

UNIVERZITA KARLOVA V PRAZE
FILOZOFICKÁ FAKULTA
ÚSTAV ANGLICKÉHO JAZYKA A DIDAKTIKY

DIPLOMOVÁ PRÁCE

Petr Janák

Multilingualism in *A Clockwork Orange* and Its Translations

Vícejazyčnost v *A Clockwork Orange* a jeho překladech

Praha 2015

Vedoucí práce

PhDr. Pavlína Šaldová, Ph.D.

Za možnost téma zpracovat, za cenné rady, trpělivost, ochotu a vstřícnost
děkuji vedoucí práce PhDr. Pavlíně Šaldové, Ph.D.

Prohlašuji, že jsem diplomovou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 1. srpna 2015

Klíčová slova

normalizace, lexikální kreativita, smyšlený jazyk, umělecké konstruované jazyky, paralelní text, nadsat, Anthony Burgess, Mechanický pomeranč, Uhrwerk Orange

Keywords

normalisation, normalization, lexical creativity, invented language, artlang, parallel text, Nadsat, Anthony Burgess, Mechanický pomeranč, Uhrwerk Orange

Abstrakt

Práce se zabývá vnitrotextovou vícejazyčností (interjazykovou heterogeností) v románu *A Clockwork Orange* britského spisovatele Anthonyho Burgesse a v jeho dvou překladech, českém a německém. Analýza 180 slov ze smyšleného jazyka nadsat, který se v románu vyskytuje, ukazuje, jak se v překladech projevuje lexikální kreativita, tedy zda a jak se lexikální kreativita originálu liší od lexikální kreativity v jeho překladech. Změny lexikální kreativity jsou dány do souvislosti s normalizací (jednou z překladových univerzálií) a s funkcemi daného smyšleného jazyka.

Formy a funkce smyšlených jazyků a především jazyka nadsat jsou stanoveny podle již existující klasifikace zavedené pro vícejazyčnost. Analýza jazyka nadsat a jeho překladových protějšků je kvantitativní a využívá konkordanční programy AntConc a ParaConc. Při analýze se zkoumá frekvence slov z nadsat, jejich rozmístění v textu a způsob, jakým je jejich význam čtenáři zprostředkován. Získané údaje jsou poté použity ke srovnání jazyka nadsat a smyšlených jazyků, které ho v českém a německém překladu nahrazují.

Analýza ukázala, že v obou překladech je smyšlených lemmat méně než v originále, přičemž v německém překladu (UO) je jich mnohem méně než v českém (MP). Český přišel celkově o 10 % lemmat z nadsat a německý o 41 %. Také rozmístění jednotlivých tvarů smyšlených slov v textu se v UO změnilo více než v MP, přičemž v UO se změny týkají většího počtu jednotlivých tvarů smyšlených slov než v MP a jsou také zásadnější. Strategie vnitrotextového překladu slov z nadsat je více dodržována v MP než v UO.

Lexikální kreativita je oslabena v obou překladech, ale mnohem více v německém, přičemž normalizace smyšlených slov vzrostla nepřímě úměrně k tomuto oslabení. Změny v překladu měly vliv na funkce jazyka nadsat. V MP nepůsobí důsledky změn průkazně – nejvíce byla patrně ovlivněna funkce utajovací a restriktivní, ale mnohé změny lze vysvětlit nepochopením originálu a nedbalou redakční prací; ostatní funkce nebyly výrazně ovlivněny. V UO byla výrazně ovlivněna funkce atmosférotvorná, charakterizační i hodnotová. Změny funkce utajovací a restriktivní nemohly být zhodnoceny, protože tato funkce byla příliš ovlivněna mnoha dalšími faktory.

Abstract

The paper explores intratextual multilingualism in *A Clockwork Orange* (ACO) by Anthony Burgess, and in two of its translations – into Czech and German. It analyses 180 words from Nadsat – the invented language in ACO – to reveal how lexical creativity is manifested in translation, i.e. whether and how lexical creativity that is present in the original text is changed in the translations. Changes in lexical creativity are linked to normalisation (a translation universal), and to the functions of the invented language.

An existing classification of forms and functions of intratextual multilingualism is applied to invented languages and, in particular, to Nadsat. The analysis of Nadsat and its counterparts in the translations is quantitative, and is conducted using the concordancers *AntConc* and *ParaConc*. It examines the frequency of Nadsat words, their distribution throughout the text, and the way their meaning is conveyed to the reader. These data are then used in the comparison of Nadsat and the invented languages that replace it in the Czech and the German translations.

The analysis shows that in both translations the number of invented lemmas is lower than in the original, and that in the German translation (UO) the number is significantly lower compared to the Czech translation (MP). In total, MP disposed of 10% of Nadsat lemmas and UO disposed of 41 %. It also showed that both translators made changes in the distribution of the invented word forms when compared to their Nadsat counterparts; the changes affect more invented word forms in UO than MP, and the changes are greater in UO. The strategy of intratextual translation of Nadsat words is used more consistently in MP than in UO.

Lexical creativity is diminished in both translations, but much more so in UO; normalisation of invented words has risen in inverse proportion to the decrease in lexical creativity. The changes in translation affected the functions of Nadsat. For MP, the results of the changes appear inconclusive: the cryptic and restrictive function was probably affected the most, but many of the changes can be ascribed to a misunderstanding of ACO and careless editorial revision; other functions were not affected significantly. In UO, the atmospheric, characterizing, and value functions were affected significantly. Changes to the cryptic and restrictive function could not be assessed, as there were too many factors involved.

Table of contents

Table of contents.....	1
List of abbreviations	3
List of tables.....	4
0 Introduction.....	5
1 Multilingualism.....	9
1.1 Definition of multilingualism	9
1.2 Some remarks on the history of and research into intratextual multilingualism in Western literature.....	11
1.3 Forms of intratextual multilingualism	12
1.3.1 Introduction.....	12
1.3.2 Primary forms	13
1.3.3 Secondary forms	17
1.3.4 Forms of Nadsat in ACO	19
1.4 Invented languages.....	21
1.4.1 Definition	21
1.4.2 A short history of invented languages in fiction.....	22
1.4.3 Functions of invented languages.....	24
2 <i>A Clockwork Orange</i> and the functions of Nadsat.....	27
2.1 Introduction.....	27
2.2 Structure of ACO	29
2.3 Setting of ACO	30
2.4 The setting as reflected in Nadsat	31
2.5 ‘External history’ of Nadsat.....	32
2.6 Reception of Nadsat.....	33
2.7 Mirroring in ACO	35
2.8 Functions of Nadsat	36
3 Textual organisation of Nadsat	38
3.1 Introducing Nadsat.....	38
3.2 Preparing the list of Nadsat words	41
3.3 Analysis.....	44
3.3.1 Frequencies	44
3.3.2 Distribution	46

3.3.3 Unfolding Meaning.....	50
3.4 Functions of Nadsat reflected through the analysis	55
4 Textual organisation of Nadsat’s counterparts in translation	57
4.1 Introduction.....	57
4.1 Concordancing	58
4.3 Analysis of the results.....	60
4.3.1 Introduction.....	60
4.3.2 Frequencies	61
4.3.3 Distribution	68
4.3.4 Unfolding Meaning.....	74
5 Conclusion	78
5.1 Summary of the analysis.....	78
5.2 Lexical creativity and normalisation.....	79
5.3 Changes in functions.....	79
5.4 Further implications	81
Sources.....	82
Résumé.....	94
Appendix I – Nadsat (ACO) lemmas alphabetically	100
Appendix II – Nadsat (ACO) lemmas by frequency	101
Appendix III – JT (MP) lemmas alphabetically	102
Appendix IV – JT (MP) lemmas by frequency.....	103
Appendix V – Nadsatsprache (UO) lemmas alphabetically	104
Appendix VI – Nadsatsprache (UO) lemmas by frequency	105

List of abbreviations

ACO	<i>A Clockwork Orange</i>
BC	before Christ
GEPCOLT	German-English Parallel Corpus of Literary Texts
JT	jazyk Týnů
KWIC	Key Word In Context
LL	log-likelihood
MP	<i>Mechanický pomeranč</i>
OED	<i>Oxford English Dictionary Online</i>
PL	primary language
SL	secondary language
ST	source text
TT	target text
UO	<i>Uhrwerk Orange</i>
US	United States of America
USSR	Union of Soviet Socialist Republics

List of tables

Table 1.1 – Primary Forms: Intratextual Multilingualism & Invented Languages.....	17
Table 1.2 – Combinability of Primary and Secondary Forms	19
Table 3.1 – Examples of Lemmas.....	43
Table 3.2 – Summary of Frequencies of Nadsat Words	44
Table 3.3 – Frequencies of the 180 Nadsat Words	45
Table 4.1 – Overview.....	60
Table 4.2 – Translation Strategies Groups – Changes in Nadsat Lemmas.....	61
Table 4.3 – Translation Strategies Groups – Groups of TT Lemmas.....	66
Table 4.4 – Translation Strategies Groups – TT Word Forms	66
Table 4.5 – Translation Strategies Groups – Changes in Nadsat Hapaxes.....	67
Table 4.6 – TT Tokens with Matching or Changed Position.....	68
Table 4.7 – Top 20 MP Lemmas with the Highest Deviation	69
Table 4.8 – Top 20 UO Lemmas with the Highest Deviation	70
Table 4.9 – Numbers of ACO Lemmas Creating Clusters (by Group)	72
Table 4.10 – ACO Lemmas Creating Clusters (by Group)	73
Table 4.11 – Intratextual Translation.....	75

0 Introduction

The aim of this paper is to explore intratextual multilingualism (Knauth 2007, 1), i.e. multilingualism embedded within a larger monolingual matrix (Sebba 2013, 98), in the novel *A Clockwork Orange* (ACO) by the British author Anthony Burgess, and in two of its translations – into Czech and German. ACO features an invented¹ language called Nadsat, which consists of non-English words created on the basis of words from other languages (mostly Russian). Its large number of words, their form, consistency in their usage, distribution throughout ACO, repetition, mirroring, analogies and wordplay make it a prominent feature of ACO.

The analysis of 180 Nadsat words will reveal how lexical creativity is manifested in translation, i.e. whether and how lexical creativity that is present in the original text is changed in the translations. The results of the analysis may – or may not – support one of the translation universals, normalisation. As such, the analysis is data-driven rather than hypothesis-driven. We will also try to assess the effect of the changes on the functions of Nadsat in ACO according to a classification proposed by Mareš for literary multilingualism. (Mareš 2003, 40–45) The analysis will be quantitative, corpus-inspired, and will use corpus linguistic methods, and it will be conducted using the concordancers *AntConc* and *ParaConc*.

The method of our exploration is quite simple: to look at the places (sentences) with words from Nadsat in the original, and at what appears at the ‘same place’ in the translations. This can basically be a word from the invented language of the translation or a word from the language of the translation.

The importance of Nadsat words for translation is twofold: first, they are prominent in ACO in a way similar to keywords, “items of unusual frequency in comparison with a reference corpus” (Scott and Tribble 2006, 55), and translation consistency is needed to successfully re-create the global meaning of ACO, and second, as invented words they are creative and require creative translations.

Anna Čermáková and Lenka Fárová look at the systematicity of Czech and Finnish translation equivalents of lexical textual keywords in *Harry Potter and the Philosopher’s Stone* (Čermáková and Fárová 2010), which is similar to our approach: they also pick words

¹ We use the term invented language following its use in Forster 1970, 88. In this use, it is a language invented and used by its author in a work of fiction. Other terms used include constructed language, fictional language, imaginary language, artificial literary language, or artlang. The definition of the term is dealt with in 1.4.1.

selected according to certain criteria from the source text (ST), and they are interested in the distribution of these words throughout ST and target texts (TT) and their translation consistency. However, as the group of Nadsat words is quite clearly delimited, we do not need to work with the concept of keyness in this paper even though some Nadsat words probably are keywords due to their high frequency in ACO and zero frequency in a hypothetical referential corpus.

Nadsat words vary greatly in their frequency, distribution, and the means by which their meaning is revealed, all of which affects the interpretation of ACO. Lexical repetition plays an important role in the global meaning of a text. (Klaudy and Károly 2000, 143–144) In informative types of texts, where proper terminology has to be used, verbatim repetition of terms cannot be avoided. (Klaudy and Károly 2000, 150) On the contrary, it is something the author uses to make the text coherent. The translator of such a text who is aware of the text-organizing function of precise, verbatim repetition is able to reproduce the global meaning of the text more precisely. We believe the role of verbatim repetition in ACO can be compared to that of informative texts, and translating Nadsat words consistently helps to reproduce the global meaning of ACO. Indeed, according to Lotfipour, lexical repetition in literary texts assumes a more important role in establishing the required literary effect than it does in other texts, and “for a translated text to be equivalent to its SL [secondary language] counterpart in terms of the texture dimension, the translator should observe its textual features.” (Lotfipour-Saedi 1997, 190)

The study of creativity in linguistics started in the early 20th century, and has been actively pursued since the latter half. (Vo and Carter 2010, 302) In the last fifteen years it has also been investigated using corpora. According to Kenny, “linguistic creativity can reside in the creation of new words, the novel use of existing words, the use of unconventional syntax or punctuation, the manipulation of typographical convention and so on.” (Kenny 2000, 94)² Lexical creativity is more accessible than other types of creativity (e.g. syntactic) because the corpus of texts does not have to be necessarily annotated. However, one of the problems with the systematic study of lexical creativity lies in finding what can be considered a creative use of lexis. Kenny uses hapax legomena to search for instances of lexical creativity. However, it was quite time-consuming for her to extract 117 hapax forms “that were deemed creative”

² Kenny studied lexical creativity in translation using GEPCOLT, the German-English Parallel Corpus of Literary Texts, which was designed for her research into lexical creativity and normalisation, and contains fourteen German literary works and their translations into English – ca. one million tokens in each language. (Kenny 2001, 111–115)

from the 42,000 hapaxes found in her (unannotated) corpus. (Kenny 2000, 95–96) She also looked for writer-specific forms using (positive) keywords, and for unusual collocations by browsing repeated clusters. These two strategies were not very fruitful, as “only ca 0.1% of the keywords uncovered were useful in a study of lexical creativity” (Kenny 2000, 98) and investigating the clusters produced only “a handful of repeated instances of creativity”. (Kenny 2000, 99)

Studying lexical creativity using hapax legomena seems complicated, and such words do not usually play an important role in the text in question. We consider Nadsat words a useful example of lexical creativity because, due to their innovativeness, they are – as with other invented languages – easily accessible in ST and traceable in TT. Moreover, they enable us to study translation consistency of creative words that is important in texts such as ACO and that cannot be studied using hapax legomena.

Kenny investigated lexical creativity in order to “find out whether or not normalization is a feature of translation.” (Kenny 2000, 94) According to Kenny, “normalization may be said to occur when translators opt for conventional target language solutions to problems posed by creative or unusual text features”; lexical normalisation is, then, “normalization at the level of individual words and collocations.” (Kenny 2000, 94) Normalisation, in this sense, belongs among the universals of translation. The terms conventionalisation and standardisation are also used. (Mauranen 2007, 40–41) Universals of translation originated from translation hypotheses proposed by Mona Baker some twenty years ago in Baker 1993, and this concept has since been developed in translation studies. The notion of “universals” raised opposition, and parallel to the term “translation universals”, other names such as “laws”, “tendencies”, or “conditioned regularities” are used. They have been studied mostly from the linguistic point of view (as opposed to the social or the cognitive point of view), benefiting from corpus linguistics. (Mauranen 2007, 33–36)

From the point of view of lexis, normalisation is manifested in translations by using “typical, common lexis instead of the unusual or the unique” (Mauranen 2007, 41) found in ST. It may include “minimization of the transfer of foreign language expressions found in the source text” and “rewriting experimental narrative in a more familiar mode” (Laviosa-Braithwaite 2008, 290), which is what we will analyse in ACO and its translations.

As pointed out above, ACO contains non-English words. That makes it classed among literature that uses multilingualism as a literary (usually stylistic) device. However, the ‘other’ language used in ACO is not an existing language but Nadsat, a language created

by Burgess for (and by) this work. Because of that, ACO is not a prototypical novel employing multilingualism. ACO belongs to a subgroup of novels that use invented languages. This subgroup is large and heterogeneous even if we limit ourselves to the so-called Western literature. Invented languages are used in various literary genres, in various ways, and to a different extent in specific texts. In many invented languages, especially within science-fiction and fantasy, some existing language can be traced. For example, Tolkien's elvish languages are influenced by Finnish and Welsh, among other languages. (Carpenter 1992, 101) However, we will not look into the lexical meaning of Nadsat words or their etymology. We will not try to analyse the structure of Nadsat lexis in general, e.g. its word classes, lexical relations within Nadsat, or the reasons why certain concepts, and not others, were rendered in Nadsat.

Before we approach the analysis of Nadsat, we outline the structure of this paper. First, the topic of literary multilingualism perhaps needs to be introduced in more detail. Chapter 1 of this paper serves as such an introduction with special emphasis on invented languages. Chapter 2 provides background information on ACO and identifies the forms and functions of Nadsat in ACO. In Chapter 3, the textual organisation of Nadsat words in the original – their frequency of Nadsat words, their distribution throughout the text, and the way their meaning is conveyed to the reader – is analysed using *AntConc*. In Chapter 4, *ParaConc* is used to contrast the systems of textual organisation of Nadsat words in ACO and the systems of textual organisation of the words of translators' substitutions for Nadsat in the Czech and German translations of ACO. In Chapter 5, the results of the analysis are summarised along with their effects on translation creativity and normalisation in the Czech and German translations.

1 Multilingualism

1.1 Definition of multilingualism

The term “multilingualism” has many more or less specific uses and definitions. Studies on multilingualism can be found anywhere from sociolinguistics to psycholinguistics to language acquisition to discourse analysis and other disciplines. The studies treat multilingualism as the use of more than one language by an individual speaker or by a group of speakers, or the capacity to do so. Multilingualism can be studied in spoken or written discourse, but, according to Sebba, “a relatively small body of linguistic research to date has concerned itself with the phenomena of written language mixing” and if so, the majority of studies “focused on interactive genres which resemble conversation”. (Sebba 2013, 98)

Nadsat seems to be similar to ‘symbiotic’ mixed languages (Matras 2009, 291), anti-languages (Halliday 1987, 164–192), secret languages, urban youth languages, or street languages. For example, Muysken’s description of relexicalization fits the character of Nadsat, where mostly core vocabulary nouns are replaced:

“the replacement of native vocabulary by words from one or more other languages. This process, like borrowing, is subject to category restrictions: it involves major class items, mostly nouns. However, unlike borrowing, it is not subject to semantic restrictions, and often involves the replacement of core vocabulary. [...] The process of relexicalization is extremely frequent in a wide variety of urban youth slangs, in jargons, secret and trade language, etc.”(Muysken 2007, 316–339)

Relexicalization, semantic inversion, and overlexicalization, which are all present in Nadsat, are conspicuous features of anti-language. In his, in many respects pioneering, essay, Mikhail Bakhtin considers languages “specific points of view on the world, forms for conceptualizing the world in words, specific world views, each characterized by its own objects, meanings and values.” (Bakhtin 2011, 291–292) The confrontation of world views, ideologies, and the norms of society leads Fowler to contemplate Nadsat, with respect to Bakhtin and Halliday, as an example of anti-language in fiction, a specialized subcode of a subculture showing an “antagonistic relationship with the norm society”. (Fowler 1981, 142) Fowler treats anti-language as a variety and a process – as a creative critique of the norm of

the society which in turn helps to create anti-language. (Fowler 1981, 150–151) This relationship, the tension between the society and anti-society, and its manifestation through Nadsat, is apparent in ACO.

However, this paper will not include an analysis of Nadsat from the sociolinguistic point of view. This means answers to questions about the characters in ACO, such as why they use Nadsat, in what situations they use Nadsat and why, how they use Nadsat, and which sociolinguistic term describes Nadsat best will be neither given nor sought. However interesting the similarities between Nadsat and these types of ‘languages’ are, we believe it is inappropriate to impose certain (socio)linguistic concepts (code-switching / code-mixing, language change, borrowing) on an invented language used in fiction for two reasons. First, invented language of this kind rarely had any real speakers, and this we consider a major drawback to any sociolinguistic research. Second, – and this we believe is true for natural languages as well, and might even be the reason why research on multilingualism has focused on spoken language – even though a literary text is a record of one use of the language (in this case two or more languages), it reflects the creativity of the author and not his natural linguistic behaviour or the behaviour of other people. According to Bakhtin, “foreign languages used in literature are stylized, artistically represented, representing the image of a language.” (Bakhtin 2011, 336)

Of course, languages presented in a literary work are – to various degrees – always “artistically represented” but as Sebba puts it, it is “not obvious” that “the term ‘code-switching’ and the related terms are applicable to written language at all”. (Sebba 2013, 98) Code-mixing in the speech of characters in a fictional world gives direct evidence only to fictional linguistic behaviour within the fictional world; Mareš also disapproves of the connection between the two worlds. (Mareš 2003, 16–17) On the other hand, multilingual texts other than literary ones, such as advertising, personal letters, and newspaper articles have been analysed from different linguistic perspectives.

Moreover, ACO is not a reliable, let alone authentic record of the use of Nadsat, because (a) everything in ACO is narrated through the perspective of the narrator, and (b) the narrator is a first-person narrator. It may well be possible, for example, that all conversation among the members of Alex’s gang is made in Nadsat, and the narrator selected only some of these words for his literary purpose. However, it has to be noted that translators (especially those without the ambition to delve into invented languages) measure Nadsat against their own knowledge of and experience with secret languages, or various argots and slangs, which is reflected in their translations.

We have tried to show that it is crucial to distinguish between the type of multilingualism that appears in fiction, and other instances of multilingualism. Within literary multilingualism, two types of multilingualism have been traditionally distinguished: intratextual and intertextual. Probably the best definition of literary multilingualism from this point of view is by Knauth:

“The term literary multilingualism primarily refers to the more or less extended mix of two or more languages in the same text, entailing a cross-cultural or experimental effect. Besides intratextual multilingualism, or mixtilingualism, there is an intertextual multilingualism between heteroglot works of different authors linked to each other in a specific way (like those of the European and Latin American corpus of Petrarchan poetry) or between the heteroglot works of the same bilingual author (like Samuel Beckett’s alternative English and French fiction and drama).

(Knauth 2007, 1)

Since Burgess wrote only in English, we focus only on intratextual multilingualism. Intratextual multilingualism is further sub-classified in Macurová and Mareš 1996, 164–165 and Mareš 2003 but for our purposes, the term intratextual multilingualism is sufficient.

1.2 Some remarks on the history of and research into intratextual multilingualism in Western literature

Having introduced the term intratextual multilingualism as a type of literary multilingualism, we will have a look at its history and the history of research into it. Literary multilingualism has been studied widely by literary scholars, who were interested mostly in poets who wrote in more than one language but did not use more languages in a single work (intertextual multilingualism). The aspect of a functional stratification of languages present in works using more than one language was studied to a lesser extent.

In his recently reprinted book, Leonard Forster (Forster 1970) provides us with a concise summary of the history of the use of different languages in literature. In 1975 the literary critic George Steiner wrote that “Leonard Forster’s *The Poet’s Tongues* introduces a large, unexplored field.” (Steiner 1998, 126)

The history of research into literary multilingualism is briefly summarised in Mareš 2003, 15–18. However, Mareš mentions mostly works on German and Czech literature, and to a lesser extent also on French and Russian, from 1927 to 1997. More recent but mostly marginal studies are reviewed in Mareš 2012, 153–172. The various ways by which

multilingualism entered literature are outlined in Forster 1970, Goetsch 1987, Mareš 2003, and Knauth 2007.

According to Mareš, providing a systematic account of the history of literary multilingualism would be an arduous if not impossible task. (Mareš 2003, 19) There are many works employing multilingualism, in different cultures and ages, and thus any history of it will always be selective and incomplete. Whether such a history is needed is a different question. According to Knauth, “Intratextual multilingualism was a marginal phenomenon during the predominantly monolingual periods of literary history from Greek and Roman antiquity until the end of the 19th century.” (Knauth 2007, 1) Nevertheless, various forms of the use of foreign languages are exemplified in studies dealing with literary multilingualism; some of them provide lists of works, which are of necessity highly selective.

From the studies available to us we may extract a short list of the well-known authors to enable readers to see the span of intratextual multilingualism. Apart from the Bible, intratextual multilingualism has been used, for example, by Aristophanes, Menander, Plautus, Seneca, Kalidasa, Raimbaut de Vaqueiras, Dante, Tifi Odasi (Macaronea), François Rabelais, Lope de Vega, John Milton, William Shakespeare, Molière, Guy de Maupassant, Leo Tolstoy, Stefan George, Thomas Mann, Rudyard Kipling, D. H. Lawrence, Marina Tsvetaeva, James Joyce, T. S. Eliot, Marcel Proust, and Ezra Pound, to name just a few. From the late 19th century, examples become too numerous to list, albeit selectively. Mareš 2003 mentions ca. forty 20th century Czech authors who use intratextual multilingualism. In all these works, multilingualism manifests itself in various forms and has various functions.

1.3 Forms of intratextual multilingualism

1.3.1 Introduction

Mareš develops an all-embracing classification of the way foreign language passages are presented in a text based primarily on a signifier-signified relation, where only the signifier – a particular stretch of text – can be accessed directly by the reader. The signified is the statement made by a participant (character) in the fictional world created by (through) the text. (Mareš 2003, 34–35)

Primary forms of the signifier are **presence**, **elimination**, **evocation** and **signalization** of multilingualism. All of these types refer to the same type of a signified, i.e. to a statement in a certain, usually foreign, language. Here we use the term secondary (embedded) language(s) (SL) for such language(s), and the term primary language (PL) for

the basic predominant (matrix) language used in the text. Primary forms constitute a scale based on how much of the secondary language that was ‘spoken’ in the fictional world is actually presented in the text, with **presence** and **elimination** of multilingualism on its two extreme ends. Secondary forms (**deformation, interference, oscillation, hybridization, construction, and simulation**) may combine with primary forms and they not only signify that the statement is made in a certain SL but also highlight some special features of the SL.³

In Mareš 2003, primary forms relate to texts where only existing languages combine but Nadsat in ACO is an invented language. Mareš considers the use of an invented language a special form of multilingualism, and treats it under one of his secondary forms, **construction** (Mareš 2003, 40). However, we want to show in this section that there is no need for a special secondary form (construction) for invented languages. We infer that an invented language can be used instead of a natural one in all the primary forms, without making any difference to the form itself. To prove this, we present, in 1.3.2, the description of the four primary forms adapted from Mareš 2003, 35–40, accompanied by our own examples, using an existing and an invented language as SL; in 1.3.3 we describe secondary forms, and focus on the combinability of secondary forms with primary forms to show that construction is set apart from the rest of the secondary forms.

1.3.2 Primary forms

1) Presence means that SL is ‘quoted’ in full in the text. The words we can read in the text are the same as the words used by the participants in the fictional world. (See Example 1.1 below.)

2) Evocation – SL is indicated by only a few expressions (words, phrases, but also distinctive syntax or spelling) embedded in PL. These expressions are usually only slightly semantically loaded, conventionalized, and they have low information value with regard to the interpretation of the depicted scene: phrases of greeting, addressing, and other means of social conversation, interjections (invectives, swearwords). Specific languages, or more precisely their literary traditions, can have a conventional set of expressions known to the general public used for characters speaking certain foreign languages. (See Example 1.1 below.)

³ Before Mareš, forms of intertextual multilingualism were studied by Goetsch 1987. Goetsch also categorizes the use of foreign languages according to the way they are presented (marked) in the text. His typology is in many respects similar to that of Mareš 2003. We can easily match Goetsch’s (unnamed) categories to Mareš’s signalization, evocation, construction, deformation and interference, and oscillation and hybridization. Two more of Goetsch’s groups are extensions of evocation and signalization according to the degree to which the secondary language permeates the text.

3) Signalization – the actual language used in the fictional world is indicated by a metalinguistic comment usually in a reporting clause or for example in a footnote. There is no SL in the text; the whole text is in PL. (See Example 1.2 below.)

4) Elimination means that in the passage of the text there is no explicit indication that SL is used in the fictional world. The reader may infer that the communication between the fictional participants does not take place in PL thanks to other information he has about them, their characteristics, their previous utterances, etc. Again, there is no SL in the text; the whole text is in PL. (See Example 1.2 below.)

To show that there is no need for a special secondary form (construction) for invented languages, we provide examples of a foreign natural language as well as examples of invented languages for all primary forms in Table 1.1. As more than one primary form may be used in one text, and can combine with more than one secondary form in it, we chose our examples from just two texts.

For the foreign secondary language, examples are taken from two dialogues in *Tortilla Flat* by John Steinbeck, where a paisano “speaks English with a paisano accent and Spanish with a paisano accent” (Steinbeck 1995, 2); for the invented one they are taken from *The Lord of the Rings* by J. R. R. Tolkien, where languages of two humanoid races, elves and orcs, are spoken. Before we present the primary forms in a tabular form, we want to explain them by commenting on some of the examples. All examples from *Tortilla Flat* are commented upon; the examples from *The Lord of the Rings* come from four unrelated passages and we comment on only one of them, because, once the forms have been explained and examples of their usage in *Tortilla Flat* given, the examples from *The Lord of the Rings* only mirror this usage. Italics in the examples is the authors’.

Example 1.1

- (a) ‘Ai, Pilon. Ai, Pablo,’ he said hazily. ‘*Que tomas?*’
(b) Pilon leaped down the bank on him. ‘*Amigo*, Jesus Maria! you are not well.’
(Steinbeck 1995, 27)

In the first line the speaker asks in SL (presence). In response, he is addressed in SL but then the answer continues in English (PL). We may suppose that Pilon’s response to drunken Jesus Maria took place in SL as a whole (evocation). However, sometimes we cannot be sure how far an evocation like this can reach: theoretically we may think of the

sentence “*Amigo*, Jesus Maria! you are not well.” as simulation of code-mixing, which is treated among secondary forms by Mareš (oscillation). In fact, these two are sometimes hard to distinguish. (Mareš 2003, 83)

Example 1.2

- (a) The policeman was saying, ‘I don’t care if I can’t understand you. You can’t sit in the gutter all day. We’ll find out about you.’
 - (b) And the boy, in Spanish with a peculiar inflection, said, ‘But, señor, I do nothing wrong. Why do you take me away?’
 - (c) The policeman saw Jesus Maria. ‘Hey, paisano,’ he called. ‘What’s this *cholo* talking about?’
 - (d) Jesus Maria stepped out and addressed the boy. ‘Can I be of service to you?’
- (Steinbeck 1995, 95)

In this dialogue, both Jesus Maria and the boy speak Spanish. The boy’s language is signaled by a metalinguistic comment (signalization). However, the fact that Jesus Maria puts his question also in Spanish is for us to infer from the situation – and perhaps from the higher register of his English, which represents Spanish in *Tortilla Flat* (elimination).

Example 1.3

‘There’s a great fighter about, one of those bloody-handed Elves, or one of the filthy *tarks*.’
(Tolkien 1993, 940)

As in Example 1.1, because of the SL word *tarks*, we may consider this whole sentence an evocation of SL (a language used by orcs) or we may claim that just one word from this secondary language was used in the midst of a speech in PL. As in example 1.1, there are reasons for both alternatives, but we will not go into them here.

Example 1.4

Frodo sat, eating, drinking, and talking with delight; but his mind was chiefly on the words spoken. He knew a little of the elf-speech and listened eagerly. Now and again he spoke to

those that served him and thanked them in their own language. They smiled at him and said laughing: ‘Here is a jewel among hobbits!’⁴

(Tolkien 1993, 96)

Up to this point all direct speech of the elves was in PL so we might think that the direct speech in this paragraph is the same. However, since Frodo listened to the elves talking in elf-speech (SL), and spoke to them in SL, we may assume the sentence ‘Here is a jewel among hobbits!’ to be spoken in the fictional world in SL as well, even though it is in PL in the text (elimination).

Table 1.1 puts side by side the forms of the foreign SL and the invented SL to show that there is no real difference in them. The first column lists the four primary forms; the second column provides a concise description of the particular form. The third and fourth column provide the examples from *Tortilla Flat* and *The Lord of the Rings* – italics in the examples is the authors’; the passage on which the assignment of the example to the corresponding form is based is in bold.

⁴ The word *hobbits* is considered an English word as it is attested in OED, where also derived words *hobbitish*, *hobbitomane*, and *hobbitry* are included.

Table 1.1 – Primary Forms: Intratextual Multilingualism & Invented Languages

	Form	Short description	Example of a foreign language (<i>Tortilla Flat</i>)	Example of an invented language (<i>The Lord of the Rings</i>)
1	Presence	SL is ‘quoted’ in full	‘Ai, Pilon. Ai, Pablo,’ he said hazily. ‘ <i>Que tomas?</i> ’	‘ <i>Elen síla lúmenn’ omentielvo</i> , a star shines on the hour of our meeting,’ he added in the high-elven speech. (Tolkien 1993, 94)
2	Evocation	few, conventionalized, expressions in SL	Pilon leaped down the bank on him. ‘ <i>Amigo</i> , Jesus Maria! you are not well.’	‘There’s a great fighter about, one of those bloody-handed Elves, or one of the filthy <i>tarks</i> .’
3	Signalization	SL identified by a metalinguistic comment	And the boy, in Spanish with a peculiar inflection , said, ‘But, señor, I do nothing wrong. Why do you take me away?’	‘Why has nothing of this been told to me before?’ he asked in the Elven-tongue . (Tolkien 1993, 374)
4	Elimination	no explicit indication of SL	Jesus Maria stepped out and addressed the boy. ‘ Can I be of service to you? ’	Now and again he spoke to those that served him and thanked them in their own language. They smiled at him and said laughing: ‘ Here is a jewel among hobbits! ’

1.3.3 Secondary forms

To the four primary forms Mareš adds six more secondary forms: **deformation, interference, oscillation, hybridization, construction, and simulation**. (Mareš 2003, 37–40) These secondary forms may combine with primary form, and they not only signify that the statement is made in a certain SL but also highlight some special features of the SL.

Deformation – is an incorrect use of language, deviating from the standard language (norm) using deformed linguistic means. The speaker of such language shows that he does not identify with the language he is speaking at present; the language is foreign to him usually to the point he is able to speak it only imperfectly.

Interference – this is in fact a special version of deformation in which a language is deformed by covertly incorporating linguistic means and structures of another language

(word for word ‘translation’ of idioms, text-construction according to grammatical principles of the other language, etc.), usually for comic effect.

Oscillation – the change in the language is presented in a way that simulates code-mixing in real communication. The speaker formulates his thoughts in one language or another freely according to the (fictional) communication needs. As an example we quote from a chapter of *Women in Love* that takes place in Austria: “Can we have Kaffee mit Kranzkuchen?” she added, to the waiter. (Lawrence 2000, 403) “Will you schuhplattern, gnädige Frau?” said the large, fair youth, Loerke’s companion. (Lawrence 2000, 411)

Hybridization – lexical units from SL are incorporated, usually in a modified form, in morphological and/or syntactic structures of PL. Hybridization often accompanies oscillation and indicates a closer interconnection between the languages.

Construction – SL is not an existing language but a made up one, created ad hoc.

Simulation – SL is simulated by imitating some of its sound or other qualities in order to give an impression that this language is used.

We can see from this outline that secondary forms are an assortment of various types of forms without any common denominator, some being closer to one another than others. We can show their heterogeneity by drawing our attention to how they can combine with primary forms. All secondary forms can combine with primary forms even though this is not explicitly stated by Mareš and we infer it from the definitions of the forms and Mareš’s examples.

Mareš shows that **construction** can combine with evocation (Mareš 2003, 53), and is aware of the fact that **deformation** can combine with those three basic forms that indicate what language is used (presence, evocation and signalization). (Mareš 2003, 38) By extension, this also holds true for **interference**, it being a special type of deformation. On the other hand, **simulation** requires a signifier that is able to simulate SL plausibly, and thus simulation can manifest itself only as evocation or presence.

The relationship between the two remaining secondary forms (**oscillation and hybridization**) is somewhat similar to that of deformation and interference, hybridization being the more specific, ‘merged’, form. Oscillation and hybridization are characterized by the fact that the manifested shifting between PL and SL is a real record of the form (real in the fictional world) used in fictional world. (Mareš 2003, 39) Because of this they combine only with presence. Presence also shows the ‘true’ form of an utterance, but the utterance

itself is monolingual, whereas oscillation is based on code-switching that takes place within an utterance.

In 1.3.2 we showed that an invented language can be used with all primary forms; in other words, construction can combine with all primary forms. From the observations above we deduce the combinability of primary and secondary forms as presented in Table 1.2. We can see that construction is the only secondary form that can combine with any primary form. This confirms our earlier decision to not treat it as a stand-alone secondary form but to treat invented languages as a special type of intertextual multilingualism.

Table 1.2 – Combinability of Primary and Secondary Forms

Secondary Form	Primary forms the secondary form combines with
Oscillation	Presence
Hybridization	Presence
Simulation	Presence, Evocation
Deformation	Presence, Evocation, Signalization
Interference	Presence, Evocation, Signalization
Construction	Presence, Evocation, Signalization, Elimination

1.3.4 Forms of Nadsat in ACO

Having established that invented languages take on the same primary forms as existing foreign languages, we look into which primary and secondary forms are adopted in ACO. The primary form of the SL in ACO is easy to identify as presence, because other primary forms require some kind of indication that SL is used, as shown in 1.3.2. There is no evidence of this kind in ACO.

To decide which secondary forms are adopted in ACO, it is useful to comment on the possible types of readers of ACO from the point of view of understanding Nadsat, so that we can distinguish between two implied readers:

1) A reader that knows Russian to the extent that he is able to uncover the Russian words used for Nadsat, and thus understand Nadsat through the connections between the meanings of the original Russian words (*орыг*) and word forms of Nadsat (*droog*). Being able to reveal the meaning of Nadsat words when he first encounters them, his reading experience approximates to the experience of reading monolingual texts. As a result, he is partially

deprived of the delight of gradually revealing the meaning of ACO and of the effects the unknown would have on him. On the other hand, his way of reading enables him to take pleasure in the combination of the languages, modifications of meanings and forms, and from discovering the method Nadsat was created with. However, Nadsat would not be so uncommon to him, the words would not stand out from the text as they do for the reader of the second type. Since he is able to realize that two languages are used in ACO (one of them in a highly modified form), the form of multilingualism used in ACO is, for him, **hybridization**. We believe that this type of implied reader was not the ideal reader Burgess had in mind when he introduced Nadsat into ACO. For example, in an interview conducted between 1971–1972, Burgess responded to the question whether he was writing “for a limited, highly educated audience” like this:

“Where would Shakespeare have got if he had thought only of a specialized audience? What he did was to attempt to appeal on all levels, with something for the most rarefied intellectuals (who had read Montaigne) and very much more for those who could appreciate only sex and blood. I like to devise a plot that can have a moderately wide appeal. [...]” (Burgess 1977, 326–327)

Burgess used Russian because he wanted his secondary language to be atemporal, freed from connotations of existing slangs and argots, but at the same time retaining the sense of secrecy and incomprehensibility of the words that is typical of argots. There are many reasons why he chose Russian but we will comment on the process of creating Nadsat only later. In our further analysis we will not use the concept of hybridization but focus on the way of reading of the second type of reader instead. On the intellectual games with meaning, puns and allusions that go across languages (with respect to ludic theory) that the first type of reader can delve into, see Coale 1981 and Oks 2009.

2) A reader who does not understand Russian. For him, the used form would be **oscillation**, English being the primary language and an unknown invented language the secondary one. He would be able to extract some meaning from the English inflexional endings that combine with unknown words or their syntactic position and collocates, but the text itself would be a challenge for him, and he would follow the path set for him by the author to decipher the meaning of the words, relying on context and the aid provided by the author. Nadsat words would be much more important for him; they would stand out of the text much more.

We have to mention at least briefly here that both readers are able to make another connection between Nadsat and English beyond the level of inflections or syntax. Some Nadsat words are similar to English words but have a completely different meaning (in real communication these would create false friends). There are more words like this in ACO than one might think. Some of Nadsat words match the form of English words completely (*rabbit, tree, starry*) some need minor adaptation (Nadsat *Bog, millicent* but English *bog, Millicent*). The reader of the second type could be confused by these meanings, while for the first reader this encourages the play with meanings even further.

To sum up, for reader 1, the form of multilingualism is **presence** and **hybridization**. The two languages involved are English and Russian (modified), and together they create Nadsat.

For reader 2, the form of multilingualism is **presence** and **oscillation**. The two languages involved are English and an invented language. In our interpretation that is used in this paper, Nadsat refers to this invented language that permeates English. (However, it is also possible that, at least in the fictional world of ACO, Nadsat is the name of the result of this combination of English and an invented language.)

1.4 Invented languages

1.4.1 Definition

So far, in trying to show that the concept of intratextual multilingualism and its forms can be also applied to invented languages, we have used the term “invented language” without defining it. It stands in fact for an “invented language used in fiction”, since new languages are created for vastly diverse reasons and purposes, and by various processes. A new language can be a sacred one (a language that emerges through mystic revelations, glossolalia, or practicing shamanism). In social communication, new languages arise as pidgins, programming languages, codes for special purposes (such as Air Traffic Control English or military codes) or international auxiliary languages (such as Esperanto). Another significant group are the outcomes of efforts to ‘find’ a divine language. Last but not least, there are languages created by art (and for art’s sake, not limited only to literature), which can be carefully structured, like Tolkien’s languages, or symbolic, like zaum. Okrent lists 500 new languages. (Okrent 2009, 298–314) The most thorough attempt to categorize all of the types mentioned above has been made Paolo Albani and Berlinghiero Buonarroti. Their categorization represented by a tree structure has up to seven layers of branching, but does

not provide us with a simple term for invented languages used in fiction. (Albani and Buonarroti 2010)

Since a large number of invented languages used in fiction are found in ‘lower genres’ of literature, especially in sci-fi and fantasy, texts on these languages usually do not come from academic sources and the terminology in this area varies greatly, depending also on the nationality of the author.

Bartlett classifies artificial language into several groups, one of which consists of languages created “for artistic use or to be part of a fictional or mythic world.” (Bartlett 2006, 488) Stockwell provides us with a definition of languages invented for artistic purposes: “invented languages that appear in literary works, as well as invented languages that appear in their own supporting worlds, and invented languages existing only for their own sake.” (Stockwell 2006, 6) However, his term “artlang” is also used in different meanings elsewhere. Bartlett’s and Stockwell’s definitions are close to our understanding of what we call “invented language used in fiction”, or, in short, invented language. However, sometimes invented languages used in fiction merge with auxiliary languages (Crystal 2010, 363), and the terms “imaginary” (Nicholls 1993, 723), “created” (Cheyne 2008, 386), “artificial” (Large 1994, 237), “fictitious” (Eco 1997, 3) and “fictional” (Okrent 2009, 284) are used for the same content. Because of this, we keep using the term “invented language” as used by Forster 1970, meaning invented for artistic use or invented for artistic purposes, even though such a definition is also problematic.

1.4.2 A short history of invented languages in fiction

According to some accounts, the first attested appearance of an invented language in fiction is the play *The Acharnians* (425 BC) by Aristophanes. (Albani and Buonarroti 2010, 373; Havliš 2009, 3) In *The Acharnians*, a minor character, Pseudo-Artabas, is trying to impersonate a Persian delegate, uttering: “Iarta name xarxana pisona satra,” (ἰαρτα νουμε ξαρξανα πισωνα σατρα in Greek) (Aristophanes 1998, 69). According to Mareš’s classification, the primary form of this sentence would in fact be **presence**, and the secondary form **simulation** and not construction, because an existing language (Persian) is imitated by the sounds, and some meaning has been ascribed to it: “the king *Artaxerxes*, and the *satrap Pissuthnes*” (Aristophanes 1998, 69). However, for us it is interesting that this sentence in an invented language has already more than one possible ‘translation’. For example, in an earlier English translation by Benjamin Bickley Rogers we find the same sentence as: “Ijisti

boutti furbiss upde rotti.” (Aristophanes 1910, 19) Our ultimate topic, the problem of translating invented languages, could be as old as invented languages themselves.

Fortunately, we may still think of Aristophanes as the first one to employ invented languages in Western literature, because he also used an invented language in his later play, *The Frogs* (405 BC). A chorus of frogs speaks in the primary language but also uses the sentence “Brekekekex koax koax!” (βρεκεκεκεξ κοαξ κοαξ in Greek) (Aristophanes 2002, 53), which is repeated and modified by the character Dionysos: “Yes, all you are is koax.” “Blast you, and your koax too.” etc. (Aristophanes 2002, 53) These sounds, while onomatopoeic to a large extent, can be seen as an invented language belonging to the languages of animals, which are later largely used in fiction. Again, these lines are translated differently by Benjamin Bickley Rogers: “Brekekekex, ko-ax ko-ax!” (Aristophanes 1989, 317), “Hang you, and your ko-axing too!”, and “There’s nothing but ko-ax with you.” (Aristophanes 1989, 319)

In the next almost two thousand years examples are scarce, but later invented languages in fiction have flourished along with imaginary voyages and utopias, especially since the 16th century.

Bakhtin mentions specific invented languages in his *Rabelais and Folk Culture of the Middle Ages and Renaissance*. Apart from the frequent use of nonexistent languages in medieval mystery plays for a seemingly primitive comic effect he points out the language “grimoire” from the 15th century farce *The Farce of Master Pierre Pathelin*, where a lawyer pretends to be ill by talking in six dialects, macaronic Latin and grimoire. (Bakhtin 2007, 444) Rabelais also used unnamed invented languages in the first dialogue between Panurge and Pantagruel, where Panurge answers in thirteen languages, three of which are invented and based on various existing languages. (Ouředník 2002, 108–112) A selection of some of the more famous works with invented languages is provided by Havliš 2009. According to Yaguello, the transformation of the imaginary voyage genre at the end of the 19th century marks the decline of invented languages in fiction. In the emerging dystopias, invented languages gain new functions. However, due to the simultaneous expansion of the international artificial languages, they only slowly infiltrated the emerging science fiction genre. (Yaguello 1991, 31–45)

1.4.3 Functions of invented languages

The functions or the effects of invented languages were not studied in detail, especially when compared to other topics concerning invented languages, such as how they were created in the first place. Some authors touch upon the subject in passing. For example, according to Yaguello, the chief aim of an invented language in Rabelais and Swift is “to amuse the reader” (Yaguello 1991, 31–32). This is of course a valid and existing function, named “comical” by Mareš,⁵ but it is neither exhausting nor the most important characterization of functions of the invented languages used by these two authors. Through some of his invented languages, Rabelais also shows the reader his reflections on the major changes in the Renaissance: the conflict between unity and plurality, the perceived nearing destruction of culture and cultural language (atmospheric function), and perhaps also the optimistic belief that these changes will not prove fatal (value function). (Ouředník 2002) Similarly, while we may agree that Václav Havel’s invented languages (ptydepe and chorukor) in *The Memorandum* have a satirical effect (Bausani 1970, 38–44), the incomprehensibility of the memo in ptydepe not only expresses both functions mentioned above (atmospheric, value) by exposing the absurdity of the state of society and its values, but it also creates a prime mover of the story (story-developing function). In her article on invented languages in fiction, Cheyne is interested in how invented languages “function within sf narratives” (Cheyne 2008, 390), but she focuses primarily on the levels at which invented languages communicate meaning, and what they reveal about their users – characters of the text. Vo and Carter feel that “such aspects as the functions of creative language, its effects on the reader/listener [...] need relevant experiments” (Vo and Carter 2010, 305) but as we are not aware of any systematic account of functions of invented languages, we will adopt Mareš’s functions of intratextual multilingualism.

Throughout history, foreign languages have been used in literature for various purposes. A summary of these purposes can be found in Elwert 1973, Goetsch 1987, and Mareš 2003. Even though Mareš (2003, 40–41) considers his classification provisional and covering only the most frequent types, it is the most developed classification we have found. Because we treat invented languages as a subtype of intratextual multilingualism, and Mareš’s functions of intratextual multilingualism are suitable for the description of functions of Nadsat in ACO, we highlight the functions of Nadsat straight away in this outline: the functions of Nadsat, which are established in Chapter 2, have already been put in bold in the

⁵ For the overview of Mareš’s functions, see below.

following list for future reference. Also we believe that all of these functions can be applied, under more or less special circumstances, to describe the functions of invented languages in general as well.

1) indicative (indiciální)

SL simply indicates that its user uses that particular language, or that the language is used in that particular setting or milieu.

2) nationality-determining (národnostně zařazovací)

Similarly, SL indicates that the user belongs to a particular nation or is the native speaker of the language.

3) documentary (dokumentační)

SL evidences the authentic wording of the statement. (For example in a footnote; this is rare in fiction.)

4) **characterizing, individually or collectively (charakterizační individuálně a skupinově)**

SL defines the user as an individual, according to his individual qualities (i.e. according to educational and professional background, communication skills and customs), as well as a member of a group or community.

5) subjectivizing (subjektivizační)

SL illustrates subjective perception and feelings, the recalling of past events, etc.

6) **atmospheric (atmosférotvorná)**

SL helps to create the big picture of the specifically national or cultural setting, emphasizing its correspondence with the experience and concepts of the reader, or its strangeness, outlandishness.

7) **value (hodnotová)**

SL represents a distinct set of values; it takes part in the polarization between what is perceived as positive and negative.

8) story-developing (dějotvorná)

The fact that SL is used may play an important role in the storyline, especially if some misunderstanding between the characters of the story is based on it.

9) **cryptic and restrictive (utajovací a restriční)**

SL blocks information completely or provides it only to those readers with relevant knowledge. On the other hand, a cryptic and incomprehensible SL stimulates the reader to search for meaning, to speculate about it, and to come up with various associations.

10) comical (komikotvorná)

The comical effect is achieved especially by the oddity and grotesqueness of expressions, deformations of language, a discrepancy between the language used and the communicative situation presented in the story, and language misunderstandings and their impact.

11) expressive (výrazová)

Graphic and sound characteristics of SL may create relationships between signifiers in the PL and SL; they may be similar, the same or contrasting, and they may seem familiar or uncommon and bizarre.

12) creative (kreativní)

Multilingualism may also help the reader to think about possible forms of language, and about the determinants and boundaries of verbal expressions and communication.

13) cultural (kulturní)

In a broader sense, language(s) used in a text integrate the text (by means of quotations, allusions, figures of speech, etc.) into a cultural and historical tradition, and also project this tradition onto the semantic structure of the text.

These functions may of course combine in one work. We have no space to provide examples of all of these functions and comment on them. However, in the next chapter we will finally speak about Nadsat and its functions in ACO, and the functions from the above list which can be ascribed to Nadsat will be identified and discussed there.

2 A Clockwork Orange and the functions of Nadsat

2.1 Introduction

In this chapter we will look at Anthony Burgess's (1917–1993) *A Clockwork Orange* and Nadsat. We will seek answers to questions which are important for understanding the character of Nadsat and its functions in ACO.

When forms of multilingualism in ACO were discussed in 1.3, we needed to distinguish between two types of reader (1.3.4). Similarly, to avoid confusion between different types of communication in fiction, we need to distinguish between two levels of communication in ACO. (Other levels, and their mixing, are of only marginal importance in ACO.) First, the narrator (implied author) communicates with the implied reader. This type of communication we call primary communication.

Second, characters communicate among themselves within the fictional world. This type of communication we call secondary. If we focus on invented language or more broadly on any instance of multilingualism, one piece of text can serve different functions in primary communication and secondary communication. For example, the intended function of the speech of Pseudo-Artabas from *The Acharnians*, “Iarta name xarxana pisona satra” (Aristophanes 1998, 69), which we have already used in 1.4.2, is to convince other characters in the play that Pseudo-Artabas is a genuine delegate from Persia. Referring to Mareš's overview in 1.4.3, this function can be described as the nationality-determining function (2) within the fictional world. (The fact that Pseudo-Artabas fails in his function, is another matter.) On the other hand, the intended function in primary communication is comical – the play mocks the politics of Athens, and this sentence is part of the satire.

Similarly, some characters in ACO – the most obvious example being Pete's wife, introduced at the end of the book (Burgess 1998, 145–146) – laugh at Alex's Nadsat but for the reader there is nothing comical about it (of course only if he does not find the whole concept of Nadsat comical in itself), and Nadsat retains its typical primary communication functions. We will focus on how Nadsat functions in primary communication. How Nadsat functions within secondary communication is discussed in Davis and Womack 2002 and Kohn 2008.

ACO was first published in 1962 in Great Britain and early in 1963 in the US; Burgess was 45 at that time. He had had his first book published only 6 years before that, in 1956, and ACO was his tenth published book. In the US, ACO was published without the last (21st) chapter, and it was soon supplemented with a Nadsat glossary. Burgess returned to ACO several times, especially after its adaptation was filmed by Kubrick in 1971. He wrote scripts for film adaptations of ACO both before and after Kubrick's film. He adapted ACO for stage in 1986 and 1990. He commented on it extensively, and his foreword to the first US edition to include the 21st chapter is often reprinted. His 1974 novel, *The Clockwork Testament, or Enderby's End*, deals with some of the topics of ACO and Burgess's experience with its film adaptation.

In a brief anonymous foreword to the first edition of ACO from 1962, we read that the story takes place in an "unspecified but not very distant future" and Alex tells it "in the peculiar slang of his generation". Even though there is no Nadsat glossary, "[i]t will take the reader no more than fifteen pages to master and revel in the expressive language of 'nadsat'; after that he has before him an easily digestible feast of picaresque villainy and social satire". (Burgess 1962, no page number) This foreword aptly summarizes the main points of our interest in this chapter.

We will be interested in the setting of ACO and the role of Nadsat in it, the motives for using Nadsat and its effects, and finally in delimiting the functions of Nadsat in ACO. However, to investigate the reception of ACO extensively, or make any claims about the meaning of ACO, is beyond the scope of this paper.

This chapter uses a number of sources which can be divided into two groups: Burgess's own published statements, and materials published on Burgess. However, reviewers, and authors of forewords, afterwords and biographies rely too much on Burgess's own accounts of his life. They also do not hesitate to present other people's hypotheses as (unascrbed) facts. Moreover, Burgess often changed his descriptions of certain events in his life, as we can see from his published sources. More importantly, Biswell, an authority on Burgess, uses unpublished documents to cast doubts on some of the questions directly related to ACO, especially Burgess's changing attitude to the 'correct ending' of ACO, the 21st chapter. (Biswell 2006) Because of these caveats, we have mainly used sources from the early stages of the ACO's reception and the latest sources represented by Biswell in the first place. We have tried to adopt a balanced approach to these sources, which may perhaps seem like sidelining Burgess's own voice, and preferring Biswell's interpretation of events.

2.2 Structure of ACO

ACO is divided into three parts, each comprising seven chapters. The first part describes a day (and two nights) in the life of a fifteen-year old hoodlum Alex, full of unprovoked acts of violence; it is narrated in the first person by this anti-hero, and it ends with Alex's arrest for murder. Part two describes his two years of imprisonment but focuses mostly on its last stage, where he undergoes experimental aversion therapy, the aim of which is to make him physically sick just thinking about violence in any form. Part three describes the misfortunes of Alex as a defenceless victim of others, mostly those he had wronged in part one of the book, and his attempted suicide followed by another stay in the hospital, ending with a reversal of the previous conditioning; its last chapter mirrors the first chapter of the book except that at the end, the now eighteen-year old Alex does no longer find violence so great a pastime, and he abandons it of his own accord, realizing slowly that he had grown up and grown out of it.⁶

The last chapter was not included in US editions until 1988, and some British editions dropped it as well. In this chapter, Alex feels he needs something different, stable, and peaceful: a wife and perhaps a son, although the son would probably go through everything Alex had gone through, and Alex would not be able to make him understand. This of course changes the 'message' of ACO – its cycle is no longer from violence to treatment and back to violence, but it is topped with Alex's understanding of his actions, and of the fact that the cycle will repeat itself in the next generation.

Burgess later claimed that the tripartite structure was very important, 21 being the number that signifies maturity, and omitting the last chapter "ruined the arithmology of the book". (Burgess 1987, vi) Morrison sees in it an allusion to the seven ages of man from Shakespeare's *As you Like It* (Morrison 1996, xx) and some critics relate the structure of the book to specific musical pieces (Ray 1981, 482; Rabinowitz 2003); the relationship between music and writing in Burgess's novels was commented upon early, not least by Burgess himself. However, in his analysis of Burgess's correspondence and the typescript of ACO submitted to his publisher Heinemann, Biswell proves that Burgess fabricated his stance only afterwards, and in fact held various opinions on the endings during his lifetime (Biswell 2006, 247–256), as can be seen for example in his 1971/1972 interview (Burgess 1977, 337–339). In the typescript, "Should we end here" is written in Burgess's hand at the end of chapter 6, part three, and an "epilogue" follows. (Biswell 2006, 251–252) The earliest

⁶ Further examples of mirroring in ACO can be found in 2.7.

surviving outline for ACO also ends with what later became chapter six of ACO. (Biswell 2012a, viii) This suggests that both editions – with or without chapter 21 – may be ‘correct’, and that the structure of ACO (and its “arithmology”) are not so stable. For our analysis we will use the edition with chapter 21 included.

2.3 Setting of ACO

Since we are ultimately interested in the possibilities and limits of translating ACO into other languages, the identification of its setting is important for an equivalent transfer of values that are attached to it.

In ACO it is not specified where and when its story takes place. However, it offers hints and allusions that were interpreted by critics and commentators in various ways. The author of the frequently reprinted *Afterword* to a 1963 US edition of ACO, Stanley Edgar Hyman, placed it in a “future England”. (Hyman 1963b, 178) His view is often repeated in less scholarly works. Some set it in even more precise locations such as a “future London”. (Jakubowski and Clute 1993, 175) Burgess commented in retrospect on the setting especially in his autobiography *You’ve had your time*, claiming that first he wanted to set his novel about youth violence and rioting in the 16th century but finally he set it in “a near future – 1970, say –” (Burgess 1991, 26). In the earliest surviving outline for ACO, Burgess set it in 1980. (Biswell 2012a, viii)

In his twelve-page *Introduction* to a 1996 British edition of ACO, the poet and literary critic Blake Morrison imagines it taking place “somewhere in Europe, circa 1972” (Morrison 1996, ix), but this seems to be a too-literal (mis)interpretation of another Burgess’s comment from *You’ve had your time*. There Burgess recalls what he thought about Kubrick’s film in 1971, which could have revived the book “set in a vague future which was already probably past”. (Burgess 1991, 244–245) However, it was past only with regard to the achievements of mankind Burgess predicted in ACO (e.g. people landing on the Moon, the development of satellite broadcasting), not because some particular date has passed. Moreover, 1960 is directly referred to in ACO, pushing the setting to a future more distant than 1970’: an old man, who is beaten up by Alex and his friends, carries with him love letters dated “right back to 1960” (Burgess 1998, 10). We may deduce his age from the fact that he wears dentures, and the group of people in the Public Library he belongs to are in their nineties (Burgess 1998, 114) – with the caveat that their age might be just an inaccurate guess by young Alex.

Together with Burgess himself and others (Dix 1971, 14; Biswell 2006, 241–242), we believe Burgess strived for an unspecified time and place, and avoided references indicating to the reader that the violent scenes and the methods to fight the violence may surface only ‘somewhere else’: that they are not universal but limited to a certain place. In 1971, Dix saw the society of ACO as being subjected to “both American and Russian intervention if not invasion”. (Dix 1971, 14) Burgess’s vision is not of a communist England but of a society in which the then superpowers, the US and the USSR, are indistinguishable in essence. In his interpretation, Mareš regards the world of ACO as divided in a bipolar manner, with the languages of the superpowers merging. (Mareš 2003, 52)

2.4 The setting as reflected in Nadsat

Indeed, the language of ACO is interspersed not only with Nadsat but with Americanisms. In April 1961, Burgess wrote in a letter that he was “fabricating with difficulty a teenage dialect compounded equally of American and Russian roots.” (Biswell 2006, 256) Biswell does not find Americanisms in the first 1961 typescript of ACO except the filler *like* (Biswell 2006, 257) but they were kept at least in Nadsat, where some Russian words had been used to replace their counterparts in what was considered American slang in 1962: *ptitsa* standing for *chick* or *kopat* for *to dig* – to enjoy or to understand (Hyman 1963b, 180). Later Burgess described Nadsat as being “compounded of the two most powerful political languages in the world – Anglo-American and Russian”, thus making the narrator “unpolitical”. (Burgess 1987, vi)

A minority of Nadsat words did not come from Russian; Biswell identifies their sources as Romany, the Lancashire dialect, and Cockney rhyming slang. (Biswell 2006, 237) Critics differ in their descriptions of the etymology of some Nadsat words but their origin is not so important for our purpose. What is important is that Nadsat further expands on the blending of the superpowers: a phenomenon Burgess returned to in his next novel, *Honey for the Bears*, which is based on his experience from a trip to Leningrad he and his wife took in 1961. Burgess interprets his own symbolism of *Honey for the Bears*: “America and Russia are the same place, and [...] the alignment of the future has nothing to do with the opposition of the superpowers, rather with their identification.” (Burgess 1991, 58) Oks shares this concept of an “interosculation of the West and the East” supported by Nadsat. (Oks 2009, 41) However, holding the view that ACO takes place in England, she believes that Nadsat contributes to the detachment of the reader from the story to such extent that it seems to take place in “some other world and is irrelevant for the reader.” (Oks 2009, 49–59)

2.5 'External history' of Nadsat

In 2.3 we mentioned that Burgess first contemplated a historical setting for his novel about youth violence. This was not realized, and Burgess states that he wrote the first version of ACO – in a modern setting but less fantastical than the published version – in late 1960, using the then slang of Teddyboys, and Mods and Rockers. However, he was afraid that the slang would soon be outdated. (Burgess 1987, v) These fears do not seem unfounded: according to Burgess, reviewers of his novel *The Doctor is Sick* – first published 21 November 1960 – criticised precisely his dated usage of dialects and slang. (Burgess 1991, 29) For Burgess, literature was meant to be permanent. (Biswell 2006, 195) In 1988 Burgess claimed that in 1960 he had realized that a future slang would have had to be invented and so put the draft aside. (Burgess 1991, 27)

This seems to contradict the idea we presented earlier in this chapter: that the connection between Russia and America was set from the beginning. We are not able to reconstruct how (and if) the idea of a future novel about violence transformed in Burgess's mind before it took some documented shape. However, it is probable that the connection between Russia and America entered the novel hand in hand with employing Russian, whenever it actually was.

Biswell begins his account of the history of ACO early in 1961, when Burgess was planning a holiday in Leningrad. Burgess was learning – or relearning (Burgess 1991, 37) – Russian when he realized that it would be possible to use Russian instead of a slang in his next novel. He then compiled a vocabulary of about 200 Russian words for his use in ACO, and the first 7 chapters (part one) of ACO were ready early in 1961 before he went to Russia. (Biswell 2006, 237–238)

Burgess used invented languages in his previous novel *Devil of a State*. He invented another language for the 1981 film adaptation of *Quest of Fire*. (Jackson 1999, 25) By intratextual multilingualism he coded obscene meanings into every volume of his *The Long Day Wanes: A Malayan Trilogy*. (Biswell 2006, 193–194) He loved linguistic games and puns, and wrote several books on linguistics. The authors he admired most were Shakespeare and Joyce; on Joyce, he wrote two books and an abridged version of *Finnegans Wake*. It is therefore only logical that in ACO “Burgess has not used Russian words mechanically, but with great ingenuity.” (Hyman 1963b, 180) Also the sound of Nadsat words was important to him, as was combining several possible meanings in one word.

Burgess thought Nadsat over for a comparatively long time. The typescript of ACO, which Burgess sent to a Heinemann editor, James Mitchie, shows his doubts about Nadsat. In the typescript, some Nadsat words are provided with a comment or are simply revised. Biswell summarizes that the tendency of the changes “seems to be towards a thickening or enriching Nadsat at the expense of Standard English”. (Biswell 2006, 250) A detailed comparison of the typescript and the published work would reveal much about the way Burgess wanted Nadsat to function in ACO.

Burgess did not have doubts like this about his other books (or at least other typescripts do not show them); Biswell infers from the typescript and correspondence that Burgess’s main concern was whether Nadsat would not discourage readers. It was not only the sheer number of Nadsat words but also the pace at which the reader was exposed to them. (Biswell 2006, 248–258) The fact that Burgess took great pains to create something more permanent, withstanding the ravages of time, and the pace in which Nadsat words are introduced into the text, explained to the reader, and used later in ACO at specific places, according to the various communicative situations in the world of ACO (Kohn 2008, 5–16), show that Nadsat became (or had been) an important – if not the essential – quality of the text.

2.6 Reception of Nadsat

How did readers and critics respond to the challenge of Nadsat? Burgess was disappointed by reviews from 1962 calling the book a “nasty little shocker” and Nadsat “a silly little joke that didn’t come off” (Burgess 1987, vii–viii), “out-of-this-world gibberish,” and “a great strain to read” (Burgess 1991, 59). However, Biswell warns again that there were also favourable reviews, and quotes from one by the influential critic and author Julian Mitchell, who considered Nadsat “an extraordinary technical feat.” (Biswell 2006, 259) Similarly, Hyman called it a “remarkable invention” and “fascinating thing”. (Hyman 1963a, 23) In any case, mainstream mixed reviews were surpassed by favourable responses from the counterculture (Biswell 2012a, xv–xvi), and in 1971, Dix already compares the popularity of ACO among American students to Tolkien’s and Huxley’s novels (Dix 1971, 5). Today, Nadsat can be described even as “vividly expressive and easily comprehensible” (Mann 2001, 85); if this were a reflection of the prevalent contemporary opinion, Burgess might have been happy that his experiment had come off.

Opinions also differ on how complicated the reading is. One part of the picture is clear from the quotations above. Nevertheless, even those who approved of ACO and Nadsat

admit that it is a demanding (yet fruitful) experience. The context and the way Nadsat words are presented play a crucial role in deciphering (and submersing into) Nadsat. (Morrison 1996, vii–xx) Nadsat seems incomprehensible to Hyman at first but “Alex translates some of it” and “Then the reader, even if he knows no Russian, discovers that some of the meaning is clear from context” and “other words are intelligible after a second context”. (Hyman 1963b, 182) Nevertheless, as early as 7 January 1963, Hyman reviewed the first US edition (published without the last chapter) and even though he praises Nadsat, he “could not read the book without compiling a glossary.” (Hyman 1963a, 23) His attempt in the review to decipher the meaning and origins of Nadsat words is unfortunate, but when his review was turned into an afterword to a later edition, his – much improved but completely unauthorized – glossary was included. Goetsch considers Nadsat accessible for readers without the glossary (Goetsch 1987, 52), and finds it natural that in some books the reader is expected to learn at least some of the words from an invented language that were already translated or otherwise explained before. (Goetsch 1987, 64)

Burgess disliked passivity, including the passivity of the reader. He wanted the reader to participate in an intellectual, linguistic game, at least to some extent. (Oks 2009, 44) One of the possible interpretations of the motives behind Alex’s violence is his response to a state of inertia, a world without feelings and without the need for self-expression. Violence is presented from the point of view of the first person narrator who is depicted as a likeable anti-hero, and Nadsat is the language the narrator and the reader share. Together with other means the narrator uses to connect with the reader, Nadsat contributes to “establish an intimate relationship” (McDougal 2003, 9–10) between them.

Burgess imagined that the reader would follow the narrator, discovering Nadsat as the story progresses, “learning minimal Russian” in the process; a glossary he called “stupid” (Burgess 1991, 257) and he considered it disruptive to the process (Burgess 1991, 38). Surprisingly, there is evidence that readers learn Nadsat words while reading the novel. Two studies used Nadsat in ACO to test vocabulary acquisition hypotheses. According to a 1978 study by Saragi, Nation and Meister mentioned by Krashen, “subjects had picked up at least forty-five words simply by reading a novel.” (Krashen 1989, 446) This test was partially replicated, and showed gains in vocabulary comparable to the results of studies that used natural languages. (Pitts, White and Krashen 1989)

The recurrence of Nadsat words in specific places is crucial for understanding Nadsat, and if the system is disrupted, so is the reading experience. Burgess repeated many times that he intended Nadsat to serve as a curtain, a mist, suppressing the violence of the book, making

it more symbolic, and “protecting the reader from his own baser instincts.” (Burgess 1991, 38) And indeed, readers find that this stylized Nadsat curbs the immediacy of certain scenes in ACO. (Žilina 2006, 13; McDougal 2003, 10–11)

2.7 Mirroring in ACO

Apart from the fact that each of the 3 parts has 7 chapters, there are other instances of repetition, mirroring, and analogy. We will list only the most obvious ones, because in most cases they are not related to Nadsat. We want to show that repetition is important not only in Nadsat but in the whole text, and this fact, reciprocally, makes the recurrence of Nadsat words the more important.

1) A couple of paragraphs at the beginning of the last chapter directly mirror the beginning of the first chapter; Alex’s behaviour in the episodes (drinking in the Korova Milkbar, beating an old man, and treating old women in the Duke of New York) is altered because he has grown up.

2) The introductory sentence ““What’s it going to be then, eh”” recurs at the very beginning of each of the three parts of ACO, and it is used fourteen times in the whole book (twelve times in the US edition).

3) The story of part three duplicates or parallels three episodes described in part one but Alex’s former victims turn into victimizers: Alex meets the old man they had beaten up, his former friend and betrayer Dim together with their common enemy of that time, and the writer F. Alexander, who is writing *A Clockwork Orange*. (On F. Alexander and his relation to Alex, see Ray 1981, and for the Joycean inspiration of this relationship, see Ingersoll 1986.)

4) After the aversion therapy, another treatment is used to reverse the effects of the first one.

5) Andy, the shopkeeper who was on friendly terms with Alex, is replaced by a younger one with no sense of classical music; Alex’s parents more or less replace Alex himself with a caring lodger.

There are also more sophisticated forms of recurrence of certain specific words in specific contexts but these are beyond the scope of this paper. For example, apart from Nadsat, which is used also by his friends, Alex speaks in his own idiolect consisting of several diverse linguistic means, such as the excessive usage of *like* or ‘old-fashioned’ English (labelled sometimes mock-Elizabethan (Davis and Womack 2002, 26) but without an accompanying linguistic analysis). His idiolect serves various functions in the novel (for an

interpretation of Alex's *like* see Davis and Womack 2002) but it is his pronominal code within the idiolect that is most related to repetition. Carson argues that the distribution of *thou*, *thee* and *thine* is ingeniously interconnected with power structures in the story of ACO. Relations of superiority and inferiority between Alex and other characters in the novel are reflected in his use of pronouns. (Carson 1976, 201–205) Naturally, if appropriate equivalents of the pronouns are not repeated in the same way in the same places in the translated texts, Alex's pronominal system is broken, and the readers lose another key to a multilayered interpretation of ACO.

2.8 Functions of Nadsat

In 1.3 we have shown that Mareš's classification of the forms of intratextual multilingualism can be adapted in a way that treats invented languages as a subtype of intratextual multilingualism, invented languages being on the same level as any other SL. In 1.4.3 we have introduced Mareš's functions of intratextual multilingualism and, as invented languages are on the same level as any other SL, the functions we considered suitable for the description of functions of Nadsat. In the previous paragraphs of Chapter 2 we have discussed several reasons for and circumstances of using Nadsat in ACO. From them, we infer several functions of Nadsat in the primary communication, and identify them with Mareš's functions in 1.4.3.

1) The most obvious is perhaps the **characterizing function**, which places its speakers among teenage brutes, and separates them from the mainstream ACO's culture not only in the eyes of other characters of the novel but in the eyes of the reader as well. (Kohn 2008, 1) Furthermore, their language picture of the world can be accessed through Nadsat.

2) In 2.3 and 2.4 we discussed the blending of the US and the USSR as the then superpowers, and the role of Nadsat in this blending, making the story apolitical, universal and symbolic. In 2.5 we discussed using a code that does not go 'out of fashion' as ordinary slangs do, enhancing the credibility of the world depicted. In both cases, Nadsat fulfils the **atmospheric function**.

3) The **value function** shifts throughout ACO. In 2.6 we saw that at first the reader is discouraged by the sheer number of unknown words. The reader may feel distance and alienation from Alex. This is a common function of invented languages spoken by extra-terrestrials in the science-fiction genre. (Cheyne 2008, 392–393) Nadsat "defamiliarizes" (Malmgren 1993, 6) the discourse and its strangeness contributes to its negative picture. Certain animosities may also have surfaced due to the Russian-based vocabulary. However,

we also mentioned that Alex is portrayed as a likeable anti-hero. As the reader follows Alex's narration, and is more and more exposed to some of the more frequent Nadsat words, learning them in the process, Nadsat paradoxically encourages him to identify with the hero of ACO. (Mareš 2003, 52)

4) In 2.6 we discussed the intricacies and benefits of reading ACO without a Nadsat glossary. Nadsat suppresses the violence of the book, and softens the effect of reading about it; the reader has to think more about the episodes and scenes where Nadsat is used, instead of just visualising them immediately. Moreover, Nadsat should not be easy to understand from the very beginning, because then it would not be the intended challenge to the reader. Here Nadsat serves the **cryptic and restrictive function**.

3 Textual organisation of Nadsat

3.1 Introducing Nadsat

In the previous chapters we have introduced Nadsat as an invented language with a specific form and specific functions in ACO. We have shown that it is essential to the interpretation of ACO. We have seen that Burgess's creativity consists not only in the use of an invented language per se but also in the way Nadsat words are presented to the reader.

In this chapter we describe Nadsat from a quantitative, linguistic point of view. The whole of Chapter 3 deals with the original text, ACO, and the original secondary language (SL), Nadsat; it provides data that are used to assess the changes in translation creativity, and possible normalisation, in the Czech and German translations.

After delimiting which words count as Nadsat in our analysis, we look into the frequency of Nadsat words, their distribution throughout the text, and the way their meaning is conveyed to the reader. These data are used in the comparison of Nadsat and the secondary languages that replace it in the Czech and the German translations conducted in Chapter 4. Chapter 4 analyses the frequencies and the distribution of SL words throughout TTs, and the way the meaning of SL words is conveyed to the reader. The results of the analysis are summarised in Chapter 5, along with their effect on translation creativity and normalisation, in the Czech and German translations.

First we present two samples from ACO to give the reader a flavour of Nadsat:

Sample 1

'WHAT'S it going to be then, eh?'

There was me, that is Alex, and my three **droogs (1/94)**, that is Pete, Georgie, and Dim. Dim being really dim, and we sat in the Korova Milkbar making up our **rassoodocks (1/2)** what to do with the evening, a flip dark chill winter bastard though dry. The Korova Milkbar was a milk-plus **mesto (1/31)**, and you may, O my brothers, have forgotten what these **mestos (2/31)** were like, things changing so **skorry (1/51)** these days and everybody very quick to forget, newspapers not being read much neither. Well, what they sold there was milk plus something else. They had no license

for selling liquor, but there was no law yet against prodding some of the new **veshches (1/72)** which they used to put into the old **moloko (1/20)**, so you could **peet (1/30)** it with vellocet or synthemesc or drenchrom or one or two other **veshches (2/72)** which would give you a nice quiet **horrorshow (1/109)** fifteen minutes admiring **Bog (1/31)** And All His Holy Angels and Saints in your left shoe with lights bursting all over your **mozg (1/2)**. Or you could **peet (2/30)** milk with knives in it, as we used to say, and this would sharpen you up and make you ready for a bit of dirty twenty-to-one, and that was what we were **peeting (3/30)** this evening I'm starting off the story with.

Our pockets were full of **deng (1/7)**, so there was no real need from the point of view of **crasting (1/14)** any more pretty polly to **tolchock (1/68)** some old **veck (1/240)** in an alley and **viddy (1/230)** him swim in his blood while we counted the takings and divided by four, nor to do the ultra-violent on some shivering **starry (1/105)** grey-haired **ptitsa (1/54)** in a shop and go **smecking (1/75)** off with the till's guts. But, as they say, money isn't everything.

(Burgess 1998, 5)

This passage is the very beginning of ACO. Nadsat words are put in bold for the purpose of this paper, and Sample 1 has been supplemented with additional data: the first number in brackets shows whether it is the first (second, third, etc.) occurrence of the type in ACO; the second number shows its absolute frequency in ACO. Sample 1 consists of 18 types (*droog, rassoodock, mesto, skorry, veshch, moloko, peet, horrorshow, Bog, mozg, deng, crast, tolchock, veck, viddy, starry, ptitsa, smeck*) represented by 22 tokens. We can make several observations that indicate what we investigate in this chapter (the numbers in brackets show the absolute frequency of the preceding word in ACO):

- 1) Frequency ranges from 2 to 240; *veck* (240), *viddy* (230), *horrorshow* (109) and *starry* (105) being in fact among the five most frequent Nadsat words.
- 2) Density of Nadsat words – Sample 1 consists of 319 tokens. Nadsat tokens account for 6.9 % of Sample 1. In total, there are 3,401 Nadsat tokens in a ca 58,563-word-book. Nadsat tokens account for 5.8 % of ACO.
- 3) Three words occur more than once in Sample 1: *mesto* (31) and *veshch* (72) twice, *peet* (30) three times.

Sample 2

So I gave it to them, and I had this shorthand *millicent* (30/48), a very quiet and scared type **chelloveck** (58/240), no real rozz at all, covering page after page after page after. I gave them the ultra-violence, the **crasting** (8/14), the *dratsing* (7/18), the old in-out-in-out, the lot, right up to this night's **veshch** (29/72) with the *bugatty* (1/1) **starry** (54/105) **ptitsa** (37/54) with the mewing *kots* (7/10) and *koshkas* (11/14). And I made sure my so-called **droogs** (45/94) were in it, right up to the *shiyah* (1/1). When I'd got through the lot the shorthand *millicent* (31/48) looked a bit faint, poor old **veck** (59/240). The top rozz said to him, in a kind type *goloss* (33/69): [...]
(Burgess 1998, 58)

Sample 2 was chosen from the end of part one of ACO; in terms of running words it is roughly at the beginning of the second third of ACO. The passage was chosen because it contains many Nadsat words introduced in Sample 1, while the setting of the scene is different from the setting of Sample 1. Sample 2 consists of 13 types represented by 15 tokens. Six types match those introduced in Sample 1 (in bold), seven are new (in italics and underlined). Again, we can make several observations:

- 1) Even though Sample 2 is taken from the beginning of the second third of ACO, two Nadsat words appear here for the first time: *shiyah* (1) and *bugatty* (1). As these words are hapax legomena, it is also for the last time as indicated by the numbers in brackets in the sample.
- 2) Even in this short sample two Nadsat words are repeated: *millicent* (48) and *chelloveck/veck*⁷ (240).
- 3) Even though we are only at the beginning of the second third of ACO, the majority of the words in Sample 2 have already appeared more than they should have if they had been distributed evenly throughout ACO.

⁷ The reason for treating *chelloveck* and *veck* as one type is explained later, in 3.2.

3.2 Preparing the list of Nadsat words

For the analysis, an on-line text of ACO was obtained, and checked against Burgess 1998. Seven obvious typos of Burgess 1998 were corrected using Burgess 2012:

- 1) So then I ~~serreeehed~~ [creeched]: ‘You filthy old soomka’, (Burgess 2012, 70)
- 2) You will have little desire to slooshy all the cally and horrible raskazz of the shock that sent my dad beating his bruised and krovvy ~~røekers~~ [rookers] against unfair like Bog in his Heaven, (Burgess 2012, 85)
- 3) Also there was Big Jew, a very ~~fast~~ [fat] sweaty veck lying flat on his bunk like dead. (Burgess 2012, 94)
- 4) He was a malenky veck, very fat, with all curly hair curling all over his gulliver, and on his spuddy nose he had very thick ~~oehkies~~ [otchkies]. (Burgess 2012, 213)
- 5) And I had to put my over-gown and toofles on and walk down the corridor to the like sinny ~~metse~~ [mesto]. (Burgess 2012, 129)
- 6) So then it seemed to me that it would not bring on the sickness and pain if I just gripped his ankles with my rookers tight round them and brought this ~~grashzny~~ [grahzny] bratchny down to the floor. (Burgess 2012, 136)
- 7) When I woke up I could ~~hear slooshy~~ [slooshy] music coming out of the wall, real gromky, and it was that that had dragged me out of my bit of like sleep. (Burgess 2012, 179)

According to Krashen, Nadsat consists of 241 words (Krashen 1989, 446); Oks’s count is “250 words and phrases” (Oks 2009, 41). Our starting point for creating the list of Nadsat words to analyse is the Nadsat glossary from the 1963 US edition of ACO (Burgess 1963). It is a simple list of Nadsat words and their translations into English.

There are 242 entries in the glossary. Five of them provide translation for two word forms (an example of the format of such an entry is: “brat, bratty – brother”); one merely refers to another entry: “veck – (see chelloveck)”. Fifty-four entries (including two of those five with two word forms) are marked with an asterisk, because they “do not appear to be of Russian origin”. (Hyman 1963b, 185)

We will not include into our sample those 54 entries supposedly not based on Russian words, even though their origin may be, in some cases, Russian in fact. These 54 entries are a heterogeneous group: the majority of them are derived from English slang (*to snuff it*),

rhyiming slang (*luscious glory, pretty polly*), neologisms formed by transparent word formation processes (*sarky*), or bordering proper nouns (names of drugs). Their etymology is not, however, our concern. The main reason for not including this group is that the glossary is influential to the extent that these words are usually not translated in the same way as the words with indisputably Russian origin. They are translated using stylistically neutral equivalents, existing slang or transparent neologisms of the target language.⁸ This leaves us with 188 entries and 191 word forms.

Apart from the above-mentioned entry “brat, bratty”, the entries that include two word forms are: “rook, rooker”, and “sloosh, slosshy” – *slosshy* being a typo, since ACO has *slooshy* instead. These entries seem to indicate that derivation ‘took place’ in Nadsat, as they include a word form that seems to consist of a base form and a derivational morpheme. However, the fact that there are three entries with more than one word form explicitly listed in the glossary does not mean that there are no other Nadsat word forms in the whole text. On the contrary, the vast majority of Nadsat word forms, both derivational and of course inflectional, are not listed in the glossary. Since for our analysis all Nadsat word forms from ACO had to be found in the end, there is no need to treat those few listed in the glossary separately from other Nadsat word forms in ACO so we will not comment on them any further.

Next, we removed from the list seven other words because they are similar to the group of 54 supposedly non-Russian words; the reasons for the removal are also similar to the reasons for the removal of those 54 words. These seven words are, first, words that may be understood in other languages (*biblio, minoota*), which could make the translator use a translation strategy different from the strategy used for ‘true’ Nadsat words, and, second, words that could have been identified as English – or English slang – by translators (*chai, clop, prod, sammy, rozz*⁹). In any case, as we will work with groups of Nadsat words according to their distribution and frequency rather than individual words, this reduction does not alter our results significantly. This leaves us with 181 entries from the original glossary.

⁸ Our preliminary analysis of the counterparts of these words in the Czech translation showed that only 30 % of them were translated using words that could not be associated, directly or indirectly, with the Czech lexis. The first fifteen German counterparts showed a similar trend.

⁹ Biswell (2012b, 206–210) provides us with notes about English slang for *sammy act, rozz*, and *merzky gets*. While the context of *sammy* in ACO is completely consistent with the slang usage (“to sam, to stand sam – to pay for drinks”), and *rozz* is directly linked to “roz(z), rozzer (British slang for policemen)”, Biswell’s note “merzky gets = filthy bastards (Cassel’s Dictionary of Slang)” is not enough to remove *merzky* from our list.

As mentioned above, Nadsat words take on various English suffixes both inflectional (e.g. *mesto* – *mestos*, *droog* – *droogs*, *bratchny* – *bratchnies*) and derivational (e.g. *droogy* used as an adverb, *rookerful* formed analogically to *handful*), usually according to English morphology. Word formation processes ‘take place’ in Nadsat (e.g. rooker-bones, itty up, skorriness, *under-veck*). Similarly to Kenny 2001, we will use unannotated (not lemmatized) texts in our analysis. It would be time-consuming and unnecessary to lemmatize ACO and its translations in their entirety because we are interested in the recurrence only of Nadsat words. For this purpose it is useful to subsume all Nadsat word forms sharing the same ‘stem’¹⁰ under one lemma. There are no complex words with two or more Nadsat ‘stems’ so a simple list of lemmas can be made. The list is based on the ACO glossary even though the forms listed there do not always represent what would be – in a traditional dictionary of an existing language – considered the lemma. In the glossary, there is even a small number of forms that do not appear in ACO in the listed (un)inflected form. However, since we are not interested in the morphology of the words, there is no space here to devise and describe a method which could be used to successfully establish the base form for every single word. Moreover, the result would be confusing as it would create an alternative ‘glossary’ with some of the words keeping their forms from the original glossary, yet others having a new form the reader is not very likely to encounter anyway.

Table 3.1 – Examples of Lemmas

	Lemma	Existing word forms
1)	bolshy	bolshy, bolshiest
2)	brat	brat, bratty, bratties
3)	cal	cally, dog-cal, cal-coloured
4)	droog	droog, droogs, droogie, droogies, droogy
5)	crast	crasted, crasting, shop-crasting

A side effect of this partial lemmatisation is that we treat *chelloveck* and *veck* as one lemma (*veck* being a ‘clipped’ form of *chelloveck*). On the other hand, we consider *spat* and *spatchka*, and *bezoomny* and *oomny* two different lemmas because the connection between them cannot be recognized in terms of English morphology. This leaves us with 180 lemmas to analyse.

¹⁰ Since Nadsat is an invented language and true word formation processes do not apply to it, we believe the term ‘stem’ cannot be used in its proper sense.

3.3 Analysis

3.3.1 Frequencies

A good tool to show the differences in the frequency and distribution of Nadsat lemmas in ACO is the concordancer *AntConc* (AntConc 3.2.4w for Microsoft Windows). For each of the 180 lemmas, we searched for its ‘stem’, or more precisely a truncated word form to make sure all its variants were among the hits. The results were checked manually (removing the few false hits – for example the results of homography between English and Nadsat words such as *tree*, *shoot*, *starry*), and the query was modified subsequently according to the preliminary results, where applicable. For example, the search command was “*bolsh*” so that we do not miss *bolshy*, *bolshiest*, or a word form with a prefix.

The table below provides us with absolute frequencies of the 180 lemmas; 60 of them occur only once in the text; almost half of the words are not repeated more than ten times in ACO. It is perhaps not surprising that there are many more low-frequency words (1–10) than those more frequent. The frequency distribution of Nadsat words shows a tendency similar to the distribution of word frequencies in a natural language (a large number of low-frequency words and a small number of high-frequency words).

Table 3.2 – Summary of Frequencies of Nadsat Words

Frequency range	No. of lemmas
101-240	5
81-100	3
71-80	6
61-70	5
51-60	5
41-50	6
31-40	5
21-30	8
11-20	21
2-10	56
1	60
TOTAL	180

The table below provides a more detailed look at the frequencies of specific Nadsat words. The table also shows that the total number of Nadsat word forms including hapaxes is 3401 or 3341 excluding hapaxes (the sum of frequencies that were multiplied by the

corresponding number of lemmas). We can also see that there are 20 words with just 2 occurrences, but for higher frequencies no significant regularities can be observed. The frequencies of the individual Nadsat words are also presented in Appendix I and Appendix II.

Table 3.3 – Frequencies of the 180 Nadsat Words

Frequency	No. of lemmas	Rank	Lemma
240	1	1	chelloveck
230	1	2	viddy
109	1	3	horrorshow
105	2	4–5	creech; starry
99	1	6	malenky
94	1	7	droog
89	1	8	rook
77	1	9	litso
76	2	10–11	glazz; sloosh
75	1	12	smeck
72	2	13–14	rot; veshch
70	1	15	gulliver
69	1	16	goloss
68	1	17	tolchock
65	1	18	cal
63	1	19	von
56	1	20	malchick
54	1	21	ptitsa
52	1	22	itty
51	2	23–24	bolshy; skorry
48	2	25–16	govoreet; millicent
47	1	27	devotchka
45	2	28–29	bezoomny; slovo
41	1	30	platties
37	1	31	krovvy
33	1	32	noga
31	3	33–35	Bog; grahzny; mesto
30	2	36–37	peet; plott
27	1	38	gazetta
25	1	39	britva
24	2	40–41	chasso; lewdies
23	2	42–43	shoom; zoobies
20	2	44–45	bratchny; moloko

19	3	46–48	carman; gromky; oddy knocky
18	3	49–51	dratsing; nochy; plenny
17	1	52	pletcho
16	2	53–54	gloopy; jeezny
15	3	55–57	baboochka; groody; nagoy
14	4	58–61	crast; koshka; nadsat; nozh
13	2	62–63	pishcha; prestoopnik
12	1	64	lomtick
10	3	65–67	kot; rabbit; razdraz
9	2	68–69	razrez; moodge
8	1	70	sharries
7	5	71–75	cheena; deng; domy; ooko; pony
6	4	76–79	brat; oozy; otchkies; toofles
5	8	80–87	cantora; chasha; forella; goober; keeshkas; kleb; lovet; shlem
4	7	88–94	cheest; grazzy; kupet; poogly; raz; voloss; yahzick
3	6	95–100	eemya; klotch; sabog; scoteena; shoot; smot
2	20	101–120	bitva; bolnoy; collocoll; gooly; gorlo; interessovat; kartoffel; knopka; kopat; mozg; pooshka; rassoodock; sladky; sloochat; sneety; spatchka; strack; tally; zvonock; zvoock;
1	60	121–180	banda; brooko; brosay; bugatty; cluve; chepooka; choodesny; dama; ded; dooby; dorogoy; dva; eeagra; gruppa; lapa; lubbilubing; maslo; merzky; messel; molodoy; morder; nadmenny; nachinat; neezhnies; nuking; odin; okno; oobivat; ookadeet; oomny; oozhassny; osoosh; platch; plesk; podooshka; pol; polezny; privodeet; pyahnitsa; radosty; raskazz; sakar; shaika; shest; shiyah; shlapa; skazat; skvat; sobirat; soomka; soviet; spat; tree; vareet; vred; yahma; yeckate; zammechat; zasnoot; zheena

3.3.2 Distribution

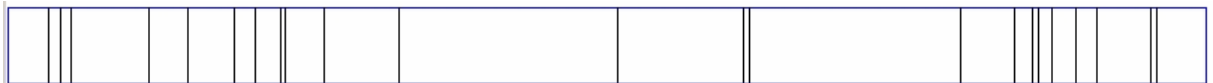
All Nadsat words might have been, theoretically, accumulated in one specific place in the text, for example at the beginning. To see whether this is the case or not, we use another tool of *AntConc*, a concordance plot, also called a dispersion plot (Baker 2006, 60). A concordance plot shows where in a text (in a file) the searched word appears. These places are connected by lexical repetition. The concordance plots revealed that there is no particular place in ACO where Nadsat words tend to accumulate. However, the way specific Nadsat words are distributed throughout the text varies, and a concordance plot is a useful tool to show these differences. Basically we can speak of words which tend to appear in clusters (in

close proximity to each other) in certain parts of the text while being absent or much less frequent in other parts, and of those that are distributed more or less evenly in the text. These two groups do not have strict boundaries, because the decision as to what constitutes a cluster is based on the interpretation of concordance plots and no quantifiable criteria are applied. We divide the words according to the overall tendency to appear in clusters, which means that even if a word with a relatively high frequency creates one or two or any small number of clusters while the rest of its occurrences is distributed evenly through the rest of the text, we do not put it in the cluster group. In the example below, the difference between the two groups is demonstrated; numbers in brackets indicate absolute frequency in ACO.

Example 3.1a – scoteena (3) (cluster)

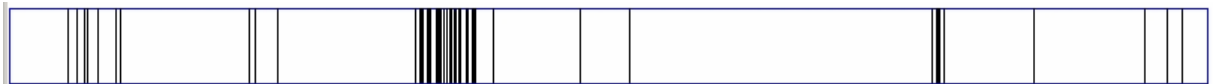


Example 3.1b – zoobies (23) (no cluster)



Naturally, words with only one occurrence do not form clusters. At the other end of the scale, words with 50 or more occurrences (24 words) do not form clusters either because there are no large gaps between individual occurrences. Generally, as the frequency of the words goes up, the clusters become less frequent in favour of a more even distribution. The most frequent word that appears in clusters is *millicent* (48); *ptitsa* (54) probably already belongs to the group of words without clusters.

Example 3.2a – millicent (48)



Example 3.2b – ptitsa (54)



However, that does not mean that words with frequency 50 or higher are distributed evenly in ACO. For example, *von* (63) is almost absent in the first third of ACO, and even the seventh most frequent Nadsat word, *droog* (94), shows large, conspicuous gaps¹¹ between some of its occurrences.

Example 3.3a – von (63)



Example 3.3b – droog (94)



Apart from *zoobies* (Example 3.1b), words with more or less even distribution include:

Example 3.4a – yahzick (4)



Example 3.4b – sharries (8)



Example 3.4c – plott (30)



From the 120 Nadsat words with at least two occurrences, 23 words (19 %) have the tendency to occur in clusters. Sorted by frequency, these words are: *millicent* (48), *krovvy* (37), *gazetta* (27), *chasso* (24), *plenny* (18), *baboochka* (15), *koshka* (14), *nozh* (14), *prestoopnik* (13), *kot* (10), *moodge* (9), *oozy* (6), *forella* (5), *lovet* (5), *shlem* (5), *voloss* (4),

¹¹ These gaps mark the episodes of Alex’s imprisonment, the aversion therapy, and his roaming the streets after being released.

eemya (3), *scoteena* (3), *bolnoy* (2), *pooshka* (2), *sneety* (2), *tally* (2), and *zvonock* (2). Concordance plots of two of them are provided below.

Example 3.5a – plenny (18)



Example 3.5b – koshka (14)



A manual analysis of the context of these words, including the accompanying determiners and pronouns, showed that the majority of the words are part of co-referential chains but all the occurrences of a specific word are only rarely part of one chain; they rather form (or contribute to) two or three shorter chains, while many of the occurrences remain outside any chain.

Many of them appear in clusters just because they are bound to certain settings (the most obvious being the prison) or topics but some of them are used in specific situations, connecting them explicitly this way. For example, *baboochka* (15) refers almost exclusively (12 occurrences) to old women drinking in a bar; *moodge* (9) seems to be synonymous with *chelloveck* (240) but it is used only in three separate scenes: twice it refers to a man Alex beats, while in the third part of the book when the ‘cured’ Alex goes through several mirror-like episodes and is beaten by his former victim and other men, he calls them “moodges” again; *forella* (5) refers exclusively to a cat lady killed by Alex. When the reader comes across these words, he might recall a previous situation when they were used and make connections between the two. We present the concordance plot and KWIC of *moodge* (9) below; the clusters are separated by an empty line.

Example 3.6 – moodge (9)



- 1) *I could never stand to see a [[moodge]] all filthy and rolling and burping and drunk*
- 2) *or like some very rude interrupting sort of a [[moodge]] making a shoom*

- 3) *with her was this chelloveck who was her [[moodge]], youngish too with horn-rimmed otchkies*
- 4) *and the horn-rimmed [[moodge]] said, dithering:*
- 5) *and this writer [[moodge]] went sort of bezoomny*

- 6) *So then I near cried, so that a very starry ragged [[moodge]] opposite me said:*
- 7) *This other [[moodge]] said:*
- 8) *But this starry old [[moodge]] was on his feet, creeching like bezoomny*
- 9) *I was like made sick by the von of old age and disease which came from these near-dead [[moodges]].*

Looking at Nadsat hapaxes, it is interesting that the sheer majority of them occur in the first half of the text; we can also say that the last third of the text contains only one Nadsat hapax. This is demonstrated in Example 3.7 below. The plot combines occurrences of all 60 Nadsat hapaxes.

Example 3.7 – concordance plot of all Nadsat hapaxes combined



3.3.3 Unfolding Meaning

The meaning of Nadsat words is hinted at by several strategies. We will briefly describe the most important ones, and we will label them Intratextual Translation, Context, Co-text, and Co-reference. In the text, they very often combine and enhance each other.

3.3.3.1 Intratextual Translation

The English equivalent is provided, usually in brackets or by “or” as in “where the warders or chassos were standing with their rifles”, for 16 words out of the 120-word Nadsat subset excluding hapaxes (13 %). Nine times it is after the first occurrence. Four times the translation is repeated again later. There is no significant difference in frequency – high-frequency as well as low-frequency words are explained this way, except for the words with

very low frequencies. Only two words with frequencies from 2 to 4 (35% of the 120-word subset) are translated intratextually; two Nadsat hapaxes are also translated intratextually.

Intratextual translation is the most reliable strategy for helping the reader to understand Nadsat, as Burgess does not want to mislead the reader by giving wrong translations for Nadsat words (the context and co-text of Nadsat words matches the context and co-text we would expect of the English equivalent). Sometimes the reader cannot be sure what exactly the English equivalent refers to. In Example 3.8 below, *pletchoes* (17) could be very big built-up shoulders, shoulders in general, or even jackets. However, the reader is usually able to narrow the meaning or make it more precise when the Nadsat word occurs again.

Example 3.8

jackets without lapels but with these very big built-up shoulders ('[[pletchoes]]' we called them)

The effect of intratextual translation is influenced by the distribution of the word in question. If the translated word belongs to a cluster (which is the case with four directly translated words), it is probable that the reader remembers the meaning throughout the cluster and, if the word is repeated often enough, even beyond. On the other hand, if there is a gap between the translated word and its next occurrence, the reader may not recall the translation. The example above showed the first occurrence of *pletcho* at the beginning of ACO. However, *pletcho* occurs only after ca 20 % of the text for the second time (without a translation).

3.3.3.2 Context

To a great extent, the meaning can be grasped from the context, gradually becoming more and more precise. This is especially true for high-frequency words, where providing suitable contexts is the most important strategy. As shown in the example below, when the reader first encounters the Nadsat word *rabbit* (10), he may not be sure about its meaning; he might for instance think it has a more narrow meaning:

Example 3.9

I heard my papapa grumbling and trampling and then ittying off to the dyeworks where he [[rabbited]],

However, when *rabbit* resurfaces some 45 words later, the reader can adjust his understanding of its meaning thanks to its context; the meaning is made clear by the English counterpart in the next sentence:

Example 3.10

Which was true, there being this law for everybody not a child nor with child nor ill to go out [[rabbiting]]. My mum worked at one of the Statemarts, as they called them, filling up the shelves with tinned soup and beans and all that cal.

Another example is *mesto* (31) from Sample 1 and Sample 2 at the beginning of this chapter, where the reader is prompted to think it denotes a milk bar or perhaps a drinking establishment, whereas later it is used to describe any venue up to the point where it enters a collocation in place of the English word *place*:

Example 3.11

I was sure I had met these millicents some [[mesto]] before.

3.3.3.3 Co-text

In ca 40% of cases a Nadsat word stands for an English word in a collocation. Co-text very much helps the reader decipher the meaning of Nadsat words. The collocability of the components varies, as shown in examples below:

Example 3.12a

But I thought that only a very dim veck would have built his [[domy]] upon sand

Example 3.12b

And now, with the [[nochy]] still molodoy, let us be on our way, O my brothers

Example 3.12c

He gripped the edge of the table and said, gritting his [[zoobies]], which were very cally and all stained with cancer-smoke:

Example 3.12d

After I had been given a nice [[chasha]] of real horrorshow coffee and some old gazettas and mags to look at while peeting it

3.3.3.4 Co-reference

When we look into the context of Nadsat words that are part of co-referential chains, we can see that anaphoric reference chains of Nadsat words can include only Nadsat words or Nadsat and English words. The presence of an English content word in the chain greatly raises the possibility that the reader can guess the meaning of the Nadsat word (of course only in case the reader is able to identify the chain).

For example, the reader can guess the meaning of *zvonock* by realizing (thanks to the overall context of the scene, and the co-occurrence with *push*) that *zvonock* refers to the bell-push:

Example 3.13a

There was a bellpush and I pushed, and brrrrrrr brrrrrr sounded down the hall inside. Alike sense of slooshying followed, as though the ptitsa and her koshkas all had their ears back at the brrrrrrr brrrrrr, wondering. So I pushed the old [[zvonock]] a malenky bit more urgent.

When the reader encounters *zvonock* the second time several sentences later (see example below), its meaning is clear even though *bellpush* is not repeated again.

Example 3.13b – zvonock (2)

--	--	--

- 1) *So I pushed the old [[zvonock]] a malenky bit more urgent.*
- 2) *when I'd rung the old [[zvonock]] pretending for help.*

We can find reference chains that include an English content word for ca 25 % of the 120 Nadsat words, but only 10 % contribute substantially to the reader's ability to decipher the meaning. If the reference chain does not contain an English content word, the fact that the Nadsat words are linked in reference chains gives the reader an extra piece of information about the Nadsat word: co-reference ensures that the information about the meaning of the Nadsat word the reader can abstract from the context of each of the word's occurrence holds true simultaneously, i.e. the same object in reality has all the characteristics abstracted from the context of the Nadsat word. Without the reference chain, the reader would not know whether the characteristics pertaining to the objects named by the Nadsat word are not mutually incompatible in one object. However, this kind of co-reference can help only a little, and needs other strategies, especially context, to have any impact at all.

3.3.3.5 Vague meaning

The Nadsat words used as examples in 3.3.3 to show the different strategies are usually easy to understand. However, the meaning of approximately one quarter (32 words) of the 120 Nadsat subset is not so easily accessible, especially in the case of low-frequency words. For example, the adverb *grazzy* (4) is not translated and the context and co-text provide only a vague description. The reader has to infer its meaning from these four occurrences, which provide only an approximate idea of its meaning. The context available to the reader is, of course, larger than in the examples below, but (to our knowledge) it does not include any additional helpful information.

Example 3.14a

[[...]] sat down, rang the bell, and waited for the boy to come. When he came, all nervous and rubbing his rookers on his [[grazzy]] apron, we ordered [[...]]

Example 3.14b

and I could viddy all the plennies sitting down slooshying the Slovo of the Lord in their horrible cal-coloured prison platties, and a sort of filthy von rose from them, not like real unwashed, not [[grazzy]], but like a special real stinking von which you only got with the criminal types, my brothers, a like dusty, greasy, hopeless sort of a von.

Example 3.14c

Then there was a crash and plop and a whish whish while the plennies picked up and dropped and lickturned the pages of their [[grazzy]] malenky hymnbooks,

Example 3.14d

Bully rang the collocoll and a waiter came in rubbing his rookers on his [[grazzy]] apron.

It's perhaps worth noticing that among the Nadsat words that are explained as vaguely as *grazzy*, the eight highest-frequency words contain five adverbs (*bolshy*, *bezoomny*, *grahzny*, *gloopy*, *razdrasz*).

3.4 Functions of Nadsat reflected through the analysis

In Chapter 2 we discussed the functions of Nadsat. The analysis of the textual organisation of Nadsat in Chapter 3 indicated how these functions are made to work.

The characterizing function of Nadsat is based on the differentiation between the fictional teen subculture Alex is a member of, and the mainstream ACO culture. The number of Nadsat words (i.e. the number of lemmas and the frequency of their word forms) is the main linguistic tool for achieving this, creating a stand-alone invented language in the process.

The atmospheric function draws its potential from the fact that an invented language is used as opposed to an existing slang. Again, the frequency of Nadsat words plays the main role. A convincing invented language requires an extensive vocabulary both with respect to the number of lemmas, and their word forms. The lower the numbers, the harder it is for the invented language to serve as something more than just an embellishment to the story. In ACO, reducing the number of invented words would diminish Nadsat's atmospheric function.

The value function is influenced by the frequency and distribution of Nadsat words, as well as the way their meaning is conveyed to the reader. On the one hand, the reader may be overwhelmed by the number of different Nadsat lemmas at the beginning of ACO (cf. Example 3.7, the distribution of Nadsat hapaxes), which causes detachment, and creates a barrier between the reader and the text. On the other hand, the reader is gradually more and more exposed to the growing number of word forms of the most frequent lemmas, which encourages him to accept the text and its narrator.

The cryptic and restrictive function is based on the process by which the reader is offered – and denied – the meaning of Nadsat words. This is carried out by the different strategies for unfolding the meaning, and the distribution of Nadsat words that enable these strategies.

4 Textual organisation of Nadsat's counterparts in translation

4.1 Introduction

In this chapter we contrast the system of textual organisation of Nadsat words in ACO (source text, ST) using the data from Chapter 3, and the systems of textual organisation of the words from the secondary language (SL) substituting Nadsat in target texts (TT). We use two translations of ACO as TTs: the only Czech translation of ACO, *Mechanický pomeranč* (MP), and the first German translation, *Uhrwerk Orange* (UO).

We look at the places in ACO where Nadsat is used and at the corresponding places in the translations of ACO. We want to see whether or not Nadsat words are translated consistently using the same expression for a particular Nadsat word in every place the Nadsat word is used in ACO. We analyse the frequencies and the distribution of SL words throughout TTs, and the way the meaning of SL words is conveyed to the reader. We also link, if possible, the differences in textual organisation in the translations to changes in the functions of the invented language used in the translations.

Our primary concern is not the actual choice of translation strategies with regard to the stylistic, or indeed lexical, features of the words used 'in place' of Nadsat. Of course, replacing Russian as the basis of Nadsat with other natural languages is tricky, and has its impact on the functions described in Chapter 2. However, not replacing Russian or using contemporary slang may have similar consequences for some target languages.

Furthermore, Burgess's 'ingenuity' in constructing Nadsat may differ from that of the translators'. Puns, 'false friends' between Nadsat and English (such as *Bog, starry*), the sound of invented words in the target text – all this is affected by the 'material' used to replace Nadsat words. And when an existing slang is mixed into the Nadsat's counterpart in TT, as it has been done in, for example, Polish translations (Lukas 2008, 25), the whole system of introducing an unknown non-obsolescent element into the text is disrupted.

These are important issues not only from the point of view of translating this particular text. However, in this paper we are interested only in the textual organisation of Nadsat's counterparts in TTs, and the subsequent changes in its functions. To put it simply, we will analyse whether

- 1) there are enough SL lemmas in TT to create a convincing invented language

- 2) these lemmas have enough word forms to permeate the text
- 3) their distribution in the text helps the reader to unfold – or prevents him from unfolding – their meaning in same way as it has been done in ACO.

4.1 Concordancing

The one and only Czech translation of ACO (*Mechanický pomeranč*, MP), by Ladislav Šenkyřík, was published five times between 1992 and 2010, always with a glossary of the Czech version of Nadsat language (i.e. “jazyk týňů”, JT). Minor changes were made to the glossary in every edition but the text itself was revised only for the 4th edition: ca 380 words omitted in the previous Czech editions were added. The text of MP was obtained from the internet and checked against the latest edition published in 2010. Typos present in that edition were not removed. The text contains 52,465 running words.

In German, there are three published translations of ACO. We will use the first one (*Uhrwerk Orange*, UO) by Walter Brumm first published in 1972. The text of UO was obtained from the internet. The text contains 62,121 running words. It does not include a glossary of the German version of Nadsat language (i.e. “Nadsatsprache”). The reader had to decipher the meaning of Nadsatsprache words from the text itself, which is the intended way of reading ACO as discussed in 2.6. In 2000, a revised version of his translation was published as *Clockwork Orange*. However, this edition was not available to us. Meanwhile, Wolfgang Krege translated ACO into German anew: his translation was first published as *Die Uhrwerk-Orange* in 1993, and in 1997 as *Clockwork Orange*. The 1997 edition was a revision, and included a glossary of “Nadsat-Sprache” with 119 entries (ca half of the English glossary). The third translation, by Ulrich Blumenbach, was published in 2013 as *Clockwork Orange: die Urfassung*.

To locate the places in ACO where Nadsat is used, and the corresponding places in its translations, we used *ParaConc* (Version 1.0) created by Michael Barlow. *ParaConc* is a multilingual concordancer, “designed primarily as a search tool designed to work with translated (i.e., parallel) texts.” (Barlow 2003, 9) *ParaConc* is able to semi-automatically align segments of text that are not pre-aligned (tagged), and it works with up to four different languages (texts) simultaneously. Now we will shortly describe how we worked with the texts. The description of the method applies to both TTs but for the sake of clarity it is presented using ST and only one of the TTs.

ST and TT were pre-processed using Microsoft Word. Paragraphs of the texts were numbered automatically, which exposed the differences between them. The original structure

of paragraphs in TT was adapted to match the structure of ST. The numbering was removed and the texts were saved in an appropriate TXT file format. ST and TT files were then loaded into *ParaConc*. This seems to be the most efficient way to align paragraphs when there is no need to preserve their structure.

The corpus was aligned by paragraphs, manually checked, aligned by segments (sentences), and manually checked again. The final alignment is not flawless, especially due to the large amount of direct speech that is represented differently in English from the way it is represented in Czech or German, but it is sufficiently precise to work with. This way we created a mono-source-language multilingual parallel corpus (Laviosa 2002, 38) or simply two bitexts or three parallel texts.

The comparison itself was done in the following way. First we searched for a truncated word form of the Nadsat word in ST in a similar way as with *AntConc* described in 3.3.1. In the Options of the Search window, context type was set to sentence; other options were default. The results of the search were sorted alphabetically using first the KWIC/Highlight option for English sentences; false hits were deleted manually and the number of the remaining hits was checked against the frequencies of Nadsat words that had been obtained using *AntConc*. When the frequencies did not match, the search query had to be modified, and false hits and the number of the remaining hits checked again.

ParaConc automatically offers possible translations of the searched word in the parallel (target) text. They can be highlighted in TT using the Hot words utility. The minimal number of hits was set to 1, and the maximum number of candidates to 12; other options were default. The list of hot words included a number of word forms of secondary language (SL) lemmas used in TT that shared a combination of characters (the ‘stems’ of the SL lemmas). With regard to the Czech translation, the number of word forms in the list was higher than in English as the words of “jazyk týňů” (JT) are more inflected. Once these word forms were selected, TT search results were sorted alphabetically using the KWIC/Highlight option again. As a result, parallel sentences that did not contain a hot word (word form) in the TT were displayed at the end of the TT list of hits.

The number of such sentences varied but we estimate up to 20% of hits were not captured by the Hot words utility. On the one hand, these hits contained word forms of ordinary translations of Nadsat words into the SL of TT that were not frequent enough to pass as hot words; on the other hand they showed where a Nadsat word has not been translated by the adequate SL word. These two groups had to be separated manually. This way, the analysis of the search results produced the total number of instances where the

Nadsat word was translated by the adequate invented word, and the total number of instances when it was translated by other means (including the list of these means) or not translated at all.

However, sometimes the SL word used in TT to stand for a particular Nadsat word was used to translate an ‘unrelated’ word, or a combination of words, used in ACO. For example, *yahzick* (4) (see Example 3.4a) is translated as *ještěk* in MP four times but *ještěk* appears in MP five times in total – the last occurrence being a translation of “going blerp blerp”, translated as “pletl mezitím ještěkem a prděl”. For this reason, the search described above had to be performed again from the point of view of TT: we searched for TT invented words identified in the first search to see whether they were used in TT not only as translations of Nadsat words but as translations of other words as well. Without this second search, frequency and distribution figures of TT would be flawed. Data on frequencies and translation equivalents were stored and categorized using Microsoft Excel spreadsheets.

4.3 Analysis of the results

4.3.1 Introduction

The analysis of the Czech and German translations showed a number of differences between the two. The first observation is the difference in length of the parallel texts in terms of running words, differences in the total number of tokens of the invented language (absolute frequency) and the density with which they permeate the text as summarised in the table below. It already shows a big difference between UO and the other two texts. The density is a basic but crude measure, as it may be affected by the different language systems. On the other hand, the low number of Nadsatsprache tokens in UO is a fact we will encounter repeatedly. The frequencies of the individual JT words are presented in Appendix III and Appendix IV, and the frequencies of individual Nadsatsprache words in Appendix V and Appendix VI.

Table 4.1 – Overview

	running words	invented language tokens	invented language density	invented language tokens per 10000 tokens
ACO	58,562	3401	5.80%	580.75
MP	52,465	3294	6.28%	627.85
UO	62,121	2379	3.83%	382.96

Since in Chapter 3 we worked with 60 Nadsat hapaxes and 120 other Nadsat words, it is not possible to describe the translation strategy used for each of them. Therefore, they were divided into several groups. The criteria for this division are mostly based on the frequency of translation equivalents of Nadsat words in TTs, and they were created post hoc: first we analysed what happened to Nadsat words during the translation into Czech and German, and only then categorised the results in order to make the categories cover both languages easily and completely. First we describe the results for the subset of 120 Nadsat lemmas (excluding hapaxes). We start with the numbers of lemmas (types) for both languages, presented in the table below.

4.3.2 Frequencies

The table below shows the relation between the subset of 120 Nadsat lemmas and their SL counterparts in both TT languages. The numbers represent the number of Nadsat lemmas (types) belonging to each category; the percentages represent the ratio between these numbers and the total number of Nadsat lemmas in the subset (120). The column called “difference” shows how many times the number of lemmas in individual groups is higher or lower in MP and UO, regardless of which one has the higher or the lower number. Taking Group 1 as an example, the number of lemmas in MP is four times the number of lemmas in UO, while in Group 2 the number of lemmas in UO is 10.67 times the number of lemmas in MP.

Table 4.2 – Translation Strategies Groups – Changes in Nadsat Lemmas

	Group	MP (Czech) No.	MP (Czech) %	UO (German) No.	UO (German) %	difference
1	Match	36	30.00 %	9	7.50 %	4.00
2	Zero	3	2.50 %	32	26.67 %	10.67
3	Change	43	35.83 %	51	42.50 %	1.19
4	Combined Change	27	22.50 %	20	16.67 %	1.35
5	Split	8	6.67 %	2	1.67 %	4.00
6	Merged	3	2.50 %	6	5.00 %	2.00
	Total	120	100 %	120	100 %	

1. **“Match”** – means that ST frequency matches TT frequency. It also means that all word forms of the lemmas of this group appear in the same sentence as their Nadsat counterparts. For example, all instances of the Nadsat word *bratchny* (pl. *bratchnies*), and no other words in ACO, are translated as *bračny* in MP:

Example 4.1 – bratchny

- (a) *Bog murder you, you vonny stinking [[bratchnies]].*
- (b) *Bog vám zakruť krkem, vy smelovatý smradlavý [[bračny]].*

2. “Zero” – means that TT frequency equals zero; all instances of the Nadsat word were translated using a word from the target primary language (Czech, German). For example, all instances of the Nadsat word *moloko* (20) were translated by the Czech word *mlíko* in MP:

Example 4.2 – moloko

- (a) *I could feel the knives in the old [[moloko]] starting to prick, and now I was ready for a bit of twenty-to-one.*
- (b) *Cítil sem, jak z vypitýho starýho [[mlíka]] vyskakujou nože a jak se mi chce rozdat si někde trochu přesilovky.*

3. “Change” – means that TT frequency has decreased (when some instances of the Nadsat word were translated by a word from the target primary language or not translated at all) or increased (when some English words were translated using words from the TT invented language). However, it also happens that both processes took place: TT frequency has decreased because one or more instances of the Nadsat word were translated by words from the target primary language, and increased because an English word was in one or more instances translated using the same word from the TT invented language. For example, the Nadsat word *bezoomny* (45) is translated 44 times as *lunatický* (*lunaticky*, *lunatik*, *rozlunatikovat* etc.) in ACO but once as *fanatik*, and three times an ordinary English word is translated as *lunatický*.

Example 4.3a – bezoomny

- (a) *And I made with a like deep bow, smiling like [[bezoomny]] but thinking all the time.*
- (b) *A udělal sem takový hluboký pukrle, smál se jak [[fanatik]], ale celou tu dobu sem urputně přemýšlel.*

Example 4.3b – madness

- (a) *And then I saw this like [[madness]] in F. Alexander's glazzies and said:*
- (b) *A pak sem uviděl ty jakoby [[lunatický]] ajka F. Alexandera a pravil sem:*

4. “**Combined change**” is similar to the previous category but the reason why TT frequency differs from ST frequency is not only the involvement of an ordinary word from the target primary language or English: at least one instance of a TT equivalent of another Nadsat word is used as a translation of the Nadsat word in question. The change in usage of one Nadsat word causes a change in usage of another Nadsat word. For example, the Nadsat word *gulliver* (70) is translated 66 times as *gulliver* in MP, once as *hedka* (which is normally used as the translation of the Nadsat word *mozg* (2) in MP), once as *kiška* (which is normally used as the translation of the Nadsat word *keeshkas* (5) in MP), and once using the Czech word *hlava*.

Example 4.4a – gulliver

- (a) *Georgie with like a cold leg of something in one rooker and half a loaf of kleb with a big dollop of maslo on it in the other, and Pete with a bottle of beer frothing its [[gulliver]] off and a horrorshow rookerful of like plum cake.*
- (b) *Jiřík měl v jedný hendce něco jako studený stehno a v druhý půl bochníku bredu s velkou hroudou másla, a Pítrš s flaškou piva, který si lil na [[hedku]] a v hendce měl velkej kus jakýhosi švestkovýho koláče.*

Example 4.4b –mozg

- (a) *All round were chellovecks well away on milk plus vellocet and synthemesc and drenchrom and other veshches which take you far far far away from this wicked and real world into the land to vidy Bog And All His Holy Angels And Saints in your left sabog with lights bursting and spurting all over your [[mozg]].*
- (b) *Všude kolem seděli hĵumaníci dobře mimo z plus-mlíka se syňágou, toldou nebo hňápce a z dalších bučí, který vás odnesou daleko předaleko z tohodle hnusnýho skutečnýho světa do země, kde lukujete Boga se všema Jeho Blahoslavenějma Andělema a Svatejma ve svý levý saboze, přičemž vám v [[hedce]] vystřelují nálože rachejtlí.*

From the point of view of MP, the JT word *gulliver* is also used in two places in MP but its counterpart is not in the corresponding place in ACO. See example below. However, this change is not the reason for inclusion in this category.

Example 4.5 – gulliver

- (a) *He was a shortish veck in middle age, thirty, forty, fifty, and he had otchkies on.*
- (b) *Byl to menší hĵumaník ve středním věku, třicet, čtyřicet, padesát, a měl na [[gulliveru]] očky.*

5. “**Split**” means that the Nadsat word has two unique equivalents in TT that are not used to translate any other Nadsat word. For example, the Nadsat word *itty* (52) is translated alternately as *guljat* and *volkovat* in MP. (However, not all instances of the ‘split’ Nadsat word are necessarily translated using one or the other equivalent.)

Example 4.6a – itty

- (a) *Let us at once [[itty]].*
- (b) *[[Odguljáme]] tam ted’ hned.*

Example 4.6b – itty

- (a) *I can walk, surely, to wherever we have to [[itty]] to.*
- (b) *Já můžu jít určitě po svejch, kamkoli musíme spolu[[odvolkovat]].*

6. “**Merged**” is similar to Group 2 – Zero because there is no unique equivalent in TT of the Nadsat word. An invented word that is used also as a translation equivalent of another Nadsat word is used as an equivalent of a ‘merged’ Nadsat word. For example, the Nadsat word *gooly* (2) is translated in the same way as the Nadsat word *itty*:

Example 4.7a – gooly

- (a) *So we [[goolied]] up to him, very polite, and I said:*
- (b) *Tak sme si to k němu pomalu [[doguljali]] a já řekl úplně zdvořile:*

Example 4.7b – gooly

- (a) *I couldn't remember what it was I wanted at first, then I remembered with a bit of a shock that I had ittied here to find out how to snuff it without pain, so I [[goolied]] over to the shelf full of reference veshches.*
- (b) *Nemohl sem si ně jak vzpomenout, co sem to tady vlastně chtěl, a pak sem si v šoku uvědomil, že sem na tohle místo odguljal, abych našel návod, jak to bezbolestně odkejhat, a tak sem [[převolkoval]] k regálu se spoustou referenčních bučí.*

Table 4.2 shows that the number of invented lemmas in UO is significantly lower compared to MP: thirty-two Nadsat lemmas (Group 2) were translated by a German lemma, and six Nadsat lemmas ceased to exist because they merged into other Nadsat lemmas (Group 6). And even though the lexicon of Nadsatsprache was slightly enriched because four Nadsatsprache lemmas stand for two Nadsat lemmas in UO (Group 5), in MP there are 16 JT lemmas for eight Nadsat lemmas. MP disposed of 6 Nadsat lemmas (Groups 2 and 6), while UO disposed of 38 (31.67 % of the Nadsat subset). Based on the number of lemmas it is safe to claim that lexical creativity is significantly lower in UO, affecting the characterizing, atmospheric, and value functions of Nadsatsprache.

This conclusion is also supported by the comparison of numbers of word forms of JT and Nadsatsprache. To compare word forms, we need a modified table of lemmas for MP and UO with the actual number of types present in MP and UO, respectively. Groups containing Nadsat lemmas that did not get into TTs were removed, and the number of types in Group 5 was doubled because in splitting Nadsat words of this group, two Nadsatsprache or JT words were created for each of the Nadsat words. While Table 4.2 showed data on Nadsat lemmas, table 4.3 shows data on Nadsatsprache or JT lemmas. The column “ $p < 0.05$ ” indicates the results of the Fisher exact test with the borderline probability value set at $p < 0.05$. The test was chosen because of low values in Group 5; we are more than 95 % confident that the differences in Groups 1 and 3 are statistically significant. The results of the log-likelihood test for Groups 1 and 3 show $p < 0.01$ (99% confidence) but also $p < 0.05$ for Group 5. (McEnery, Xiao and Tono 2006, 55–56) The column “difference 2” compares the number of lemmas adjusted for the different total number of lemmas used to translate the original Nadsat subset into Nadsatsprache and JT, respectively: it compares the percentages instead of numbers of lemmas. This is because the number of JT lemmas used to translate the original Nadsat subset has risen slightly to 122, whereas the number of Nadsatsprache lemmas dropped quite dramatically to 84.

Table 4.3 – Translation Strategies Groups – Groups of TT Lemmas

	Group	MP No.	MP %	UO No.	UO %	p<0.05	difference 2
1	Match	36	29.51 %	9	10.71 %	yes	2.75
3	Change	43	35.25 %	51	60.71 %	yes	1.72
4	Combined Change	27	22.13 %	20	23.81 %	no	1.08
5	Split	16	13.11 %	4	4.76 %	no	2.75
	Total	122	100 %	84	100 %		1.00

Table 4.4 shows numbers of Nadsatsprache/JT tokens in MP a UO in the individual groups. The column “LL” indicates the results of the log-likelihood test for MP and UO sets. The level of statistical significance (probability value) for all groups is less than 0.001 ($p < 0.001$); we are more than 99.9% confident that the difference is statistically significant. (McEney, Xiao and Tono 2006, 55–56) “Difference 2” compares normalised frequencies.

While frequencies of 29.51 % of types matched in MP, we can see that they comprised only 9.5 % of tokens. However, in UO the difference is even higher – only 1.96 % of tokens match. The table also shows the total number of JT and Nadsatsprache tokens in the subset. The corresponding subset of 120 Nadsat words in ACO has 3341 tokens as mentioned in 3.3.1. While the 120 Nadsat subset was translated using 122 types in MP and 84 types in UO (i.e. using 101.67 % and 70% lemmas, respectively, compared to the subset), the number of tokens is lower for both TTs. Compared to ACO, MP uses 97 % word forms and UO 70 % word forms. If we use these numbers to calculate type/token ratios for the three texts, and order them from the lowest, we can see that UO has the lowest ratio, ACO comes second, and MP has the highest ratio. This means MP does not repeat invented words as often as ACO.

Table 4.4 – Translation Strategies Groups – TT Word Forms

	Group	MP No.	MP %	UO No.	UO %	LL	difference 2
1	Match	308	9.50 %	46	1.96 %	141.84	4.85
3	Change	1457	44.93 %	1350	57.50 %	42.39	1.28
4	Combined Change	1110	34.23 %	938	39.95 %	13.48	1.17
5	Split	368	11.35 %	14	0.60 %	305.1	19.03
	Total	3243	100 %	2348	100 %		

Nadsat hapaxes can be divided into the same groups as words with more occurrences (Table 4.5). In MP, 45 hapaxes (75 %) of the sixty Nadsat hapaxes in ACO belong to Group 1, as they are translated by a JT hapax legomenon in MP (1:1 ratio). Seven were translated by a Czech equivalent, and five by a more frequent JT equivalent. The change in translation of

each of the three remaining words is the same: one occurrence of another Nadsat word is translated by the JT word used to translate the hapax in question. As a result, the counterparts of these three Nadsat hapaxes occur in MP twice.

In UO, only 17 (28%) hapaxes belong to Group 1. Thirty-four Nadsat words were translated by a German equivalent (or not translated at all), and two words merged into a more frequent Nadsatsprache equivalent. The four words in Group 3 changed their frequency because they were used to translate an English word or were used in phrases the translator added to UO (these phrases cannot be traced to any corresponding phrase in ACO, they elaborate on its content). However, their frequency did not change that much: three of them occurred twice and one of them three times. The remaining three words in Group 4 (equivalents of *merzky*, *oozhassny*, and *sumka*) changed in a similar way as hapaxes of this group in MP but there are some differences: the Nadsatsprache word *Sumka* occurs four times in UO, because it is used to translate each of the Nadsat words *cheena* (7), *ptitsa* (54), *baboochka* (15) and *sumka* (1) once. The Nadsatsprache word *merzki* (4) is used to translate the both Nadsat hapaxes *merzky* (1) and *oozhassny* (1), and the Nadsat word *von* (63); its last occurrence is due to another phrase the translator added. These changes are summarized in the table below.

Table 4.5 – Translation Strategies Groups – Changes in Nadsat Hapaxes

	Group	MP No.	MP %	UO No.	UO %
1	Match	45	75 %	17	28.33 %
2	Zero	7	11.67 %	34	56.67 %
3	Change	0	0 %	4	6.67 %
4	Combined Change	3	5 %	3	5 %
5	Split	0	0 %	0	0 %
6	Merged	5	8.33 %	2	3.33 %
	Total	60	100 %	60	100 %

To sum up, the translator of MP deprived JT of 20 % of hapaxes and 5% of lemmas with higher frequency, while the translator of UO deprived Nadsatsprache of 60 % of hapaxes and 31.67 % of the other lemmas. Pertaining to the whole 180-word Nadsat lexicon, these numbers are 10 % for MP and 41 % for UO. Removing 41 % of Nadsat lemmas from UO has a particular impact on the atmospheric function.

4.3.3 Distribution

During the *ParaConc* analysis we recorded how much the frequency of TT counterparts in Group 3 and 4 decreased and increased (how many tokens of each TT lemma exist in TT, and how their number was reached). Compared to their Nadsat counterpart frequencies, some TT lemmas only gained tokens, some only lost tokens, and some TT lemmas both gained and lost tokens. (See the description of Group 3 – Change, in 4.3.2.) We use these numbers to measure the change in the distribution of the invented language in TTs.

The following table shows the numbers for Groups 3 and 4 counted together, focusing on the extent of changes made in the distribution of tokens in MP and UO. Removed tokens represent Nadsat word forms that were not translated using the corresponding SL counterpart in TT; matching tokens are those that were. Added tokens represent situations when a word form of the SL lemma in question was used in TT in place of an ordinary target PL expression or the appropriate JT/Nadsatsprache expression. The number of different tokens is the sum of removed and added tokens. Deviation is the ratio between the different tokens and all tokens involved (different tokens divided by the sum of matching and different tokens).

Coincidentally, the changes in frequencies that belong to these two groups involved an almost identical number of Nadsat lemmas (though the actual lemmas in these groups are by no means identical in the two TTs). In MP, the changes involved slightly more frequent lemmas (on average), as the number of tokens is higher. However, even though the translator of UO worked with fewer tokens, he removed or added almost more than twice the number of tokens compared to the MP translator. More precisely, he changed the position of word forms pertaining to the 71 Nadsat lemmas in such a way that 25.37 % of the tokens he worked with do not match the original position of Nadsat word forms, whereas it is only 10.06 % for MP. In other words, ca 75 % of the tokens the UO translator worked with match the original position, while ca 90 % of the tokens the MP translator worked with match the original position.

Table 4.6 – TT Tokens with Matching or Changed Position

	types involved	ACO tokens involved	removed tokens	matching tokens	added tokens	TT tokens	different tokens	deviation
MP	70	2609	158	2451	116	2567	274	10.06 %
UO	71	2366	377	1989	299	2288	676	25.37 %

However, Table 4.6 is only a crude measure of the difference. For example, five occurrences of *malchick* (56) and five occurrences of *oozy* (6) were not translated by its

Nadsatsprache counterpart in UO. Both changes are included in the same way in the table above but the impact of these changes on the structure of Nadsatsprache is very much different: *oozy* (6) almost ceased to exist in UO (its counterpart is *Lolli* with only one occurrence), while the lower frequency of *Malitschickiwick* (51) in UO had hardly any effect at all.

As comparing concordance plots of all Nadsat lemmas would be too cumbersome and would not yield transparent results, we used the same type of data as in Table 4.6 to have a closer look at the differences in distribution. This time, however, we used data obtained for every lemma from Groups 3 and 4 individually. The lemmas were then arranged according to deviation. In the tables below we present the top twenty lemmas (out of 70 and 71, respectively) from each TT according to deviation. To adjust for the different number of ACO tokens involved in MP and UO, we present total sums converted to 100 ACO tokens.

Table 4.7 – Top 20 MP Lemmas with the Highest Deviation

	ACO lemma	ACO tokens involved	MP lemma	removed tokens	matching tokens	added tokens	MP tokens	different tokens	deviation
1	gorlo	2	gorlo	0	2	11	13	11	84.62%
2	moodge	9	házbend	6	3	0	3	6	66.67%
3	smot	3	smotrovat	2	1	0	1	2	66.67%
4	deng	7	many	3	4	1	5	4	50.00%
5	kartoffel	2	potejtovej	1	1	0	1	1	50.00%
6	kot	10	ketek	5	5	0	5	5	50.00%
7	strack	2	fír	1	1	0	1	1	50.00%
8	tally	2	teřich	0	2	2	4	2	50.00%
9	ooko	7	ír	2	5	2	7	4	44.44%
10	razrez	9	riповat	0	9	7	16	7	43.75%
11	oozy	6	čejn	0	6	4	10	4	40.00%
12	klootch	3	špér	1	2	0	2	1	33.33%
13	mozg	2	hedka	0	2	1	3	1	33.33%
14	raz	4	tajm	0	4	2	6	2	33.33%
15	spatchka	2	spačka	0	2	1	3	1	33.33%
16	zvook	2	džvenk	0	2	1	3	1	33.33%
17	razdraz	10	zapsetovat	3	7	0	7	3	30.00%
18	keeshkas	5	kiška	0	5	2	7	2	28.57%
19	dratsing	18	šlág	1	17	5	22	6	26.09%
20	grazzy	4	sojlovatý	1	3	0	3	1	25.00%
	TOTAL	109		26	83	39	122	65	
	TOTAL per 100 ACO tokens			23.85	76.15	35.78	111.93	59.63	

Table 4.8 – Top 20 UO Lemmas with the Highest Deviation

	ACO lemma	ACO tokens involved	UO lemma	removed tokens	matching tokens	added tokens	UO tokens	different tokens	deviation
1	oozy	6	Lolli	5	1	0	1	5	83.33%
2	yahzick	4	Schlapper	3	1	1	2	4	80.00%
3	lomtick	12	Lomtick	9	3	0	3	9	75.00%
4	keeshkas	5	Kischkas	2	3	6	9	8	72.73%
5	sabog	3	Sabog	1	2	4	6	5	71.43%
6	von	63	Sung	45	18	0	18	45	71.43%
7	smot	3	smottete	0	3	7	10	7	70.00%
8	veshch	72	Wetsch	47	25	4	29	51	67.11%
9	shoot	3	Durak	2	1	0	1	2	66.67%
10	pishcha	13	motschka	7	6	1	7	8	57.14%
11	deng	7	Deng	0	7	9	16	9	56.25%
12	bitva	2	Rampferei	1	1	0	1	1	50.00%
13	collocoll	2	Kollokoll	1	1	0	1	1	50.00%
14	mozg	2	Mozg	0	2	2	4	2	50.00%
15	poogly	4	puglig	1	3	2	5	3	50.00%
16	zvonock	2	Zvonock	1	1	0	1	1	50.00%
17	shoom	23	Schum	10	13	2	15	12	48.00%
18	razrez	9	rizrazzen	2	7	4	11	6	46.15%
19	jeezny	16	Dschizny	6	10	2	12	8	44.44%
20	razdraz	10	razdraz	2	8	4	12	6	42.86%
	TOTAL	261		145	116	48	164	193	
	TOTAL per 100 ACO tokens			55.56	44.44	18.39	62.84	73.95	

From these tables we can derive several observations:

- 1) The number of different tokens per 100 ACO tokens in MP is lower than in UO (and, conversely, the number of matching tokens higher).
- 2) Frequencies of lemmas with the most significant changes are lower in MP than UO: only three lemmas exceed the frequency of 10 compared to seven lemmas in UO; in UO there are 2 lemmas with high frequency (rank 13–14 and 19). The difference is mostly due to these two lemmas but even without them, MP changes would still be made to lemmas with lower frequencies than those of UO.
- 3) The change is caused mostly by adding word forms in MP and removing them in UO. Nadsatsprache has not only far fewer lemmas than JT (see Group 2 in Table 4.2, and Table 4.3) but also fewer word forms of the top 20 lemmas in Groups 3 and 4 per 100

ACO tokens. (This holds true for the whole Group 3 and 4 but the data are not presented here.)

- 4) Deviation shows that the number of different tokens in MP falls more sharply, indicating that changes below the top 20 threshold are not so significant. While in MP the 25 % deviation is reached by No. 20, *grazzy* (4), it is only No. 40 in UO that has deviation below 25 %; the deviation of the last word from Groups 3 and 4 is 1.47 % in MP and 8 % in UO.
- 5) Seven low-frequency ACO lemmas (35 %, in bold) appear in both tables. This indicates that these words forced both translators to use alternative translation strategies; the tables also show that they worked with the majority of these words in opposite ways.

To illustrate how the levels of deviation changed the distribution of the lemmas in question, we compare several concordance plots of Nadsat words with corresponding JT and Nadsatsprache plots. In UO, matching occurrences have a slightly different position from the other two texts because of the different length (number of tokens) of UO. Unfortunately, some occurrences are so close to each other that they are impossible to distinguish.

Example 4.8a – zapsetovat (7) – MP



Example 4.8b – razdraz (10) – ACO

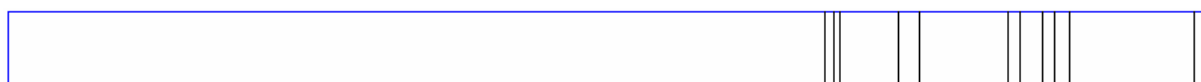


Example 4.8c – razdraz (12) – UO



Example 4.9a – dratsing (18) – ACO



Example 4.9b – šlág (22) – MP**Example 4.10a – jeezny (16) – ACO****Example 4.10b – Dschizny (12) – UO**

In Chapter 3 (3.3.2) we discussed 23 words that tend to appear in clusters in ACO. Looking at their TT counterparts we found out that due to the different translation strategies used by the translators, changes in this group of words cannot be meaningfully generalized. While more than a half of the words match (Group 1) in MP, it is just 13% in UO:

Table 4.9 – Numbers of ACO Lemmas Creating Clusters (by Group)

	Group	MP	UO
1	Match	12	3
2	Zero	1	6
3	Change	7	7
4	Combined Change	2	3
5	Split	1	2
6	Merged	0	2
	Total	23	23

The more specific table below shows that seven lemmas belong to the same group in both TTs, and the distribution of three of them changed (*gazetta* (27), *noz* (14), *oozy* (6)), but these changes have little effect on the reading of TT. On the other hand, when we look at *baboochka* (15), *moodge* (9), and *forella* (5) described in Chapter 3, we can see that the distribution of *baboochka* (15) and *forella* (5) in MP matches the distribution in ACO, while *moodge* (9) and *forella* (5) were merged with other Nadsatsprache words in UO.

Table 4.10 – ACO Lemmas Creating Clusters (by Group)

Lemma	ACO Frequency	MP Group	UO Group
millicent	48	Change	Zero
krovvy	37	Match	Change
gazetta	27	Change	Change
chasso	24	Split	Change
plenny	18	Change	Combined Ch.
baboochka	15	Match	Combined Ch.
koshka	14	Change	Match
nozh	14	Change	Change
prestoopnik	13	Match	Combined Ch.
kot	10	Change	Split
moodge	9	Combined Ch.	Merged
oozy	6	Change	Change
forella	5	Match	Merged
lovet	5	Match	Zero
shlem	5	Match	Change
voloss	4	Zero	Zero
eemya	3	Match	Zero
scoteena	3	Match	Split
bolnoy	2	Match	Match
pooshka	2	Match	Match
sneety	2	Match	Zero
tally	2	Combined Ch.	Zero
zvonock	2	Match	Change

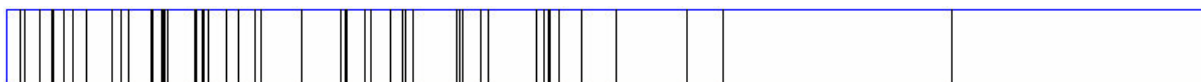
In UO, *forella* (5) is translated by the German *Mutter* and *Reff*, and by the Nadsatsprache *Babuschka* (22) (twice) and *Titsa*, which are normally used for the Nadsat *baboochka* (15) and *ptitsa* (54). The evocative power of the Nadsat word crumbled into several ‘common’ Nadsatsprache words, and the specific role the cat lady plays in the novel is not reflected in Nadsatsprache.

The Nadsat *Baboochka* (15), which refers almost exclusively (12 occurrences) to old women drinking in a bar in ACO, is translated as *Babuschka* (fourteen-times) and once as *Sumka* (4), which is used for the Nadsat *soomka* (1) and three other instances of other Nadsat words. However, the Nadsatsprache *Babuschka* (22) is also used instead of the Nadsat *forella* (5) (twice), *ptitsa* (54) (three-times), and the English word *woman* and two pronouns. The situation is similar: in UO, “Babuschka” refers, in addition, to old female shopkeepers in two different scenes.

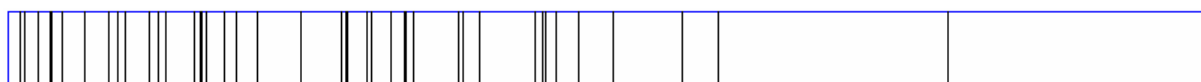
In UO, *moodge* (9) is translated by the Nadsatsprache *Veck* (279), which is normally used for the Nadsat *chelloveck* (240), five times; the remaining occurrences are translated by German words. In MP, *moodge* (9) is translated by the MP *házbend* (3) when referring to the married writer of *A Clockwork Orange*, and by other JT and Czech words in other instances.

To contrast the distribution of Nadsat hapaxes with their counterparts in MP and UO, we can use concordance plots. We can see how the differences in frequencies described in 4.3.2 affected distribution: at the beginning of UO, the reader is not exposed to so many invented words as in ACO and MP.

Example 4.11a – Nadsat Hapaxes



Example 4.11b – JT Hapaxes



Example 4.11c – Nadsatsprache Hapaxes



4.3.4 Unfolding Meaning

In Chapter 3 we briefly discussed how the meaning of Nadsat words can be deciphered while reading the text. In contrasting ACO and TTs we focus on intratextual translation, one of the strategies described there, as the analysis of other strategies is beyond the scope of this paper.

In 3.3.3.1 we identified 16 Nadsat words out of the 120-words subset (13 %) and two Nadsat hapaxes that are intratextually translated into English in ACO. The translation is provided usually in brackets or by “or”. Now we look at whether these Nadsat-English translations are transposed as JT-Czech and Nadsatsprache-German translations.

Table 4.11 – Intratextual Translation

	ACO lemma	ACO frequency	Translated more than once (in ACO)	Translated Occurrence(s) (in ACO)	Translated in MP	Translated in UO
1	rook	89	no	1.	yes	n/a
2	litso	77	no	1.	yes	yes
3	glazz	76	no	3.	yes	n/a
4	millicent	48	yes	1., 11.	no	n/a
5	Bog	31	yes	2., 4., 5., 20., 23., 24.	yes	not 4., 23.
6	chasso	24	yes	2., 4., 11.	yes	not 11.
7	nochy	18	no	6.	yes	yes
8	pletcho	17	no	1.	yes	yes
9	nagoy	15	yes	1., 12.	yes	not 12.
10	nadsat	14	yes	1./2., 7.	yes	not 7.
11	deng	7	no	5.	no	yes
12	ooko	7	no	1.	no	n/a
13	oozy	6	no	2.	yes	no
14	shlem	5	no	2.	yes	no
15	yahzick	4	yes	1., 3.	yes	no
16	kartoffel	2	no	1.	no	n/a
17	pyahnitsa	1	no	1.	yes	yes
18	morder	1	no	1.	yes	no
	TOTAL		6		14	5

The table shows which words are translated more than once in ACO, and which occurrence(s) are translated. Further it shows which of the words are translated in MP, and UO – “yes” means that all occurrences intratextually translated in ACO are intratextually translated in TT; “no” means that none were. The situation in UO is more complicated: five Nadsat words marked “n/a” were not translated by a Nadsatsprache word at all (they belong to Group 2 – Zero; see Table 4.2) so their meaning is not explained to the reader by a Nadsatsprache-German translation; when only some occurrences of a Nadsatsprache equivalent were not intratextually translated in UO, the occurrences are indicated by numbers in the last column. Now we will have a closer look at the specific words that are not translated or translated only partially. There are four words in MP that are not translated:

Example 4.12 – millicent (48)

- (a) *the armed [[millicents]] or rozz patrols weren't round there much*
 (b) *ozbrojení [[policajti]] nebo hlídky rožů sem moc nechodili*

The first occurrence of *millicent* (48) is translated by the Czech word *policajt* to explain the meaning of the English slang word *rozz*, which is translated as *rož* in MP. The eleventh occurrence of *millicent* is correctly translated by the JT word *millicjant* but the English *rozz*, which serves as its explanation, is again translated by the word *rož*, so from the viewpoint of the Czech reader, two unknown words are used to explain each other's meaning.

Example 4.13 – deng (7)

- (a) *'And what will you do,' I said, 'with the big big big [[deng]] or money*
 (b) *„A co budete dělat,“ řekl sem, „s tím velikým prevelikým [[katrem]] neboli penězma,*

The fifth occurrence of *deng* (7) is not translated by its usual JT counterpart *many* but by *katr*, which is normally used in MP as the translation of the Nadsat word *cutter*. *Cutter* is explained by a translation “there was only a malenky bit of cutter (money, that is)” in ACO, and correctly translated as “byla jen bitková trocha katru (peněz, aby bylo jasno)” in MP. (*Cutter* is not part of the 180-word sample of Nadsat, because it is marked by an asterisk in the ACO glossary.) As a result, *katr* is explained by a JT-Czech translation twice in MP, while *many* is not explained this way.

Example 4.14 – ooko (7)

- (a) *I cracked this veck who was sitting next to me and well away and burbling a horrorshow crack on the [[ooko]] or earhole, but he didn't feel it*
 (b) *sem praštil toho hjanika, co seděl vedle mě, a pěkně sem mu ji usadil na [[ajko]] nebo možná na ucho, ale on to vůbec necítil*

The first two occurrences of *ooko* (7) are translated as *ajko* in MP, which is normally used for the Nadsat word *glazz* (76). It seems the translator thought *ooko* means *eye*, and invented the JT work *ajko* to translate it, while translating *glazz* by the Czech word *oko*. When he later came across the line in ACO intratextually translating *glazz* (76) into English (“began to swing it beautiful in the eyes or [[glazzies]]”), he started using *ajko* as a translation of *glazz* because of its meaning. He used *ajko* for both *ooko* and *glazz*, and only after he came across a line reading “and creeched loud in my ooko”, he invented another JT word, *ír*, and has used it to translate *ooko* from that time on.

Example 4.15 – kartoffel (2)

(a) *looked like whipped-up [[kartoffel]] or spud with a sort of a design made on it with a fork.*

(b) *měly barvu [[bramborový]] kaše, do který někdo vyryl vidličkou vzorek.*

The first occurrence of *kartoffel* is translated by the Czech word *bramborový*, and the Nadsat-English translation is not translated into Czech.

In UO, the situation is not as complicated as in MP. All Nadsat words that have been intratextually translated in ACO in more than one occurrence (and have their Nadsatsprache counterpart in UO) have some of these repeated translations removed from UO (*Bog, chasso, nagoy, nadsat*). This seems to indicate that the translator believed that the meaning of a word that has already been explained by intratextual translation does not need to be repeated. The reason why three other words were not explained by an intratextual translation (*oozy, yahzick, morder*) is that the occurrence that was accompanied by the English translation in ACO is not translated by a Nadsatsprache word in UO. Such a situation is in fact the same as the situation with *kartoffel* (2) in ACO (Example 4.15). The English translation of the last word, *shlem*, was simply not translated into German.

5 Conclusion

5.1 Summary of the analysis

The analysis of the 120 ACO lemmas showed that the number of invented lemmas in UO is significantly lower compared to MP: MP disposed of 6 Nadsat lemmas, while UO disposed of 38 (5 % and 31.67 % respectively). It also showed that the position of Nadsatsprache words changed more than the position of JT words: word forms of 29.51 % of JT lemmas are always used in the same sentence as their Nadsat counterparts, compared to just 10.71 % of Nadsatsprache lemmas. This group of lemmas represents 9.5 % of JT word forms and 1.96 % of Nadsatsprache word forms.

The analysis of ACO hapaxes showed that the number of invented hapaxes in UO is significantly lower compared to MP: MP disposed of 12 Nadsat hapaxes, while UO disposed of 36 (20 % and 60 % respectively), and that the position of Nadsatsprache hapaxes changed more than the position of JT hapaxes: 75 % of JT hapaxes are always used in the same sentence as their Nadsat counterparts, compared to 28.33 % of Nadsatsprache hapaxes. In total, MP disposed of 10 % of Nadsat lemmas and UO disposed of 41 %.

Both translators made changes in the distribution of JT/Nadsatsprache word forms when compared to their Nadsat counterparts; the changes involved 70 JT lemmas and 71 Nadsatsprache lemmas. In UO, 25.37 % of the tokens of this group do not match the original position of Nadsat word forms, whereas it is only 10.06 % for MP. In other words, ca 75 % of Nadsatsprache tokens from this group match the original position, while ca 90 % of the JT tokens from this group match it. Changes in the distribution of the word forms of individual lemmas from this group are greater in MP than in UO, as shown by the top 20 lemmas with the highest deviation from ACO.

The denotative meaning of several words whose Nadsat counterparts tend to appear in clusters in ACO was altered by changes made to the distribution of their word forms in MP and UO.

At the beginning of UO, the reader is not exposed to as many hapaxes as he is in ACO or MP. The strategy of intratextual translation of Nadsat words is used more consistently in MP than in UO. While there are 14 words that follow the strategy in MP, in

UO there are only 5. Four other words that are intratextually translated in ACO in more than one occurrence have some of these repeated translations removed from UO.

5.2 Lexical creativity and normalisation

In her research, Kenny considers lexical items to be creative in the target language “if they are not already known to the researcher, not attested in standard target language lexicographical sources, and/or not present in the comparative data supplied by the British National Corpus.” (Kenny 2001, 140) Applied to ACO, we consider JT and Nadsatsprache words to be creative in MP and UO if they are used in place of their Nadsat counterparts. The lower the number of invented lemmas and word forms in TT, the lower the lexical creativity. Perhaps counterintuitively, changes in the distribution of JT/Nadsatsprache word forms, including their additional occurrences in places without a Nadsat counterpart in ACO, are not considered creative compensation for some losses of other Nadsat counterparts in TT as they disturb repetition patterns and the overall textual organisation of TT described in Chapter 2. Based on the numbers obtained by the analysis, we can conclude that lexical creativity is diminished in both translations, but much more so in UO.

According to Kenny, “normalization occurs when translations contain fewer instances of abnormal or creative language features than would be expected on the basis of their respective source texts alone.” (Kenny 2001, 66) Lexical normalisation in translation is then deemed to have taken place according to “whether or not the translator has matched a lexically creative form or collocation in the source text with an equally creative form or collocation in the target text.” (Kenny 2001, 140) From the point of view of a single text, this definition seems to place lexical creativity in opposition to normalisation: loss of lexical creativity leads to a higher level of normalisation. Therefore it seems that both translations were normalised. Following Kenny, the level of normalisation of invented words can be assessed by the same data provided by the analysis for those invented words TT disposed of. For example, we can say that 20 % of Nadsat hapaxes were normalised in MP, and 60 % of Nadsat hapaxes were normalised in UO.

5.3 Changes in functions

When we try to assess how the functions of Nadsat described in Chapter 2 were influenced by the changes, we cannot avoid being subjective in the interpretation of the results.

For MP, the results appear inconclusive. Even though MP disposed of 10 % of Nadsat lemmas and made changes in the distribution of 70 lemmas which occur more than once in ACO, the removed lemmas were mostly hapaxes, and the greatest changes affected mostly less frequent words (Table 4.7). The reduction in the number of invented words did not diminish JT's characterizing and atmospheric functions. Furthermore, the reader was exposed to invented language word forms comparable in number to those in ACO, and not many hapaxes were removed from the beginning of ACO; the value function was not affected significantly. The cryptic and restrictive function was probably affected the most, but many of the changes in intratextual translation can be ascribed to a misunderstanding of ST and careless editorial revision. The cryptic and restrictive function was also affected by splitting a number of not so infrequent Nadsat lemmas, which might have added to the confusion.

However, these functions may have been altered significantly by lexical choices made by the translator when creating JT. The ambiguous nature of the Czech translation was reflected in contemporary reviews: the translation was criticised mostly for using both Russian and English words as the basis of Nadsat (Dudek 1993, 12; Mandys 1993, 10) but other reviewers held positive (Kosatík 1993, 57) or at least not negative (Beneš 1993, 24) view of MP. Doležalová noticed that “foreign” words are less frequent in MP than in ACO. (Doležalová 1996, 247) This may correspond to the lower number of JT lemmas in MP. On the other hand, MP was also considered “overstuffed by words from Nadsat” (Flemr 1993, 52), which may correspond to the higher density of JT word forms compared to ACO (Table 4.1).

Speaking about UO, it is possible to claim, based on the number of Nadsatsprache lemmas, that lexical creativity is significantly lower in UO, affecting all of the mentioned functions. Removing 41 % of Nadsat lemmas from UO has an impact on the atmospheric and characterizing functions. Changes in the number of Nadsatsprache lemmas and the frequency of their word forms were linked to changes in the gist of the novel. (Mannewitz 1999, 86–88; Haberl 2009, 102–103) The value function was affected to the greatest extent by the removal of 60 % of Nadsat hapaxes and the lower density of Nadsatsprache word forms. The reader is not as overwhelmed by unknown expressions at the beginning of UO, and his overall exposure to the more and more familiar way of narration that is permeated by Nadsatsprache expressions is also less intensive compared to ACO. Changes to the cryptic and restrictive functions cannot be assessed, even subjectively, as there are too many factors involved: on the one hand, the meaning of many originally Nadsat words is made much clearer to the reader, because it is conveyed in German or German slang instead of Nadsatsprache; on the

other hand, only five words out of eighteen are intratextually translated in UO in the same way as they are in ACO, and the repetition of intratextual translation is not used to the same extent. Other factors include the overall lower number of Nadsat words, which may help the reader to decipher the meaning of Nadsatsprache words in sentences where German is used to translate other Nadsat words; the lower number of hapaxes, which might have confused the reader; changes in the denotative meaning of some Nadsatsprache words; and the higher extent of the changes in distribution of Nadsatsprache as quantified by deviation (Table 4.8).

5.4 Further implications

The present paper showed that invented languages can be used to quantify how creativity changes in translation. Focusing on invented languages, corpora – which have been used mostly for exposing typical, conventional patterns (Stewart 2000, 73–75) – can serve as a device to partially answer the question “how to measure, describe, and handle this creativity” (Mahlberg 2007, 221). Of course, invented languages are a specific, and perhaps peripheral, part of the creative use of language and their contribution to creativity studies is limited. However, as they are easy to uncover in the source text, they bypass the difficulties in identifying them mentioned by Kenny (2000, 95–99).

Sources

ALBANI and BUONARROTI 2010

Albani, Paolo and Berlinghiero Buonarroti. *Dictionnaire des langues imaginaires*. Paris: Les Belles Lettres, 2010.

ARISTOPHANES 1910

Aristophanes. *The Acharnians of Aristophanes*. London: George Bell & sons, 1910. 8. May 2013 < <https://archive.org/details/cu31924026467823>>.

ARISTOPHANES 1989

Aristophanes. "The Frogs". *The Peace. The Birds. The Frogs*. London: Harvard University Press, 1989.

ARISTOPHANES 1998

Aristophanes. "Acharnians." *Acharnians. Knights*. Ed. Jeffrey Henderson. Cambridge: Harvard University Press, 1998.

ARISTOPHANES 2002

Aristophanes. "Frogs" *Frogs. Assemblywomen. Wealth*. Ed. Jeffrey Henderson. Cambridge, London: Harvard University Press, 2002.

BACHTIN 2007

Bachtin, Michail Michajlovič. *François Rabelais a lidová kultura středověku a renesance*. Praha: Argo, 2007.

BAKER 1993

Baker, Mona. "Corpus Linguistics and Translational Studies." *Text and Technology: In Honour of John Sinclair*. Ed. Mona Baker, Gill Francis, and Elena Tognini-Bonelli. Amsterdam: John Benjamins Publishing Company, 1993. 233–250.

BAKER 2006

Baker, Paul. *Using corpora in discourse analysis*. London: Continuum, c2006.

BAKHTIN 2011

Bakhtin, Mikhail “Discourse in the Novel.” *The Dialogic Imagination*. Ed. M. Holquist. Austin: University of Texas Press, 2011. 259–422.

BARLOW 2003

Barlow, Michael. *ParaConc: A Concordancer for Parallel Texts. (Draft 3/03)*. [Houston], 2003. 8 June 2015

<<http://web.archive.org/web/20120131140950/http://www.athel.com/paraconc.pdf>>.

BARTLETT 2006

Bartlett, P. O. “Artificial Languages.” *Encyclopedia of Language and Linguistics. Vol. 1*. Ed. Keith Brown and Anne Anderson. 14 vols. Oxford: Elsevier, 2006.

BAUSANI 1970

Bausani, Alessandro. *Geheim- und Universalsprachen: Entwicklung und Typologie*. Stuttgart: Kohlhammer, 1970.

BENEŠ 1993

Beneš, Jan. “Mechanický pomeranč. Stále živá Burgessova vize.” *Reflex* 4.12 (1993): 24.

BISWELL 2006

Biswell, Andrew. *The Real Life of Anthony Burgess*. London: Picador, 2006.

BISWELL 2012a

Biswell, Andrew. “Introduction.” Anthony Burgess. *A Clockwork Orange. The Restored Edition*. Ed. A. Biswell. New York: W. W. Norton, 2012. vii–xxiii.

BISWELL 2012b

Biswell, Andrew. “Notes.” Anthony Burgess. *A Clockwork Orange. The Restored Edition*. Ed. A. Biswell. New York: W. W. Norton, 2012. 205–212.

BURGESS 1962

Burgess, Anthony. *A Clockwork Orange*. London: Heinemann, 1962.

BURGESS 1963

Burgess, Anthony. *A Clockwork Orange*. New York: Norton, 1963.

BURGESS 1977

Burgess, Anthony. "Interview with John Cullinan." *Writers at Work: Fourth Series*. Ed. George Plimpton. London: Secker and Warburg, 1977.

BURGESS 1987

Burgess, Anthony. "A Prefatory Word." *A Clockwork Orange: A Play with Music*. London: Hutchinson, 1987. v–xi.

BURGESS 1991

Burgess, Anthony. *You've had your time: the second part of the confessions*. New York: Grove Weidenfeld, 1991.

BURGESS 1998

Burgess, Anthony. *A Clockwork Orange*. London: Penguin Books, 1998.

BURGESS 2012

Burgess, Anthony. *A Clockwork Orange. The Restored Edition*. London: W. W. Norton, 2012.

CARPENTER 1992

Carpenter, Humphrey. *J R R Tolkien. A Biography*. London: Grafton, 1992.

CARSON 1976

Carson, Julie. "Pronominalization in A Clockwork Orange." *Papers on Language & Literature* 12.2 (1976): 200–205.

CHEYNE 2008

Cheyne, Ria. "Created Languages in Science Fiction." *Science Fiction Studies* 35.3 (2008): 386–403.

COALE 1981

Coale, Samuel. "The Ludic Loves of Anthony Burgess." *Modern Fiction Studies* 27.3 (1981): 453–463.

CRYSTAL 2010

Crystal, David. "Artificial languages." *The Cambridge encyclopedia of language*. Ed. David Crystal. 1 vol. Cambridge: Cambridge University Press, 2010.

ČERMÁKOVÁ and FÁROVÁ 2010

Čermáková, Anna and Lenka Fárová. "Keywords in Harry Potter and their Czech and Finnish Translation Equivalents." *InterCorp: Exploring a Multilingual Corpus*. Ed. František Čermák, Aleš Klégr, and Patrick Corness. Praha: Nakladatelství lidové noviny, 2010. 177–188.

DAVIS and WOMACK 2002

Davis, Todd F. and Kenneth Womack. "'O my brothers': Reading the Anti-Ethics of the Pseudo-Family in Anthony Burgess's *A Clockwork Orange*." *College Literature* 29.2 (2002): 19–36.

DIX 1971

Dix, Carol M. *Anthony Burgess*. Harlow: Longman Group, 1971.

DOLEŽALOVÁ 1996

Doležalová, Petra. "'Vy gražny bračny...': Mechanický pomeranč Anthonyho Burgesse v českém překladu." *Čeština doma a ve světě* 4.4 (1996): 246–248.

DUDEK 1993

Dudek, Petr. "Porouchaný mechanismus pomeranče." *Literární noviny* 4.5 (1993): 12.

ECO 1997

Eco, Umberto. *The Search for the Perfect Language*. London: Fontana, 1997.

ELWERT 1973

Elwert, Wilhelm Theodor. "L'emploi de langues étrangères comme procédé stylistique." *Das zweisprachige Individuum und andere Aufsätze zur romanischen und allgemeinen Sprachwissenschaft*. Ed. Wilhelm Theodor Elwert. Wiesbaden: Steiner, 1973. 228–256.

FLEMR 1993

Flemr, Jan. *Burgessův "Mechanický pomeranč": pokus o interpretaci*. Praha, 1993. M.A. thesis.

FORSTER 1970

Forster, Leonard. *The Poet's Tongues: Multilingualism in Literature. The de Carle Lectures at the University of Otago 1968*. London: Cambridge University Press in association with University of Otago Press, 1970.

FOWLER 1981

Fowler, Roger. *Literature as Social Discourse*. Bloomington: Indiana University Press, 1981.

GOETSCH 1987

Goetsch, Paul. "Fremdsprachen in der Literatur. Ein typologischer Überblick." *Dialekte und Fremdsprachen in der Literatur*. Ed. Paul Goetsch. Tübingen: Narr, 1987. 43–68.

HABERL 2009

Haberl, Michael. *Übersetzungskritik: A Clockwork Orange. Der Roman von Anthony Burgess in Gegenüberstellung zu zwei deutschen Übersetzungen*. Wien, 2009. M. A. thesis. 16 June 2013 <http://othes.univie.ac.at/6727/1/2009-09-22_0026725.pdf>.

HALLIDAY 1987

Halliday, Michael Alexander Kirkwood. *Language as social semiotic: the social interpretation of language and meaning*. London: Arnold, 1987.

HAVLIŠ 2009

Havliš, Jan. "Konstruované fiktivní jazyky, fenomén fantastické literatury." Studentská vědecká konference, Ústav české literatury a knihovnictví, Filozofická fakulta Masarykovy univerzity, Brno 16. října 2009. 8 May 2015

<www.phil.muni.cz/wucl/home/veda/projekty/havlis.doc>.

HYMAN 1963a

Hyman, Stanley Edgar. "Anthony Burgess's Clockwork Oranges." *The New Reader* 7 Jan 1963: 22–23.

HYMAN 1963b

Hyman, Stanley Edgar. "Afterword." Burgess, Anthony. *A Clockwork Orange*. New York: Norton, 1963.

INGERSOLL 1986

Ingersoll, Earl. "Burgess' A CLOCKWORK ORANGE." *The Explicator* 45.1 (1986): 60–62.

JACKSON 1999

Jackson, Kevin. "Real Horrorshow: A Short Lexicon Of Nadsat." *Sight and Sound* 9.9 (1999): 24–27.

JAKUBOWSKI and CLUTE 1993

Jakubowski, Maxim and John Clute. "Burgess, Anthony." *The Encyclopedia of Science Fiction*. Ed. John Clute and Peter Nicholls. 1 vol. New York: St. Martin's Press, 1993.

KENNY 2000

Kenny, Dorothy. "Lexical Hide-and-Seek: Looking for Creativity in a Parallel Corpus." *Intercultural Faultlines: Research Models in Translation Studies I : Textual and Cognitive Aspects*. Ed. Maeve Olohan. Manchester: St. Jerome Publishing, c2000. 93–104.

KENNY 2001

Kenny, Dorothy. *Lexis and creativity in translation: a corpus-based study*. Manchester: St. Jerome Publishing, 2001.

KLAUDY and KÁROLY 2000

Klaudy, Kinga and Krisztina Károly. "The Text-organizing Function of Lexical Repetition in Translation." *Intercultural Faultlines: Research Models in Translation Studies I : Textual and Cognitive Aspects*. Ed. Maeve Olohan. Manchester: St. Jerome Publishing, c2000. 143–159.

KNAUTH 2007

Knauth, Alfons K. "Literary Multilingualism I: General Outlines and Western World." Comparative Literature: Sharing Knowledges for Preserving Cultural Diversity. Encyclopedia of Life Support Systems (EOLSS). 12 Feb 2013.
<<http://www.eolss.net/Sample-Chapters/C04/E6-87-07-04.pdf>>.

KOHN 2008

Kohn, Liberty. "Antilanguage and a Gentleman's Goloss: Style, Register and Entitlement To Irony in A Clockwork Orange." *eSharp*: 1.11 (2008): 1–27. 8 May 2015
<http://www.gla.ac.uk/media/media_81291_en.pdf>.

KOSATÍK 1993

Kosatík, Pavel. "Z Mažňáka Dušínem (a zase zpátky)." *Ikarie* 4.5 (1993): 56–57.

KRASHEN 1989

Krashen, Stephen. "We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis." *The Modern Language Journal* 73.4 (1989): 440–464.

LARGE 1994

Large, A[ndrew]. "Artificial Language." *The encyclopedia of language and linguistics. Vol. 1*. Ed. R. E. Asher and J. M. Y. Simpson. 10 vols. Oxford: Pergamon Press, 1994.

LAVIOSA 2002

Laviosa, Sara. *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam: Rodopi, 2002.

LAVIOSA-BRAITHWAITE 2008

Laviosa-Braithwaite, Sara. "Universals of translation." *Routledge encyclopedia of translation studies*. Ed. Mona Baker and Kirsten Malmkjær. 1. vol. London: Routledge, 2008. 288–291.

LAWRENCE 2000

Lawrence, D. H. *Women in Love*. London: Penguin Books, 2000.

LOTFIPOUR-SAEDI 1997

Lotfipour-Saedi, Kazem. "Lexical Cohesion and Translation Equivalence." *Meta* 42.1 (1997): 185–192.

LUKAS 2008

Lukas, Katarzyna. "Übersetzung als Entwurf einer Zukunftssprache. Zu den Übersetzungen des Romans *A Clockwork Orange* von Anthony Burgess ins Deutsche und Polnische." *Studia Germanica Gedanensia* 1.16 (2008): 11–28.

MACUROVÁ and MAREŠ 1996

Macurová, Alena and Petr Mareš. "Soužití a střet. Ke stylu vícejazyčného textu." *Stylistyka* 3.1 (1996): 164–182.

MAHLBERG 2007

Michaela Mahlberg. "Corpus stylistics: bridging the gap between linguistic and literary studies." *Text, discourse and corpora: theory and analysis*. Ed. Michael Hoey. London: Continuum, 2007. 219–246.

MALMGREN 1993

Malmgren, Carl. "The Languages of Science Fiction: Samuel Delany's *Babel-17*." *Extrapolation* 34.1 (1993): 5–17.

MANDYS 1993

Mandys, Pavel. "Násilím na násilí." *Český deník* 1. Mar 1993: 10.

MANN 2001

“Burgess, Anthony. (1917–1993), Great Britain” *The Mammoth Encyclopedia of Science Fiction*. Ed. George Mann. 1 vol. London: Robinson, 2001.

MANNEWITZ 1999

Cornelia Mannewitz. “Anthony Burgess’ ‘A Clockwork Orange’: Zur lexikalischen Systematik im Original und in Übersetzungen.” *Interlinguistik und Lexikographie. Beiträge gehalten auf der 8. Jahrestagung der Gesellschaft für Interlinguistik e.V., 6. - 8. November 1998 in Berlin* (1999): 82–88.

MAREŠ 2003

Mareš, Petr. „Also: nazdar!“ *Aspekty textové vícejazyčnosti*. Praha: Karolinum, 2003.

MAREŠ 2012

Mareš, Petr. *Nejen jazykem českým. Studie o vícejazyčnosti v literatuře*. Praha: Filozofická fakulta Univerzity Karlovy, 2012.

MATRAS 2009

Matras, Yaron. *Language contact*. Cambridge: Cambridge University Press, 2009.

MAURANEN 2007

Mauranen, Anna. “Universal tendencies in Translation.” *Incorporating Corpora: The Linguist and the Translator*. Ed. Gunilla Anderman and Margaret Rogers. Clevedon: Multilingual Matters, 2007. 31–48.

MCDUGAL 2003

McDougal, Stuart Y. “‘What’s It Going to Be Then, Eh?’ Questioning Kubrick’s Clockwork.” *Stanley Kubrick’s A Clockwork Orange*. Ed. Stuart Y. McDougal. New York: Cambridge University Press, 2003. 1–18.

MCENERY, XIAO and TONO 2006

McEnery, Tony, Richard Xiao, and Yukio Tono. *Corpus-based language studies: an advanced resource book*. London: Routledge, 2006.

MORRISON 1996

Morrison, Blake. "Introduction." Anthony Burgess. *A Clockwork Orange*. London: Penguin, 1996. vii–xxiv.

MUYSKEN 2007

Muysken, Pieter. "Mixed Codes." *Handbook of multilingualism and multilingual communication*. Ed. Peter Auer and Li Wei. Berlin: Mouton de Gruyter, 2007. 316–339.

NICHOLLS 1993

Nicholls, Peter. "Linguistics." *The Encyclopedia of Science Fiction*. Ed. John Clute and Peter Nicholls. 1 vol. New York: St. Martin's Press, 1993.

OKRENT 2009

Okrent, Arika. *In the land of invented languages: Esperanto rock stars, Klingon poets, Loglan lovers, and the mad dreamers who tried to build a perfect language*. New York: Spiegel & Grau, 2009.

OKS 2009

Oks, Marina. "The Rebus of 'Nadsat,' or, A Key To *A Clockwork Orange*." *Textual Intricacies: Essays on Structure and Intertextuality in Nineteenth and Twentieth Century Fiction in English*. Ed. Christiane Bimberg and Igor Volkov. Trier: WVT Wissenschaftlicher Verlag Trier, 2009. 37–56.

OUŘEDNÍK 2002

Ouředník, Patrik. "Jak se Pantagruel seznámil s Panurgem, jehož pak miloval po celý život. (O třech imaginárních jazycích Françoise Rabelaise)." *Souvislosti* 13.2 (2002): 107–117.

PITTS, WHITE and KRASHEN 1989

Pitts, Michael, Howard White and Stephen Krashen. "Acquiring Second Language Vocabulary through Reading: A Replication of the Clockwork Orange Study Using Second Language Acquirers." *Reading in a Foreign Language* 5.2 (1989): 271–275.

RABINOWITZ 2003

Rabinowitz, Peter J. “‘A Bird of Like Rarest Spun Heavenmetal’ Music in *A Clockwork Orange*.” *Stanley Kubrick’s A Clockwork Orange*. Ed. Stuart Y. McDougal. New York: Cambridge University Press, 2003. 109–130.

RAY 1981

Ray, Philip E. “Alex Before and After: A New Approach to Burgess’ ‘A Clockwork Orange’.” *Modern Fiction Studies* 27.3 (1981): 479–487.

SCOTT and TRIBBLE 2006

Scott, Mike and Christopher Tribble. *Textual patterns: key words and corpus analysis in language education*. Amsterdam – Philadelphia: Benjamins, 2006.

SEBBA 2013

Sebba, Mark. “Multilingualism In Written Discourse: An Approach To The Analysis Of Multilingual Texts.” *International Journal Of Bilingualism* 17.1 (2013): 97–118.

STEINBECK 1995

Steinbeck, John. *Tortilla Flat*. London: Arrow Books, 1995.

STEINER 1998

Steiner, George. *After Babel: aspects of language and translation*. Oxford: Oxford University Press, 1998.

STEWART 2000

Stewart, Dominic. “Conventionality, Creativity and Translated Text: The Implications of Electronic Corpora in Translation.” *Intercultural Faultlines: Research Models in Translation Studies I : Textual and Cognitive Aspects*. Ed. Maeve Olohan. Manchester: St. Jerome Publishing, c2000. 73–91.

STOCKWELL 2006

Stockwell, P[eter]. “Invented language in Literature.” *Encyclopedia of Language and Linguistics*. Vol. 6. Ed. Keith Brown and Anne Anderson. 14 vols. Oxford: Elsevier, 2006.

TOLKIEN 1993

Tolkien, J. R. R. *The Lord of the Rings*. London: HarperCollins Publishers, 1993.

VO and CARTER 2010

Vo, Thuc Anh and Ronald Carter. "What can a corpus tell us about creativity?" *The Routledge Handbook of Corpus Linguistics*. Ed. Anne O'Keeffe and Michael McCarthy. London: Routledge, 2010. 302–316.

YAGUELLO 1991

Yaguello, Marina. *Lunatic Lovers of Language. Imaginary Languages and Their Inventors*. London: The Athlone Press, 1991.

ŽILINA 2006

Žilina, Miloslav. "Nad Mechanickým pomerančem Anthonyho Burgesse." *Literární noviny* 13 Nov 2006. 13.

Résumé

Diplomová práce se zabývá vnitrotextovou vícejazyčností (interjazykovou heterogeností) v románu *A Clockwork Orange* (ACO) britského spisovatele Anthonyho Burgesse a dvou jeho překladech, českém a německém. V románu se vyskytuje smyšlený jazyk nadsat založený převážně na ruském lexiku, který je nápadným kreativním prvkem románu. Práce analyzuje, jak se v překladech projevuje lexikální kreativita, tedy zda a jak se lexikální kreativita originálu liší od lexikální kreativity v jeho překladech. Změny lexikální kreativity jsou dány do souvislosti s normalizací (jednou z překladových univerzálií) a s funkcemi daného smyšleného jazyka. Analýza jazyka nadsat a jeho překladových protějšků je kvantitativní a využívá konkordanční programy AntConc a ParaConc. Při analýze se zkoumá frekvence slov z nadsat, jejich rozmístění v textu a způsob, jakým je jejich význam čtenáři zprostředkován. Získané údaje jsou poté použity ke srovnání jazyka nadsat a smyšlených jazyků, které ho v českém a německém překladu nahrazují. Práce sestává z úvodu, čtyř kapitol a závěru.

V první kapitole je představen koncept vícejazyčnosti literárního díla a smyšlených jazyků. Existující klasifikace forem a funkcí vícejazyčnosti v literárním díle (Mareš 2003) je nově aplikována na smyšlené jazyky. Smyšlené jazyky tedy nemusí být považovány za speciální formu vícejazyčnosti, jak je tomu u Mareše, ale mohou být naopak klasifikovány stejným způsobem jako běžné „cizí“ (sekundární) jazyky ve vícejazyčném literárním díle.

Primární formou sekundárního jazyka nadsat je dle této klasifikace tzv. prezence vícejazyčnosti, tedy situace, kdy slova uvedená v sekundárním jazyce v textu díla přímo odpovídají promluvě pronesené ve fikčním světě.

Sekundární forma jazyka nadsat se liší podle toho, o jakém typu implikovaného čtenáře (adresáta) uvažujeme. Pro čtenáře, který ovládá ruštinu do té míry, že je schopen rozluštit význam slov z nadsat, je sekundární formou tzv. hybridizace, tedy zapojení lexikálních jednotek sekundárního jazyka do morfologických a/nebo syntaktických struktur bázevého jazyka díla (primárního jazyka). Takový čtenář pak čte ACO podobně, jako by se jednalo o jednojazyčný text. Tento způsob čtení ale v této práci zkoumán není.

Druhým typem implikovaného čtenáře je čtenář, který ruštinu neovládá. Pro něj je sekundární formou tzv. oscilace. Při oscilaci je změna jazyka prezentována způsobem, který napodobuje míšení kódů ve skutečné komunikaci. Postava tak promlouvá jedním nebo

druhým jazykem podle svých (fikčních) komunikačních potřeb, přičemž v tomto případě jde o oscilaci mezi angličtinou (jakožto primárním jazykem) a neznámým smyšleným jazykem (jakožto sekundárním jazykem). Práce předpokládá, že koncept míšení kódů lze v pravém slova smyslu aplikovat pouze na přirozené jazyky, a proto s ním nepracuje.

V závěru první kapitoly jsou uvedeny funkce smyšlených jazyků. Vzhledem k tomu, že neexistuje jiná podobně detailní klasifikace přímo pro smyšlené jazyky, práce využívá Marešovu klasifikace funkcí vícejazyčnosti.

Druhá kapitola se zabývá funkcemi jazyka nadsat v primární komunikaci (tj. komunikaci mezi autorem a čtenářem, resp. podavatelem a adresátem). Nejprve podává základní informace o románu, ději a jeho struktuře.

Román byl poprvé vydán v roce 1962 ve Velké Británii a v roce 1963 v USA. Americká vydání ovšem až do roku 1988 obsahovala pouze 20 kapitol, o jednu méně než britská. Burgess a další později považovali britské vydání za jediné správné, ovšem Biswell dokazuje, že obě verze byly v době vzniku románu spíše rovnocenné (Biswell 2006 a 2012a). Pro analýzu bylo nicméně použito vydání britské s 21 kapitolami. V románu je představena fiktivní subkultura mladých výtržníků, s níž je spojeno právě používání jazyka nadsat. Jednou z funkcí nadsat je tedy funkce charakterizační.

Kapitola dále stručně popisuje místo a čas, v nichž se román odehrává, a okolnosti a především důvody vzniku jazyka nadsat i jejich vliv na jeho podobu a jeho funkce – děj ACO se odehrává v záměrně nespécifikovaném místě a čase, a román tak získává obecnou platnost, kterou nadsat dále prohlubuje tím, že se v něm prolínají jazyky obou tehdejších velmocí. Smyšlený jazyk Burgess zvolil také proto, že je nadčasový, narozdíl od slangu, který rychle zastarává, což bylo Burgessovi dříve vytýkáno. Obě tyto vlastnosti se spojují v atmosférotvornou funkci jazyka nadsat.

Burgess se obával, že velký počet neznámých smyšlených slov a tempo, s jakým je nadsat čtenáři prezentován, budou čtenáře odrazovat. K tomu také skutečně došlo, ale pozitivní ohlasy postupně převážily nad negativními. Rozdílné názory existují také na to, zda se ACO čte lehce nebo obtížně. Na začátku románu si čtenář může od jazyka nadsat právě kvůli velkému množství neznámých slov udržovat odstup, ale tím, jak je s ním čím dál tím více v kontaktu, se původní odcizení proměňuje v identifikaci s hlavním hrdinou i jeho jazykem. Hodnotová funkce jazyka nadsat se tedy v průběhu textu mění.

K americkému vydání začal být záhy připojován slovníček jazyka nadsat (a dalších méně známých slov), ovšem původním Burgessovým záměrem bylo, aby se čtenář nadsat

postupně učil. Nadsat měl podle něj také částečně skrývat či zmírňovat násilí, které je v románu zobrazováno. V tomto smyslu plní nadsat funkci utajovací a restriktivní. Aby se čtenář mohl nadsat skutečně učit, jsou mnohé výrazy z něj často opakovány. V ACO hraje důležitou roli i opakování motivů či analogie mezi jednotlivými scénami.

Třetí kapitola analyzuje nadsat v ACO kvantitativně, pomocí konkordančního programu AntConc. Nejdříve je popsáno, jakým způsobem byl vytvořen seznam slov z nadsat, která jsou dále analyzována, a z jakých důvodů do něj některá slova byla zahrnuta a jiná ne. Analyzováno je 180 lemmat, z nichž 60 jsou hapax legomena. Při analýze se zkoumá jejich frekvence, rozmístění v textu a způsob, jakým je jejich význam čtenáři zprostředkován. Těmto 180 lemmatům (typům) odpovídá 3401 výskytů (tokenů). ACO obsahuje 58563 slov, a slova z nadsat tak tvoří 5,8 % slov v románu.

Absolutní frekvence slov z nadsat se pohybuje v rozmezí od 1 výskytu do 240 výskytů. Distribuce frekvence slov (rozložení četností) má podobnou tendenci, jakou lze nalézt v přirozených jazycích – velké množství slov s nízkou frekvencí výskytu a malé množství slov s vysokou frekvencí.

AntConc disponuje nástrojem Concordance Plot, který zobrazuje rozmístění jednotlivých výskytů hledaného lemmatu v textu. Analýza rozmístění výskytů (tokenů) slov z nadsat v ACO neukazuje, že by se na některých místech výskyty kumulovaly. Samostatná analýza hapax legomena (60 lemmat) ovšem ukazuje, že naprostá většina z nich se vyskytuje v první polovině textu, a že v poslední třetině textu se vyskytuje pouze jedno hapax legomenon. Jednotlivá slova z nadsat jsou ovšem v textu pochopitelně rozmístěna různě.

Slovní tvary některých lemmat z nadsat se vyskytují spíše ve shlucích na jednom či více místech v textu, zatímco jiné se vyskytují víceméně rovnoměrně, resp. tak, že shluky netvoří. Přestože přechod mezi oběma skupinami je pozvolný, lze říci, že 23 lemmat z nadsat má tendenci se shlukovat. Většina z nich vytváří koreferenční řetězce, ovšem jednotlivé výskyty bývají jen zřídka součástí jediného řetězce – spíše se vytváří dva nebo tři kratší řetězce, přičemž mnoho výskytů zůstává i tak mimo ně. Shluky často přímo souvisí s určitým prostředím, v němž se děj odehrává, nebo s určitým tématem či scénou. Díky opakování určitých slov v podobných situacích si čtenář tyto situace snáze spojí a také lépe pochopí význam příslušného slova z nadsat.

V ACO je význam slov z nadsat čtenáři přiblížen různými způsoby. Nejjednoznačnější je vnitrotextový překlad, kdy po slovu z nadsat hned následuje anglický

ekvivalent. Vnitrotextově jsou přeložena 2 hapax legomena (z 60) a 16 lemmat s více než jedním výskytem (tedy 13 % ze této skupiny 120 slov).

Čtenář může pochytit význam slova z nadsat také dalšími způsoby, pomocí kontextu, ko-textu nebo koreferenčních řetězců, které obsahují i anglická slova. U 32 slov z nadsat s více než jedním výskytem (tedy cca jedné čtvrtiny z nich) může čtenář význam odhadnout jen obtížně. V této skupině jsou především slova s nižší frekvencí, která nejsou vnitrotextově přeložena, a kterým čtenář plně neporozumí ani pomocí kontextu, ko-textu či koreference.

Výše zkoumané vlastnosti jazyka nadsat jsou v závěru třetí kapitoly vztaženy k jeho funkcím. Pro charakterizační a atmosférotvornou funkci je podstatný počet lemmat smyšleného jazyka a jejich frekvence. Pro hodnotovou funkci je to frekvence a rozmístění slov nadsat v textu a také způsob, jakým je čtenáři význam slov z nadsat zprostředkován. Pro utajovací a restriktivní funkci je podstatné, do jaké míry je čtenáři význam slov z nadsat zprostředkován.

Ve čtvrté kapitole jsou údaje získané v kapitole třetí použity ke srovnání jazyka nadsat a smyšlených jazyků, které ho v českém a německém překladu nahrazují. Výchozím textem českého překladu je 4. vydání Mechanického pomeranče v překladu Ladislava Šenkyříka (MP). Výchozím textem německého překladu je první vydaná verze Uhrwerk Orange v překladu Waltera Brumma (UO). MP obsahuje 52465 slov, z toho 3294 slov je ze smyšleného jazyka týnů; tato slova tedy tvoří 6,27 % celého textu. UO obsahuje 62121 slov, z toho 2379 slov je ze smyšleného jazyka Nadsatsprache; tato slova tedy tvoří 3,83 % celého textu. Pro srovnání originálu a obou překladů byl použit konkordanční program ParaConc. V ParaConcu je možné téměř automaticky zarovnávat paralelní texty. Všechny tři texty byly nejprve upraveny v programu Microsoft Word, aby odstavce obou překladů odpovídaly originálu. Texty pak byly nahrány do ParaConcu a takto vzniklý korpus byl postupně zarovnán podle odstavců a podle vět.

Při vyhledávání překladových protějšků k lemmatům z nadsat bylo využito funkce Hot words, která automaticky nabízí pravděpodobné protějšky hledaného výrazu (často několik tvarů jednoho slova). Věty, v nichž program žádné zvolené „hot word“ nenašel, byly analyzovány manuálně, aby se zjistilo, jakým způsobem bylo hledané lemma z nadsat přeloženo v cílovém textu. U každého lemmatu z nadsat se takto zjistilo, v kolika případech bylo dané lemma v cílovém textu přeloženo odpovídajícím výrazem (jehož slovní tvar se

v takovém případě mezi „hot words“ nedostal), v kolika případech jiným výrazem (včetně toho, o jaký výraz se jedná), či zda nebylo přeloženo vůbec.

Toto vyhledávání bylo následně provedeno v obráceném směru – ke každému takto identifikovanému slovu ze smyšleného jazyka cílového textu byly vyhledány odpovídající výrazy v textu zdrojovém (v ACO), aby se zjistilo, zda dané slovo ze smyšleného jazyka cílového textu nebylo použito i pro překlad jiných výrazů z ACO. Např. slovo *yahzick* se ve zdrojovém textu vyskytuje čtyřikrát a v MP jsou všechny jeho výskyty přeloženy slovem *ještěk*, ale slovo *ještěk* se v MP vyskytuje celkem pětkrát – pátý výskyt je překladem běžného anglického výrazu, což se ukázalo jako častý jev v obou překladech. Údaje o frekvencích a překladových protějšcích byly kategorizovány pomocí programu Microsoft Excel.

Při analýze byla opět zkoumána frekvence slov ze smyšlených jazyků obou cílových textů, jejich rozmístění v textu a způsob, jakým je jejich význam čtenáři zprostředkován. Získané údaje byly porovnávány s údaji o jazyce nadsat ze třetí kapitoly práce. Rozdíly v počtu slov ze smyšlených jazyků byly zjišťovány na úrovni lemmat a slovních tvarů. Analýza ukázala, že v obou překladech je smyšlených lemmat méně než v originále, přičemž v německém překladu (UO) je jich mnohem méně než v českém (MP). Český přišel celkově o 10 % lemmat z Nadsat a německý o 41 %.

Rozdíly v rozmístění jednotlivých tvarů smyšlených slov v textu byly popsány podle toho, kolik příslušných tokenů se v cílovém textu shoduje, tedy se nachází na stejném místě (ve stejné větě) jako ve zdrojovém textu, kolik jich v cílovém textu chybí a kolik jich v něm naopak přibýlo. Tato odchylka pak byla kvantifikována jako podíl součtu chybějících a přidaných tokenů a celkového počtu tokenů, tedy součtu chybějících, přidaných a shodných tokenů. Rozmístění jednotlivých tvarů smyšlených slov v textu se v UO změnilo více než v MP, přičemž v UO se změny týkají většího počtu jednotlivých tvarů smyšlených slov než v MP a jsou také zásadnější.

Změna způsobu, jakým je význam slov smyšlených jazyků čtenáři zprostředkován, byla posuzována na změnách ve vnitrotextovém překladu dvou hapax legomena a 16 lemmat s více než jedním výskytem. V MP nebyla vnitrotextově přeložena čtyři lemmata. V UO nebyla vnitrotextově přeložena také čtyři lemmata, ale kromě toho nebylo pro dalších pět lemmat vůbec použito slovo ze smyšleného jazyka, proto k nim neexistuje ani vnitrotextový překlad. Další čtyři lemmata jsou ve zdrojovém textu vnitrotextově přeložena opakovaně, ale v UO je vnitrotextový překlad uveden vždy pouze u jednoho výskytu. Strategie vnitrotextového překladu slov z Nadsat je tedy více dodržována v MP než v UO.

Závěr práce shrnuje výsledky analýzy a vlivy změn na lexikální kreativitu, normalizaci a funkce smyšleného jazyka v cílových textech. Lexikální kreativita je oslabena v obou překladech, ale mnohem více v německém, přičemž normalizace smyšlených slov vzrostla nepřímo úměrně k tomuto oslabení. Změny v překladu měly vliv na funkce jazyka nadsat. V MP nepůsobí důsledky změn průkazně – nejvíce byla patrně ovlivněna funkce utajovací a restriktivní, ale mnohé změny lze vysvětlit nepochopením originálu a nedbalou redakční prací; ostatní funkce nebyly výrazně ovlivněny. V UO byla výrazně ovlivněna funkce atmosférotvorná, charakterizační i hodnotová. Změny funkce utajovací a restriktivní nemohly být zhodnoceny, protože tato funkce byla příliš ovlivněna mnoha dalšími faktory.

Appendix I – Nadsat (ACO) lemmas alphabetically

lemma	freq.	lemma	freq.	lemma	freq.	lemma	freq.
baboochka	15	gooly	2	neezhnies	1	scoteena	3
banda	1	gorlo	2	noga	33	shaika	1
bezoomny	45	govoreet	48	nochy	18	sharries	8
bitva	2	grahzny	31	nozhh	14	shest	1
Bog	31	grazzy	4	nuking	1	shiyah	1
bolnoy	2	gromky	19	oddy knocky	19	shlapa	1
bolshy	51	groody	15	odin	1	shlem	5
brat	6	gruppa	1	okno	1	shoom	23
bratchny	20	gulliver	70	oobivat	1	shoot	3
britva	25	horrorshow	109	ookadeet	1	skazat	1
brookoo	1	interessovat	2	ookoo	7	skorry	51
brosat	1	itty	52	oomny	1	skvat	1
bugatty	1	jeezny	16	oozhassny	1	sladky	2
cal	65	kartoffel	2	oozy	6	sloochat	2
cantora	5	keeshkas	5	osoosh	1	sloosh	76
carman	19	kleb	5	otchkies	6	slovo	45
chasha	5	klootch	3	peet	30	smeck	75
chasso	24	knopka	2	pishcha	13	smot	3
cheena	7	kopat	2	platch	1	sneety	2
cheest	4	koshka	14	platties	41	sobirat	1
chelloveck	240	kot	10	plenny	18	soomka	1
chepooka	1	krovvy	37	plesk	1	soviet	1
choodessny	1	kupet	4	pletcho	17	spat	1
cluve	1	lapa	1	plott	30	spatchka	2
collocoll	2	lewdies	24	podooshka	1	starry	105
crast	14	litso	77	pol	1	strack	2
creech	105	lomtick	12	polezny	1	tally	2
dama	1	lovet	5	pony	7	tolchock	68
ded	1	lubbilubbing	1	poogly	4	toofles	6
deng	7	malenky	99	pooshka	2	tree	1
devotchka	47	malchick	56	prestoopnik	13	vareet	1
dobby	1	maslo	1	privodeet	1	veshch	72
domy	7	merzky	1	ptitsa	54	viddy	230
dorogoy	1	messel	1	pyahnitsa	1	voloss	4
dratsing	18	mesto	31	rabbit	10	von	63
droog	94	millicent	48	radosty	1	vred	1
dva	1	molodoy	1	raskazz	1	yahma	1
eegra	1	moloko	20	rassoodock	2	yahzick	4
eemya	3	moodge	9	raz	4	yeckate	1
forella	5	morder	1	razdraz	10	zammechat	1
gazetta	27	mozg	2	razrez	9	zasnoot	1
glazz	76	nadmenny	1	rook	89	zheena	1
gloopy	16	nadsat	14	rot	72	zoobies	23
goloss	69	nagoy	15	sabog	3	zvonock	2
goober	5	nachinat	1	sakar	1	zvook	2

Appendix II – Nadsat (ACO) lemmas by frequency

lemma	freq.	lemma	freq.	lemma	freq.	lemma	freq.
chelloveck	240	carman	19	poogly	4	lubbilubbing	1
viddy	230	gromky	19	raz	4	maslo	1
horrorshow	109	oddy knocky	19	voloss	4	merzky	1
creech	105	dratsing	18	yahzick	4	messel	1
starry	105	nochy	18	eemya	3	molodoy	1
malenky	99	plenny	18	klootch	3	morder	1
droog	94	pletcho	17	sabog	3	nadmenny	1
rook	89	gloopy	16	scoteena	3	nachinat	1
litso	77	jeezny	16	shoot	3	neezhnies	1
glazz	76	baboochka	15	smot	3	nuking	1
sloosh	76	groody	15	bitva	2	odin	1
smeck	75	nagoy	15	bolnoy	2	okno	1
rot	72	crast	14	collocoll	2	oobivat	1
veshch	72	koshka	14	gooly	2	ookadeet	1
gulliver	70	nadsat	14	gorlo	2	oomny	1
goloss	69	noz	14	interessovat	2	oozhassny	1
tolchock	68	pishcha	13	kartoffel	2	osoosh	1
cal	65	prestoopnik	13	knopka	2	platch	1
von	63	lomtick	12	kopat	2	plesk	1
malchick	56	kot	10	mozg	2	podooshka	1
ptitsa	54	rabbit	10	pooshka	2	pol	1
itty	52	razdraz	10	rassoodock	2	polezny	1
bolshy	51	moodge	9	sladky	2	privodeet	1
skorry	51	razrez	9	sloochat	2	pyahnitsa	1
govoreet	48	sharries	8	sneety	2	radosty	1
millicent	48	cheena	7	spatchka	2	raskazz	1
devotchka	47	deng	7	strack	2	sakar	1
bezoomny	45	domy	7	tally	2	shaika	1
slovo	45	ooko	7	zvonoock	2	shest	1
platties	41	pony	7	zvook	2	shiyah	1
krovvy	37	brat	6	banda	1	shlapa	1
noga	33	oozy	6	brooko	1	skazat	1
Bog	31	otchkies	6	broosat	1	skvat	1
grahzny	31	toofles	6	bugatty	1	sobirat	1
mesto	31	cantora	5	chepooka	1	soomka	1
peet	30	chasha	5	choodessny	1	soviet	1
plott	30	forella	5	cluve	1	spat	1
gazetta	27	goober	5	dama	1	tree	1
britva	25	keeshkas	5	ded	1	vareet	1
chasso	24	kleb	5	dobby	1	vred	1
lewdies	24	lovet	5	dorogoy	1	yahma	1
shoom	23	shlem	5	dva	1	yeckate	1
zoobies	23	cheest	4	eegra	1	zammechat	1
bratchny	20	grazzy	4	gruppa	1	zasnoot	1
moloko	20	kupet	4	lapa	1	zheena	1

Appendix III – JT (MP) lemmas alphabetically

lemma	freq.	lemma	freq.	lemma	freq.	lemma	freq.
ajko	75	ganka	2	lubilubikovat se	1	sojlovatý	3
ancintní	1	gard	24	lukovat	215	sovjet	1
babooshka	15	gejmka	1	lunatický	47	spačka	3
bajat	54	gejt	1	maladaja	1	stoulnout	11
balnoj	2	gorlo	13	many	5	stupído	15
batl	2	gražny	30	milicjant	47	sumka	1
bejk	1	gromkový	18	nadměnný	1	šajka	1
bel	2	grupa	1	najf	16	šary	8
bérový	15	gud	1	najt	17	šlág	22
bitkový	84	guljat	28	nížnys	1	šlapa	1
blažka	37	gulliver	69	nojz	22	šlema	5
bog	32	házbend	3	očky	6	šnout	1
bolšácký	45	hedka	3	odinočky	21	špér	2
bračny	20	hendka	89	odlidovali	1	šugr	2
bred	5	hepenout se	2	oulđaný	106	tajm	6
brejny	1	hírovat	75	ovošnout se	2	tejl	1
bresty	15	hjúmaník	210	pejpry	26	teřich*	4
brosít	1	horibl	2	píply	26	tolčok	67
brotek	6	houla	1	pišča	12	tolkovat	46
buč	71	houm	7	plac	31	trí	1
bygin	1	chorošný	102	plečo	18	tu	1
cány	23	ikspenzivní	1	pleník	19	tufle	6
čajina	46	ímja	3	pležr	1	týn	15
časó	1	interesovat*	2	plot'	30	vajfka	1
čejn	10	ír	7	ponímat	8	van	1
čina	8	ješték	5	popugat	5	verd	41
čipuka	1	júsfl	1	potejtový	1	vípovat	1
čuděsný	1	kantora	5	prestúpnik	13	vojs	69
decidovat se	2	kápek	5	ptica	52	volkovat	24
drajovat	1	karman	17	purčejnout	4	vyklínsovat	2
drajvovat	2	kečnout	5	pylou	1	windou	1
dresy	40	ketek	5	rabbitit	10	zakilovat	1
drím	2	ketka	15	rejzr	25	zapsetovat	7
drinkat	30	kiška	7	remkový	1	zgrabovat	1
drunkard	1	knopka	2	ričatý	1	zharmovat	1
dygovat	2	kolokolčik	2	ripovat	16	* Unfortunately, these two words should not have been included in the list of invented words, as they are used in their original Czech meaning in MP.	
džvenk	3	kúkovat	1	robnout	1		
fancy	1	kviklý	51	saboga	3		
fejs	79	láfat se	63	silykový	1		
fír	1	lajf	17	skotyna	3		
fisíz	65	lapka	1	skrímat	106		
forela	5	lejdýz	1	slípnout si	1		
frendík	94	lips	80	smajlnout	4		
fútka	35	litlko	1	smel	64		
ganga	1	lomtyk	10	smotrovat	1		

Appendix IV – JT (MP) lemmas by frequency

lemma	freq.	lemma	freq.	lemma	freq.	lemma	freq.
lukovat	215	odinočky	21	smajlnout	4	ikspenzivní	1
hhumaník	210	bračny	20	teřich*	4	jůsfl	1
ouldaný	106	pleník	19	džvenk	3	kúkovat	1
skrímat	106	gromkový	18	házbend	3	lapka	1
chorošný	102	plečo	18	hedka	3	lejdýz	1
frendík	94	karman	17	ímja	3	litlko	1
hendka	89	lajf	17	saboga	3	lubilubikovat se	1
bitkový	84	najt	17	skotyna	3	maladaja	1
lips	80	najf	16	sojlovatý	3	nadměnný	1
fejs	79	ripovat	16	spačka	3	nížnýs	1
ajko	75	babooshka	15	balnoj	2	odlídovali	1
hírovat	75	bérový	15	batl	2	pležr	1
buč	71	bresty	15	bel	2	potejtový	1
gulliver	69	ketka	15	decidovat se	2	pylou	1
vojs	69	stupído	15	drajvovat	2	remkový	1
tolčok	67	týn	15	drím	2	ričatý	1
físíz	65	gorlo	13	dygovat	2	robnout	1
smel	64	prestůpník	13	ganka	2	silykový	1
láfát se	63	pišča	12	hepenout se	2	slípnout si	1
bajat	54	stoulnout	11	horibl	2	smotrovat	1
ptica	52	čejn	10	interesovat*	2	sovjet	1
kviklý	51	lomtyk	10	knopka	2	sumka	1
lunatický	47	rabbitit	10	kolokolčik	2	šajka	1
milicjant	47	čina	8	ovošnout se	2	šlapa	1
čajina	46	ponímat	8	špér	2	šňout	1
tolkovat	46	šáry	8	šugr	2	tejl	1
bolšácký	45	houm	7	vyklínsovat	2	trí	1
verd	41	ír	7	ancintní	1	tu	1
dresy	40	kiška	7	bejk	1	vajfka	1
blažka	37	zapsetovat	7	brejny	1	van	1
fútka	35	brotek	6	brostit	1	vípoval	1
bog	32	očky	6	bygin	1	windou	1
plac	31	tajm	6	časo	1	zakilovat	1
drinkat	30	tufle	6	čipuka	1	zgrabovat	1
gražny	30	bred	5	čuděsný	1	zharmovat	1
plot'	30	forela	5	drajovat	1	* Unfortunately, these two words should not have been included in the list of invented words, as they are used in their original Czech meaning in MP.	
guljat	28	ještěk	5	drunkard	1		
pejpry	26	kantora	5	fancy	1		
píply	26	kápek	5	fír	1		
rejzr	25	kečnout	5	ganga	1		
gard	24	ketek	5	gejmka	1		
volkovat	24	many	5	gejt	1		
cány	23	popugat	5	grupa	1		
nojz	22	šlema	5	gud	1		
šlág	22	purčejznout	4	houla	1		

Appendix V – Nadsatsprache (UO) lemmas alphabetically

lemma	freq.	lemma	freq.	lemma	freq.
Babuschka	22	Malitschickiwick	51	Spatschka	2
Banda	1	Maslo	1	stari	103
bezumnie	51	merzki	4	Sumka	1
Bog	28	Mesto	33	Sung	18
bolnoy	2	molodoi	2	Titsa	52
bolschig	45	Moloko	26	tollschock	59
Bratschni	21	Motschka	7	troms	1
Bratti	6	Mozg	4	Tschasso	23
Britva	23	Murke	3	Tschelloveck	279
Deng	16	Nadsat	10	Tschipoka	1
Dewotschka	49	nagoi	15	tschudesny	1
dorogoi	3	Noga	29	vredde	1
dratsen	15	Notschi	17	Wetsch	29
Droog	88	Nozh	12	zammechat	1
Dschizny	12	Okno	1	Zubi	29
Durak	1	Otschky	6	Zvonock	1
ehn	1	parz	1		
Gazetta	26	pitschen	25		
gespattet	1	Platties	44		
glupig	17	Plenni	21		
Goloss	70	Pletscho	17		
govorit	57	Plotti	26		
graznig	32	polezni	1		
gromkig	22	Prestupnik	11		
Grudies	13	puglig	5		
Gruppa	1	Puschka	2		
Guber	7	Radosty	2		
Gulliver	74	Rampferei	1		
horrorchau	91	razdraz	12		
Kischkas	9	rizrazzen	11		
Kleb	5	Sabog	6		
Klutsch	2	Sakar	1		
Knopka	3	Schaikas	1		
Kollokoll	1	Schlapper	2		
Kontora	5	Schlemmie	6		
Koschka	14	Schum	15		
Kot	9	skorri	47		
krasten	11	Skutina	2		
Krowy	34	skvattete	1		
kupetten	4	sladki	3		
Litso	84	Slovo	46		
Lolli	1	sluschten	73		
Lomtick	3	Smeck	86		
lubbilubben	1	smottete	10		
malenki	96	sparten	2		

Appendix VI – Nadsatsprache (UO) lemmas by frequency

lemma	freq.	lemma	freq.	lemma	freq.
Tschelloveck	279	Koschka	14	tschudesny	1
stari	103	Grudies	13	parz	1
malenki	96	Dschizny	12	Gruppa	1
horrorchau	91	Nozh	12	lubbilubben	1
Droog	88	razdraz	12	Maslo	1
Smeck	86	krasten	11	ehn	1
Litso	84	Prestupnik	11	Okno	1
Gulliver	74	rizrassen	11	polezni	1
sluschten	73	Nadsat	10	Sakar	1
Goloss	70	smottete	10	Schaikas	1
tollshock	59	Kischkas	9	skvattete	1
govorit	57	Kot	9	Sumka	1
Titsa	52	Motschka	7	gespattet	1
Malitschickiwick	51	Guber	7	troms	1
bezumnie	51	Schlemmie	6	vredde	1
Dewotschka	49	Sabog	6	zammechat	1
skorri	47	Bratti	6		
Slovo	46	Otschky	6		
bolschig	45	puglig	5		
Platties	44	Kontora	5		
Krowy	34	Kleb	5		
Mesto	33	Mozg	4		
graznig	32	kupetten	4		
Wetsch	29	merzki	4		
Noga	29	Lomtick	3		
Zubi	29	Knopka	3		
Bog	28	sladki	3		
Plotti	26	Murke	3		
Gazetta	26	dorogoi	3		
Moloko	26	Schlapper	2		
pitschen	25	Klutsch	2		
Britva	23	Skutina	2		
Tschasso	23	bolnoy	2		
gromkig	22	Puschka	2		
Babuschka	22	Spatschka	2		
Bratschni	21	molodoi	2		
Plenni	21	Radosty	2		
Sung	18	sparten	2		
Notschi	17	Lolli	1		
Pletscho	17	Durak	1		
glupig	17	Rampferi	1		
Deng	16	Kollokoll	1		
Schum	15	Zvonock	1		
dratsen	15	Banda	1		
nagoi	15	Tshipoka	1		

