaries and a developing corpus, and most of all a vibrant speech commu-

⁶² This section was previously published in a changed form as Stria 2015c.

nity makes it possible for Esperanto to be studied in the LWV paradigm. The data specifically required to describe the LWV according to Bartmiński (2012a), Grzegorzczkova (1999) and in EUROJOS (2008) are as follows: (a) system (including hyponyms, opposites, synonyms, derivatives, collocations and proverbs), (b) dictionary definitions, and (c) real-life instances from texts, corpora and questionnaires (see section 1.4).

Although Esperanto is treated here as natural and meets all the theoretical prerequisites postulated by the EUROJOS team, there are constraints to the task of studying the LWV of the language. First of all, Esperanto has barely been studied within this paradigm. It is therefore difficult to establish a starting point. A further reason for this is changes in the Esperanto movement. The importance of one linguacultural domain and not the other is clearly dependent on the present socio-cultural situation within the movement. This situation has changed a lot during the years: from *finvenkismo* – an intellectual programme claiming that Esperanto should be the only international language (from *fin venko*, ‘final victory’), through *raŭmismo* (see section 4.2), up to the currently prevalent trend of collaboration with international institutions for language rights, linguistic equality and justice and even direct support for minority languages. Secondly, the ESL has not occupied itself with bi- and multilingualism. The possible transfer of concepts from native languages and cultures makes it difficult to judge if there are any domains coherent and typical of Esperanto. Thirdly, Bartmiński (2012a: 71) claims the content adequacy of profiled notions can be achieved in several ways, among them “using questionnaires with native speakers”. However, non-native Esperanto speakers constitute the great majority of users, and the *denaskuloj* are not norm givers. A relevant problem here is also who the mythical native speaker is and how they might be defined. In Esperanto, the notion “native speaker” translates as *denaska parolanto* (‘from-birth speaker’). It is, however, a well-known phenomenon that speakers gaining one language may lose the one they

have spoken from birth (language attrition) or in fact never fully acquire it. A familiar example is Joanna Krupa; a Polish-born celebrity now living in the US, whose first language is Polish, but which she never mastered, abandoning it for English – now her “indigenous”⁶³ language. Is she a native speaker of Polish or of English, or both, or none? In my opinion, none really, at least in the sense in which ‘native speaker’ is popularly used, namely a (mostly) monolingual speaker-from-birth of a language which they have mastered (fully and efficiently acquired). Another example is given by Miner (2011) – of a Croatian-born man, who spent his youth in Germany and then migrated to the US, never completely acquiring any of those three and considering English his “best” language (more examples of language attrition and partial acquisition in Pavlenko 2011b). Therefore, for the purposes of future research it is proposed here that a native speaker of Esperanto be defined as an advanced (to exclude language attrition cases) speaker-from-birth (truly “native”, acquiring language in family) continuously using the language throughout their adult life. What follows is, of course, that in the study a reasonable lower age limit should be established, for example 18 years (to exclude partial acquisition). Those and other prerequisites can be tested in a personal-data questionnaire.

Therefore, studies of Esperanto should, in general, primarily consult advanced speakers, active in the movement. However, there is no certainty that the answers provided by them will not be influenced by their L1 (or any other language used at sufficiently high level). The study conducted by Kou-

⁶³ This word used by Miner (2011) to differentiate between speakers-from-birth and those who have learnt a language in a speech community (‘native’ in Miner’s wording). He translates this opposition of ‘from-birth’ – ‘native’ as *denaska* – *indigena* into Esperanto and the latter will be here followed also in English. That is, Esperanto *denaska* will be denoted as ‘native’ (according to the etymology of both) and *indigena* will be translated as ‘indigenous’ to signify a language in which a speaker has the greatest competence.

tny (2010) suggests that this might be the case. It is certainly worth repeating with a larger and better-balanced sample of respondents and a different set of questions.

Consequently, as the study of Koutny has shown, the multilingualism of Esperanto speakers leads to an assumption that their worldview will contain elements transferred from their native languages, and a multilingualism survey should appear as an additional part of any LWV study of the language (see also chapter 6). Therefore, one of the research stages must contain various tasks designed to detect these elements and determine the impact of native languages on Esperanto. Ideally, to simplify the task, advanced users of Esperanto at the same time being indigenous speakers of the same language should be chosen as the respondents. However, finding a sufficient number of such speakers would definitely pose a challenge. Another extremely important stage of LWV research would be the study of a small group of native/indigenous Esperanto speakers in order to determine whether they present a consistent Esperanto-dependent worldview, as well as a study of the specific worldview pertaining to the culture of Esperanto, created in the communicative practice of the Esperantist community. This stage would also show whether Esperanto culture has a strong foothold in the community, in spite of not being taught to children at schools, as is traditionally done in the case of national languages.

The results of Koutny (2010), together with the theoretical analysis carried out in previous chapters, led to the formulation of the following research questions:

Q1: Will the LWV of non-native users be taken from their native language (L1)?

Q2: Is there a homogeneous, culturally embedded LWV of Esperanto, understandable for non-native speakers?

Q3: Is there a homogeneous LWV among native / indigenous speakers of Esperanto?

Answers to these questions can be obtained in a procedure comprising the following research tasks: (a) establishing a list

of survey questions, (b) preparing questionnaires and conducting surveys among respondents, (c) choosing and grouping questionnaires according to established parameters, (d) comparative analysis of questionnaires filled out by non-native respondents, (e) analysis of questionnaires filled out by native speakers.

The proposed project should cover a better-balanced group of respondents than those of Koutny (2010). The study sample should include **at least** approx. 125 advanced users of Esperanto. It is proposed that the data should be collected by selecting four equal groups of native speakers of different languages distant from each other in order to compare the results statistically. The first four groups of 25 persons each are to be as follows (with example languages; the researcher can choose otherwise):

culture language	European	non-European
Indo-European	Polish, German, French	Hindi, Farsi
non-Indo-European	Hungarian, Finnish	Chinese, Japanese

Table 11 Samples according to language

Preferably, all four groups should be used. However, the choice of languages should depend on the knowledge of those languages of the researchers and the resources.

The fifth group included in the project should be at least 20 (ideally 25) indigenous speakers of Esperanto. The diversification of the research pool according to language spoken would allow for determining the impact of language and culture on categorisations. It seems that 25 respondents from each of the groups is a number large enough to control for idiosyncrasy, while small enough to successfully carry out the project within a reasonable period. Any subsequent project should obviously strive for as large a number of respondents as possible.

The questions have to be written wholly in Esperanto to prevent transfer of structures from another language on the responses in the questionnaires. Again, for simplification, the responses (where possible) should be compared with LWVs known from the L1s of the users (this procedure was not used by Koutny), which will determine the scope of from L1. Such a procedure requires establishing the LWV of each language if ready-made analyses are not available for comparison or, alternatively, consulting experts on those languages or indigenous speakers. Subsequent studies should also include translation tasks for domains of interest designed to determine the extent of transferring.

THE TASKS

The first stage of the project is establishing a list of survey questions. The questionnaire should consist of an introductory part with questions on personal details, taking into account sex, age, native language, and the command of Esperanto and other languages (e.g. how often do you use Esperanto?). The remaining part of the questionnaire should only be accessible to the advanced users (of C1 level at least, indigenous speakers included). The rest of the questions should relate to the linguistic worldview of the respondents.

The questionnaires may be distributed traditionally in paper form (during congresses) or via the Internet through national Esperanto associations, mailing lists and other media, for example social networking sites. The latter method seems to be quicker and allows for reaching a larger number of respondents. After collection of the data (until the minimum of 25 questionnaires of each of four languages are collected, with at least 20 indigenous speakers in group five) there should be a brief period of quantitative analysis, that is, statistical categorisation of language, sex and age, as well as coding the responses, that is, grouping them by descriptors. The answers relating to syntax and morphology should be segmented into basic sentence parts and morphemes, and compared with

grammatical schemes occurring in languages of the control group (i.e. three unrelated languages of the greatest number of respondents). For example, the sentence *Mi devus fari ĝin* ('I should do it') is coded as S – Aux – V – O and compared with the Polish scheme (S) – Aux – O – V *Powiniennem to zrobić*. As the patterns of Standard Average European languages are largely similar, the inclusion of at least one non-Indo-European language and one Indo-European from outside Europe is advisable, if not necessary. The answers concerning vocabulary, collocations and idioms should be coded according to domains and descriptors (cf. e.g. the grouping in Bartmiński 2012a: 186). The LWVs of the language groups shall later be statistically compared.

After the division, the questionnaires filled in by non-native speakers of Esperanto should be analysed qualitatively. If most responses are matching (i.e. mutually compatible throughout, or compatible in terms of descriptors), it could be assumed that Esperantists of different languages and cultures share a coherent and consistent LWV (which might confirm the strength of Esperanto culture and a weak version of the Sapir-Whorf hypothesis). In the opposite case, that is, the lack of agreement (i.e. partial agreement or none at all), the responses grouped according to the respondents' native language should be compared with worldviews known from those L1s in order to determine whether there is any conceptual transfer from native languages to Esperanto.

Here a translation task can also be used. A very popular method for studying conceptualisation in bilinguals involves measuring response times in naming tasks or eye-tracking (see Heredia & Cieślicka 2015: 121) but this would be very cumbersome for research in Esperanto; it would be almost impossible to gather enough speakers of Esperanto and one other (the same) language in a laboratory. Therefore, so-called off-line tasks should be used instead, for example a translation task (Saygin 2001) or supplementation tasks (cf. Heredia & Cieślicka 2015: 118–122).

Those parts that do not have their counterparts in the native languages could be considered as belonging to the worldview of Esperanto. This task has basically two stages: S1, where only strictly cultural concepts (e.g. the importance of green, the term “citizen” and the five-pointed star, and idioms) are examined, as well as S2, where cross-linguistically varying domains are compared (e.g. colours, categorisation of animals and plants and the definitional characteristics of astronomical objects). The analysis of the questionnaires completed by native speakers should be carried out in the same manner. If the worldviews of these respondents are not uniform, individual responses should be compared with the second (or “strongest”) native language of the respondents.

SAMPLE QUESTIONS

The first part of the questionnaire should be questions on personal details, comparable to, or based on Li et al. 2006 and Marian et al. 2007.

The second part of the questionnaire should only be accessible to advanced users (C1 level at least, and indigenous speakers, excluding speakers on low and intermediate levels where transfer seems to be high). The chief objective in the selection of questions is taking into account different levels of language, that is, vocabulary, idioms, morphology and syntax as proposed in Anusiewicz et al. (2000), Bartmiński (2012a) and Grzegorzczkova (1999) (see also section 1.4). To establish the list of questions the assumptions of Lucy (1992a; 1992b) and Rosch (1978) (as an addition to the original Lublin LWV theory) as well as the preliminary study of Koutny (2010) may also be used.

The questionnaire should contain open-ended questions: “complete”, “name X”, definitional, the ‘but’ test, sentence transformation, and similar. Sample questions could be as follows:

1. Kian koloron havas? [what is the colour of:]
 - Suno [the sun]

- Sablo [sand]
 - Papriko [a bell pepper]
 - Vulpo [a fox]
 - Ĉielo [the sky]
2. Priskribu mallonge (se eblas, per unu – du vortoj)
[Describe briefly (if possible, use one – two words)]:
- Li ĉiam estis aktiva esperantisto, sed ĵus forlasis la movadon. (Kion li faris?) [He was always an active Esperantist, but has just left the movement. (What has he done?)]
 - Kiam ili estas inter esperantistoj, ili ofte parolas en sia denaska lingvo. (Kion ili faras?) [When they are among Esperanto speakers, they often speak in their native language. (What do they do?)⁶⁴]

The first question is designed to test whether Esperanto assigns a specific colour to each term, or whether these colours are borrowed from native languages. The domain of colours is known to vary cross-linguistically, and therefore provides a relevant test of linguistic transfer. The second question verifies knowledge of Esperanto culture and idioms related to it. Other questions can include (cf. Bartmiński 2012a: 132–148, 178–198): filling in the blanks (collocation test), questions about personifications (whether the sun is male, female, a child, etc.), prototypicality of plants and animals (e.g. “list five birds, plants, vegetables”), grammatical transformations (semantic compositionality test; see Koutny 2012: 119), the ‘but’ test (“Complete the sentence: ‘John is an Esperantist but...’” or assessment of the acceptability of ready-made presuppositions, e.g. “John is an Esperantist but he doesn’t speak Esperanto at all”). The ques-

⁶⁴ Expected answers are: *kabei(ĝ)i*, from the initials of Kazimierz Bein, who was a very well-known Esperantist until disappearing without giving any reasons, and *krokodili* ‘to crocodile’, i.e. speak one’s native language when Esperanto is supposed to be used.

tions should be designed not only to control the variation in basic cognition domains but also to take into account the symbols commonly associated with the Esperanto movement (green, five-pointed star, citizenship), as well as idioms and proverbs developed in the community and understood only in it (the selection of questions should be based on (Fiedler 1999), the most comprehensive scholarly work on Esperanto phraseology).

The data should be coded according to descriptors (e.g. coral, crimson, burgundy – descriptor: red) and subjected to statistical analysis. This coding can be explained using the example of the question: “According to you, what is a real Esperantist like?” Responses should then be grouped into domains (e.g. social aspect / ideological aspect / physical aspect, etc.) in which the keywords will be placed (e.g. respect, altruism / equality / green) extracted on the basis of the responses (e.g. *respects others and helps them selflessly / believes that all are equal / dresses in green*). This method has been repeatedly used in the ESL and in the EUROJOS project (Bartmiński 2012a; EUROJOS 2008). In the case of heterogeneity of the worldview, the answers should be compared with the already known worldviews of the native languages of the respondents in order to capture elements transferred from L1 to Esperanto. Those parts that are not found in the native languages may be considered as belonging to the worldview of Esperanto.

CONCLUSIONS

The results of such a project would be of both theoretical and practical importance. Firstly, the application of the LWV framework to a constructed language could improve the understanding of the development of languages in general and the ways of conceptualising the world using language. Moreover, the weak version of the Sapir-Whorf hypothesis could be tested.

Secondly, the results would also contribute to the understanding of how native/indigenous users of different languages

speaking the same language (in this case Esperanto) in certain social situations produce new culturally embedded meanings.

The last research question (Q3) concerns whether Esperanto indigenous speakers present a consistent LWV. This question posits a fundamental linguistic problem. Not only are native Esperanto speakers never monolingual, and therefore subject to the influence of another language (or several other languages), but also they do not always remain in close and constant contact with the speech community, which allows for questioning the consistency of their LWV. What is more, many a time Esperanto is not the person's dominant language, and with the passage of time, their use of it may even decline (see section 4.2). Note that native speakers never set the standards of Esperanto, which means they do not fulfil the same function as native speakers of (and in) other languages (cf. Fiedler 2012). A negative answer to the Q3 question could explain the ways of formation of the linguistic categorisations in multilingual users staying in non-Esperanto environment, while a positive one might indicate the validity of the weak version of the Sapir-Whorf hypothesis.

It is significant that such a project could be a contribution to the further, detailed study of Esperanto and pioneer research into other constructed languages (e.g. Ido, Interlingua, and Klingon) provided they develop a functioning speech community.

6. Research problems based on the example of a pilot study

The assumptions presented in sections 5.3 and 5.4 were partly tested in a pilot study conducted from July to September 2015 during the 100th World Esperanto Congress (25 VII – 1 VIII 2015, Lille, France) and through the mailing lists of Interlinguistic Studies at Adam Mickiewicz University (Poznań, Poland). The term ‘pilot’ should be understood here not as a feasibility study or a trial run, but rather as “the pre-testing or ‘trying out’ of a particular research instrument” (van Teijlingen & Hundley 2001).

The questions were designed by the author of the present book in collaboration with Professor Iлона Koutny and partly based on her previous research (Koutny 2010) as well as on Bartmiński 2007 and 2012a. The aim was also to compare the answers to the two questionnaires, and establish if there have been any particular changes after a period of about ten years (Koutny’s questionnaire was administered in 2004).

The analysis by Koutny (2010) was based on a questionnaire containing five questions on personal details and nine complex questions concerning the linguistic worldview. The questions related to both the grammar and the vocabulary of Esperanto. A hundred speakers of nineteen languages responded, of which four persons declared themselves native speakers of Esperanto. The study showed that Esperanto users may follow their native languages when naming colours of objects or assessing to which class (e.g. animate/inanimate) a notion belongs. At the same time, as Koutny (2010: 298, 300) implies, some cultural concepts are consistently recognised within the community, i.e. are understandable only to those familiar with Esperanto culture and actively participating in the community and in such a way form a specific Esperantic worldview.

The present questionnaire contains 16 complex open questions in Part I (the LWV questions) and 7 complex ques-

tions on personal data in Part II (for the full version in Esperanto and its translation into English refer to Appendix: The questionnaire). In Part I, the participants were asked to answer questions about the prototypicality of plants and animals (list 5 wild animals, 5 vegetables etc.), linguistic stereotypes (personifications of the sun, life etc., symbolic values of plants and animals and colour stereotypes), lexicalisations and collocations. Part II consisted of detailed questions verifying the level of Esperanto against the declared level.

Thirty questionnaires were collected; moreover, two native speakers of Esperanto were asked to participate. The general guidelines are to use samples constituting 10% of the future sample; however, Hertzog (2008) suggests at least 25 participants per group for testing of the instrumentation. The present pilot study was not conducted to include balanced groups of respondents; therefore, 32 questionnaires were gathered, as a number greatly exceeding the required 10% of the sample for full study (at least 125 respondents) and large enough to allow for selecting smaller control groups. It is also a number giving at least partial coherence of the results.

The importance of conducting a pilot study before a full-scale study is conducted cannot be underestimated. Van Teijlingen & Hundley (2001) give a list of reasons for conducting pilot studies, of which the following seem the most important for the present case:

- testing adequacy of research instruments
- assessing the feasibility of a full-scale survey
- identifying logistical problems which might occur using proposed recruitment approaches and study methods
- estimating variability in outcomes to help determining sample size
- developing the research question

As will be shown, the questionnaire approach proved to be a valid method for testing the LWV in Esperanto. Nevertheless, the initial assumptions underlying the study need to be

partly altered. Misleading or redundant questions need to be changed or eliminated, the recruitment approach adjusted to fit the modern communication media and research questions modified according to the preliminary findings.

A set of procedures was proposed for conducting a questionnaire-based study by Peat et al. (2002: 123, quoted in van Teijlingen & Hundley 2001):

- administer the questionnaire to pilot subjects in exactly the same way as it will be administered in the main study
- ask the subjects for feedback to identify ambiguities and difficult questions
- record the time taken to complete the questionnaire and decide whether it is reasonable
- discard all unnecessary, difficult or ambiguous questions
- assess whether each question gives an adequate range of responses
- establish that replies can be interpreted in terms of the information that is required
- check that all questions are answered
- re-word or re-scale any questions that are not answered as expected
- shorten, revise and, if possible, pilot again.

The present study has clearly shown that the most effective way to distribute the questionnaires is through web-based surveys. The processing of paper and e-mail distributed questionnaires is laborious and error-prone. Feedback was obtained through e-mail and as comments in the margins. This, as well as points 3 – 9, will be discussed in the present section along with the results obtained.

6.1 The respondents

The study sample was 32 respondents, of which 26 were male (81.25%) and 6 female. The statistics reflect neither the distribution in the general population nor the distribution in the

movement. Typically, more than 70% of Esperantists who participate in the studies are male (thus in Galor & Pietiläinen 2015: 29; cf. Piron 1989b: 166f.). Future studies should ensure higher participation of women to better mirror the actual structure of the movement. 30 respondents with higher education participated in the study and only two with secondary school education (although they either are pursuing or had pursued further education; private conversation). This number as well as the age distribution (presented in Figure 11) is similar to those found in the movement. One female did not give her age.

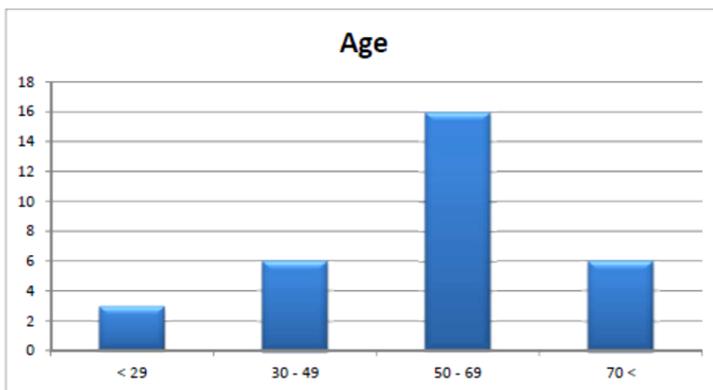


Figure 11 Age distribution among the participants

Nineteen respondents are currently employed (with at least one also being a student); of those not working, only one is a student, while 12 are retired. Ten of the participants are teachers, three are academic teachers, one a school headmaster and five are office workers. Three persons are connected to languages and linguistics: one is an editor, one a linguist by profession and one a translator. Only one person is a blue-collar worker. There are two lawyers, two accountants, an

engineer and a programmer; one respondent is an MD and two persons are medical personnel⁶⁵.

The mother tongues of the respondents are presented below in Figure 12. The numbers do not add up to 32, as there are several native speakers of more than one language. Two native speakers of Swedish use Hungarian and Esperanto as their mother tongues; one German speaker is fluent also in Low German and one in the Moselle Franconian dialect; the Farsi speaker considers also the Gilaki dialect as his native language, and one of the Portuguese speakers has learned Spanish from birth.

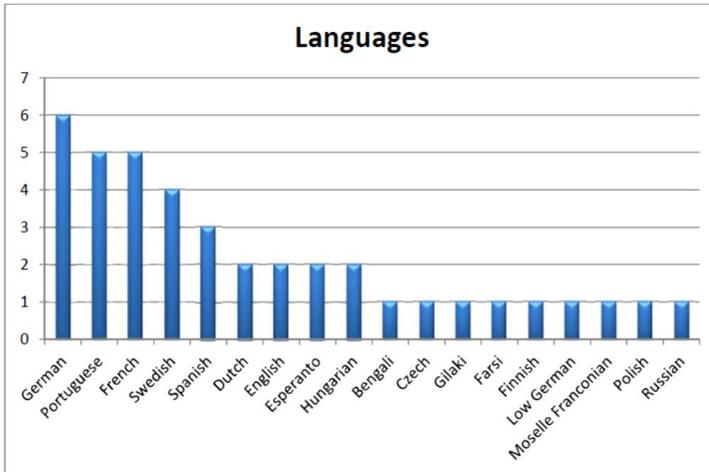


Figure 12 Native languages of the respondents

Of all the respondents, 31 declared knowing Esperanto to at least C1 level in some of the areas (speaking, listening, reading, and writing). Only one person (a native speaker of Esperanto) declared a lower overall level. Nineteen use Espe-

⁶⁵ In some cases, it was necessary to choose one of the given professions. The professions are given also for the retirees.

ranto every day, eight several times a week, four several times a month and only one less often (Figure 13).

As mentioned in 5.4, it is advisable to restrict the sample of respondents only to those who use Esperanto more often than once a month and to an advanced level (at least C1). Only such users could be expected to consistently demonstrate a special Esperanto worldview. Likewise, indigenous speakers using the language actively should be chosen. In the present study, all respondents were calculated as either being advanced or using Esperanto sufficiently often. Participants using Esperanto rarely or with lower declared level attest to the validity of our previous assumptions.

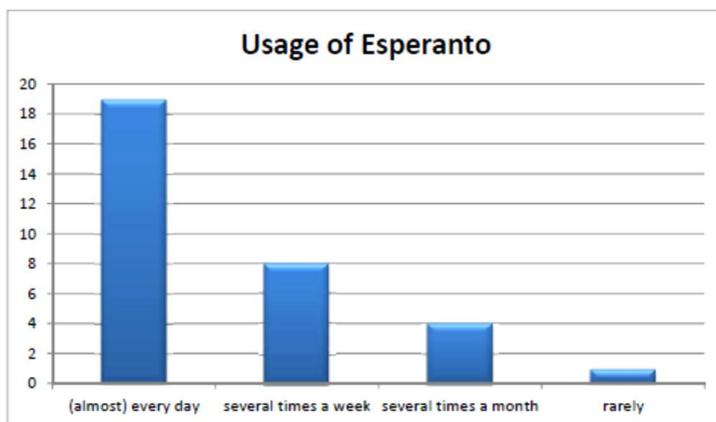


Figure 13 Usage of Esperanto among the respondents

The numbers are partly reinforced by the answers to further questions. Seven persons admit to listening to the radio and/or watching TV in Esperanto (which is a surprisingly large number, given the scarcity of Esperantic programmes). 22 read for pleasure in this language (more than in English, with the number being 15 and only two native speakers of Esperanto and English) and 10 read for work (however, 18 do so in English). As many as 28 write to friends and 23 write formally. 29 use Esperanto in conversations and 26 on the Internet

(again, more than in English, the numbers being 14 and 18; see Figure 14). Some of the numbers are almost as high as in case of native languages (Figure 15).

Of all the participants, 16 would choose to read a book in Esperanto if the original language was not known to them (only 7 in English, 7 in German, and 4 in French). As many as 19 would read it in their native language (Figure 16). One person commented that they would base their choice on the quality of the translation rather than on the language available.

The participants were also asked to assess in which language they do the following activities most often: count and do simple arithmetic, dream, express feelings, pray and swear. Esperanto had the highest count in expressing feelings (22), which is undoubtedly connected to the fact, that it is mostly used in informal communication (see Figure 17). Eighteen people do not pray; Esperanto had the highest count in this category right after native languages.

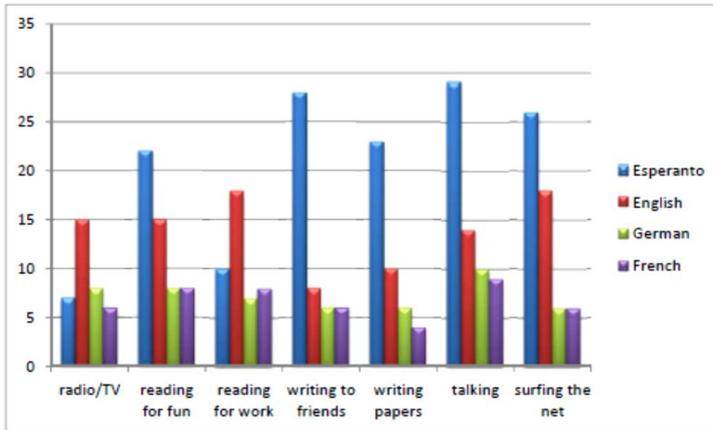


Figure 14 Usage of four most popular languages according to activities (Q II.5)

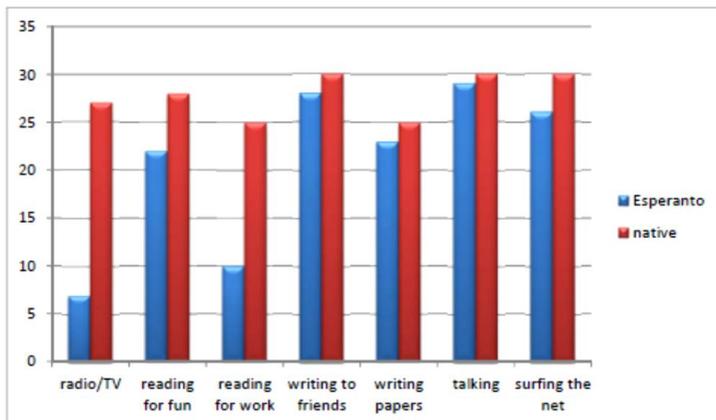


Figure 15 Usage of Esperanto vs. native languages according to activities (Q II.5)

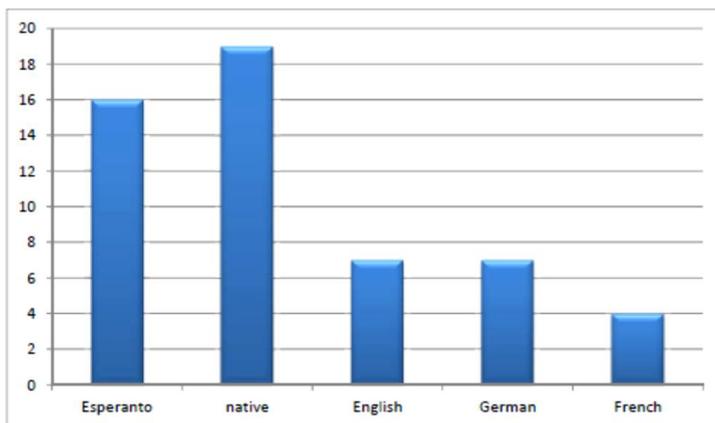


Figure 16 In which language would you read a book if the original language was not known to you?

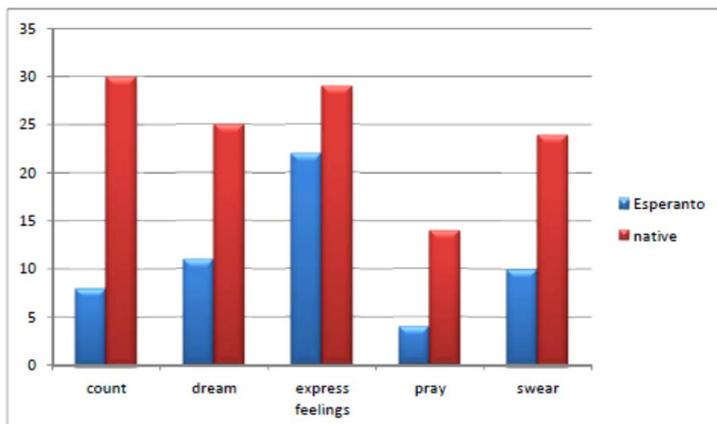


Figure 17 Usage of Esperanto vs. native languages according to activities (Q II.7)

Such detailed questions allowed for controlling the validity of the respondents' self-assessments in question II.3. A comprehensive part on personal data serves the purpose of selecting a desired pool of respondents, that is, according to what has been said in section 5.4, indigenous speakers of Esperanto and Esperantists actively using the language on an advanced level.

Visibly, the respondents use Esperanto quite often and in diverse fields. However, the language does not seem to be as well established as the respondents' native languages. Basic activities such as counting and dreaming take place mostly in native languages, although Esperanto is always on the second place (that is, before other foreign languages).

The answers may also partly serve as an explanation for the difficulties in filling in some of the questions in part I, that is, they show that Esperantists may have problems with, for example, listing plants and animals, as they do not usually engage in conversations about these topics. Another reason would be lack of schooling in Esperanto; topics typically taught in national languages are not present in the life of Esperantists.

6.2 The questions

The questions were designed to cover several areas of interest. They concern typical cognitive categories (colours in 1 and 11 as well as plants and animals in 2, 5 and 12), cross-culturally varying symbolic values of animals and plants (questions 10 and 13 pertaining both to cognitive domains and to culturally laden linguistic expressions), and grammatical-semantic categories (grammatical gender in 9 and collocations in 14), Esperanto culture (4) and finally stereotypes as understood by the ESL (6, 7 and 8)⁶⁶. Such a large array of domains results from the fact that Esperanto has not been sufficiently studied in the cognitive or the LWV paradigm. Therefore, the present questionnaire contains a much larger array of domains than usual and does not profoundly explore any particular domain. The need to identify domains specific to Esperanto arises not only in the present pilot study; such research will need to be continued in subsequent questionnaires as postulated in section 5.4.

The order of questions proved to be effective. We wanted to avoid any influence of the previous questions on further responses (e.g. Q2 is separated from Q5 and Q12). Questions 6-8 were grouped together on purpose, to see whether Esperantists connect cultural symbols (Q7) with a “true” Esperantist (Qs 6 and 8).

The translations are the closest possible to the Esperanto original.

COLOUR STEREOTYPES (Q1 AND Q11)

Let us now take a closer look at the results according to the areas. Esperanto uses eight basic colour names: *blanka* ‘white’, *nigra* ‘black’, *flava* ‘yellow’, *ruĝa* ‘red’, *verda*

⁶⁶ I am not going to discuss questions 3, 15 and 16 designed by Professor Koutny alone. Similarly, these questions are removed from the appendix.

‘green’, *blua* ‘blue’, *bruna* ‘brown’ and *griza* ‘grey’. Other colours are either expressed as compounds in the form of root + *-kolora* (‘X-coloured’, such as *rozkolora* ‘pink’, literally ‘rose-coloured’) or as ambiguous adjectives (e.g. *roza*, either ‘rosy’ or ‘pink’).

In Q1 the respondents were asked to indicate the colour of the sun, the moon, the sky, fire and sea (as cognitively the same but culturally different and symbolic), grass, light-coloured hair, a fox, sand (cognitive control group), and finally envy (as culturally laden). Multiple answers were allowed.

The **sun**, according to the respondents, is yellow (23 answers, 72%, of which 19 gave only this answer and one respondent wrote that it is chiefly yellow, but other – unspecified – colours may also appear), yellowish white⁶⁷ (1 as *flavablanka*) or yellowish red (1 as *flaveruĝa*), white (5), red (4, of which 2 gave this answer as the only one), golden (1 answer as *ora* and 1 as *orkolora*) and orange (1 as *oranĝa*). It would be well advised to include in the study more speakers of Asian languages, where the sun is culturally coded as red, to investigate if such answers appear also in Esperanto. The results of Koutny (2010) show the same colours (complex forms calculated as two colours, e.g. ‘yellowish white’ as 1 ‘yellow’ and 1 ‘white’), however in more varied proportions.

The **moon** received two main answers: ‘white’ (16) and ‘yellow’ (10). One person wrote that “it depends”, three that the moon is silver, one that it is silverish yellow (or yellowish silver, i.e. *arĝentflaveta*), further four that it is grey (of which one ‘light grey’). Red and orange appeared only one time each.

The **sky** is evidently blue for most (30 answers as well as two synonyms: 2 times ‘sky-blue’ and 1 time ‘azure’). Three

⁶⁷ Compounded colour names (*flavruĝa* or *flava-ruĝa*, ‘yellow-red’), names of type adverb + adjective (*flave ruĝa*, ‘yellowish red’ or ‘yellow-red’ diversely written) and with diminutive suffix *-et-* (*flaveta*, ‘yellowish’) are all normalised and translated as ‘X-ish’ (e.g. yellowish) for simplicity’s sake.

persons answered ‘grey’, two that “it depends”; one answered ‘white’ and one ‘rosy’ (*rozea*).

The **sea** was described as blue by 28 with further synonyms: 2 ‘greyish blue’, 1 ‘dark blue’, 1 ‘bluish’ (*glauka* from Greek *glaukos* designating light blue, greyish blue or blue-green) and 1 ‘sea blue’ (altogether 33 answers coded under the descriptor ‘blue’). Eight respondents think the sea is green and one that it is green-blue (which would be expected of speakers of Asian languages not present in the pool). Two persons answered ‘black’ and one ‘grey’.

Flames are red for 23 respondents; however yellow for 8, orange for 6 (only one as *oranĝkolora*), blue for 4 and white for 2. One person wrote ‘(light) yellowish red’ (*hele flavoruĝa*) and one ‘fire-coloured’ (*fajra*).

For most prompts, one answer appears as the main choice. However, the respondents often indicate more than one colour, being aware of the changing nature of the object which they are being asked about. No correlation between the answers and the culture of the respondents may be found – on one hand because the number of participants of a non-European-type culture was not sufficient, and on the other because of the diversity of the answers received.

The only prompt with a clear answer is **grass** with 32 respondents giving ‘green’ as its colour. There appear also ‘brown’ (2), ‘yellow’ (2) and ‘yellowish brown’ (1). Such consistency would therefore be expected to be repeated in Q11.5.

Light-coloured hair, a **fox** and **sand** received the most varied answers. The colour of foxes is obviously connected with the species occurring in a specific area. Nevertheless, the number of different similar colours which could be coded under the descriptor ‘russet/ginger’ might indicate that many Esperantists simply do not know the term in Esperanto (i.e. *rufa*, given only 3 times; one person explicitly wrote she lacks the term in Esperanto). Therefore, a **fox** is described as brown (11), red (5), russet/ginger (3), brown-red or red-brown (3), yellowish red (1), orange (1), cream-coloured (*kremkolora*, 1)

and greyish yellow (*grizflaveta*, 1). 5 participants answered ‘grey’, 2 ‘black’ and one ‘white’. **Light-coloured hair** is mostly described as blonde (16); however many other possibilities occur, including ‘white’ (8), ‘yellow’ (7), ‘grey’ (6) or “any, if dyed” (one empty). **Sand** is yellow for 9, white for 7, light-brown for 5; thirteen other possibilities follow (including one empty, ‘khaki’, ‘ashen’, ‘sandy’ etc.).

The last prompt was **envy**. It is green for 7 and yellow for another 7. Several comments appear, e.g. “no colour, but in my native language one can get green of envy” (French speaker) or “green, metaphorically” (English speaker). Envy cannot claim any colour for 5 participants. 4 answered ‘black’, of which 2 are native speakers of Swedish in which jealousy⁶⁸ is black (jealousy, *svartsjuka*, ‘black illness’ vs. envy, *avundsjukan*, etym. ‘illness of not liking’). One person answered *gala* ‘gall, biliary’. Seven other possibilities appear, including 3 empty answers, red (3, with one ‘dark red’), purple (2) and even 1 grey. Such a correlation with the native language was anticipated in abstract concepts.

The answers appearing most frequently were also expected to appear (in reverse) in Q11. For example, if most respondents wrote that the sun is yellow in Q1, it could be expected that ‘the sun’ would be the most frequent answer to prompt 4 in Q11. This is the case, even though the questions were separated to avoid the priming effect. However, ‘the sun’ gets a much lower count (only 9) than predicted. It was also assumed that ‘green’ would be immediately connected with Esperanto as the colour of hope, the Esperanto flag, the five-pointed star symbolising the language and the language itself. This, however, was not entirely borne out in the questionnaire.

White is the colour of snow (17), the moon (3) and 13 other items (including paper, pigeon/dove, death, aspirin and eye whites). **Grey** is the colour of mice (4). Other answers appear sporadically (two empty ones, ‘old person/old man’s

⁶⁸ Esperanto also differentiates between jealousy, *jaluzo* and envy, *envio*.

hair' 3 times, 'an unimportant person' 1 time, 'chagrin' 1 time etc.). However, if several answers are grouped in the domain weather (with descriptors **clouds/cloudy day**, **rain/rainy day** and **winter**), the number of such answers amounts to 8.

The highest count in the prompt **black** was 'night' (19). It is also the highest number among all prompts in this question. 5 respondents answered 'coal' and 3 'raven/crow'. There appear 11 other possibilities with 'god/the most perfect one' being the most interesting. The respondent is Iranian. One answer is a typical collocation – 'black hole'.

Yellow is the colour of the sun (9), sunflowers (6), lemons (4) and bananas (4). 11 other answers appear (1 empty), including 'envy' (German speaker) and 'gall/bile' (Portuguese speaker; according to him envy is black).

'Grass' has the highest count (16; 19 when grouped with 'lawn', 'field' and 'spring grass') in the prompt **green** although one would expect 'Esperanto' ('star' – implicitly 'Esperanto star' – answered 3 participants, including one native speaker). 8 answers are connected to trees and leaves ('forest', 'tree leaves' etc.). A German speaker wrote that it is hope that is green.

Blue received, not surprisingly, two main answers: the sky (18) and the sea (11, including the very specific 'Caspian Sea'). One person – an engineer – answered 'copper (II) sulphate'. This is an example of the so-called scientific worldview. One person answered 'drunken'. This is a case of transferring from the native language – here German, in which *blau* is a metaphor for 'inebriated'.

Red is the colour of blood (13) and tomatoes (5). One answer is a general football collocate – 'red card' and one a typical Swedish one – 'red house' (a native speaker of Swedish; culturally laden *röd stuga*, or even as *den röda stugan*, 'the red house', a summerhouse usually painted red). 10 other answers follow, e.g. 'apple' (3) and 'fire' (3). The last two prompts are **brown** and non-basic **purple** (given as *viola* in the questionnaire; one person commented that the correct form should be

violkolora). Both receive a range of answers: **brown** ‘chocolate’ (4) and ‘bear’ (4) as well as 4 empty ones, while **purple** mostly ‘violet’ (9) and ‘flower’ (6) as well as 4 empty ones. It is worth noting that brown features as the colour of Nazi shirts in two answers, as “two steps to God” as the answer of the Farsi speaker and “the colour of a donkey running away” as the answer of a Portuguese speaker. These are certainly culturally influenced.

It may be noted that in many cases colours seem to be ascribed to everyday objects without any symbolic values (white painkillers, yellow bananas, blue jeans, red pillow etc.). However, there appear also similitudes known from native languages and cultures (general Europe-descended as “white as snow” or culture specific such as the Swedish “red house”). Esperanto-specific expressions do not appear apart from ‘green star’ three times.

PLANTS AND ANIMALS (Q2, Q5 AND Q12)

Various plants and animals were the topic of Qs 2, 5 and 12. In Q2, the participants were asked to list 5 animals and 5 plants. As expected, the most popular answers were domesticated animals: ‘dog’ (22) and ‘cat’ (21); but the 3rd position is occupied jointly by ‘lion’ (10) and ‘horse’ (10). Fifth on the list is ‘cow’ (8). As for plants, the numbers are lower: ‘rose’ (21), ‘tulip’ (10), ‘grass/herb’ (8), ‘tree’ (6) and ‘fir tree’ (*abio* can also signify a Christmas tree in general) (6). Such answers as *salato* (incorrect for ‘lettuce’, which in Esperanto is *laktuko*; both 2 times) and *ĉajotarbo* (literally ‘chayote tree’, i.e. *Sechium edule*, an edible plant belonging to the gourd family *Cucurbitaceae* native to Mesoamerica, popular in Brazil; it is a sprawling plant, not a tree) are interesting. There appear several plants not native to Europe, which is not surprising, given that several of the respondents are non-European.

Q5 focuses on plants: trees, vegetables and fruits. The five most popular trees are oak (18), fir (13), birch (12), pine (11)

and apple tree (10). One person gave ‘grapevine’ (not a tree, obviously), which might stem from the fact that Esperanto allows the *-ujo* ending, meaning a plant on which an X grows (*pomujo/pomarbo*, ‘apple tree’, *mangujo/mangarbo*, ‘mango tree’ etc.). There appear also trees from outside Europe, such as jacaranda, mahogany (supposedly *Swietenia* trees), araucaria or avocado tree (respondents from outside of Europe). Among vegetables, the first four places are occupied by ‘carrot’ (18) and ‘potato’ (18), ‘cabbage’ (12), ‘cucumber’ (9) and ‘onion’ (8) jointly with ‘tomato’ (8). Interestingly, ‘tomato’ appears also 3 times as a fruit. 4 participants listed *salato*, while 3 *laktuko* (‘lettuce’, see above). Fruits are listed as follows: ‘apple’ (28), ‘pear’ (21), ‘orange’ (18), ‘banana’ (16) and ‘cherry’ (*ĉerizo*, Polish *czereśnia*; 9) with ‘grapes’ (9). The Czech speaker listed ‘sour cherry’ (*acida ĉerizo*, Polish *wiśnia*) as the Czech language distinguishes between those two types. Both among vegetables and fruits there appear exotic (for Europeans) species.

Q12 is devoted to birds and wild animals. The first five places are occupied by the following birds: eagle (21), sparrow (17), swallow (14), pigeon (9 times as *kolombo*, however a Swedish speaker listed *dovo*, Swedish *duva*, English *dove* and a Portuguese speaker *palomo*, Spanish *paloma* – both incorrect in Esperanto and clearly borrowed from known languages), seagull (8) and duck (8). Wild animals are lion (20), tiger (17), wolf (16), bear (12) and elephant (11). Answers with such high numbers would be expected to occur also in Q2; nevertheless, the only wild animal making it to the first five is the lion (10). As the sixth answer in Q2 appears ‘elephant’ with only 6 listings.

The respondents’ spelling errors and borrowings from native or other known languages show that this domain is not sufficiently known among Esperantists. The answers vary to a great degree also because of the participants’ very different geographical and cultural backgrounds. It may be assumed that some convergence (i.e. the popularity of the answer ‘dog’) is

due to the common cognitive basis, that is, the prevalence of particular species. Idiosyncrasies originate, in turn, from environmental differences.

SYMBOLIC VALUES OF ANIMALS AND PLANTS (Q10 AND Q13)

The linguistic worldview of the respondents concerning animals and plants was tested in Q10 and Q13. In Q10, the participants were asked to name animals that are typically connected with such features as fidelity (loyalty), courage and strength. One person refused to answer the question due to personal beliefs, saying that she does not approve of stereotypical judgements (Swedish native speaker). **Courage** is typically connected with lions (24). **Stupidity** received 17 answers ‘donkey’, 3 ‘hen’, 3 ‘cow’ (2 as *bovino*, ‘cow’ and 1 as *bovo*, ‘cattle’), 2 empty and seven other answers 1 occurrence each, of which one was ‘tapir’, an animal typically considered stupid in Brazil (*anta* may serve as an insult in Brazilian Portuguese). **Fidelity** is symbolised by dogs (26). **Obstinacy** (stubbornness) received a full spectrum of answers: ‘donkey’ (17), ‘bull’ (as *virbovo* or *taŭro*, 3; one person listed ‘cow’ and one ‘cattle’), ‘goat’ (2 as *kapro*, 1 as *ibekso*, ‘wild goat/ibex’), ‘mule’ (2; 1 ‘mule’, 1 ‘molly mule’), and five other.

Speed (rapidity) is symbolised by the hare (9), the cheetah (5), the gazelle (4), the panther (4) and other. **Hard work** (laboriousness, diligence) received only five answers with ‘ant’ as the dominant one (20); ‘bee’ (7), ‘horse’ (5), ‘cattle’ (3) and ‘beaver’ (1). One can be **cunning** as a fox (26) but also a cat (2), a hare (1), a snake (1), a jackal (1) and a wolf (1). **Strength** is a trait commonly ascribed to bears (11) and elephants (10). Other answers were ‘bull’ (5, but also one ‘cattle’), ‘horse’ (4), ‘lion’ (3) and ‘buffalo’ (1).

The respondents’ usage of animal names is consistent with the Esperanto system: the basic form denotes a whole species (i.e. *bovo*, ‘cattle’), while male/female distinction is

made through addition of either the feminine suffix *-in-* (*bo-vino*, ‘cow’) or the lexical morpheme *vir-* (*virbovo*, ‘bull’).

Q13 concerned the metaphoric value of various plants. The respondents were asked to supply plant names to which they ascribe the following features: slim/slender, small, hard, strong, tall, sour/acidic, and stupid. Two respondents did not answer this question, saying that “this is out of their knowledge”. Apart from ‘strong’ and ‘sour’, most prompts received an empty answer as the most frequent one. **Stupid** is not usually connected to any plant by 22. Two participants answered *štīpo*, ‘log’, which comes from Zamenhof’s works. Two answers are particularly of interest: *sekalfungo*, ‘*Claviceps purpurea*, rye fungus’ listed by a former farmer (specialised viewpoint) and *sambuko*, ‘*Sambucus*, commonly known as elderberry’ given by the Farsi speaker, who commented that he transferred the answer from his native culture.

Slim/slender received 8 empty responses. 4 participants answered *betulo*, ‘birch’, 3 *palmo*, ‘palm tree’ and 6 connected this feature with diverse legumes (the common bean 2 times, the broad bean 1, the broad bean’s pole 1, the common bean’s stalk 1, and a stalk in general 1). 3 respondents chose bonsai as a symbol of being **little** (9 empty) and another 3 the daisy, *lekanto*. As **hard** (10 empty) is given the oak (4) and the nut (5) or the coconut (2). Two persons listed ebony (one using in a compound an incorrect form *-holzo* from the German *Holz* instead of the Romance-derived *ligno*). **Strong** (6 empty), again, is the oak (16). **Tall** (7 empty) is the pine (6), the fir tree (4) and the redwood (3 as *sekvojo* and 1 – in quotation marks – as *ruĝarbo de Kalifornio*). 15 respondents listed the lemon as **sour** (6 empty); 4 the sorrel, *okzalo*.

Q13 (“Which plants are the symbols of ...?”) was one of the most difficult questions for the respondents. They complained that it was too complicated and that it took too much time to find an answer.

The respondents had much less difficulty in answering questions about animals. It seems that various animal names

collocate quite strongly with different attributes. Some similes received a high result regardless of language (the dog as a symbol of fidelity or the lion symbolising courage). However, most of the respondents speak an SAE language or come from a Europe-based culture. The consistency of the results should therefore be tested again with a more diversified sample of participants. Again, the wide array of responses comes definitely from environmental conditions (cf. questions above).

GRAMMATICAL GENDER (Q9)

The metaphorical gender of objects and concepts caused the most problems (Q9). Even though many respondents speaking languages with no grammatical gender tried to fill in the prompts, their listings were quite inconclusive. The Russian speaker wrote explicitly that in answering the question he followed the distinctions present in his native language. The same case (i.e. transferring the native language's gender to Esperanto) is visible in the answers of the Polish speaker. Moreover, also in languages with grammatical gender present (German, French, Spanish), some prompts were not ascribed any specific gender and were rather treated as objects or left empty⁶⁹.

In fact, only the first four prompts (the sun, the moon, life and death) received responses to the greatest degree congruent with the grammatical gender system present in the native languages. It may be assumed that these words have the biggest symbolic value and therefore are consequently recognised as male/female. Other sets of answers were surprisingly inconsistent. 'Fork', for example, is feminine both in German and in Czech (*die Gabel, vidlička*). The Czech speaker and one German speaker ascribed masculinity to it, contrary to expectations, while the remaining German speakers left out the gender entirely.

The **sun** is female according to 8 participants, including 5 German speakers (feminine *die Sonne*) and 1 Czech

⁶⁹ In this case answers with no gender defined explicitly were treated as empty, even if some characteristics were given.

speaker (contrary to its neutral gender in Czech). However, the sun is male for 9 respondents, including 1 German (contrary to the feminine form in German), 2 Spanish and 2 Portuguese speakers (one of them speaking both natively; masculine *el sol, o sol*) and 1 French (masculine *le soleil*).

The **moon** is female for 12 respondents, i.a. 3 Portuguese and 2 Spanish speakers (feminine *a luna, la luna*), 2 French (feminine *la lune*) and 1 German speaker (contrary to the masculine form in German). It is male for 7 participants, i.a. 4 German speakers (masculine *der Mond*).

Koutny's study (2010: 301) revealed that the answers to the question about the gender and animacy of the sun, the moon, life and death show great convergence with the grammatical gender in the native languages. In the present questionnaire, the answers were much more inconclusive, although similar.

COLLOCATIONS AND LEXICALISATIONS (Q14)

Q14 concerned collocations and compounds. The respondents could freely fill in any word they could come up with, in contrast to the questionnaire by Koutny (2010) where several ready answers were given to choose from in Q3. We repeated some of the prompts present in Koutny 2010, that is 'illness', 'question', 'bell/chilli pepper' and 'hospital' (the first three in Q3 "Which of the words would you preferably use" and the last one in Q5 "Say in other words").

Illness was *grava*, 'grave' for 12, *serioza* 'serious' for 3 and *severa* 'severe' for 2. *Forta* 'strong', *akra* 'sharp, acute' and *akuta*, 'acute' appeared one time each. In Koutny 2010, *grava* gained 60%, while *serioza* 33%. The third possibility was to use *peza* 'heavy', but only five such answers appeared. In the present questionnaire, this answer was never used. Many respondents chose to give an adjectival form denoting a specific illness, e.g. 'stomach', 'heart', 'flu'.

The second gap was "John's grandfather remained in _____ [where? not at home]". It was expected that

most respondents would use *malsanulejo*, a compound meaning ‘a place for non-healthy people’, i.e. hospital. This was the case: 17 participants used *malsanulejo*. However, 12 chose the word *hospitalo*. Only one used *kuracejo*, ‘place for treating’ and two ‘nursing home’ (a compound, *maljunulejo*). In Koutny 2010, 89% gave *malsanulejo* as a synonym of *hospitalo*, and only 3 participants *kuracejo*.

The respondents were expected to fill in the third gap with an equivalent of the English ‘food/foods’⁷⁰. 21 of them used *manĝaĵo*, ‘meal, dish’, and only 4 *manĝo* ‘food; eating’; 1 person used *manĝaro*, ‘a collection of foods’. 5 other responses appeared (specific dishes). Three participants violated the agreement between the previous singular adjective and the following word they should use, using the plural form. Such mistakes could suggest that *manĝaĵo* is treated rather as ‘foods/a dish’ than ‘food’ in general, which also complies with the Esperanto system, where the *aĵo* ending implies concretisation of abstract ideas.

Analogically to gap 1, in gap 4 “the _____ [who? works there], who is called Molina” we wanted to investigate, whether *doktoro* would appear quite as often along *kuracisto* as *hospitalo* along *malsanulejo*. This is not the case, as *kuracisto* seems to be quite well entrenched in the language as ‘a physician’, while *doktoro* would rather be interpreted as a title (either academic or honorific). Only 1 person answered *doktoro*, while as many as 21 *kuracisto* (including three in the feminine form, *kuracistino*). 9 respondents chose either *flegistino* (7, ‘female nurse’) or *flegisto* (2, ‘male nurse’). Visibly, even in Esperantoland there is the traditional role division present: most participants think a doctor is male, while a nurse is female. Interestingly, the task clearly mentions in its last part that Molina is a male (“Sinjoro Molina”). However, so

⁷⁰ Mistakes such as no accusative ending or use of the plural instead of the singular form were normalised and treated as correct. Forms in the nominative are given here.

many uses of the female forms suggest that most participants ignored this information and focused on their worldview while filling in the gap. The high number of male forms cannot be ascribed solely to the fact that Molina was given as a male. Some speakers might have used *kuracisto* as a generic form including male and female doctors.

Collocations were again tested in gaps 5, 6 and 7. **Papriko** (more often as *kapsiko*) may be *spica*, ‘spicy’ (7), *akra*, ‘sharp’ (as in e.g. Polish *ostra papryka*, 7), *forta*, ‘strong’ (3), *pik(ant)a*, ‘spicy, piquant’ (3 *pika* and 2 *pikanta*). Koutny (2010: 300, 305) offered the following answers: *akra* (34%), *pikanta* (29%), *pika* (20%) and *forta* (11%). Some respondents wrote that none of the suggestions suited them. The participants in this study chose *spica* (not appearing in Koutny 2010) as often as *akra* (both 22%). A question, **demando**, one may *fari*, ‘make’ (17; 53%), *starigi*, ‘set up’, lit. ‘make stand’ (11; 34%), *levi*, ‘raise’⁷¹ (5) and *demandi*, ‘demand, ask’ (2). This confirms the results of Koutny (2010: 300), who reported the following: *fari* (54%), *starigi* (34%), *meti*, ‘put’ (9%, not present here) and *doni*, ‘give’ (sporadically, not present here). In the Esperanto corpus (<http://tekstaro.com/>) there are 273 contexts with the form *demandon* (accusative). *Fari*, *meti*, *starigi* and *levi* (in diverse forms) appear altogether 89 times, of which *fari* collocates 58 times (65%), *meti* and *starigi* 13 each (14%) and *levi* only 5 (6%). A photograph, **foto**, collocates most often with *fari* (24; 75%) and *preni*, ‘take’ (5; 15.5%).

Importantly, the derived forms *kuracisto* and *malsanulejo* are lexicalised and appear much more often than non-derived borrowings. Such synthetic morphological forms appear also in the study by Koutny (2010) regardless of the type of L1. More frequent use of such complex forms as *malsanulejo* confirms the view presented in section 4.2 that many Esperanto

⁷¹ This is probably because *levi la demandon* collocates very often in the sense of ‘raise an issue’.

users would rather introduce vocabulary by endogenous means.

Both studies confirm also that collocations are not as fixed as in native languages. Parallel forms occur under the influence of similar constructions in other languages. However, there is evidence that some forms are beginning to take over (e.g. *fari demandon* in contrast to *starigi demandon*). This corroborates the assertions from sections 3.3 and 4.1 that Esperanto is nearing natural as it changes spontaneously and is not fully codified.

ESPERANTO CULTURE (Q4)

Question 4 was designed to give an answer to research question 2 “Is there a homogeneous, culturally embedded LWV of Esperanto, understandable for non-native speakers?” The following cultural concepts were expected as responses:

- *kabe(iĝ)i*, stop being active as an Esperantist; from the initials of Kazimierz Bein, who was a very well-known Polish Esperantist until disappearing without giving any reasons
- *kongresa edzino*, ‘congress wife’, i.e. a partner only for the time of the congress
- *krokodili*, ‘to crocodile’, i.e. speak one’s native language when Esperanto is supposed to be used
- *finvenkist(in)o*, a person believing in the “final victory” of Esperanto (*fina venko*)
- *ĝisostulo*, ‘to-the-bone guy’, i.e. a die-hard Esperantist
- *volapukaĵo*, something incomprehensible, senseless; from the name of a rival planned auxiliary language Volapük
- *homaranismo*, a philosophy developed by L. L. Zamenhof based on the ethic of reciprocity

In several cases, we received more than one answer; however, only 31 respondents filled in this question. Three out of these were also tested in Koutny 2010: *kabeiĝi*, *finvenkisto* and *kongresa edzino*. The respondents were asked to supply their definitions.

Kabei (an intransitive verb) was given 26 times and *kabeiĝi* (a reflexive form) 4 times. This cultural concept is thus known to 29 out of 31 who responded. The only ones who paraphrased the prompt, not knowing the expected answer, were the two *denaskulinoj* ('stopped being active' and 'lost interest'). This bears out the results of Koutny (2010: 298), who writes that 8% of the respondents were unable to define the word (with 2 out of 4 *denaskuloj*).

Kongresa edzino and *kongresedzino* appeared 12 times (only 37.5%). 4 participants did not respond at all. Non-fixed synonymous forms appeared several times, e.g. *kongresa amatino* (3, 'a congress lover'), *okaza kunulino* (1, 'an occasional companion') and *kongresa samlitano* (1, 'a congress bed-sharing woman'). There were also forms attesting to the creative power of compounding/derivation in Esperanto: *leĝerulino* (1, 'an easy-going girl') and *unusemajnolino* (1, 'a woman for one week'). One response was clearly a word play based on another culturally embedded Esperanticism: *eterna amkomencantino* (1, 'an eternal love-beginner') is a reference to *eterna komencanto* 'an eternal beginner', i.e. a person who participates in the Esperanto community for a long time, but still does not speak Esperanto well. Two persons responded *papilia amo*, 'butterfly love', which is an expression meaning fickle and unserious love (known from works of Zamenhof). It is however possible that more respondents passively know the expected expression. In Koutny 2010, the comprehensibility of this fixed expression was much higher (18% did not understand, 9% misunderstood; 3 out of 4 *denaskuloj* did not know it).

All 31 respondents know the expression *krokodili*. This is not unexpected, as this word is also widely known outside Esperantoland as an example of original Esperanto culture and linguistic creations.

The answer *finvenkistino* was given 25 times (78%). 3 participants responded *esperantistino* (with adjectives *ortodoksa* and *idealisma*). The answers contribute to the stereo-

typical view of an Esperantist (see below) as hopeful (*esperplena*, 1), idealistic (1), but also naïve (1) and dreaming (*revemulino*, ‘a female dreamer’, 1). The number of correct responses is a little lower than that in Koutny 2010: 3% were not able to define the word and further 9% misunderstood.

Finvenkistino appears also one time in the next prompt. The highest score was *ĝisostulo* (15; 47%). The two native speakers wrote ‘my father’ (without consulting each other) – this shows the distance that native speakers not participating in the culture have to those active in the structures. Other answers also contribute to the stereotype of an active Esperantist: *fanatikulo* (‘a fanatic’, 2), *militantema* (‘fighting for, waging war’, 1), *movadano* (‘the Movement member’, 1), *kredanta aktivulo* (‘a believing activist’, 1), *fundamentisto* (‘a fundamentalist’, 1; a reference to the *Fundamento* from 1905) and *Esperanto batalanto* (‘an Esperanto fighter’, 1; also a culturally embedded phraseological unit).

Volapukaĵo appeared only 10 times. 8 participants responded *strangaĵo*, ‘something strange’. Other synonyms to ‘strange, odd’ also appeared. One respondent used an idiomatic expression from his own native languages (Czech): *hispana vilaĝo*, ‘a Spanish village’ and one from Zamenhof’s collection of proverbs *ĥina scienco*, ‘Chinese science’.

Homaranismo is known to 20 respondents. 2 added *hilelismo* (an earlier name for *homaranismo*). There appeared several answers not directly connected to Esperanto: ‘Christianity’ (2), ‘world peace’ (1), ‘a global human family’ (1) and ‘universal fraternity’ (1). However, one respondent wrote *interna ideo*, ‘an internal idea’ – this is a well-known quotation from the 1912 congress speech of Zamenhof, in which he states that the internal idea of Esperanto is to remove the barriers between peoples and to promote fraternity. In addition, there appear cynical answers: ‘a utopia’, ‘Santa Claus’ and ‘something that will never happen’.

In this question also, synthetic derivations and compounds appear often, regardless of the type of the L1 of the partici-

pants (similar to those in European languages as *revemulino*, which might be compared to *marzycielka* in Polish and *Träumerin* in German or highly compositional semantically as *samlitanino*).

The results from Q4 show the strength of Esperanto culture among active Esperantists, even though it is not taught as a part of an educational system. Although some respondents did not answer according to expectations, many responses draw from Zamenhof's literary works and his collection of proverbs, Esperanto sayings and cultural keywords (e.g. *papilia amo* and *Esperanto batalanto*)

They also confirm the assumption that native speakers of Esperanto will not be able to recognise some cultural concepts if not being active Esperantists. Such is the case here: the only concept known to both *denaskulinoj* is 'to crocodile'.

STEREOTYPES (Q6, Q7 AND Q8)

Questions 6, 7 and 8 concerned the stereotype of an Esperantist as seen by Esperantists themselves. I will present the results in the following order: 8 (what is a true Esperantist like?), 6 (the 'but' test) and finally 7 (cultural symbols). The results were interpreted along the lines of Bartmiński (2012a: ch. 14).

Q8 inquired about a "true" Esperantist, combining a prescriptive (ideal, exemplary) view and a descriptive (typical) view (see more in section 1.4). One person did not respond, asking if there can be "not true" Esperantists. One person specifically divided his response into "typical" and "ideal" quoting Melnikov (Melnikov (1992) introduces the concept of 'a typical Esperantist' taking part in culture, however the concept differs from that of 'an average Esperantist') and giving his own opinion.

The participants provided 27 answers (synonymous expressions are normalised and counted as one; Table 12). The most frequent answers were 'uses the language' (11), 'knows the language well' (9), and 'works for the benefit of Esperanto' (9). The stereotypisation index of the two most frequent

features is 29.85, which is a very high value (Si, see Bartmiński 2012a: 182).

domain	No. of features	sum of features frequencies	No. of frequencies	descriptors
linguistic	3	22	11	uses the language
			9	knows the language well
			2	multilingual
cultural/connected to the movement	5	20	9	works for the benefit of Esperanto
			4	takes part in Esperanto culture
			3	meets other Esperantists
			3	supports the movement in some way
			1	interested in Esperanto
ideological	4	11	6	<i>homaranisto</i> / lives according to the internal idea

			2	peace loving, pacifist
			2	idealistic
			1	works for equal rights
psychological	6	6		stubborn optimistic curious studious open likes travelling
social	5	6		tolerant (2) talkative respectful good-hearted hospitable
everyday life	2	2		educated hard working

Table 12 The features of a “true” Esperantist arranged by aspect

Question 6 required of the respondents to imagine a stereotypical Esperantist and supply a contrasting feature. Thirty different answers were given and only one clearly three times (‘still an eternal beginner’). Other features needed to be grouped according to domains. The table below shows the presupposed features (e.g. those imagined to be characteristic of an Esperantist). Out of the 6 most frequent ones, 4 features coincide with those of a ‘true’ Esperantist. The new features ‘improves knowledge of Esperanto’ and ‘is not a fanatic’ can be explained by the fact that the question used the modifier “ardent”, thus making the participants think rather of an ‘ideal’

Esperantist. The feature ‘cat lover’ is an inside joke (as there are regular cat lovers’ meetings during Esperanto congresses) and an ironic observation rather than an automatic presupposed response.

domain	No. of features	sum of features frequencies	No. of frequencies	descriptor
linguistic	3	12	7	knows Eo well
			3	improves knowledge of Eo
			2	uses the language
cultural/connected to the movement	4	14	7	is not a fanatic
			4	participates in the movement
			3	works for the benefit of Esperanto
social	4	5		his wife is an Esperantist (2) stiff, aloof social proud of being an Esperantist
psychological	1	1	1	a cat lover

Table 13 The features presupposed in the formula ‘John is an ardent Esperantist but...’

In question 7, the participants were asked to list cultural symbols of an Esperantist (Table 14 Symbols of an Esperantist). We aimed rather at concepts of symbolic value rather than physical ones, as Esperantists form a very heterogeneous group of diverse cultures and traditions. As expected, the most listings went to the green star (26), followed by the green flag/banner (20) and the anthem *La Espero* (14). Altogether 24 responses were given. Such a high listing for the green star was expected; however, all the more surprising is the fact that ‘green’ did not collocate with ‘star’ in Q11 more than 3 times.

domain	No. of features	sum of features frequencies	No. of frequencies	descriptor
objects/symbols	8	63	26	green star
			20	green banner
			5	ZEOj (Zamenhof/Esperanto objects, e.g. statues)
			4	crocodile
			3	Latin and Cyrillic E (<i>melono</i>)
			3	the colour green
			1	the word “Esperanto”
			1	<i>Plena Ilustrita Vortaro</i>
culture	6	31	14	<i>La Espero</i> – E-o anthem
			5	literature
			5	Zamenhof
			3	Esperanto culture
			2	<i>La Fundamento</i>
			2	Esperanto itself

the movement	5	12	5	congresses
			4	Movement structures
			1	<i>Akademio de Esperanto</i>
			1	<i>UEA</i> logo
			1	<i>Pasporta Servo</i>
ideology	3	3		equal rights <i>homaranismo</i> internal idea
places	2	2		places important in the history of the movement Montevideo (i.e. where the resolution in support of Esperanto was passed by UNESCO in 1954)

Table 14 Symbols of an Esperantist

The view of an Esperantist emerges as a uniform well-developed set of features. The self-stereotype is well established and it revolves around the language. The most frequent features in the ‘but’ test corroborate the view of a “true” Esperantist as speaking the language fluently and being active for the benefit of Esperanto and the movement. The symbols of an Esperantist are the symbols of the movement and the language itself: the green star, the green banner and the anthem.

6.3. Conclusions

The results of the present pilot study allow for developing a new more feasible questionnaire and point to specific areas that should be further tested. Several assumptions from section 5.4 were also confirmed. However, all conclusions should be drawn with caution.

The prototypicality of plants and animals (Qs 2, 5 and 12) proved to be a difficult task to be studied. On one hand, Esperantists do not typically concern themselves with this area, and therefore often do not know the needed names. In several cases, they clearly transfer from their native languages. On the other hand, due to the geographical distribution of Esperantists, this task may only be sensible to conduct in groups that are more compact; nevertheless, this would doubtless only demonstrate worldviews transferred from native languages. Similarly, colours (Qs 1 and 11) are known to vary cross-linguistically and cross-culturally. The pool of respondents did not contain enough speakers of languages and cultures from outside Europe to confirm the assumption, that Esperantists do not categorise colours in the same way. However, the importance of green was shown in Q7 related to Esperanto culture.

Moreover, the metaphorical values of animals and plants in Q10 and Q13 differed across languages. The sample should undoubtedly be widened, as in the previous cases, to confirm the assumption that Esperantists do not have a consistent worldview when it comes to simple cognitive domains. Grammatical gender (Q9) seems to be borrowed from native languages as well.

The above cases proved also to be most time absorbing and complicated. A French speaker left out Qs 4-11, explaining that they were difficult and time-consuming (participant from Lille). Two persons skipped Q13. This shows that such questions should be avoided, as not giving an Esperantic worldview and not feasible. Additionally, the questionnaire should be certainly administered in an online-based form, to facilitate answering and processing the results.

The questions pertaining to Esperanto itself proved most valuable. They attest that native speakers of Esperanto do not necessarily understand the concepts stemming from the culture that arose around the language. What is more, they show that the stereotypes that active Esperantists have of themselves are sufficiently consistent and worthy of further investigation.

This pilot study permitted tentative answers to our previous research questions; non-native Esperantists indeed transfer their LWVs from their native languages in several domains; there exists a homogeneous, culturally embedded LWV of Esperanto, understandable to non-native speakers. The third research question (Is there a homogeneous LWV among native/indigenous speakers of Esperanto?) must, however, be tested on a larger group of *denaskuloj*.

In conclusion, the areas in which the native LWV is transferred are:

- cognitive categories
 - colour stereotypes
 - prototypicality of plants and animals
- cross-culturally varying symbolic values of animals and plants
- grammatical-semantic categories
 - grammatical gender
 - collocations and lexicalisations

However, the last category does not show such a strong influence of vernaculars on Esperanto as the previous ones. Although several alternatives of collocates may emerge as calques from native languages, the process of stabilisation of one or two equivalent Esperanto forms is observable.

A LWV typical for Esperanto and presented consistently is visible in the following areas:

- cultural concepts related to the Movement
- stereotype of an Esperantist

In general, the LWV of Esperanto is more limited than those of ethnic languages, yet it demonstrably exists. Evidently, this observation confirms the strong influence of culture on language. Nevertheless, the reverse influence is also present: in Esperanto, there exist such concepts as *krokodili* or *volapukaĵo*; “a true Esperantist” is one speaking the language

fluently on an everyday basis; synthetic forms appear in Esperanto regardless of the background of the speaker.

It is advisable that a future questionnaire follows the basic guidelines presented in 5.4, that is includes at least 125 respondents in four groups from different cultures and languages as well as one group of native speakers of Esperanto. More tasks designed to control the extent of transferring should also be added. Ideally, Osgood's semantic differential should be devised before further testing stereotypical judgements.

Final remarks

The boundaries of artificiality

It has been shown that the class of artificial languages is not a homogenous one. There is a great diversity of such systems both from the historical and from the typological perspective. The only unifying factor is the genetic one: all artificial languages are created on purpose by (a) known creator(s) in a relatively short period of time (Chapter 2). However, if a larger set of properties is applied to distinguish “natural” from “artificial”, the procedure will result in a class of borderline cases.

Various types of languages were assessed in terms of Hockett’s design features, Lyons’ classes of naturalness, Svadost’s levels of deliberate influence and modified lists of features according to Duličenko and Liu (Chapter 3). These assessments disprove the traditional binary division into artificial and natural, showing that all languages lay on a continuum between artificiality and naturalness. The most natural language among artificial languages is Esperanto – in most respects it may be treated on par with natural ethnic languages (a more detailed description was presented in Chapter 4). The least natural appear to be programming languages. This may seem surprising, in view of the fact that they incorporate features from natural languages; their limitations stem mainly from the fact that they serve as a means of communication between two different species, i.e. humans and computers.

The present work demonstrates that the existing genetic definitions of artificial languages do not provide a sufficient basis for determining which languages may constitute useful material for linguistic explorations. Rejecting all artificial constructs on the ground of their origin restricts heavily the scope of linguistics. Examining artificial languages could, after all, reveal the limitations of the human language faculty. A close analysis of a language transforming from artificial into natural (i.e. Esperanto) could also explain the nature of language.

Based on the analysis conducted, it is possible to formulate the following conclusions:

1. artificial languages as defined genetically do not constitute a homogenous group
2. the natural/artificial dichotomy is disproved and the scale of naturalness is proposed instead
3. the language which may be treated as natural is clearly Esperanto.

A linguistic worldview for artificial languages

The importance of the subject and the hierarchical structure of the facets within a concept may be recognised as the most basic tenets of the linguistic worldview theory in the Ethnolinguistic School of Lublin's framework. The Ethnolinguistic School of Lublin assumes that a linguistic worldview is moulded within a speech community and a culture. Therefore, as has been shown, it is of utmost importance to identify the subject of an artificial language, the origin of the facets and their connection to culture and identity.

The **subject** from whose perspective a given text is produced varies strongly across artificial languages. Because Esperanto may be considered natural, as developing in a speech community, the community itself is the subject. Other artificial languages cannot boast a speech community at all, or only a very small one. In such cases, the safest assumption is to identify the subject as the author of the language in question. Therefore, an author's personal beliefs will always be revealed in the semantic categories of a language.

The **facets** clearly depend on the subject. Esperanto has developed its own culture and a speech community independent of Zamenhof's views, and therefore the facets are structured according to the contemporary social agreement. Universal schemes, in turn, as lying almost opposite to Esperanto, must be examined as reflecting the personal convictions and experiences of their authors. Such constructs rely often on the author's

knowledge of other languages. This seems to be the case of international auxiliary languages, which were meant to follow (at least semantically) their source languages very closely.

The problem of **culture** is the most intricate one. Again, it is closely connected to the subject. Universal languages of the 17th and 18th c., as based on the Aristotelian categorisation, must be treated as stemming from the Greco-Roman tradition permeating European science. Artistic languages stem from two sources: they are a response either to aesthetic needs, and therefore a language is overbuilt with a cultural layer, which substantiates the language, or for more practical requirements, that is, they are created to reflect the cultural reality of an imaginary world. The connection of international auxiliary languages to culture may be twofold: the worldview in these languages may stem from their source languages embedded in particular cultures, or from the interconnecting cultures of the speakers.

Such an understanding of the role the subject plays in the reconstructions of the linguistic worldview for any language could mean that only languages with a sufficiently large number of speakers may be of interest to linguistic worldview researchers. To exclude such a situation, Beaugrande and Dressler's **criteria of textuality** proved useful. These criteria helped decide whether an artificial language can be treated as language or rather as text. In connection with the features presented in chapter 3 the criteria allowed for evaluating artificial languages on a scale from a fully developed language, through semi-language and texteme (i.e. text model) to a text (i.e. a product of a specific author). For example, formal languages and linguistic reconstructions were assessed as textemes (matrices for language) and universal languages as texts (non-used, non-useable constructs analysable as artistic inventions). This hierarchy allows for including all artificial languages in analyses, assuming that some of them are artistic creations which can be studied similarly to poetry or other literary texts.

As can be seen, identifying the subject of an artificial language is crucial for further analyses. From the outcome of the investigation, it is possible to conclude that Esperanto is the only artificial language whose subject is not its author but its speech community. It is a model case of how a deliberately created language may evolve naturally (i.e. in similarity to ethnic living languages). An important issue emerging from the present theoretical analysis is that although Esperanto is nearing the natural pole on the continuum, its community still remains a community of L2 speakers, which is a crucial factor affecting any study of the linguistic worldview. Therefore, it is proposed that a specially designed enquiry into bilingualism is included to assess its impact on the worldview present in Esperanto.

The theoretical analyses conducted in this book revealed that the diversity of artificial languages requires a diversity of perspectives in approaching them using the Ethnolinguistic School of Lublin's framework. Any future study must include cultural and linguistic transfer, as the users of artificial languages are all multilingual.

The pilot study presented in ch. 6 revealed many problems stemming from the fact that Esperantists are not a homogenous group. However, this only confirms our previous assertions: firstly, that Esperanto is almost natural, in that its norms are negotiated and develop spontaneously; secondly, that many areas of the worldview present in Esperanto are strongly influenced by the speakers' native languages and cultures; and finally, that the hard core of the movement presents a coherent and consistent worldview based on the sense of belonging to the community, its culture and norms. Moreover, it showed that the ESL framework has to be widened through additional studies of the LWV of multilinguals.

Further research into the linguistic worldview of Esperanto and in the long term other artificial languages is desirable to extend our knowledge of the human faculty of language and the interdependence of language and culture revealed in semantic categories.

References

- Abramowicz, M., Bartmiński, J. and Chlebda, W. 2009. Językowo-kulturowy obraz świata Słowian na tle porównawczym. Założenia programu „A”. *Etnolingwistyka* 21, pp. 341–342.
- Adams, M. 2011. The Spectrum of Invention. In: Adams, M. ed. *From Elvish to Klingon: exploring invented languages*. Oxford, New York: Oxford University Press, pp. 1–16.
- Akademio de Esperanto 1975. AKTOJ de la AKADEMIO 1968-1974 - Baza Radikaro Oficiala. Available at: <http://www.akademio-de-esperanto.org/aktoj/aktoj2/-bro.html> [Accessed: 12 October 2015].
- Albani, P. and Buonarotti, B. 1994. *Aga magéra difira: dizionario delle lingue immaginarie*. Bologna: Zanichelli.
- Allwood, J. 1983. Kan man tänka oberoende av språk? In: Teleman, U. ed. *Tal och tanke*. Lund: Liber.
- Allwood, J. 2004. Språk och världsbilder. In: Ilie, C. ed. *Language, Culture, Rhetoric: Cultural and Rhetorical Perspectives on Communication. Papers from the ASLA Symposium in Örebro, 6-7 November 2003*. Stockholm: ASLA.
- Almqvist, I. and Sägval Hein, A. 1996. Defining ScaniaSwedish – a controlled language for truck maintenance. In: *Proceedings of the 1st International Workshop on Controlled Language Applications*. pp. 159–164.
- Andrzejewski, B. 1989. *Wilhelm von Humboldt*. Warszawa: Wiedza Powszechna.
- Anusiewicz, J. 1999. Problematyka językowego obrazu świata w poglądach niektórych językoznawców i filozofów niemieckich XX wieku. In: Bartmiński, J. ed. *Językowy obraz świata*. 2nd rev. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, pp. 261–289.
- Anusiewicz, J., Dąbrowska, A. and Fleischer, M. 2000. Językowy obraz świata i kultura. Projekt koncepcji badawczej. *Językowy obraz świata i kultura*. *Język a Kultura* 13, pp. 11–44.

- Athanasopoulos, P. 2009. Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition* 12(01), pp. 83–95.
- Athanasopoulos, P. 2011. Colour and bilingual cognition. In: Cook, V. and Bassetti, B. eds. *Language and bilingual cognition*. New York, Hove: Psychology Press, pp. 241–262.
- Barandovská-Frank, V. 1995. *Enkonduka lernolibro de interlingvistiko*. Sibiu: Editura Universităţii.
- Barandovská-Frank, V. 2003. Gehören Plansprachen zur Sprachplanung? *Grundlagenstudien aus Kybernetik und Geisteswissenschaft (grkg/Humankybernetik)* 44(2).
- Barandovská-Frank, V. 2011. Spracherfindung und Nationalsprache: das Beispiel El Glheþ Talossan. In: Fiedler, S. ed. *Spracherfindung und ihre Ziele. Beiträge der 20. Jahrestagung der Gesellschaft für Interlinguistik*. Berlin: Gesellschaft für Interlinguistik, pp. 33–50.
- Baron, N.S. 1994. Computer languages. In: Asher, R. E. ed. *The Encyclopedia of language and linguistics*. Oxford: Pergamon Press.
- Bartlett, P.O. 2009. Artificial Languages. In: Brown, K. and Ogilvie, S. eds. *Concise encyclopedia of languages of the world*. Amsterdam; Boston: Elsevier.
- Bartmiński, J. 2007. Kryteria ilościowe w badaniu stereotypów językowych. In: *Stereotypy mieszkają w języku: studia etnolingwistyczne*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, pp. 72–84.
- Bartmiński, J. 2012a. *Aspects of cognitive ethnolinguistics*. Reprint. Zinken, J. ed. London; Oakville, CT: Equinox.
- Bartmiński, J. 2012b. *Językowe podstawy obrazu świata*. 4th ed. Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- Bartmiński, J. 2013a. Rola etymologii w rekonstrukcji językowego obrazu świata. *LingVaria* 16(2), pp. 233–245.
- Bartmiński, J. 2013b. The Cognitive Definition as a Text of Culture. In: Gład, A., Danaher, D., and Łozowski, P. eds. *The Linguistic Worldview: Ethnolinguistics, Cognition, and Culture*. London: Versita, pp. 161–180.

- Bartmiński, J. and Chlebda, W. 2008. Jak badać językowo-kulturowy obraz świata Słowian i ich sąsiadów? *Etnolingwistyka* 20, pp. 11–27.
- Bartmiński, J. and Niebrzegowska-Bartmińska, S. 2012. *Tekstologia*. 1st ed. Warszawa: Wydawnictwo Naukowe PWN.
- Bausani, A. 1970. *Geheim- und Universalsprachen: Entwicklung u. Typologie*. Stuttgart, Berlin, Köln, Mainz: Kohlhammer.
- Beaugrande, R.-A. de and Dressler, W.U. 1981. *Introduction to text linguistics*. London, New York: Longman.
- Bergen, B.K. 2001. Nativization processes in L1 Esperanto. *Journal of Child Language* 28(03), pp. 575–595.
- Blahuš, M. 2011. Toki Pona – eine minimalistische Plansprache. In: Fiedler, S. ed. *Spracherfindung und ihre Ziele. Beiträge der 20. Jahrestagung der Gesellschaft für Interlinguistik*. Berlin: Gesellschaft für Interlinguistik, pp. 51–56.
- Blanke, D. 1985. *Internationale Plansprachen: eine Einführung*. Berlin: Akademie-Verlag.
- Blanke, D. 1989. Planned languages – a survey of some of the main problems. In: Schubert, K. ed. *Interlinguistics: Aspects of the Science of Planned Languages*. Berlin, New York: Mouton de Gruyter, pp. 63–87.
- Blanke, D. 1997. The Term ‘Planned Language’. In: Tonkin, H. ed. *Esperanto, Interlinguistics, and Planned Language*. Lanham: University Press of America, pp. 1–20.
- Blanke, D. 2001. Vom Entwurf zur Sprache. In: Schubert, K. ed. *Planned languages: from concept to reality*. Brüssel: Hogeschool voor Wetenschap en Kunst, pp. 37–89.
- Brown, R. 2014. Glossopoeia [Online]. Available at: <http://www.carolandray.plus.com/Glosso/Glossopoeia.html> [Accessed: 7 January 2015].
- Carlevaro, T. 1989. Planned auxiliary language and communicative competence. In: Schubert, Klaus ed. *Interlinguistics: Aspects of the Science of Planned Languages*. Berlin, New York: Walter de Gruyter, pp. 173–187.
- Chlebda, W. 2013. Can Polish Ethnolinguistics Become a Philological Keystone of the Humanities? In: Głaz, A.,

- Danaher, D., and Łozowski, P. eds. *The Linguistic World-view: Ethnolinguistics, Cognition, and Culture*. London: Versita, pp. 25–37.
- Coleman, J. 2006. Design Features of Language. In: Brown, K. ed. *Encyclopedia of Language and Linguistics*. 2nd ed. Oxford: Elsevier, pp. 471–475.
- Columbia Encyclopedia, 6th ed. 2014. Available at: <http://www.encyclopedia.com/> [Accessed: 23 May 2015].
- Corsetti, R. 2010. Saŝiĝo de la vortaro en kreolaj lingvoj: la ekzemplo de Esperanto. In: Blanke, D. and Lins, U. eds. *La arto labori kune: festlibro por Humphrey Tonkin*. Rotterdam: Universala Esperanto-Asocio, pp. 373–385.
- Corsetti, R. 2012. The movement for Esperanto: between creolization and the Report Grin / La movado por Esperanto: inter kreoliĝo kaj la Raporto Grin. *InKoj. Philosophy & Artificial Languages* 3(1), pp. 58–78.
- Couturat, L. and Léau, L. 1903. Histoire de la langue universelle [Online]. Available at: <https://archive.org/details/histoiredelalang00coutuoft>.
- Cram, D. 1992. Language universals and 17th century universal languages schemes. In: Subbiondo, J. L. ed. *John Wilkins and 17th-century British Linguistics*. Amsterdam, Philadelphia: J. Benjamins.
- Dnghu Association 2007. Indo-European – About Indo-European language revival [Online]. Available at: <http://dnghu.org/en/indo-european%20revival/> [Accessed: 22 May 2015].
- Duliĉenko, A. 1989. Ethnic language and planned language. In: Schubert, Klaus ed. *Interlinguistics: Aspects of the Science of Planned Languages*. Berlin, New York: Mouton de Gruyter, pp. 47–61.
- Duliĉenko, A. 1990. *Международные вспомогательные языки [International auxiliary languages]*. Tallinn: Valgus.
- Duliĉenko, A. 2001. Planlingvo: inter lingvokonstruo kaj etna lingvo. In: Fiedler, S. and Liu, H. eds. *Studoj pri interlingvistiko: Festlibro omaĝe al la 60-jariĝo de Detlev*

- Blanke = *Studien zur Interlinguistik: Festschrift für Detlev Blanke zum 60. Geburtstag*. Dobrichovice (Praha): Kava-Pech, pp. 109–115.
- Duličenko, A. 2006. *En la serĉado de la mondlingvo: aŭ interlingvistiko por ĉiuj*. Kaliningrad: Sezonoj.
- Dutilh Novaes, C. 2012. *Formal languages in logic: a philosophical and cognitive analysis*. Cambridge: Cambridge University Press.
- Eco, U. 1997. *The search for the perfect language*. London: Fontana Press.
- Encyclopædia Britannica 2014. Available at: <http://www.britannica.com> [Accessed: 23 May 2015].
- EUROJOS 2008. Available at: <http://ispan.waw.pl/default/pl-projekty-naukowe/eurojos> [Accessed: 16 January 2015].
- Farris, M.A. 1994. Sign language research and Polish Sign Language. *Lingua Posnaniensis* 36, pp. 13–36.
- Ferguson, C.A. 1968. Language development. In: Fishman, J., Ferguson, C. A., and Das Gupta, J. eds. *Language problems of developing nations*. New York: Wiley and Sons, pp. 27–36.
- Fettes, M. 2005. Artificial Languages. In: Strazny, P. ed. *Encyclopedia of linguistics*. New York: Fitzroy Dearborn.
- Fiedler, S. 1999. *Plansprache und Phraseologie: empirische Untersuchungen zu reproduziertem Sprachmaterial im Esperanto*. Frankfurt am Main, New York: Peter Lang.
- Fiedler, S. 2006. Standardization and self-regulation in an international speech community: the case of Esperanto. *International Journal of the Sociology of Language* 2006(177), pp. 67–90.
- Fiedler, S. 2007. Phraseology in planned languages. In: Burger, H. ed. *Phraseologie: ein internationales Handbuch zeitgenössischer Forschung = Phraseology: an international handbook of contemporary research*. Berlin, New York: Mouton de Gruyter.
- Fiedler, S. 2011. Literarische Spracherfindungen aus interlinguistischer Sicht: die englischsprachigen Autoren J. Swift, G. Orwell und J.R.R. Tolkien und ihre fiktionalen Sprachen. In:

- Fiedler, S. ed. *Spracherfindung und ihre Ziele. Beiträge der 20. Jahrestagung der Gesellschaft für Interlinguistik*. Berlin: Gesellschaft für Interlinguistik, pp. 9–31.
- Fiedler, S. 2012. The Esperanto denaskulo: The status of the native speaker of Esperanto within and beyond the planned language community. *Language Problems & Language Planning* 36(1), pp. 69–84.
- Fiedler, S. 2015. Esperanto Phraseology. *Interdisciplinary Description of Complex Systems* 13(2), pp. 250–263.
- Frommer, P. Learn Na'vi [Online]. Available at: <http://learnnavi.org/> [Accessed: 18 February 2015].
- Galor, Z. 2001. Esperanto-movado: ĉu perdita komunumo? In: Fiedler, S. and Liu, H. eds. *Studoj pri interlingvistiko: Festlibro omaĝe al la 60-jariĝo de Detlev Blanke = Studien zur Interlinguistik: Festschrift für Detlev Blanke zum 60. Geburtstag*. Dobrichovice (Praha): Kava-Pech, pp. 504–523.
- Galor, Z. and Pietiläinen, J. 2015. *UEA en konscio de esperantistoj*. Dobrichovice (Praha): Kava-Pech.
- Głaz, A., Danaher, D. and Łozowski, P. 2013. *The Linguistic Worldview: Ethnolinguistics, Cognition, and Culture*. London: Versita.
- Gledhill, C. 1998. *The grammar of Esperanto: a corpus-based description*. München: Lincom Europa.
- Gobbo, F. 2012. Alan Turing creator of Artificial Languages / Alan Turing kreinto de Artefaritaj Lingvoj. *InKoj. Philosophy & Artificial Languages* 3(2), pp. 181–194.
- Gobbo, F. 2014. Interlingvistiko por Esperantistoj. La malfacila tasko klasifiki planlingvojn. *Esperanto* (7–8), pp. 164–165.
- Grzegorzczkova, R. 1999. Pojęcie językowego obrazu świata. In: Bartmiński, J. ed. *Językowy obraz świata*. 2nd rev. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, pp. 39–46.
- Grzeszczak, M. 2009. Definicja kognitywna pojęcia DEMOKRACJI w języku polskim. *Etnolingwistyka* 21, pp. 69–84.

- Guz, M. 2012. *Językowy obraz świata u wybranych przedstawicieli lingwistyki niemieckiej, amerykańskiej i polskiej*. Poznań: Wydawnictwo Rys.
- Hendriks-Hermans, J. 1999. Klingon and its users: a sociolinguistic profile. MA thesis. Tilburg University.
- Heredia, R.R. and Cieślicka, A.B. 2015. *Bilingual figurative language processing*. New York: Cambridge University Press.
- Herring, J. 2005. Syntactic and Lexical Changes in Esperanto: A Quantitative and Corpus-Based Survey. Available at: http://clt.osu.edu/mclc/paper/syntactic_herring.pdf [Accessed: 23 February 2014].
- Hertzog, M.A. 2008. Considerations in determining sample size for pilot studies. *Research in nursing & health* 31(2), pp. 180–191.
- Higley, S.L. 2007. *Hildegard of Bingen's unknown language: an edition, translation, and discussion*. New York: Palgrave Macmillan.
- Hockett, C.F. 1960. The origin of speech. *Scientific American* 203, pp. 88–96.
- Ivanov, D. 2014. Summary of LdP's distinctive features [Online]. Available at: <http://lingwadeplaneta.info/en/svodka.shtml> [Accessed: 18 May 2015].
- Jackson, H. 2011. Invented Vocabularies: The Cases of Newspeak and Nadsat. In: Adams, M. ed. *From Elvish to Klingon: exploring invented languages*. Oxford; New York: Oxford University Press, pp. 49–74.
- Jansen, W. 2010. Esperanto: een taal van mij, een taal om van te houden / Esperanto: lingvo malfremda, lingvo aminda. *In-Koj. Philosophy & Artificial Languages* 1(1), pp. 12–34.
- Janton, P. 1993. *Esperanto: Language, Literature, and Community*. Tonkin, H. ed. Albany, N.Y.: State University of New York Press.
- Jurkowski, M. 1986. *Od Wieży Babel do języka Kosmitów: o językach sztucznych, uniwersalnych i międzynarodowych*. Białystok: Krajowa Agencja Wydawnicza.

- Kamasa, V. 2014. Techniki językoznawstwa korpusowego wykorzystywane w krytycznej analizie dyskursu. *Przeгляд. Przegląd socjologii jakościowej* 10(2), pp. 100–117.
- Kennaway, R. 2010. Artificial Languages. In: Malmkjær, K. ed. *The Linguistics Encyclopedia*. 3rd ed. London, New York: Routledge.
- Kimura, G.C. 2009. Esperanto als Minderheitensprache: Eine sprachsoziologische Betrachtung. In: Fiedler, S. ed. *Esperanto und andere Sprachen im Vergleich. Beiträge der 18. Jahrestagung der Gesellschaft für Interlinguistik*. Berlin: Gesellschaft für Interlinguistik, pp. 11–24.
- Kimura, G.C. 2010. La kornvala kaj Esperanto: entreprenoj similaj? In: Blanke, D. and Lins, U. eds. *La arto labori kune: festlibro por Humphrey Tonkin*. Rotterdam: Universala Esperanto-Asocio, pp. 171–177.
- Kimura, G.C. 2012. Esperanto and minority languages: A sociolinguistic comparison. *Language Problems & Language Planning* 36(2), pp. 167–181.
- Kirillov, V. 2012. Лингва де планета (всепланетный язык) [Lingwa de planeta (a world language)] [Online]. Available at: <http://journal.spbu.ru/?p=8660> [Accessed: 17 May 2015].
- Knowlson, J. 1975. *Universal language schemes in England and France, 1600-1800*. Toronto; Buffalo: University of Toronto Press.
- Koutny, I. 2009. Esperanto im Rahmen der Sprachtypologie. In: Fiedler, S. ed. *Esperanto und andere Sprachen im Vergleich. Beiträge der 18. Jahrestagung der Gesellschaft für Interlinguistik*. Berlin: Gesellschaft für Interlinguistik, pp. 117–130.
- Koutny, I. 2010. Esperantlingva bildo de la mondo. In: Blanke, D. and Lins, U. eds. *La arto labori kune: festlibro por Humphrey Tonkin*. Rotterdam: Universala Esperanto-Asocio, pp. 290–305.
- Koutny, I. 2012. From semantic networks to dictionary structures. *Język, Komunikacja, Informacja* VII, pp. 115–128.

- Krägeloh, C. and Neha, T.N. 2014. Lexical expansion and terminological planning in indigenous and planned languages. *Language Problems & Language Planning* 38(1), pp. 59–86.
- Large, A. 1985. *The artificial language movement*. Oxford; New York; London: B. Blackwell; A. Deutsch.
- Large, A. 1994. Artificial languages. In: Asher, R. E. ed. *The Encyclopedia of language and linguistics*. Oxford: Pergamon Press.
- Lazard, G. 2001. On the grammaticalization of evidentiality. *Journal of Pragmatics* 33(3), pp. 359–367.
- Li, P., Sepanski, S. and Zhao, X. 2006. Language history questionnaire: A Web-based interface for bilingual research. *Behavior research methods* 38(2), pp. 202–210.
- Libert, A.R. 2000. *A priori artificial languages*. München: Lincom Europa.
- Libert, A.R. 2008. *Daughters of Esperanto*. München: Lincom Europa.
- Libert, A.R. 2013. The Representation of Uralic Languages in Artificial International Auxiliary Languages. *Journal of Universal Language* 14(1), pp. 117–142.
- Lindstedt, J. 2006. Native Esperanto as a test case for natural language. In: Suominen, M. ed. *A Man of Measure: Festschrift in Honour of Fred Karlsson on His 60th Birthday*. SKY Journal of Linguistics. pp. 47–55.
- Lins, U. 1988. *Die gefährliche Sprache: die Verfolgung der Esperantisten unter Hitler und Stalin*. Gerlingen: Bleicher.
- Liu, H. 2001. Creoles, Pidgins, and Planned Languages. In: Schubert, K. ed. *Planned languages: from concept to reality*. Brussel: Hogeschool voor Wetenschap en Kunst.
- Liu, H. 2006. Neutrality of international languages. *Journal of Universal Language* 7(2), pp. 37–64.
- Łozowski, P. 2013. Language Vis-à-vis Culture in Jerzy Bartmiński's Cognitive Ethnolinguistics. In: Głaz, A., Danaher, D., and Łozowski, P. eds. *The Linguistic Worldview*:

- Ethnolinguistics, Cognition, and Culture*. London: Verita, pp. 351–370.
- Lucy, J.A. 1992a. *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*. Cambridge, New York: Cambridge University Press.
- Lucy, J.A. 1992b. *Language diversity and thought: a reformulation of the linguistic relativity hypothesis*. Cambridge, New York: Cambridge University Press.
- Lyons, J. 1977. *Semantics*. Reprint. Cambridge: Cambridge University Press.
- Lyons, J. 1991. *Natural language and universal grammar*. Cambridge, New York: Cambridge University Press.
- Machado, A. 2004. An index of naturalness. *Journal for Nature Conservation* 12(2), pp. 95–110.
- Maćkiewicz, J. 1999. Co to jest ‘językowy obraz świata’. *Etnolingwistyka* 11, pp. 7–24.
- Malmkjær, K. 2002. Artificial Languages. In: Malmkjær, K. ed. *The Linguistics Encyclopedia*. 2nd ed. London, New York: Routledge.
- Manaris, B., Pellicoro, L., Pothering, G. and Hodges, H. 2006. Investigating Esperanto’s statistical proportions relative to other languages using neural networks and Zipf’s law. In: *Proceedings of the 24th IASTED international conference on Artificial intelligence and applications*. Innsbruck, Austria: ACTA Press, pp. 102–108.
- Mannewitz, C. 2003. Science-Fiction-Sprachen im Internet. In: Blanke, D. ed. *Plansprachen und elektronische Medien. Beiträge der 12. Jahrestagung der Gesellschaft für Interlinguistik*. Berlin: Gesellschaft für Interlinguistik, pp. 40–53.
- Mannewitz, C. 2009. Sprachplanung im Internet: das Projekt Slovio. In: Fiedler, S. ed. *Esperanto und andere Sprachen im Vergleich. Beiträge der 18. Jahrestagung der Gesellschaft für Interlinguistik*. Berlin: Gesellschaft für Interlinguistik, pp. 157–164.
- Marcus, S. 2004. Formal Languages: Foundations, Prehistory, Sources, and Applications. In: Martín Vide, C., Mitrana,

- V., and Păun, G. eds. *Formal Languages and Applications*. Studies in fuzziness and soft computing. Berlin, Heidelberg: Springer-Verlag.
- Marian, V., Blumenfeld, H.K. and Kaushanskaya, M. 2007. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50(4), pp. 940–967.
- Maurus III, B.V., 2014. The Sanskrit experience and language planning. *Język, Komunikacja, Informacja* 9, pp. 58–70.
- McKinnon, K. 2000. *An Independent Academic Study on Cornish*.
- Melnikov, A. 1992. *Specifaj kulturaj scioj de la esperantista kvazaŭetno kaj ilia respeguliĝo en la koncerna lingv(a)ŝo*. Rostov-na-Donu.
- Miner, K. 2011. The impossibility of an Esperanto linguistics / La neebleco de priesperanta lingvoscienco. In *Koj. Philosophy & Artificial Languages* 2(1), pp. 26–51.
- Murawski, R. 2012. *Filozofia matematyki: zarys dziejów*. 4th ed. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.
- Niebrzegowska-Bartmińska, S. 2015. O profilowaniu językowego obrazu świata. *Poradnik językowy* (1), pp. 30–44.
- Nöth, W. 1995. *Handbook of semiotics*. Bloomington, Indianapolis: Indiana University Press.
- Okrand, M., Adams, M., Hendriks-Hermans, J. and Kroon, S. 2011. Wild and Whirling Words: The Invention and Use of Klingon. In: Adams, M. ed. *From Elvish to Klingon: exploring invented languages*. Oxford, New York: Oxford University Press, pp. 111–134.
- Okrent, A. 2010. *In the land of invented languages: a celebration of linguistic creativity, madness, and genius*. New York: Spiegel & Grau Trade Paperbacks.
- O'Regan, G. 2012. *A brief history of computing*. 2nd ed. London: Springer.
- Pajdzińska, A. 2013. The Linguistic Worldview and Literature. In: Głaz, A., Danaher, D., and Łozowski, P. eds. *The*

- Linguistic Worldview: Ethnolinguistics, Cognition, and Culture*. London: Versita, pp. 41–60.
- Parkvall, M. 2010. How European is Esperanto? A typological study. *Language Problems & Language Planning* 34(1), pp. 63–79.
- Patrick, P.L. 2002. The Speech Community. In: Chambers, J. K., Trudgill, P., and Schilling-Estes, N. eds. *The handbook of language variation and change*. Oxford: Blackwell.
- Pavlenko, A. ed. 2006. *Bilingual minds: Emotional experience, expression, and representation*. Clevedon [England], Buffalo, NY: Multilingual Matters.
- Pavlenko, A. ed. 2011a. *Thinking and Speaking in Two Languages*. Multilingual Matters.
- Pavlenko, A. 2011b. Thinking and Speaking in Two Languages: Overview of the Field. In: Pavlenko, A. ed. *Thinking and Speaking in Two Languages*. Multilingual Matters, pp. 237–257.
- Peterson, D.J. 2011. Case Assignment and Verb Classes in Dothraki. Available at: <http://conference.conlang.org/lc-c4/speakers/djplcc4.pdf> [Accessed: 18 February 2014].
- Peterson, D.J. About Dothraki [Online]. Available at: <http://www.dothraki.com/about-dothraki/> [Accessed: 18 February 2015].
- Piron, C. 1989a. A few notes on the evolution of Esperanto. In: Schubert, K. ed. *Interlinguistics: Aspects of the Science of Planned Languages*. Berlin, New York: Mouton de Gruyter.
- Piron, C. 1989b. Who are the speakers of Esperanto? In: Schubert, K. ed. *Interlinguistics: Aspects of the Science of Planned Languages*. Berlin, New York: Mouton de Gruyter.
- Pogonowski, J. 2012. Kilka uwag o intuicji matematycznej. *Filozofia Nauki* 78(2), pp. 107–113.
- Portnow, J. 2011. Gaming Languages and Language Games. In: Adams, M. ed. *From Elvish to Klingon: exploring invented languages*. Oxford, New York: Oxford University Press, pp. 135–159.

- Pressey, C. 2015. Chris Pressey's Lingography | Cat's Eye Technologies [Online]. Available at: http://catseye.tc/node/Chris_Pressey%27s_Lingography [Accessed: 19 May 2015].
- Prorok, K. and Głaz, A. 2013. The Cognitive Definition of Iron (Żelazo) in Polish Folk Tradition. In: Głaz, A., Danaher, D., and Łozowski, P. eds. *The Linguistic World-view: Ethnolinguistics, Cognition, and Culture*. London: Versita, pp. 181–198.
- Quijada, J. 2011. Ithkuil FAQs [Online]. Available at: <http://www.ithkuil.net/faqs.html> [Accessed: 17 May 2015].
- Rogers, S.D. 2011. *A dictionary of made-up languages : from Adúnaic to Elvish, Zaum to Klingon – the Anwa (real) origins of invented lexicons*. Avon: Adams Media.
- Romaine, S. 2009. Pidgins and Creoles. In: Brown, K. and Ogilvie, S. eds. *Concise encyclopedia of languages of the world*. Amsterdam, Boston: Elsevier.
- Romaine, S. 2011. Revitalized Languages as Invented Languages. In: Adams, M. ed. *From Elvish to Klingon: exploring invented languages*. Oxford, New York: Oxford University Press, pp. 185–225.
- Rónai, P. 1969. *Der Kampf gegen Babel: oder Das Abenteuer der Universalsprachen*. München: Ehrenwirth.
- Rosch, E. 1978. *Cognition and categorization*. Lloyd, B. B. ed. Hillsdale: Lawrence Erlbaum Associates.
- Sakaguchi, A. 1998. *Interlinguistik: Gegenstand, Ziele, Aufgaben, Methoden*. Frankfurt am Main, Berlin, Bern, New York, Paris, Vienna: Lang.
- Sakaguchi, A. 2003. Dychotomia „sztuczny” i „naturalny” z perspektywy języka międzynarodowego esperanto. *Scripta Neophilologica Posnaniensia V*.
- Salmon, V. 1972. *The works of Francis Lodwick: a study of his writings in the intellectual context of the seventeenth century*. London: Longman.
- Sandler, W. 2009. Sign Languages. In: Brown, K. and Ogilvie, S. eds. *Concise encyclopedia of languages of the world*. Amsterdam, Boston: Elsevier.

- Saygin, A.P. 2001. Processing figurative language in a multi-lingual task: Translation, transfer and metaphor. In: *Proceedings of Corpus-Based & Processing Approaches to Figurative Language Workshop, Corpus Linguistics*. Lancaster: Lancaster University.
- Schubert, K. 1989. Interlinguistics – its aims, its achievements, and its place in language science. In: Schubert, K. ed. *Interlinguistics: Aspects of the Science of Planned Languages*. Berlin, New York: Mouton de Gruyter, pp. 7–44.
- Schubert, K. 2001. Gestaltete Sprache. Plansprachen und die regulierten Sprachen der internationalen Fachkommunikation. In: Schubert, K. ed. *Planned languages: from concept to reality*. Brussel: Hogeschool voor Wetenschap en Kunst, pp. 223–257.
- Schubert, K. 2011. Zum bewussten Eingreifen in die Sprache. In: Brosch, C. and Fiedler, S. eds. *Florilegium Interlinguisticum: Festschrift für Detlev Blanke zum 70. Geburtstag*. Frankfurt am Main: Peter Lang, pp. 47–60.
- Skalická, J. 2005. *Jan Amos Komenský (Comenius): Panglotia*. Toruń: Esperanto-Societo ‘Flamo’.
- Slaughter, M. 1982. *Universal languages and scientific taxonomy in the seventeenth century*. Cambridge, New York: Cambridge University Press.
- van Steenberg, J. 2008. Classificatie van kunsttalen [Online]. Available at: <http://steen.free.fr/classificatie.html> [Accessed: 7 January 2015].
- van Steenberg, J. 2013. History of Interslavic [Online]. Available at: <http://steen.free.fr/interslavic/history.html> [Accessed: 18 May 2015].
- Stria, I. 2013. Classifications of artificial languages. *Język, Komunikacja, Informacja* 8, pp. 125–132.
- Stria, I. 2015a. Esperanto as a Natural Language. *Język, Komunikacja, Informacja* 10, pp. 32–42.
- Stria, I. 2015b. Esperanto speakers – an unclassifiable community? In: Malec, W., Rusinek, M., and Sadowska, A. eds. *Challenging Ideas and Innovative Approaches in Ap-*

- plied Linguistics*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 175–189.
- Stria, I. 2015c. Linguistic worldview of Esperanto – a questionnaire method. In: Kosecki, K. and Badio, J. eds. *Empirical Methods in Language Studies*. Łódź Studies in Language. Frankfurt am Main: Peter Lang, pp. 41–51.
- Svadost, E.P. 1968. *Как возникнет всеобщий язык? [How arises a universal language?]*. Moskva: Nauka.
- Święczkowska, H. 1998. *Harmonia linguarum: język i jego funkcje w filozofii Leibniza*. Białystok: Wydawnictwo Uniwersytetu w Białymstoku.
- Tabakowska, E. 2013. A Linguistic Picture, Image, or View of ‘Polish Cognitive Studies’. In: Głaz, A., Danaher, D., and Łozowski, P. eds. *The Linguistic Worldview: Ethnolinguistics, Cognition, and Culture*. London: Versita, pp. 321–335.
- van Teijlingen, E.R. and Hundley, V. 2001. The importance of pilot studies. *Social Research Update* 35(4) Available at: <http://sru.soc.surrey.ac.uk/SRU35.html>.
- Tokarski, R. 2001. Słownictwo jako interpretacja świata. In: Bartmiński, J. ed. *Współczesny język polski*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, pp. 343–370.
- Tolkien, J.R.R. 1983. A Secret Vice. In: Tolkien, C. ed. *The Monsters and the Critics and Other Essays*. London, Boston: Allen & Unwin, pp. 198–223.
- Tomaszewski, P. 2004. Polski Język Migowy (PJM) – mity i fakty. *Poradnik językowy* (6), pp. 59–72.
- Underhill, J.W. 2009. *Humboldt, Worldview and Language*. Edinburgh: Edinburgh University Press.
- Wahlgren, Y. 2004. Klingon as Linguistic Capital. BA thesis. Lund University.
- Wandel, A. 2015. How Many People Speak Esperanto? Esperanto on the Web. *Interdisciplinary Description of Complex Systems* 13(2), pp. 318–321.

- Waringhien, G. and Duc Goninaz, M. eds. 2002. *La nova plena ilustrita vortaro de esperanto*. Paris: Sennacieca Asocio Tutmonda.
- Weiner, E.S.C. and Marshall, J. 2011. Tolkien's Invented Languages. In: Adams, M. ed. *From Elvish to Klingon: exploring invented languages*. Oxford, New York: Oxford University Press, pp. 75–109.
- Wilbur, R. 2011. Modality and the Structure of Language: Sign Languages Versus Signed Systems. In: Marschark, M. and Spencer, P. E. eds. *The Oxford handbook of deaf studies, language, and education*. 2nd ed. New York: Oxford University Press, pp. 350–366.
- Wilcox, S. 2015. Signed Languages. In: Dąbrowska, E. and Divjak, D. eds. *Handbook of Cognitive Linguistics*. Berlin: de Gruyter Mouton, pp. 667–688.
- Wood, R.E. 1979. A Voluntary Non-Ethnic, Non-Territorial Speech Community. In: Mackey, W. F. and Ornstein, J. eds. *Sociolinguistic Studies in Language Contact*. Berlin, New York: De Gruyter Mouton.
- Yule, G. 2010. *The study of language*. 4th ed. New York: Cambridge University Press.
- Zamenhof, L.L. 1887. *Język międzynarodowy: przedmowa i podręcznik kompletny: por Poloj*. Reprint. Białystok: Książnica Podlaska im Łukasza Górnickiego.
- Zuckermann, G. and Walsh, M. 2011. Stop, Revive, Survive: Lessons from the Hebrew Revival Applicable to the Reclamation, Maintenance and Empowerment of Aboriginal Languages and Cultures. *Australian Journal of Linguistics* 31(1), pp. 111–127.
- Żuk, G. 2010. Językowy obraz świata w polskiej lingwistyce przełomu wieków. In: Karwatowska, M. and Siwiec, A. eds. *Przeobrażenia w języku i komunikacji medialnej na przełomie XX i XXI wieku*. Lublin: Wydawnictwo Drukarnia Best Print.

Appendix: The questionnaire

PARTO I

1. Skribu, kian koloron havas...!

suno _____ luno _____
 ĉielo _____ sablo _____
 vulpo _____ maro _____
 fajro/flamoj _____ herbo _____
 helaj haroj _____ envio _____

2. Listigu po kvin bestojn kaj plantojn, kiuj tuj venas en vian kapon!

_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

3. Donu sinonimojn (similsignifajn vortojn)!⁷²

4. Priskribu mallonge (se eblas, per unu – du vortoj)!

- Li ĉiam estis aktiva esperantisto, sed ĵus forlasis la movadon. (Kion li faris?)

- Ili renkontiĝas nur dum kongresoj. Tio estas efemera amafero. (Kio ŝi estas?)

- Kiam ili estas inter esperantistoj, ili ofte parolas en sia denaska lingvo. (Kion ili faras?)

⁷² Questions 3, 16 and 16 have been removed. See also comment 57.

- Ŝi kredas, ke Esperanto finfine estos konata de ĉiuj kiel monda helplingvo (Kio ŝi estas?)

- Tre fervora, persista kaj dogmema esperantist(in)o (Kio li/ŝi estas?)

- Io nekomprenebla kaj bizara (Kio ĝi estas?)

- Li postulas, ke ĉiuj rigardu kaj amu ĉiulandajn homojn kiel siajn fratojn (Je kio li kredas?)

5. Listigu po kvin arbojn, legomojn, fruktojn, kiuj tuj venas en vian kapon!

_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

6. Kompletigu!

Johano estas arda esperantisto, sed _____

7. Kiuj estas kulturaj simboloj de esperantistoj? Listigu!

8. Kia estas vera esperantisto?

9. Imagu, ke la ĉi-subaj estas personoj. Listigu iliajn trajtojn (ekz. vira/ina, ida... + aspekto)!

- suno _____
- luno _____
- vivo _____
- morto _____
- forko _____
- tranĉilo _____
- kulero _____
- ŝipo _____
- aŭto _____
- sunfloro _____

10. Kiuj bestoj estas la simboloj de ...?

kuraĝo: _____ rapideco: _____
 malsaĝeco: _____ laboremo: _____
 fideleco: _____ ruzeco: _____
 obstineco: _____ forto: _____

11. Kio/kiu havas tian koloron? Kompletigu “blanka /verda ktp. (kiel) ...”!

- blanka _____
- griza _____
- nigra _____
- flava _____
- verda _____
- blua _____
- ruĝa _____
- bruna _____
- viola _____

12. Listigu po kvin birdojn kaj sovaĝajn bestojn!

_____	_____
_____	_____
_____	_____
_____	_____

13. Kiuj plantoj estas la simboloj de ...?

svelta (kiel): _____ alta (kiel): _____

malgranda: _____ acida: _____

malmola: _____ malsaĝa: _____

forta: _____

14. Enmetu vortojn!

Pro _____ [kia?] malsano la avo de Johano restis en _____ [kie? ne hejme]. Johano vizitis lin preskaŭ ĉiutage kaj kunportadis bongustan _____: diversajn legomojn, fruktojn kaj iom da ne grasa viando. Antaŭ kelkaj tagoj dum la inspekto la _____ [kiu? laboras tie], kiu nomiĝas Molina, diris, ke la avo ne povas manĝi _____ [kiajn?] paprikojn, ĉar ili malbone influas la stomakon. Johano _____ la demandon, ĉu la avo povas do manĝi pipron. Evidentiĝis, ke ne. Post du semajnoj la avo povos reveni hejmen. Johano kaj lia avo decidis _____ foton tiam kaj sendi ĝin al Sinjoro Molina por montri, ke la avo nun finfine povas manĝi pipron.

15. Diru per unu vorto!**16. Klarigu mallonge!**

.....

PARTO II**1. Bv. kompletigi kaj meti krucon al la ĝusta respondo en la kvadraton!**aĝo (en jaroj): _____ sekso: virino viro loĝlando: _____edukado: baza (1-8 klasoj) meza (ĉ. 12 klasoj) supera (universitato, altlernejo) laborrilato: laboras emerito senlaborulo studento ne laboras (ekz. dommastrino)

profesio (nuna aŭ estinta): _____

2. denaska(j) lingvo(j)

3. konataj lingvoj kun nivelo (taksu la nivelon de ĉiu agado laŭ la suba skalo skribante la ciferon; skribu kiom da jaroj vi lernis kaj/aŭ, se vi havas atestilon aŭ trapasis oficialan ekzamenon, skribu E):

1 ap- 2 iomete 3 funkcie 4 bone 5 tre 6 kvazaŭ
 enaŭ (A2) (B1) (B2) bone denaske
 (A1) (C1) (C2)

lingvo	jaroj (+ekz.)	legado	skribado	parolado	aŭskultado

4. Kiom ofte vi uzas viajn lingvojn?

1 (preskaŭ) ĉiutage 2 kelkfoje semajne 3 kelkfoje monate 4 tre malofte

lingvo	cifero	lingvo	cifero	lingvo	cifero

5. Aldonu la nomojn de lingvoj, en kiuj vi plej ofte estas engaĝita en la sekvaj agadoj:

- aŭskulti radion / televidi _____
- legi por amuziĝo _____

- legi por laboro _____
- skribi al geamikoj _____
- skribi artikolojn _____
- interparoli _____
- retumi _____

6. En kiu lingvo vi preferus legi tekston disponeblan en ĉiuj lingvoj konataj al vi? Supozu, ke la originalo estis skribita en lingvo, kiu estas nekonata al vi.

7. En kiuj lingvoj vi kutime:

- kalkulas? _____
- sonĝas? _____
- esprimas sentojn? _____
- preĝas? _____
- blasfemas? _____

The questionnaire in English⁷³

PART I

1. What is the colour of...?

the sun _____ the moon _____

the sky _____ sand _____

a fox _____ sea _____

fire/flames _____ grass _____

light-coloured hair _____ envy _____

2. List animals and plants (five each) that come to your mind immediately!

_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

3. Give synonyms!

4. Describe briefly (if possible, use one – two words)!

- He was always an active Esperantist, but has just left the movement. (What has he done?)

- They meet only during congresses. This is a short-lived love affair. (What is she?)

- When they are among Esperanto speakers, they often speak in their native language. (What do they do?)

⁷³ The translations are the closest possible to the Esperanto original.

- She believes that Esperanto will ultimately be known by all as a world auxiliary language (What is she?)

- A very keen, persistent and dogmatic Esperantist (What is he/she?)

- Something incomprehensible and bizarre (What is it?)

- He requires that all consider and love men of all countries as their brethren (What does he believe?)

5. List trees, vegetables and fruits (five each) that come to your mind immediately!

_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

6. Complete!

John is an ardent Esperantist, but _____

7. What are the cultural symbols of Esperantists? List!

8. What is a true Esperantist like?

9. Imagine that the following are people. List their traits (for example male/female, a child + their appearance)!

- the sun _____
- the moon _____
- life _____
- death _____
- a fork _____
- a knife _____
- a spoon _____
- a ship _____
- a car _____
- a sunflower _____

10. Which animals are the symbols of...?

courage: _____ speed/rapidity: _____
stupidity: _____ hard work/diligence: _____
fidelity/loyalty: _____ cunning: _____
obstinacy/stubbornness: _____ strength: _____

11. What / who is of this colour? Complete: “white/green etc. (as)...”!

- white _____
- grey _____
- black _____
- yellow _____
- green _____
- blue _____
- red _____
- brown _____
- purple _____

12. List birds and wild animals (five each)!

_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

2. native language(s)

3. known languages with their level (evaluate the level of each activity according to the scale below placing the appropriate number; write how many years you have studied and / or if you have a certificate or passed an official exam, write E):

1 barely (A1) 2 a little (A2) 3 functionally (B1) 4 good (B2) 5 very good (C1) 6 native-like (C2)

language	years (+e)	reading	writing	speaking	listening

4. How often do you use your languages?

1 (almost) everyday 2 several times a week 3 several times a month 4 very rarely

lan- guage	num- ber	lan- guage	num- ber	lan- guage	num- ber

5. Name languages in which you are most often engaged in the following activities:

- listening to radio / watching TV _____
- reading for fun _____
- reading for work _____
- writing to friends _____
- writing articles/papers _____
- talking _____
- surfing the Internet _____

6. In which language would you prefer to read a text available in all languages known to you? Assume that the original was written in another language, which is unknown to you.

7. In which languages do you usually:

- count and do simple arithmetic? _____
- dream? _____
- express feelings? _____
- pray? _____
- swear? _____

Summary in Esperanto / Resumo

INVENTI LINGVOJN, INVENTI MONDOJN

AL LA LINGVA BILDO DE LA MONDO POR ARTEFARITAJ LINGVOJ

La libro celis determini kiel malsamas artefaritaj lingvoj, kiuj faktoroj distingas naturajn lingvojn de artefaritaj lingvoj kaj kiel eblas tiujn lingvojn esplori uzante la ilojn de kultura lingvistiko, specife de la Lublina Skolo (Pollando).

La titola semantika koncepto de *lingva bildo de la mondo* (kutime pole mallongigita kiel JOS, ‘językowy obraz świata’), bazita al verkoj de Humboldt kaj disvolvita en Pollando ĉefe de Jerzy Bartmiński, estas uzata por studi naturajn etnajn lingvojn. Ĝuse oni ankaŭ proponas komparajn studojn. La sola verko konata al la aŭtorino de tiu ĉi libro, kiu diskutis enketilan studon de lingva bildo de la mondo por artefarita lingvo, nome Esperanto, estas artikolo de Koutny (2010). Tiu ĉi libro estas ĝis nun la sola provo apliki la koncepton de JOS al pli larĝa spektro de artefaritaj (ne-etnaj) lingvoj el teoria perspektivo.

En la **ĉapitro 1. “Lingva bildo de la mondo”** la koncepto de JOS (“lingvo-enradikiĝinta interpreto de realeco, kiu povas esti esprimita en la formo de juĝoj pri la mondo”; Bartmiński 2012: 23) kaj ĝia historia evoluo estas klarigitaj. Bartmiński opinias, ke ĉiuj lingvoj interpretas la mondon siamaniere kaj, ke prudenta analizo de lingvaj esprimoj povas malkovri kategoriojn, kiujn kreis lingva komunumo. La teorio kaj metodiko disvolvita kiel parto de la Lublina Etnolingvistika Skolo estas prezentataj kune kun diversaj alternativaj difinoj de JOS kaj praktikaj konsekvencoj de la adopto de iu el ili. Klarigitaj al la legantoj estas ŝlosilaj konceptoj kiel *subjekto*, *perspektivo* kaj *facetoj*. Ĉiuj verkoj de la Skolo traktas etnajn lingvojn kiel nature evoluantajn en lingvokomunumoj.

En la **ĉapitro 2. “Lingvoj artefaritaj”** prezentataj estas la titolaj lingvoj el historia kaj tipologia perspektivo. Mi ankaŭ diskutas iliajn difinojn. En ĉapitro 2. estis montrite, ke artefaritaj lingvoj estas difinitaj kiel lingvoj kreitaj intence de specifa(j) persono(j) (iam kiel komuna karakterizaĵo aldonite

estas „kreitaj dum relative mallonga periodo”). Tiuj difinoj klasifikas lingvojn surbaze de la deveno, dividante ilin en du homogenajn klasojn: artefaritaj kaj naturaj. Tamen, kiel montras la analizo en ĉapitro 2., tiuj klasoj ne estas unuformaj. Eĉ la historia superrigardo sufiĉas por konstati, ke malsamaj tipoj de artefaritaj lingvoj evoluis en 5 ĉefaj etapoj:

- aprioraj en la 17a jc.,
- aposterioraj en la 19a jc.,
- formalaj en la 20a jc.,
- artaj en la 20a jc.,
- interretaj ekde 90aj jaroj de la 20a jc.

Tiuj ĉi lingvoj diferencas ankaŭ laŭ la strukturo kaj origino de la materialo, laŭ la celo de kreado kaj ankaŭ funkcie.

Detale la problemo de binara divido inter lingvoj naturaj kaj artefaritaj estis ekzamenata en la **ĉapitro 3. “Natura aŭ artefarita”**. La analizo helpe de kvar larĝaj aroj de trajtoj ebligis identigi klason de limokazoj:

- piĝinoj,
- revigligitaj (*revitalised*) kaj revivigitaj (*revived*) lingvoj,
- kontrolitaj/minimumaj lingvoj,
- signolingvoj kaj signaj lingvoj,
- lingvaj rekonstruoj.

Lingvoj estis taksataj koncerne la difinajn trajtojn de homa lingvo laŭ Hockett, koncerne la klasojn de natureco laŭ Lyons, la nivelojn de celkonscia influo laŭ Svadost kaj modifitajn listojn de trajtoj de naturaj lingvoj laŭ Duliĉenko kaj Liu. Tiu ĉi proceduro ebligis al mi renversi la tradician dividon inter artefaritaj kaj naturaj lingvoj, montrante, ke ĉiuj lingvoj estas metitaj sur la skalon inter tiuj du polusoj. Laŭanalize la menciita absoluta divido ne nur povas esti konsiderata erara, sed ankaŭ malutila al la disvolviĝo de lingvistiko kiel scienco kiu provas malkovri la mekanismojn malantaŭ lingvaj kategorigoj per forigo de valora esplormaterialo, nome artefaritaj lingvoj.

La plej natura el artefaritaj lingvoj estas Esperanto. Al Esperanto kiel lingvo evoluanta nature estas dediĉita la **ĉapitro 4 “Esperanto kiel transira kazo”**. Estas montrite, ke Esperanto evoluas nature, kvankam ĝi estas lingvo artefarite kreita. Tiu ĉi ĉapitro koncentriĝas pri la socilingvistika situacio de uzantoj de la internacia helplingvo kompare al lingvoj tradicie konsideritaj naturaj. Grava konkludo estas, ke la komunumo de Esperanto-uzantoj estas dualingva komunumo. Plie, eĉ denaskaj parolantoj estas ĉiam almenaŭ dulingvaj.

La **ĉapitro 5. “Lingva bildo de la mondo por artefaritaj lingvoj”** estas teoria provo apliki la paradigmon de la Lublina Etnolingvistika Skolo al artefaritaj lingvoj kaj kelkaj limokazoj. Mi skizas eblajn limigojn de tiu aliro surbaze de la ŝlosilaj konceptoj de JOS aplikataj al diversaj specoj de lingvoj.

La Lublina Skolo supozas, ke la JOS formiĝas en parolkomunumo enradikiĝinta en la kulturo; tiu komunumo estas la **subjekto** de la esprimado kaj kreas la strukturon (forme de facetoj) de la koncepto. Tio signifas, ke la identigo de subjekto estas ege grava por, ke la Lublina metodiko povu esti aplikata al iu ajn artefarita lingvo.

Tia kompreno de funkcioj de la subjekto povus signifi, ke nur lingvoj kun sufiĉe granda nombro da parolantoj povus esti interesaj kiel esplormaterialo al lingvistoj uzantaj la koncepton de JOS. Por forigi tian situacion la tekstecaj kriterioj de Beaugrande kaj Dressler (1981) pruvigis helpemaj. Artefaritaj lingvoj estis taksitaj tra tiuj kriterioj kaj hierarkiigitaj de plene evoluinta lingvo tra duon-lingvo kaj tekstemo (*texteme*; “teksta modelo” laŭ Bartmiński & Niebrzegowska-Bartmińska 2010) ĝis teksto (produkto de specifa aŭtoro). Tiu hierarkio ebligas inkluzivi ĉiujn artefaritajn lingvojn en la analizo de JOS, supozante, ke kelkaj el ili estas artaj kreaĵoj, kiuj povas esti esplorataj nur kiel poezio aŭ aliaj literaturaj tekstoj.

La **ĉapitro 5**, ankaŭ proponas esploron de JOS en Esperanto. Levitaj en tiu kunteksto estas la problemoj de multlingveco kaj semantika nebuleco. Tio igas, ke ajna studo de JOS en Esperanto devus konsideri la problemon de reciproka influo

de lingvo kaj kulturo en la respondantoj. La propono estas do, ke la enketoj enhavu demandojn por determini la influon de aliaj lingvoj konataj al enketitoj sur Esperanto.

La **ĉapitro 6. “Esploro de problemoj sur la ekzemplo de pilota studo”** prezentas la rezultojn de la pilota studo pri la lingva bildo de la mondo de 32 altnivelaj Esperanto-parolantoj (inkluzive du denaskulinojn). Mi pruvas, ke kvankam Esperanta JOS estas pli limigita ol tiu en etnaj lingvoj, ĝi klare ekzistas en la sekvaj terenoj:

- kulturaj konceptoj asociitaj kun la Movado;
- stereotipo de Esperantisto.

La studo malkaŝis multajn problemojn pro la fakto, ke Esperanto-parolantoj ne estas homogena grupo kaj precipe tiu de multlingvuloj. Sed samtempe ĝi konfirmis la hipotezon antaŭe starigitan: unue, ke Esperanto alproksimiĝas naturecon; due, ke multaj terenoj de la nuna JOS en Esperanto estas forte influitaj de denaskaj lingvoj kaj kulturoj de parolantoj; kaj fine, ke la kerno de la Movado prezentas koheran bildon de la mondo bazitan sur la sento de aparteno al la komunumo, ĝia kulturo kaj normoj.

Teoriaj analizoj en tiu ĉi libro montris, ke artefaritaj lingvoj ne estas homogena grupo kaj iliaj diversaj karakterizaĵoj necesigas diversajn perspektivojn, kiam oni provas uzi la paradigmon de la Lublina Skolo. La lingvo, kiu povas enkadre de tiu paradigmo esti traktata kiel natura, estas sendube Esperanto. Plia studo de ĉi tiu lingvo, kaj en estonteco eble ankaŭ aliaj artefaritaj lingvoj, helpos pligrandigi la konojn pri la homa kapablo uzi lingvon.



Ida Stria – doktor nauk humanistycznych w zakresie językoznawstwa stosowanego, magister filologii szwedzkiej oraz etnolingwistyki (UAM). Wykładowca akademicki. Bada językowy obraz świata w języku szwedzkim oraz esperanto.

Jak do tej pory języki sztuczne odrzucane były jako przedmiot badań współczesnego językoznawstwa w kontraście do języków naturalnych. Problematyka konstruowania języków podejmowana była w pracach historycznych, a także interlingwistycznych. Wobec tego zamierzeniem monografii *Idy Strii* stała się ocena możliwości zastosowania metodologii tzw. Lubelskiej Szkoły Etnolingwistycznej do semantycznej analizy szeroko pojmowanych języków sztucznych oraz zaprezentowanie ograniczeń wynikających z takiego podejścia.

Autorka wprowadza czytelnika w problematykę językowego obrazu świata, historię i typologię języków sztucznych i wreszcie przedstawia studium empiryczne. Zaprezentowane analizy wskazują, że kontrast pomiędzy językami naturalnymi a sztucznymi zaciera się. Wśród tych ostatnich szczególnym przypadkiem jest język esperanto, który bez wątplenia może być traktowany jak naturalny w ramach Lubelskiej Szkoły Etnolingwistycznej. W ostatniej części rozprawy autorka podejmuje się zatem opisu językowego obrazu świata esperantystów i wykazuje istnienie silnie zakorzenionych stereotypów kulturowo-językowych w tej społeczności. Postuluje jednocześnie uwzględnienie transferu z innych języków znanych respondentom a także zewnętrznych wpływów kulturowych we wszelkich badaniach języków sztucznych w ramach paradygmatu lubelskiego.



DOI: 10.14746/9788394760915