Moravians in Prague: A Sociolinguistic Study of Dialect Accommodation in the Czech Republic

Volume 1

James Wilson

September 2007

Thesis submitted for the degree of Doctor of Philosophy

Department of Russian and Slavonic Studies, University of Sheffield

Abstract

The thesis reports on the linguistic accommodation of 39 university students from Moravia (the eastern half of the Czech Republic) living in Prague, Bohemia (the western part of the Czech Republic). In Bohemia, the informants' highly-localized native dialects and Standard Czech (SC) – a semi-artificial, archaic and primarily non-spoken standard with no native speakers – are both stigmatized, although for different reasons. Consequently, it has been 'hypothesized' that speakers of Moravian dialects living in Bohemia quickly reduce the frequency of or avoid stigmatized variants of their localized vernaculars and converge towards the host dialect, Common Czech (CC). Although a non-standard variety, CC is a semi-prestigious koine that is socially unrestricted throughout Bohemia and parts of western Moravia and is, according to some linguists, assuming the role of a national vernacular. However, the 'contact hypothesis' is based solely on introspective data and is ideologically driven, insomuch as it is the product of linguists who promote the social and geographical spread of CC.

The present study is the first attempt to systematically describe the results of dialect contact between speakers of CC and Moravian dialects and to test the above hypothesis. To my knowledge, it is the first systematic variationist account of language variation in the Czech Republic. The study combines a quantitative analysis of six linguistic variables with both qualitative and ethnographic research and it identifies to what extent speakers of Moravian dialects living in Prague assimilate CC forms, what route their accommodation takes, and which variants of the host variety are most likely to acquired or rejected. A primary aim of the study is to describe the impact of a set of independent social variables on speakers' assimilation of CC forms. Special attention is accorded to speakers' sex, region of origin, length of residence in the host community and network integration.

Contents

Volume 1

Volume 2

Tables

Figures

Acknowledgements

First of all, I would like to thank the Arts and Humanities Research Council (AHRC), The Leverhulme Trust and the University of Sheffield for their financial support, without which this study could not have been undertaken. In 2002, I received a University of Sheffield Fee Bursary (1 + 3 Incentive Awards Scheme), which covered tuition fees for both my MA and PhD degrees. In 2004, I was awarded both a Leverhulme Trust Study Abroad Studentship, which allowed me to spend a full year carrying out fieldwork in Prague and in addition covered costs for recording equipment and transcriptions, and an AHRC Studentship that covered maintenance costs in 2005 – 2006.

I would also like to thank my supervisors – Professor Neil Bermel at the Department of Russian and Slavonic Studies (University of Sheffield) and Professor Joan Beal at the National Centre for English Cultural Tradition (University of Sheffield) – for their advice and suggestions throughout the course of the PhD and for their patient reading of early drafts of the thesis. I am deeply grateful to Docent Jiří Nekvapil for his help and advice during my year in Prague and for affording me the opportunity to present my research at Charles University in Prague to groups of undergraduate and graduate students, whose comments about my research were most helpful. I would also like to express my appreciation to those at the Czech Language Institute, Academy of Sciences of the Czech Republic who transcribed some of the interviews – especially to Martin Havlík, who was kind enough to check the Czech-language examples that are used in the thesis.

I would like to take this opportunity to thank all my friends and acquaintances at Kajetánka (the university hall of residence at which the fieldwork was carried out) for their help and support, in particular Kateřina Piklová for her assistance in the

initial stages of recruiting the informants. I am especially indebted to my fiancée Dr Markéta Hnátková, both for her willingness to record the sociolinguistic interviews on my behalf despite a hectic work schedule and often at very short notice and / or following a night shift at the hospital and for her understanding and patience when interviews were cancelled at the last minute. Last, but by no means least, a special thank you goes to all my informants for their time and enthusiasm in both the recorded interviews and in our informal meetings, and, most of all, for making the fieldwork fun.

Transcription conventions

With the aim of making the thesis accessible to both Bohemists and (socio)linguists interested either in the Czech language situation or in dialect contact, I use both simplified Czech phonetic transcription and symbols of the International Phonetic Alphabet (IPA). The latter are used in discussions of phonology and pronunciation or in cases where detailed transcription is required. In all other cases, I adhere to a simplified system of transcription according to which speech is transcribed using characters from the standard Czech orthography, as used in other Bohemistic studies and by transcribers at the Czech Language Institute, Academy of Sciences of the Czech Republic (*Ústav pro jazyk český, Akademie věd České republiky*). For readers who are unfamiliar with Czech, the Czech orthographic characters are listed in tables alongside their IPA representations and a list of basic pronunciation rules is provided on the following pages. Although the Czech writing system is not entirely phonetic, in which every sound is realized exactly how it is written, Czech spelling is much closer to pronunciation than, say, that of English. Therefore, having consulted the relevant materials, readers with no prior knowledge of Czech should be able to recognize the Czech-language examples cited in the thesis. All the Czech-language examples including individual words and phrases are translated into English.

Czech consonants and their IPA phonemic representations

C	IPA	No	Description
b	b	102	voiced bilabial plosive
c	t͡s	211	voiceless alveolar fricative
č	t͡ʃ	213	voiceless postalveolar affricate
d	d	104	voiced alveolar plosive
ď	ɟ	108	voiced palatal plosive
f	f	128	voiceless labiodental fricative
g	g	110	voiced velar plosive
h	ɦ	147	voiced glottal fricative
ch	x	140	voiceless velar fricative
j	j	153	voiced palatal approximant
k	k	109	voiceless velar plosive
l	l	155	voiced alveolar lateral approximant
m	m	114	voiced bilabial nasal
n	n	116	voiced alveolar nasal
ň	ɲ	118	voiced palatal nasal
p	p	101	voiceless bilabial plosive
q	k	109	voiceless velar plosive
r	r	122	voiced alveolar trill
ř	r̝	122 + 429(̝)	voiced alveolar trill fricative
s	s	132	voiceless alveolar fricative
š	ʃ	134	voiceless postalveolar fricative
t	t	103	voiceless alveolar plosive
ť	c	107	voiceless palatal plosive
v	v	129	voiced labiodental fricative
w	v	129	voiced labiodental fricative
x	ks	109 + 132	voiceless velar plosive + voiceless alveolar fricative
z	z	133	voiced alveolar fricative
ž	ʒ	135	voiced postalveolar fricative

C = consonant
IPA = IPA symbol
No = IPA number

Czech vowels and their IPA phonemic representations

1. Short vowels

V	IPA	No	Description
a	a	304	open front unrounded vowel
e	ɛ	303	open-mid front unrounded vowel
i	ɪ	310	near-close near-front unrounded vowel
o	o	307	close-mid back rounded vowel
u	u	308	close back rounded vowel
y	ɪ	310	near-close near-front unrounded vowel

2. Long vowels

V	IPA	No	Description
á	aː	304	open front unrounded vowel
é	ɛː	303	open-mid front unrounded vowel
í	iː	301	close front unrounded vowel
ó	oː	307	close-mid back rounded vowel
ů	uː	308	close back rounded vowel
ú	uː	308	close back rounded vowel
ý	iː	301	near-close near-front unrounded vowel

V = vowel

Notes

The graphemes <i> and <y> are both realized [ɪ] and their long counterparts <í> and <ý> are both pronounced [iː]. Although orthographically distinct, long <ú> and <ů> have the same pronunciation: [uː].

Short vowels are distinguished from their long counterparts only by the duration of their articulation, with the exception of long [iː], which is noticeably higher than short [ɪ] (see the vowel inventory on the following page).

The grapheme <ě> is realized [jɛ]: IPA 153 + 303.

Besides the five short (a, ɛ, ɪ, o, u) and five long (aː, ɛː, iː, oː, uː) vowels Czech has

seven glide diphthongs in Czech (**D** = diphthong):

D	IPA	D	IPA	D	IPA	D	IPA
aj	aj	oj	oj	au	au	ou	ou
ej	ɛj	uj	uj	eu	ɛu		

Vowel inventory of Czech

[from Dankovičová 1999: 70]
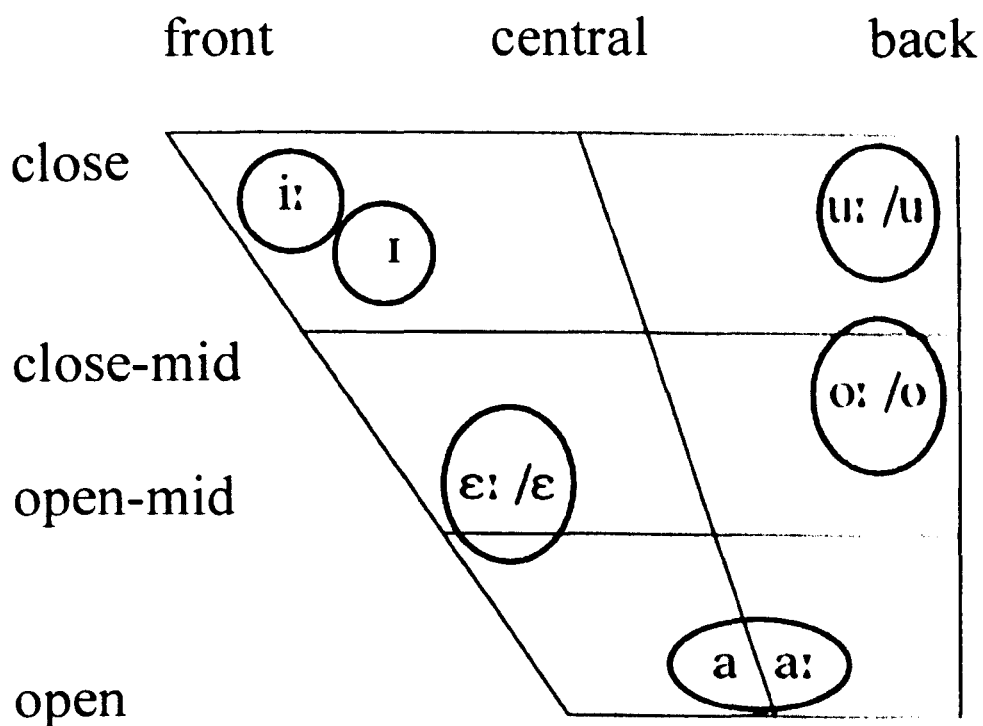
Some basic pronunciation rules

● The consonants <d>, <n> and <t> are realized [ɟ], [ɲ] and [c] before the vowels <i>

and <í>: *vtip* 'joke' [fcɪp].

● Czech has two syllabic liquids: [l̩] and [r̩]. This means – to the bemusement of many

foreigners who study Czech – that Czech has many words without vowels, two of the

longest being *scvrnkls* [st͡svr̩ŋkl̩s] from the verb *scvrknout* 'to flick away' and

čtvrthrst [ʃtvr̩dɦir̩st] 'a quarter of a handful'. Also note the tongue-twister *strč prst*

skrz krk [str̩t͡ʃ pr̩st skr̩s kr̩k] 'stick your finger through your throat'.

● Before the velars /g/ and /k/ the consonant <n> is pronounced [ŋ], as in the example

banka 'bank' ['baŋka].

● In Standard Czech (and most other dialects), stress is always placed on the first

syllable. In prepositional phrases, the preposition takes the stress (if it contains a

vowel), as in the example *do banky* 'to the bank' ['do baŋkɪ].

● Some Czech consonants undergo voice assimilation. Voice assimilation in Czech is

often not transcribed, unless of course it is a feature under study. In Standard Czech,

voiced consonants undergo a process called 'devoicing' when they are positioned

either at the end of a word or before a voiceless consonant in a consonant cluster.

Therefore, *chléb* 'bread' is realized [xlɛːp] not [xlɛːb], unless the following word

beings with a voiced consonant *chléb byl starý* 'the bread was old' ['xlɛːb bɪl 'stariː],

and *hubka* 'sponge' is always pronounced ['ɦupka] and not *['ɦubka]. Alternatively, if a voiceless consonant precedes a voiced consonant in a consonant cluster or over a word boundary, then the opposite process – 'devoicing' – occurs: *fotbal* 'football' is therefore realized ['fodbal] and not *['fotbal]. In the majority of cases, in Standard Czech, voice assimilation is in the backward direction and is often termed 'regressive assimilation'. The above processes are commonly observed across the voiced – voiceless pairs that are presented in the table below:

Voiced	Voiceless	Voiced	Voiceless
b /b/	p /p/	g /g/	k /k/
d /d/	t /t/	h /ɦ/*	ch /x/*
ď /ɟ/	ť /c/	v /v/*	f /f/
(dz) /d͡z/*	c /t͡s/	z /z/	s /s/
(dž) /d͡ʒ/*	č /t͡ʃ/	ž /ʒ/	š /ʃ/

* Although at the end of a word or before a voiceless consonant /ɦ/ is realized [x], as in the example *lehký* 'easy' ['lɛxkiː], in the opposite direction, /x/ is pronounced [ɣ]: *kdybych byl* ... 'if I were ...' ['kdɪbɪɣ bɪl].

* (dz) and (dž) are placed in parentheses because these digraphs are not part of the Standard Czech writing system. However, [d͡z] and [d͡ʒ] are allophones of /t͡s/ and /t͡ʃ/, occurring in examples such as *léčba* 'treatment' ['lɛːd͡ʒba] and *otec byl doma* 'father was at home' ['otɛd͡z bɪl 'doma].

* While /v/ undergoes voice assimilation in word-final position or before a voiceless consonant, as in the example *vchod* 'entrance' [fxot], it does not cause voicing of preceding voiceless consonants: *svatý* 'holy' ['svatiː] or *tvůj* 'your(s)' [tvuːj].

* The consonant <ř> /r̝/ also undergoes devoicing in the positions mentioned above, as in the examples *keř* 'shrub' and *věřte* 'believe'. To my knowledge, an IPA phonetic symbol for its voiceless realization has not been established and I use therefore makeshift [r̝ʃ], whereby /r/ and /ʃ/ are realized in close succession: [kɛr̝ʃ] and ['vjɛr̝ʃtɛ]. Additionally, /r̝/ assimilates progressively, undergoing devoicing if preceded by a voiceless consonant, as in the examples *tři* 'three' [tr̝ʃɪ] and *střevo* 'intestine' ['str̝ʃɛvo].

* In some words, voice assimilation in the consonant cluster <sh> is also progressive and in Bohemia <sh> is often realized as [sx]. The most common example is in the phrase *na shledanou* 'goodbye' ['na‿sxlɛdanou]. In some parts of Moravia, regressive assimilation is typical in this consonant cluster ['na‿zɦlɛdanou] and in some words only regressive assimilation is observed, regardless of the speaker's origin: *shora* 'from above' is pronounced ['zɦora], never *['sxora].

* The phonemes /j/, /l/, /m/, /n/, /ɲ/ and /r/ are all voiced and do not have a voiceless counterpart. There is a considerable amount of regional (and probably idiolectal) variation with respect to their effect on preceding voiceless consonants.

* The grapheme <x> exists only in words of foreign origin and has both voiced and voiceless realizations, pronounced as either [ks] (*relaxace* 'relaxation') or [gz] (*existovat* 'to exist').

Abbreviations and symbols

The following abbreviations and symbols are used throughout the thesis.

Abbreviations

CC – Common Czech

ColC – Colloquial Czech

SC – Standard Czech

PLC – Prague Linguistic Cirle

I1 – interview 1

I2 – interview 2

VI – *v*-insertion (in tables)

ÉR – *é*-raising (in tables)

ÝD – *ý*-diphthongization (in tables)

PU – paradigm unification (in tables)

LT – *l*-truncation (in tables)

GN – gender neutralization (in tables)

Symbols

1. Slashed brackets /.../ – phonemic transcription.

2. Square brackets [...] – (1) phonetic transcription and (2) translations of works written in Czech (in the main body of the thesis).

3. Angled brackets <...> – graphemes.

4. An asterisk (*) – (1) grammatically incorrect or impermissible forms or (2) historically reconstructed / posited forms.

5. The symbol > is used in discussions of statistical significance to indicate 'more than' (i.e. p > 0.05 denotes that the probability is more than 0.05).

6. The symbol < is used in discussions of statistical significance to indicate 'less than' (i.e. p < 0.05 denotes that the probability is less than 0.05).

7. The symbol >< 'versus' is used to compare features or words from different varieties or languages (e.g. SC být >< CC bejt = SC být versus (or compared to) CC bejt).

8. An arrow → is used to indicate (phonological, vocalic, etc.) shifts (e.g. /o/ → /vo/ = *o* becomes *vo*).

1 *Introduction*

1.1 General introduction

The present study, which addresses the linguistic accommodation of Moravian migrants in Bohemia, is the first attempt to systematically describe the results of dialect contact between speakers of *obecná čeština* 'Common Czech' (hereafter, CC) and speakers of the more localized Moravian dialects. To my knowledge, it is the first systematic variationist account of language variation in the Czech Republic. CC, a majority but non-standard vernacular (variously described as a 'koine', 'interdialect' or 'common language'), is spoken throughout Bohemia (the western half of the Czech Republic) and some linguists report that it is a second central variety of the Czech national language and is assuming the role of a national vernacular (Sgall and Hronek 1992, Sgall et al. 1992). In this study, I examine the influence of CC on the linguistic behaviour of a group of university students from Moravia and Czech Silesia[1] (the eastern part of the Czech Republic) living at a hall of residence in Prague (situated roughly in the centre of Bohemia). A map of the Czech Republic is given in Figure 1. Other maps are included in the Appendices section of the thesis (§ 1.1, 1.2).

---

[1] For practical reasons, the terms 'Moravia' and 'Moravian' are used to denote the whole eastern part of the Czech Republic, including the area of Silesia which is located within the Czech Republic. Likewise, the noun 'Moravian' denotes an individual from this region. Terms such as 'Silesia' or 'East Moravian' are used only in situations when phenomena that are characteristic of a specific region are described.
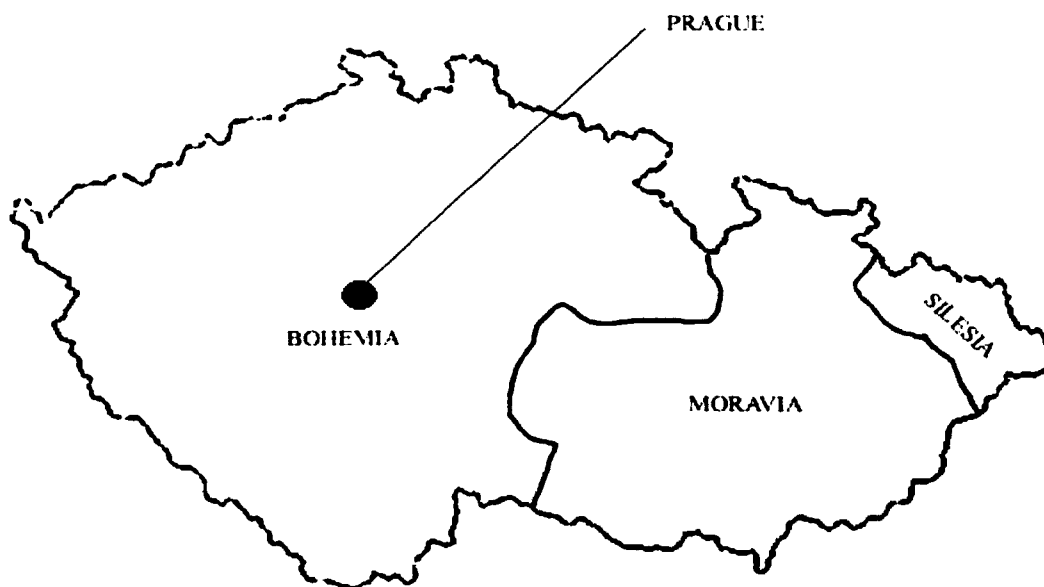
Figure 1.1 *Outline map of the Czech Republic*

The study draws on methodological frameworks that have been tried and tested predominantly in Western societies with a fundamentally different sociolinguistic profile than that of the Czech Republic. A triangular approach is used, whereby a quantitative analysis of several linguistic variables is combined with both qualitative data and an ethnographic study of informants' social and linguistic behaviour, which was accomplished through participant observation at a university hall of residence and surrounding areas during the academic year 2004 – 2005. A sample of 39[2] informants stratified according to sex, region of origin, length of residence in the host[3] community and network integration participated in the study

---

[2] The results presented in Chapter 7 are based on the accommodation of 37 of the 39 informants. Two informants who participated in the study are from western Moravia, where CC is spoken as the native dialect. Their linguistic behaviour is analyzed separately (§ 7.4).

and the data were elicited in two recorded interviews. A more comprehensive account of the fieldwork strategies that were employed is given in chapter 5.

Although the accommodation of adult speakers migrating to other speech communities in the same nation state has received little scholarly attention, at least in the case of first-generation migrants, and is overshadowed by the formation of new varieties in the literature on dialect contact (§ Chapter 2), this particular case of dialect contact is especially interesting in view of the historical development and current function, role and prestige of the varieties in contact. The introduction of an archaic and fossilized standard in the late eighteenth / early nineteenth century after almost two centuries of enforced Germanization caused a wide gap between the prescribed standard and actual language use.[4] In Bohemia, this resulted in a semi-diglossic situation whereby CC, a variety that is socially unrestricted over a large geographical area, supplants the codified variety, *spisovná čeština* 'Standard Czech' (hereafter, SC) in various semi-official domains. SC can be succinctly described as an archaic, artificial and relatively weak standard that has, despite lay perceptions, no native speakers and is perceived as unnatural or as carrying tones of social pretentiousness throughout the areas where CC is spoken as a mother-tongue variety (§ Chapter 3). Prescriptive language planning and regulation in the years following the inception of the archaic standard, the top-down regulatory approach of the former

---

[3] The term 'host' ('host community', 'host dialect' / 'variety, host culture') is used throughout this study to describe the speech community to which the informants have migrated. Other frequently used terms such as 'target' or 'new' were considered inappropriate for the given contact situation. The term 'target' used in the literature on second language acquisition implies that a speaker aspires to acquire a given variety, whereas acquisition in this instance need not be the result of a conscious decision. The term 'new' is also problematic, since it is often used in the literature on new-dialect formation, describing a variety that has emerged due to prolonged contact between speakers of regionally different varieties, often in the second generation (Kloferová 2001, Kerswill and Williams 2000).

[4] The term 'actual language use', which is used throughout this study, denotes the type of language that Czech speakers use in informal, spontaneous communication.

Communist regime, and, in the modern era, the interference of language ideologies[5] and out-dated approaches to the study of the vernacular (§ Chapter 3) have meant that the present language situation exhibits similar problems to those of almost two centuries ago.

There has been much hypothesizing with respect to the varieties of Czech, in particular the role and function of SC or CC, the sociolinguistic domains in which these two varieties prevail and the switching between these two varieties that occurs frequently in informal and sometimes (semi-)formal speech. However, the varieties of Czech have been accorded very little empirical and systematic research – owing to reasons discussed in Chapter 3, the inductive approach has a much stronger tradition in Czech linguistics than empirical investigation (Neustupný and Nekvapil 2003) – and there is still too little material that can allow us to accurately describe the present state of the Czech language situation. Moreover, as well as being based almost exclusively on introspection and linguists' own intuitions, much of the material that we have at our disposal is heavily influenced by linguists' own views on the contemporary language situation with respect to the role and function of SC and CC and to how the linguistic status quo should be maintained. And, in this respect, different camps of linguists hold fundamentally different views (§ 3.2).

There is a striking lack of material on the non-standard varieties of Czech. It was traditionally assumed in Czech linguistics that all matters of language planning and regulation could be successfully resolved within the framework of the standard

---

[5] Terms such as 'language ideology' and 'ideologically influenced' are used several times in the present study and it is important, therefore, to define how they are to be understood – especially since the term 'ideology' has a specific sub-meaning in post-communist societies and 'ideologically influenced' may be construed as 'influenced by the former regime'. In the present study, any reference to 'language ideologies' denotes linguists' subjective and impressionistic stance on the Czech language situation; thus, assertions that are 'ideologically driven' are conditioned by linguists' attitudes towards contemporary Czech. Conflicting stances on the Czech language situation within the Czech linguistic community are outlined in Chapter 3. For a discussion of the various interpretations of the term 'ideology', see Woolard (1998: 5-7).

language, and the spoken language has thus been severely neglected. Exceptions to this are descriptive accounts of traditional dialects that have undergone intensive dialect levelling and koineization[6] and for the most part are today obsolescent. While in Western (socio)linguistics expansive varieties that are the result of dialect contact and mixing have become the main focus of dialectological investigation, varieties such as CC or the 'interdialects' (§ 3.8) that are taking shape in Moravia have until recently been seen by many Czech linguists as interfering with the dichotomy of standard – traditional dialects, which is now considered somewhat idealized (Winford 2003). Thus, although at various times there have been calls for linguists to research the vernacular using modern and objective methods (Sgall 1963, Sgall and Trnková 1963, E. Eckert 1993, Čermák and Sgall 1997), both pre- and post-1989 when the boundaries of research into non-standard speech were relaxed (§ 3.4), these aims have not been fully implemented (Sgall 1994).

The theme of the present study has also been the object of introspective analysis. Proponents of CC as a national vernacular, in particular the Bohemian linguist Petr Sgall and his associates (§ Chapter 4) have speculated about the outcomes of contact between speakers of CC and Moravian dialects. According to part of what I term the 'contact hypothesis' (§ Chapter 4), speakers of Moravian dialects who move to Bohemia quickly drop features of their localized vernaculars and assimilate CC forms, rather than shifting in the direction of SC (Sgall and Hronek 1992, Sgall et al. 1992). Conversely, the same authors have speculated that accommodation in the opposite direction – speakers of CC living in Moravia (the second part of the contact hypothesis) – does not involve Bohemian migrants

---

[6] The terms 'levelling' and 'koineization' are sometimes confused. Levelling implies the loss of irregularity. Koineization describes a process whereby two (or more) dialects in contact result in mixing and the emergence of new compromise forms that are often simpler and more regular than those of the input dialects.

assimilating features of the host dialects to any great extent, but rather that Bohemians who move to Moravia help diffuse CC beyond its heartland. However, we must be cognizant of the effects of linguists' stance on the Czech language situation. Because, in the opinion of the above authors, SC is a hyperlectal variety incapable of meeting its users' full range of communicative needs and CC is regarded as a socially unrestricted variety that is assuming the function of a nationwide vernacular, it is easy to understand why they predict that CC will be the dominant variety in this contact situation. With the exception of two studies (Bachmannová 1996, Jonášová 2001) reporting on the accommodation of very small groups of speakers from specific parts of Moravia – two and three informants, respectively – this part of the contact hypothesis has not been challenged. Therefore, we shall see in the present study whether the hypothesis that Moravians who live in Bohemia accommodate to CC rather than to SC is correct by analyzing the linguistic behaviour of representative sample of informants from the three major dialect regions in Moravia (§ 3.11).

1.2 Aims and objectives

The underlying research question is whether or not Moravians who have migrated to Bohemia 'really' converge towards CC rather than accommodating to SC, and thus one of my principal aims is to test this part of the henceforth unsubstantiated contact hypothesis outlined above. The study makes an important contribution to several of the post-1989 objectives that were advanced for study into the vernacular (see E. Eckert 1993: 3-4; see also 3.4) by assessing the role, function and prestige of SC, CC and Moravian dialects in informal communication between university students from different dialect regions of the Czech Republic. According to introspective data (§ Chapters 3 and 4), speakers of Moravian dialects are, on the one hand, expected to

adapt their speech habits in all encounters outside the immediate circle of their family and friends in an upwards direction, allegedly towards SC (Balhar 1995). On the other hand, however, it has been suggested that in Bohemia SC in informal communication is perceived negatively as unnatural, stilted and inappropriate and is accordingly avoided in all but the utmost formal speech acts (Sgall and Hronek 1992).[7] At least, this is the case for native speakers of CC. The boundaries where speakers use the standard in place of their regional dialect or vice-versa are not entirely clear and it is important for linguists to investigate in which sociolinguistic domains and situations SC and CC (or other non-standard varieties) prevail. That is, under which circumstances SC is perceived as stilted or inappropriate and in which situations CC is, or is becoming, the dominant variety. In the present study, I test: (1) whether SC is functionally flexible enough for speakers to use it as a conversational code in this informal but out-group communicative situation; or (2) whether CC is in fact socially unmarked and functionally superior to other non-standard varieties and may be used by non-native speakers as a kind of lingua franca.

In addition to trying to answer the general question *Do speakers of Moravian dialects who live in Bohemia use CC?*, special attention is directed towards uncovering the various language-internal and extralinguistic forces that motivate accommodation (or non-accommodation). Rather than proposing a model of migration-induced standardization, it is suggested instead that in-migrants will converge towards a non-standard regiolect which they otherwise perceive negatively, because the standard is not fully developed and is avoided in informal communication

---

[7] The assertion that SC is 'perceived negatively' in Bohemia by native speakers of CC needs to be treated with caution. First, it should be pointed out there has been no systematic research into speakers' attitudes towards the varieties of Czech, thus existing material is anecdotal. Second, many Czech speakers do not necessarily distinguish successfully between SC and CC forms; therefore, many speakers erroneously consider themselves to speak to SC (see Čermák 1993: 29). I would suggest instead that several forms which are labelled by linguists as SC are perceived by many non-specialists as bookish or stilted and are thus generally avoided in all but the most formal types of communication.

by native speakers of the host variety. Many of the general principles that have been advanced for dialect accommodation and second-dialect acquisition are tested in the present contact situation (§ Chapter 6). I examine among other things why some linguistic phenomena are assimilated more readily than others, which factors accelerate, delay or inhibit the adoption of specific features and whether accommodation takes place along a specific route. This part of the analysis deals with what I label in this study as 'variable-specific' factors, that is, the various language-internal and extralinguistic factors that might encourage or accelerate or impede or even prevent the adoption of the linguistic forms. The second, and most important, part of the study addresses the relationship between the assimilation of the dependent linguistic variables and a set of external independent variables (sex, region of origin, length of residence, network integration) that are linked directly to the individual informant. These are collectively termed 'speaker-specific' factors. The general aims of the study are summarized below:

a) 'Variable-specific' considerations:

1. Do speakers of Moravian dialects living in Bohemia accommodate to CC and to what extent do informants assimilate CC forms?

2. Is speakers' accommodation complete or partial? What types of accommodation are observed?

3. Do informants' attempts to accommodate result in hyperadaptation and / or the emergence of intermediate forms?

4. What degree of inter- and intra-speaker variation is present?

5. Which CC variants are acquired the most (or the least)? Are some variants acquired more than others and, if so, why?

6. Does SC play an active part in this contact situation? Or are SC forms avoided in favour of their CC equivalents?

b) 'Speaker-specific' considerations:

1. Does informants' region of origin in any way influence their accommodation? Are there significant differences in the accommodation of informants from the three dialect regions under study?

2. Are there any sex-related differences in the adoption of CC variants: do men accommodate more than women or vice-versa?

3. Do informants assimilate more CC forms the longer they live in the host community? Is the adoption of CC variants gradual or are most features acquired rapidly in the first year or two? Is there a point at which acquisition stabilizes or stops?

4. Does informants' level of integration in the host community influence their linguistic behaviour? Are the most integrated individuals the highest acquirers of CC forms? Can network integration as a quantifiable sociolinguistic variable reliably predict innovative language use?

5. Do informants' attitudes towards the host variety have a direct effect on speakers' linguistic behaviour? And can the relationship between language attitudes and language use be meaningfully explained?

6. Which of the above social constraints is most influential in terms of shaping informants' linguistic behaviour? To what extent do the independent social variables interact in controlling speakers' linguistic behaviour?

1.3 Overview of the thesis

I offer here a brief statement of the content of each of the forthcoming chapters. Chapter 2 is a brief overview of dialect accommodation and the dialect contact framework. It is concerned with the description of the forces behind dialect accommodation, both short-term and long-term and it examines the literature on dialect contact between speakers of mutually intelligible but regionally different varieties. Chapter 3 looks at the Czech language situation. This includes an historical account of the factors that have contributed to today's complex quasi-diglossic situation in Bohemia, a comprehensive description of CC and an insight into language evolution in Moravia. In chapter 4, the 'contact hypothesis' is examined and its principal problems are explained. Chapter 5 constitutes an overview of the methodology and fieldwork strategies that were used in collecting data. This includes a section on social network analysis and its application in studies of dialect and language maintenance. Chapter 6 provides a description of the linguistic variables that were analyzed and it presents universal theories that have been advanced to explain the reasons why some linguistic phenomena are assimilated more quickly and

successfully than others and what social constraints can impact on speakers' accommodation. In the latter section of the chapter, some general principles on accommodation and second-dialect acquisition are applied to the present contact situation. In Chapter 7, the quantitative analysis is discussed in detail and the results are tabulated and interpreted. I also include here a comprehension section on informants' overt representations and evaluations of the host variety, which is very interesting from the perspective of folklinguistics. Finally, in chapter 8 I bring the study to a close and present my conclusions.

2 *Accommodation theory and dialect contact*

2.1 Accommodation theory: some introductory remarks

According to the speech or communicative accommodation theory framework (SAT /

CAT),[8] which was introduced by Giles and his associates in 1973, in all contact

situations and in all social interactions between two individuals of different social or

regional dialects it is expected that speakers will in some way modify their speech

styles as a means of attaining specific goals (Giles 1973: 90). Beebe and Giles (1984:

8-9) comment that speakers accommodate in order to: (1) evoke their interlocutors'

social approval; (2) increase communicational efficiency, that is, mutual

intelligibility; and (3) maintain positive social identities. This type of accommodation

is known as 'convergent accommodation' or 'convergence' and implies that speakers

will attempt to bring their speech styles closer to those of their interlocutors and

minimize differences at various linguistic levels, for example, accent, speech rate,

vocal intensity, and so on. On the other hand, speakers may wish to distance

themselves from the interlocutor and accentuate the differences in their speech in

order to 'develop, maintain, or stress social or personal identity ... or to demarcate the

ingroup from the relevant outgroup' (Hinskens, Auer and Kerswill 2005: 6). This type

of behaviour is known as 'divergent accommodation' or 'divergence'. In light of more

recent research, it has been argued that rather than shifting in the direction of the

immediate interlocutor, speakers converge more towards the speech patterns believed

to be characteristic of the host community. That is, they converge towards

'stereotypical *persona* or mental representation (model) of a social group' (Auer and

---

[8] Giles and his associates used the term 'accent mobility' (see Giles 1973) in earlier studies of accent convergence and divergence. The term 'communicational accommodation' was developed later when it became apparent that accommodation is manifested not only in speakers' modifications to their 'accents', but also in a number of other ways such as gesture and posture, pauses, and jokes.

Hinskens 2005: 337) or they adapt their speech 'to how they believe others in the situation would best receive it' (Giles and Smith 1979).

First of all, let us differentiate between 'social psychological' accommodation and 'linguistic' accommodation. Hinskens, Auer and Kerswill (2005: 6) comment that social psychological accommodation need not result in 'actual linguistic accommodation', while linguistic accommodation is 'the linguistic manifestation' of speakers converging to or diverging from their interlocutors. The former draws on four social socio-psychological theories: similarity-attraction, social exchange, casual attribution, and intergroup distinctiveness (see Giles and Smith 1979: 47-53 for a comprehensive description of these theories) and is the line of enquiry inspired by Giles and associates, which evaluates upward or downward style shifting in interpersonal situations in which there is some kind of social asymmetry between the participants. Traditionally, according to the social psychological approach towards linguistic shifting, informants' speech habits are measured impressionistically by lay judges' evaluations and no linguistic analysis is employed. As Coupland (1984: 51) describes it, 'it is established practice to collect evaluations of tape-recorded spoken texts from groups of "linguistically naïve" subjects who are asked to rate these texts on a number of scales, linguistic as well as non-linguistic'. He also comments that the setting of the interaction is typically in language laboratories and is thus unnatural. Obviously, this is problematic for sociolinguists in terms of validity.

Sociolinguists, on the other hand, look to objectively measure accommodation and make generalizations about accommodation theory based on the quantification of linguistic styles that are recorded in naturalistic settings. This second type of accommodation is known as 'linguistic' accommodation. One of the first works to incorporate a linguistic study of accommodation is Coupland's (1984) analysis of the

linguistic behaviour of Sue, an assistant at a travel agency in Cardiff, Wales. Coupland analyzed the distribution of variants of four phonological variables in recordings of 51 clients (all from Cardiff) talking to his informant Sue (also from Cardiff). All of the speakers who participated in the study are described as having 'regionally accented speech'. According to accommodation theory, Sue was expected to converge to the speech habits of her interlocutors in order to: (1) evoke their social approval; and (2) ensure mutual intelligibility. The results are striking: Sue consistently accommodated to the phonological behaviour of her clients, both in an upward and downward direction, and Coupland argues that 'the percentage of variants in Sue's speech provide a reasonably reliable index of the socioeconomic status and educational background of her interlocutors, just as the percentages of those forms in the clients' own speech are able to do' (1984: 60-61). Hoffmannová and Müllerová (2000: 25) identified similar patterns of accommodation in a study of Czech doctor-patient conversations. Presenting the data elicited from one of the doctors whose speech was analyzed, they illustrate that the doctor used more regional forms when he was speaking to a young child, increased the number of SC forms when he was examining a university professor and used more specialist medical terms when he was speaking to a patient who worked as a veterinary surgeon.

The kind of accommodation we have looked at so far is often set in various institutional settings and is motivated by social or professional differences between the individuals in contact. Other interactional situations where there is some kind of social asymmetry between interlocutors are also reported in the literature. These include interactions between professionals and laymen: for example, in courtroom discourse (Linell 1991) and in communicative exchanges between doctor and patient (Street 1991, Hoffmannová and Müllerová 2000). Social psychologists of language

and sociolinguistics have also researched interethnic discourse (Bourhis and Giles 1977); native–non-native interactions (Zuengler 1991); conversations between young and old speakers (Hoffmannová and Müllerová 2002); and interactions between workers in various areas of the service sector or information services and their customers (Coupland 1984, Hoffmannová and Müllerová 2000, Alexová and Davidová 2004).

The above studies all concern short-term manifestations of convergent and divergent accommodation (both sociopsychological and linguistic). However, besides analyzing shifts in linguistic behaviour in short-term encounters between speakers as in the examples given above, linguists also look to identify the effects of long-term accommodation and migrant- or contact-induced change. Analyses of long-term accommodation are not restricted to situations where there are discernable social differences in participants' accents; instead, the accommodation of speakers of regionally different varieties is measured in view of prolonged contact. According to the definition advanced by Trudgill (1986), long-term accommodation characterizes dialect adjustments that are maintained regardless of the interlocutor and the setting, and is a precursor to dialect levelling. That is, prolonged contact with speakers of the host variety results in permanent structural change, and 'short-term accommodation becomes long-term accommodation as soon as it permanently affects the accommodating speakers' (Auer and Hinskens 2005: 335). When speakers of mutually intelligible but distinct dialects of the same language meet the outcome is linguistic accommodation, initially of the short-term character described above, while in the long term, this contact may become 'routinized' or 'permanent' (Britain and Trudgill 1999: 245).

Trudgill states that accommodation 'within' a speech community is different from accommodation 'beyond' the speech community (1986: 12). In-community (social) accommodation involves 'altering the frequency of usage of particular variants of variables over which the speaker already has control', whereas the accommodation process of a speaker who moves to another speech community often requires adopting variants of the host variety. Trudgill (1986: 40) argues that the initial modifications we make to our speech habits may become permanent adjustments over time and if 'accommodation, through the adoption of a feature from an alien linguistic variety, is frequent enough, then that feature may become a permanent part of a speaker's accent or dialect, even replacing original features'. Trudgill also points out that the chances of accommodation becoming permanent are also influenced, and made more likely, by favourable attitudinal factors (1986: 39).

2.2 The dialect-contact framework

The dialect-contact framework is a relatively recent concept in sociolinguistics, developing from the 1980s onwards. Rather than being an extension of language contact, dialect contact 'is a phenomenon typologically different from language contact because it does not involve speakers learning a new language, either wholly (giving rise to varying degrees of bilingualism) or in a restricted sense (typically resulting in lexical borrowing, but without any other changes)' but instead 'items can be mixed apparently at will and with minimal loss of intelligibility' (Kerswill and Williams 2002: 82). Most of the literature on language contact and language mixing has been devoted to contact between distinct language or linguistic systems (Siegel 1985: 357) and until the 1980s studies of contact situations between speakers of the same language were by and large limited to short-term encounters between speakers

of socially different dialects (see above). Long-term accommodation and the results of prolonged dialect contact between speakers of regionally different varieties of the same language in informal communicative situations where there is no discernable socioeconomic or professional between interlocutors, on the other hand, has received far less attention. In fact, until recently it has been largely neglected within the variationist paradigm.

The lack of research into such contact situations can be attributed to several factors. One of the reasons is that traditional perspectives on language concentrated primarily on the classical and somewhat idealized standard-with-dialects model. Winford (2003: 1) argues that traditionalists and purists strongly opposed dialect contact and mixing: they regarded varieties that emerged out of dialect contact and mixing, which are often used over a larger area than the contributing dialects, as 'corrupt' and therefore often downgraded them. Second, the main focus within traditional dialectology has typically been on dialect or language maintenance in tight-knit, highly-localized groups of non-mobile speakers (see, for example, the Survey of English Dialects[9] (1962-1971)). Linguists were preoccupied with identifying a 'pure' dialectal form of the language spoken by elderly speakers, with 'isolation and immobility' (Chambers 2002: 117), while they generally overlooked the 'contaminated' varieties used by younger and / or more mobile speakers. Following its inception, variationist sociolinguistics generally continued trends developed within traditional dialectology by targeting the maintenance of localized

---

[9] The Survey of English Dialects (SED) was conceived by Eugen Dieth of the University of Zurich and Harold Orton of the University of Leeds in 1946. Fieldwork was carried out from 1951 to 1961 and the results were published between 1962 and 1978. Informants who participated in the study were predominantly non-mobile, older, rural males (NORMs) and the data were elicited via the use of a questionnaire. Recordings were taken in 313 localities in four regions: the north, the south, and the east and west Midlands.

vernaculars, predominantly in urban, rather than rural areas and looking to correlate variation with numerous independent social variables.

As L. Milroy (1987a: 2) comments, sociolinguistic research has been heavily dependent on earlier large-scale dialectological studies and work carried out within the Labovian variationist paradigm until recently 'may be seen as an explicit modification of dialectological methods', while mixed or supralocalized varieties were not accorded a great deal of scholarly attention. The Labovian variationist paradigm did not take into account the effects of dialect contact, and the speech community was viewed as an autonomous entity 'where the effects of contact obscure important structural patterns' (L. Milroy 2002: 4). Variationists viewed the speech community as an idealized entity, comprised solely of native speakers and isolated from the effects of contact. This is certainly true of early sociolinguistic studies. As Kerswill (1994: 26) states, Labov in his New York study (1966) and Trudgill in analysis of the English spoken in Norwich (1974) focused solely on the linguistic behaviour of 'natives' of given speech communities; he talks of a paradox whereby the majority of speech communities that have been studied are socially heterogeneous but 'the explicit models that have been proposed exclude that heterogeneity' (Kerswill 1993: 34). Labov later (1972b) admits that such homogeneous speech communities are 'myths' and points out that systematic variation is observed even in the most remote localities, giving the example of Gauchat's study of the French vernacular in Charmey, Switzerland (1905), which correlates linguistic variability with the independent social variables sex and age and is, therefore, credited by some sociolinguists as being the first sociolinguistic study (Chambers 2003: 16).

Nowadays, however, there are fewer opportunities to study isolated dialects and 'dwindling social relevance in doing so' (Chamber 2002: 177), and language

contact and convergence is attracting much more interest. From the 1980s onwards linguists began to realize the importance of dialect contact in the process of language variation and change, and methodologies to effectively observe the linguistic behaviour of socially, geographically and occupationally mobile speakers have been and continue to be developed. Over the last two decades, the study of dialect contact has grown and is the focus of many dialectologists – chiefly thanks to the contributions of Peter Trudgill. Trudgill's seminal work *Dialects in Contact* (1986) 'provided a framework for a considerable and steadily expanding research agenda' (L. Milroy 2002: 3), which is of practical and theoretical value not only within the variationist paradigm but for sociolinguists more generally. Trudgill looks at dialect contact in a number of societies, in particular at the various dialects of English that were involved in contact in Australia, Canada and the Falkland Islands. He also takes into consideration the dialects of Hindi transplanted to Trinidad, Mauritius and Fiji, and analyzes dialect contact in Norway and Sweden, and, citing data from various studies, he highlights processes such as long-term accommodation, dialect levelling, koinéization, and new-dialect formation. Trudgill shows that in contemporary society, as a result of widespread movement of the population, internal and transnational migration, industrialization, gentrification of the countryside and other such factors, speech communities are becoming more heterogeneous, networks that were previously highly-localized and tight-knit are being loosened and as a result more varieties that are the result of dialect contact and mixing will emerge. His work has since provided the impetus for several other original, innovative and exciting works that identify the various procedural stages in, and results of, dialect contact.

2.3 Research in the dialect-contact framework

The dominant and principal line of enquiry within the dialect-contact framework is the study of 'new-dialect formation'. New-dialect formation is described as 'the emergence of distinctive, new language varieties following the migration of people speaking mutually intelligible dialects to linguistically near-"virgin" territory' (Kerswill and Williams 2005: 1023-1024). Linguists are particularly interested in the advanced stages of dialect contact, in particular the emergence of the relatively stable varieties that are made up of various forms of the dialects that were part of the initial mix and intermediate forms that were not part of any of the varieties in contact but are the result of contact between them. Therefore, a great deal of linguistic attention has been accorded to the stabilized varieties that are the outcome of dialect contact and mixing in new towns such as Milton Keynes (Kerswill and Williams 2000; 2002; 2005), in the English Fens (Britain 1997; 2002, Britain and Trudgill 1999) in former British colonies (Trudgill 1986; 2004, Trudgill et al. 1998; 2000), in Hindi-speaking areas of the Indian subcontinent (Trudgill 1986) and in German language islands (Rosenberg 2005).

With regard to Czech, much attention has been directed towards new-dialect formation (*formování běžné mluvy*) in the 'Sudeten' lands that were repopulated after the Second World War (Balhar and Pallas 1963, Kloferová 1987; 1995; 1997; 2000, Jančák 1997; 2001). Kloferová has carried out extensive research across three generations of speakers in towns and villages in the Šumperk and Bruntál areas of north Moravia, which due to its socio-demographic profile is particularly appealing to linguists. The ethnic mix of migrants here is particularly heterogeneous: the region was inhabited by Czechs who were forced to leave the region following the Munich Agreement (*Mnichovská dohoda*) in 1938, migrants from nearby towns and villages,

as well as from more distant parts of the Czech and Slovak Republics, re-emigrants from Czech-speaking communities in Russia and Romania, Germans that for whatever reason had been allowed to stay after the end of the Second World War and later Romany Gypsies and immigrants from Macedonia and Greece. New-dialect formation here differs markedly from new-dialect formation in the Sudeten lands in Bohemia (Jančák 1997: 239; 2001: 214). In the latter case, the initial migrants were predominantly from Bohemia, who with minor regional differences used the same variety (CC); therefore, already by the second-generation the situation had stabilized and the emergent variety is considered identical to the CC spoken elsewhere in Bohemia (Jančák 1997).

However, although there is already an impressive amount of literature on short-term contact between speakers of socially different dialects and an accumulating body of research on the final stages of dialect contact, in particular koineization and new-dialect formation, there is a striking lack of studies that can be located between these polar ends of the continuum. That is, there is virtually no research on the inherently heterogeneous results of long-term accommodation in the first generation, where speakers have migrated to other speech communities. This is somewhat surprising, insomuch as national and transnational migration between speech communities is an increasingly common and even unavoidable in many contemporary societies, in particular in Western societies where most sociolinguistic research is being carried out. Large-scale migration can be traced back to the middle of the nineteenth century in most European countries and coincides with the rise of industrialization (Auer and Hinskens 1996: 19). Initially, migration chiefly involved movements of individuals from the surrounding rural hinterland into large urban centres, but nowadays both national and international migration are characteristic for

most Western societies, both from the countryside into the towns and increasingly vice-versa (counter-urbanization). Such intensive migration has created a 'world in which there is vast language and dialect contact' (Tagliamonte and Molfenter 2007: 649) and this undoubtedly leads to language change. Moreover, studying the accommodation of first-generation migrants can give us a better understanding of the processes involved in both levelling and the formation of new or intermediate dialects.

The linguistic behaviour of migrants who are culturally and ethnically similar to members of the host community has so far been touched on in only a very marginal fashion. Existing studies of migrant communities have typically analyzed the linguistic behaviour of speakers who are culturally or ethnically very different from members of the host variety (Fought 1999, Evans 2004). Furthermore, in the few works that do look at second dialect acquisition, a process by which individuals who have been transported from one region to another acquire a second dialect of the same language (Chambers 1992: 674), linguists have focused primarily on the speech of adolescents or pre-adolescents (Payne 1980, Chambers 1992, Roberts and Labov 1995, Trudgill 1986, Roberts 1997, Starks and Bayard 2002, Tagliamonte and Molfenter 2007). Chambers (1992), for example, studied the accommodation of six Canadian youngsters between the ages of nine and 17 who had moved to England. Tagliamonte and Molfenter conducted a similar study, analyzing the linguistic behaviour of Tagliamonte's three children who were also living in England and had all moved there from Canada under the age of five. Trudgill (1986) cites the results of a six-month longitudinal analysis of the accommodation of seven-year-old twins, Debbie and Richard, who moved from Reading, in the south of England to Sydney, Australia. Here 15 phonological forms that differentiate English English from

Australian English are analyzed. Starks and Bayard's (2002) study is an analysis of postvocalic /r/ in the speech of four children born in New Zealand to North American parents. And Payne (1980), Roberts and Labov (1995) and Roberts (1997) all analyzed adolescents' and pre-adolescents' acquisition of phonologically simple and complex variables in Philadelphia. Accommodation after the critical age[10] has received far less scholarly attention. This is perhaps because linguists presume that migrants will fail to successfully assimilate all features of the target variety or that they will tend to gravitate towards the standard. Indeed, Trudgill writes that children are 'much more rapid and complete accommodators than adults' (Trudgill 1986: 31), and Tagliamonte and Molfenter claim that children appear to be 'the only sector of the population' capable of successfully acquiring a second dialect (2007: 650).

Of the studies that address the accommodation of adult speakers, an interesting pattern has been observed whereby variants of the host variety are assimilated in a relatively fixed order. An example is Trudgill's (1986) study of the accommodation of speakers of English English living in America, which includes a *post hoc* description of his own accommodation. His informants were mainly academics and his data are based on informal observations of their linguistic behaviour at conferences and during lectures. Although he admits that his study is not a systematic attempt to quantitatively analyze the linguistic behaviour of a representative sample of informants, Trudgill concluded that he had identified a fixed route whereby speakers assimilated features of the host variety in the same order. This proposal was termed the 'fixed-route hypothesis' and Trudgill later tested it in other contact situations (see below).

---

[10] Chambers (2003: 179) argues that after the age of seven children rarely succeed in fully mastering a second dialect. Payne (1980) found that even very young children failed to acquire complex variants if their parents were not native speakers of the local dialect.

In a pilot study, Shockey (1984) identified a similar fixed route of accommodation in the opposite direction: Americans living in England. She studied the linguistic behaviour of two male and two female speakers who had lived in England for between eight and 27 years. The informants were speakers of Midwest or Californian dialects and all were teaching staff at the University of Exeter. Like Trudgill, Shockey also analyzed changes in her own speech. She comments that 'Americans who had been in Britain for some time had adopted certain British features of pronunciation' to the point where 'they sounded like Americans to British ears and like British people to Americans' (1984: 87). She also emphasizes that the 'transplanted individuals' assimilated British features 'in a similar direction'. For example, the principal adjustments in her long-term informants were: (1) the diphthong /ou/ as in *boat* was fronted from [oʊu] to [əʊ]; (2) they realized /o/ in *hot* mostly as [ɒ] rather than [ɑ] – for this variable there is a greater level of accommodation than in the opposite direction (see above); and (3) they used [t] and [d] in positions where speakers of their native dialects use [ɾ] as in *ladder* and *latter*. Shockey accorded special attention to the last variable, insofar as it would 'be easier to judge auditorily' and she could 'be sure of the subjects' original behaviour in their native dialects'. In conclusion, Shockey argues that the motivation for accommodating was mainly to attain communicational efficiency rather than to gain the social approval of the speakers of the host variety (1984: 92-93).[11]

Trudgill (1986: 24-28) lists 'further evidence on the regularity of the accommodation process' based on the analysis of the linguistic behaviour of Swedish women living in Bergen, Norway, presented in Nordenstam (1979). In this case,

---

[11] Shockey argues that words such as *matting*, *heating* and *putty*, realized with [ɾ], might not be intelligible to some speakers of English English.

lexical and morphological differences – where the two varieties differ the most – are examined in greater detail. The conclusion is that Nordenstam's informants acquire variants of the host variety in a relatively fixed order. Since lexical differences between the two varieties are highly salient, the first stage of accommodation involves the adoption of lexemes. In contrast to the Anglo-American accommodation, where phonological differences are more marked, the next step is the acquisition of morphological variants. Therefore, although the cases which we have examined are from different language societies, there does seem to be a trend whereby accommodation follows a particular route. That said, in order to make generalizations about the route accommodation takes, we need to carry out systematic research in several speech communities and to analyze the linguistic behaviour of larger groups of speakers.

To my knowledge, there is a paucity of systematic accounts of long-term, migration-induced dialect accommodation of speakers of mutually intelligible, but regionally different varieties of the same 'language' (within the same nation state) after the critical age. Exceptions include Bortoni-Ricardo (1985), Kerswill (1994), Matter and Ziberi (2001) and a number of sociolinguistic studies of internal migrants in Nordic countries (reported in Kerswill 1994: 74-75). The subjects of these studies are rural dialect speakers from tight-knit communities who have migrated to large urban centres and migration in these cases has typically been accompanied by a sharp transition in way of living from the base to the host community. Bortoni-Ricardo looked at a community of rural migrants in Brasilia, Kerswill studied the linguistic behaviour of rural Stril migrants living in Bergen, Norway and Matter and Ziberi observed the accommodation of speakers who had moved to the capital, Berne from a geographically peripheral area of the Swiss Alps, and whose regional Alemannic

Alpine dialects cause problems of intelligibility. Interestingly, the societies where the above studies have been carried out are comparable to the Czech language situation: Swiss German in terms of diglossia and Norwegian in view of the coexistence of two central varieties (although both central varieties in the Norwegian case are considered standard).

3 *The Czech language situation*

3.1 Historical background

Until the early seventeenth century the Czech standard dialect, based on the dialect of central Bohemia from the thirteenth century onwards, was a highly-developed, fully-functional and cultivated variety, which enjoyed almost three centuries as a written language and which reached its cultural height of prosperity in the late Humanist period (Renaissance). Besides Old Church Slavonic, Czech is reputed to be the oldest Slavonic language used for official purposes (Kopečný 1949: 15, Daneš 1996: 116).[12] However, its evolution was disrupted following a series of religious feuds between the Protestant noblemen of Bohemia and the Roman Catholic Habsburgs. The growing political unrest escalated when two senior Habsburg officials were defenestrated at Prague Castle in 1618[13] and this act of defiance instigated the Thirty Years War (1618-1648), during which the Czech rebellion was brutally crushed after just two years, at the battle of White Mountain (*Bílá Hora*) in 1620. This led to 'the beginning of a long process of both the national and the religious disintegration of the Czech nation' (Grygar-Rechziegel 1990: 12). German was gradually imposed as the 'official' language; Czech was stripped of several prestige functions and its usage was narrowed. Several intellectuals, including many scholars who were responsible for language planning and maintaining the linguistic status quo were either direct victims of the brutality (Sgall et al. 1992: 168) or were forced to emigrate following the re-Catholicization of Czech lands (Cuřín 1985: 58). In short, the Czech lands were almost stripped of their identity.

---

[12] Neustupný and Nekvapil write that by the end of the fourteenth century 'Czech was a stylistically highly elaborated language which had penetrated to the domains of administration and ideology' (2003: 226).

[13] This event took place on 23 May 1618 and is known as the 'second defenestration of Prague'.

Nonetheless, contrary to popular belief, the Czech language continued to develop and was not completely removed (Stich 1993, Nebeská 2003: 104). During the Baroque, a period often characterized as an age of national decline (*doba úpadku*) or a 'Period of Darkness' (Neustupný and Nekvapil 2003), the standard language receded dramatically, having been stripped of official functions; however, the spoken language, the base for today's CC, continued to develop. It was confined by and large to rural areas where it was used as the everyday means of communication by the peasantry; the social elite in towns and cities generally preferred to speak German, Latin or French. Auty emphasizes that Czech continued to be used in letters, religious texts and private journals, and that Czech books were still printed, although 'their number diminished to a mere trickle' (1976: 83); it was largely removed, however, from schools, the sciences, the humanities, law, and administration (Neustupný and Nekvapil 2003: 227). The Baroque period is also characterized by several developments within the spoken language that were influenced either by changes taking place before 1620 or by Germanization. The vocabulary was augmented with a stock of new words, predominantly with an influx of Germanisms and regional forms (Starý 1995: 61-62). However, Daneš argues that these new lexical acquisitions and the attempts at language planning and regulation at the time were 'amateurish' (1996: 218).

The re-emergence of the Czech language began during the National Revival (*Národní obrození*), which started around 1770 and continued into (and beyond) the mid-nineteenth century. Intellectuals at the time attributed great importance to the language, many viewing it as the most important element of the Czech national identity, which represented the nation's glorious past (Starý 1993: 82). The foundations for the codification of the contemporary standard language were set –

unintentionally – by Josef Dobrovský, a leading linguist and historian. Dobrovský was interested primarily in comparative Slavonic philology from a historical perspective and his work *Ausführliches Lehrgebäude der böhmischen Sprache*, a grammar of Czech written in German, was intended as a description of the Czech language of the Humanist era.[14] It was not intended as the normative grammar that would later be used as the main tool for codifying today's SC (Sgall and Hronek 1992: 98), since Dobrovský himself considered a full revival of Czech beyond possibility.

Nevertheless, it was to be Dobrovský's grammar that the next generation of revivalists (led by Josef Jungmann and his associates), optimistic about a full re-emergence of the Czech language, used as their tool for the codification of contemporary SC. The revivalists were faced with two choices: (1) to reinstate the language of the late Humanist era, a 'Golden Age' of Czech culture and prose, an age of prosperity and great expectations; or (2) to codify and regulate the 'crude' language which had developed during the Baroque, an age of alleged national decline, and which, according to Neustupný and Nekvapil (2003: 228), was 'underdeveloped in many respects and could not easily serve either as a national symbol or as the tool of communication in a society aspiring to enter the age of modernisation'. Not surprisingly, they chose the former. The codified standard was highly conservative. Innovative features that had gained currency over a wide geographical area and which Dobrovský had listed in his grammar were dropped in favour of more conservative

---

[14] Dobrovský also published other important works, including *Bildsamkeit der slavischen Sprache*, which looked at word formation in the Slavonic languages, *Deutsch – böhmisches Wörterbuch*, a German-Czech dictionary and *Geschichte der böhmischen Sprache und Literatur*, an historical overview of Czech language and literature.

features that in the spoken language were either obsolete or obsolescent.[15] This included the rejection of certain developments in the phonology which Dobrovský listed, such as forms with prothetic /v/ *vokno* (SC *okno* 'window') and personal pronouns *von* (SC *on* 'he'), *vona* (SC *ona* 'she') and forms with diphthongized /ɛj/ such as *mlejn* (SC *mlýn* 'mill').[16] Bookish forms of the type *mladá knížata se sešla* 'the young princes met' or *celá města byla zničena* 'entire towns were destroyed' were also codified, although there had been a merger in the spoken language between neuter and feminine adjectives, pronouns and the past tense of verbs and writers of this era used innovative agreement patterns such as *mladé knížata se sešly* and *celé města byly zničeny*.

In light of more recent research, it is clear that the choice to reinstate the language of the Humanist period was made primarily on ideological grounds, not because the spoken variety was functionally restricted (see, for example, Nebeská 2003: 103). Starý (1993: 82) believes the revivalists' decision was influenced predominantly by extralinguistic factors, arguing that at the time 'linguistic issues mirrored issues of socio-political life' and 'typological characteristics of languages were correlated with certain human values, which of course corresponded to certain values on an axiological scale'. Therefore, positive values were assigned to 'highly developed' flective languages as opposed to 'crude' or 'basic' non-flective languages. Gammelgaard (1999: 33) comments that during the National Revival the language

---

[15] Although the revivalists can be criticized for their overly conservative codification of SC forms, their efforts to revitalize the vocabulary with various foreign borrowings and neologisms were very successful, especially with respect to the introduction of new technical terms.

[16] Nebeská (2003: 103) states that CC forms of the type listed above were frequently observed in religious texts written in the Baroque period. Furthermore, they were observed even in texts of Moravian and Silesian authors. Kopečný (1949: 15) cites that of the shifts *ó* → *ů*, *é(ie)* → *ý/í*, *ú* → *ou* and *ý* (or following sibilants *í*) → *ej* that had taken place in the spoken language, only *ó* → *ů* and *ú* → *ou* (apart from in word-initial position) were fully accepted into the codification. The shift *é* → *ý(í)* was codified only partially and *ý* → *ej* was not admitted into the codification at all.

was much more than 'a mere tool of communication' and that 'its very existence, lexical capacity and beauty were considered as the very raison d'être of the Czech nation'. Grygar-Rechziegel also emphasizes that the language was an essential tool in resurrecting the national consciousness and culture and that it 'should serve all cultural needs of the reawakened Czech population' (1990: 13).

Taking this into consideration, it is not difficult to understand why the revivalists codified Humanist Czech: it was a highly developed flective standard with a rich literary tradition, and was higher on the axiological scale than German, the language of the former occupiers. In addition, the spoken language had been heavily influenced by Germanisms and this was obviously deemed inappropriate. Kopečný (1992: 33) views the reinstatement of an archaic code as the means of everyday communication as a major 'impediment' to the evolution of Czech, regarding the reforms inspired by Dobrovský's grammar as the 'first artificial intrusion' into the Czech language, which had previously evolved 'naturally'. In his opinion, the situation was made worse by future generations of reformers, starting from Jan Gebauer and continuing until the present day. Stich (1993: 92-99) also questions the widely propagated belief that the eighteenth century was a period of decay for SC. He argues that Dobrovský's chapter on the decay of Czech was based exclusively on the development of Czech literature, whereas the actual language situation, about which little was known at the time, was not mentioned.[17] However, Sgall (2002: 6) points out that the conservative nature of the codification established by Josef Jungmann and his associates was motivated by the necessity to maintain linguistic unity with Slovakia, and if we accept this interpretation, the revivalists' decision to instate an

---

[17] Conversely, other grammars that were published at the time (Nudožerský 1603, Rosa 1672, and later Doležal 1746, Pohl 1756, and Pelcl 1795) did acknowledge changes that had taken place in the phonology and morphology.

archaic standard might have been made more out of necessity than choice.[18]

The revivalists' decision, regardless of its motivation, had far-reaching consequences. It created a wide gap between SC and actual language use in informal situations and speakers were in effect forced to learn their own language in the way they would learn a foreign language (Neustupný and Nekvapil 2003: 228-229). In an attempt to achieve what is regarded by some linguists as 'short-term national prestige', the revivalists effectively imposed an archaic standard with no real chance of fulfilling its users' range of communicative needs. Consequently, there was an imbalance between the language that was spoken on the streets and the codified standard that was prescribed in grammars and style manuals.

3.2 Approaches to language regulation

Following the introduction of what Sgall et al. (1992: 173) describe as an 'artificially established norm' back into the mainstream, the task of establishing linguistic stability was assumed by the protagonists of two strongly opposing trends: first the purists and, later, members of the Prague Linguistic Circle (hereafter, PLC). Both had fundamentally different ideas with respect to language planning and regulation. Purist interventionism promoted historical authenticity, national genuineness, aesthetic plausibility (see Starý 1993: 81); the purists aspired to a 'pure', 'genuine', 'refined', and 'beautiful' standard to be used under all circumstances, from a scholarly debate to an informal chat, and thus linguistic variability was precluded. Purism overtly prescribed and proscribed linguistic forms, even if this resulted in a loss of function or meaning. If two expressions could have been used in the same context, the purists construed this to mean that one of these expressions must be wrong, and Dickens

---

[18] In addition, Kopečný (1949: 16) argues that the revivalists themselves spoke mostly German or Latin and, therefore, did not have a feel for the changes that had taken place in the spoken language during the Baroque period.

(1994: 24) argues that because of purist interventionism 'Czech retained a large number of quite unnecessary anachronisms'. This meant that, instead of the situation improving, the gap between the prescribed standard and actual language use became wider.

Members of the PLC, on the other hand, placed emphasis on a functional approach to language, using what they called 'elastic stability' (*pružná stabilita*), that is, a gradualist approach to language whereby new lexical items and constructions are gradually incorporated into the codification. They believed that different communicative situations require different functional styles and they developed the 'theory of language cultivation', according to which SC could be made functionally and stylistically flexible enough to meet users' full range of communicative needs. Nebeská (2003: 91-92) states that the theory is based on several assumptions. Their central criterion was the opposition between 'standard (use)' (*spisovnost*) and 'non-standard (use)' (*nespisovnost*) that permitted the use of variable forms as standard, as opposed to the purists' 'right' (*správnost*) – 'wrong' (*nesprávnost*) opposition that admitted only one form. According to this theory, the standard language is functionally complete and serves all the communicative needs of the 'educated speaker', as well as fulfilling functions that cannot be fulfilled by non-standard varieties. The PLC predicted that the language use of the 'educated speaker' would serve as an example and that Czech speakers would use SC as their everyday means of communication and gradually give up non-standard forms, and that eventually SC would be used spontaneously rather than consciously. However, the PLC did not accord attention to the processes that were occurring outside the standard language and the vernacular was ignored.

Starý (1993; 1995) criticises the PLC's approach to language cultivation,

insofar as it is: '(1) framed by the idea of standard language, (2) it is doctrinal, and (3) it is interventionist by its nature' (1993: 80). He argues that the goals of PLC were in many respects similar to the goals of their staunch rivals, the purists, insofar as both were unwilling to accept that relevant sociolinguistic processes take place outside the standard language (1993: 81). This was a cardinal error in both approaches, since we know that the standard variety is constantly affected or even threatened by rival varieties that influence its evolution (Haugen 1997: 348). Nebeská (2003: 104-105) ascribes the failure of the PLC's theory of language cultivation to their failure to take into account important changes in the country's social structure at the time: the relaxation of official and traditional institutions led to a shift in moral values, including a relaxation in social and linguistic behaviour. The prestige of SC decreased and informal, spontaneous behaviour and communication became more fashionable. The PLC's efforts did see the standard move in the direction of the spoken language, but not to the extent that its protagonists had anticipated.

3.3 Research into spoken Czech

In comparison with most other Slavonic languages, Czech has been accorded a considerable amount of scholarly research and has attracted a great deal of linguistic interest from outside the Czech Republic. However, for various reasons, some of which can be explained as a direct or indirect result of what ensued after the battle of White Mountain in 1620, the spoken language has been neglected and has been overshadowed by the standard dialect (Schmiedtová 1995: 84). E. Eckert (1993: 7), for example, believes that Czech linguists 'who consider the standard language variety the backbone of the Czech system have been preoccupied with it to the point of neglecting the spoken language'. We could argue that while there is an extensive

literature on the standard dialect and the traditional rural dialects spoken in Moravia, we know comparatively little about the levelled varieties spoken in urban communities in Moravia or about the precise role, function and areal distribution of CC (Hronek 1987: 89). As Sgall and Trnková (1963: 28) comment, Czech is typically analyzed at polar ends of the language continuum, from the perspective of written texts and with respect to the highly-localized dialects that have receded considerably in recent years. It is not investigated, however, in terms of the language that is spoken in everyday, spontaneous communication. Although the authors expressed this view almost half a century ago, it is still relevant with respect to contemporary trends in Czech linguistics and it is clearly reflected in lay speakers' knowledge of language variation and the stratification of the Czech national language. Bayerová-Nerlichová states that the term *obecná čeština* (CC) is unknown to most Czechs (2004: 182-183). She argues that, although Czechs are taught at school about SC and Moravian dialects,[19] CC is often omitted from teachers' descriptions of the Czech language situation and the label *obecná čeština* is frequently attached to any non-standard or substandard variety or linguistic feature, including slang, argot and highly-localized regional speech.

Most of the systematic attempts to describe CC are in fact the work of foreign scholars or Czech émigré linguists. E. Eckert (1993: 8) points out that the non-functionality, or in E. Eckert's words, the 'primacy' of the standard dialect was first brought into question by foreign Bohemists and Czech linguists publishing outside the Czech Republic (Vey 1946, Širokova 1954; 1955, Kučera 1955 and later). Vey (1946), a French linguist, was the first to systematically describe the morphology of

---

[19] 'Moravian dialects' in the sense of the traditional dialects used in Moravia as opposed to the regiolects (interdialects) that have taken shape in recent years. Thus, in terms of actual language use, the general public is ill-informed of the actual language situation in both Bohemia and Moravia.

CC and he compared the Czech situation to spoken French, Širokova (1955) reported on the Czech language situation in the Russian journal *Voprosy jazykoznanija* and described CC as a central variety of the national language, and Kučera (1955; 1958; 1973) set the foundations for empirical research into spoken Czech with his statistical analysis of the distribution of CC forms. Kučera's empirical research is accredited with paving the way to defining the actual language situation and his research is considered objective, since some Czech linguists consider that it is not burdened by *a priori* or biased ideological stances that prevail within linguistics in the Czech Republic (Hronek and Sgall 1999: 184-185). More recently, foreign linguists have been active in researching CC elements in literary texts (Short 1992, Townsend 1993, Gammelgaard 1997, Bermel 2000, Maglione 2001) and to a lesser extent in analyzing variation in speech in specific types of public discourse (Maglione 2003, Hedin 2005).

Bayerová-Nerlichová points out that the greater interest in CC from outside the Czech Republic could stem from the 'fundamentally different' attitudes of foreign Bohemists to, and their areas of interest in, the problems of the Czech language situation (2004: 174). Hronek (1987: 90) also emphasizes differences in the perceptions and objectives of Czech and foreign linguists, arguing that, while on the one hand, CC as a non-standard majority vernacular constitutes something unusual or not typical of the foreign scholars' native speech communities, on the other hand, for Czech linguists the standard language is the cultivated variety of their national language, which should be revered and used as much as possible. While some linguists have welcomed the contributions of foreign Bohemists, who often approach the situation from different viewpoints, others have questioned the work of outsiders. Jelínek (1963: 48), for instance, is convinced that 'only Czech linguists', who rely not only on empirical research but also on their native and complete knowledge of the

standard and other non-standard varieties, can make valid contributions with respect to the problems associated with the Czech language situation. Lutterer, in a review of Kučera's research, argues that, although not hindered by the subjective evaluation of linguistic phenomena, Kučera's work is nevertheless hampered by an incomplete knowledge of the complex Czech language situation, which 'only a native speaker can possess' (1964: 295).[20]

Thus, to an extent, the early attempts of foreign linguists to describe spoken Czech and their research on CC were overlooked by home linguists, being conveniently passed off as the result of an incomplete knowledge of the Czech language situation. It was a controversial article written by the Czech linguist Petr Sgall and published in the Russian linguistic journal *Voprosy jazykoznanija* (1960) that brought Spoken Czech and CC to forefront of domestic linguistic discussions. Sgall argued that SC was not capable of meeting its users full range of communicative needs, he promoted CC as a majority vernacular and a central variety of the Czech national language and he advocated a rapprochement between SC and CC in a bid to reduce the gap between the prescribed codified standard and actual language use; he also cast doubt on Colloquial Czech (§ 3.6.2) as an independent language variety.[21] Initially, Sgall was taken to task by his colleagues and his views with regard to the Czech language situation received severe criticism. The pervasive code-switching between SC and CC, which Sgall described based partly on his own observations and intuitions and partly on the empirical results Kučera had presented, was dismissed as

---

[20] In an earlier review (1957: 300), Daneš views Kučera's work positively as a valid contribution to Czech linguistics, in particular with regard to the methodology Kučera employed to elicit the data. However, he emphasizes that the study is based on 'restricted' and 'not entirely reliable' material, listing some of Kučera's examples as inappropriate. Again, the selection of these 'bad' examples is ascribed to an incomplete knowledge of the Czech language situation.

[21] Sgall did not believe that SC could be made more flexible via the establishment of ColC, which other linguists such as Bělič saw as a means of reducing the gap between the codified standard and the actual language that was spoken in informal discourse. Instead he suggested that this could be achieved by a 'democratization' of SC, which would involve the codification of supra-local (national) CC forms.

intellectual slang. That is, the frequent switching between SC and CC identified in Bohemians' speech was not considered the product of unconscious or unmonitored speech or representative of the vernacular of the majority of Bohemians, but was said to be characteristic only of the 'deliberate concoctions of educated speakers' (Micklesen 1978: 440), particularly of certain academics at the Philosophical Faculty at Charles University, who consciously use this affected and exaggerated style of speaking.[22] Nevertheless, Sgall's article led to an impassioned debate in the Czech journal *Slovo a slovesnost,* which lasted three years and whose participants included among others Jaromír Bělič, František Daneš and Alois Jedlička; it was brought to a close by Bohuslav Havránek in 1963 (for a thorough and lucid discussion of this debate on SC and CC, see Cvrček 2006: 35-49). The debate was generally inconclusive. On the one hand, it was established that the language situation in Bohemia differed markedly from the situation in Moravia. It was also made clear that further amendments to the rigid SC codification were necessary and that further research into spoken Czech, in particular into CC, which had still not been systematically described or defined (Skalička 1962: 201), needed to be carried out. On the other hand, Hronek (1987: 90) argues that the results the discussion yielded were not adequately pursued and certain linguists gradually returned to the classic two-dimensional model of the Czech language, viewing the situation as a clear-cut case of standard with dialects and according CC the same status as the locally-restricted dialects in Moravia.[23]

In spite of the above, the debate did lead to two much-needed lines of enquiry into spoken Czech. The first was led by the linguist Jaromír Bělič and its aim was to

---

[22] Micklesen, however, points out that only 13 of Kučera's 23 informants can be classed as 'intellectuals' (1978: 445).

[23] Hedin (2005: 13) argues that after the debate had ended, there was no discussion of CC for 'a considerable period' and CC became a 'non-issue in Czech linguistics'.

investigate actual language use in large towns and cities. Bělič believed that city speech had been overshadowed in Czech linguistics by research in rural locations and he considered the analysis of the more heterogeneous Czech used in urban environments essential in describing actual language use (1962; 1968). Material was gathered using methods typical of traditional dialectology, in particular questionnaire-based surveys, and the results showed that CC serves as the means of everyday means of communication not only in Prague but also in other towns in Bohemia (see, for example, Brabcová 1973, Dejmek 1976; 1981; 1987, Jančáková 1987). The second line of enquiry looked to analyze actual language use in naturalistic settings and leaned heavily towards Kučera's earlier statistical analyses, which were used as a point of departure. The fieldwork was intended to be carried out based on elicitation techniques described in Sgall (1963) and Sgall and Trnková (1963), which mirrored variationist approaches to language use that were being used by Labov and his associates at around the same time in the West. The suggestions for the study of actual language use presented in Sgall (1963) and Sgall and Trnková (1963) are summarized below:

1. Research into spoken Czech should be carried out in as many locations as possible in all dialect regions, throughout Bohemia and Moravia in both urban and rural speech communities and in the Sudeten areas repopulated after the Second World War.

2. Data should be elicited in a range of communicative situations and informants should be categorized according to sex, age, level of education, employment and other social criteria.

3. Data should be elicited using carefully planned field methods that draw and build on existing methodologies and should be elicited using a combination of quantitative, qualitative and ethnographic techniques.

4. Data analysis should be carried out using the most up-to-date methods and relative percentages should not be taken as proof of the use of a particular linguistic feature. Instead, statistical analyses should be performed on the elicited data in order to make valid generalizations. Additionally, it is necessary to highlight which external variables or their combination influence speakers' linguistic behaviour and to what extent the different external variables are interrelated.

5. Research into actual language use should look first to analyze those linguistic phenomena that are important in terms of changes to the codification.

6. In carrying out the fieldwork researchers should not be influenced by their native intuitions and pre-determined views of the Czech language situation; they should be aware of, and able to control, their native cultural assumptions to ensure that their observations and perceptions are not in any way biased or prejudiced.

The given line of enquiry into spoken Czech is clearly orientated towards a variationist approach. There is mention of combining the quantitative study of linguistic variables with ethnographic data, emphasis is placed on studying linguistic

variation in relation to various social constraints and the necessity for using complex

statistical tests, including multivariate analysis, to test the validity and significance of

the elicited data, is also discussed. This latter line of enquiry, however, was

unfortunately not realized to its full potential. It produced only a few publications, the

most cited of which is Kravčišinová and Bednářová's analysis of the language use of

79 speakers from various backgrounds and from different parts of Bohemia – which

in truth does not meet all the criteria described above. It is alleged that

sociolinguistically orientated research into spoken Czech was discontinued after the

'normalization' process which began in 1968 when, as Gammelgaard (1999: 37)

states, 'free discussion concerning the status of CC could not take place' and several

linguists researching CC were 'expelled or downgraded from their posts'. Thus, the

period from 1968 until 1989 can be described as a time of stagnation with regard to

the investigation into spoken varieties of Czech and the status of CC, at least at

language institutions within the Czech Republic.[24]

3.4 Research into spoken Czech in the post-communist era and Czech sociolinguistics

As we have established, sociolinguistically-orientated research into Czech (and other

Slavonic languages) was, according to some linguists, more or less prohibited under

the former regime. E. Eckert (1993: 5) emphasizes that linguists were not allowed to

'investigate language usage in various settings, study speakers' language attitudes,

---

[24] Sgall (2002: 5) states that during the above period empirical fieldwork into CC was carried out solely by foreign Bohemists and the only piece of systematic research on CC is Hammer's (1985) PhD dissertation that examines code-switching in Spoken Prague Czech. However, Nekvapil (personal communication) ascribes the lack of research into the vernacular to the fact that Czech sociolinguistics is grounded in the Prague School tradition, which was concerned predominantly with SC and its cultivation. He suggests that sociolinguistic research into the vernacular did not take root in the Czech Republic – although this line of enquiry was in fact possible – since linguists were preoccupied with other sociolinguistic issues concerning the standard dialect.

record spontaneous conversations, or research unedited materials'. The level of academic research into the vernacular therefore lags severely behind that of most Western societies. In 1987, Nekvapil and Chloupek wrote that '... in principle there is no sociolinguistics in Czechoslovakia ... sociolinguistics is not the subject of lectures in any university, in no institute of linguistics has a department of sociolinguistics been established, and not a single manual of sociolinguistics published in Czechoslovakia is the work of a Czech or Slovak specialist' (1987: 7). The situation has improved since then and introductory courses to sociolinguistics, which tackle the theoretical issues behind the discipline, are currently being offered at Czech universities.[25] The main focus in Czech sociolinguistics is the branch of interactional sociolinguistics (including code-switching, conversation management, intercultural communication) that is associated with the work of John Gumperz. Sociolinguistically-orientated research is generally based on a qualitative and ethnographic methodological approach and variationist and quantitative analyses associated with the Labovian paradigm are rare. Thus, while we can no longer argue that there is no sociolinguistics in the Czech Republic, we can argue that a 'variationist' tradition has still not taken root there.

The collapse of the former regime in 1989 meant that linguistic issues that were previously considered taboo subjects could be freely investigated. Spoken Czech could be studied using innovative methods developed in Western sociolinguistics and collaboration between Czech and foreign Bohemists was made easier. E. Eckert cites

---

[25] The modules at Charles University focus primarily on sociolinguistic problems that relate directly to the Czech language situation (code-switching, diglossia, language planning) and reflect contemporary trends in Czech sociolinguistics. Jiří Nekvapil, who was the first to teach a systematic course in sociolinguistics in Prague, also runs a sociolinguistics reading group for graduate students, at which in addition to discussing selected sociolinguistic texts participants have the opportunity to present their own research and attend guest seminars from visiting researchers.

seven primary objectives that the new 'sociolinguistic' research should fulfil (1993: 4); these are listed below:

1. Introduction of sociolinguistically and functionally orientated research into spoken language varieties.

2. Study of new forms of linguistic expression brought about by the new freedom of speech.

3. New updated edition of the SC codification.

4. Investigation of the actual impact of the expansion of CC.

5. Reevaluation of the role and function of SC.

6. Cultivation of SC in elevated functions.

7. Study of Czech language islands abroad as well as the language of returning Czech emigrants.

These objectives are addressed in E. Eckert's edited volume *Varieties of Czech: Studies in Czech Sociolinguistics*, which was published in 1993 at roughly the same time as the works by Sgall and Hronek (1992) and Sgall et al. (1992) that look to describe the contemporary role, function and areal distribution of CC. Hronek and Sgall (1999) also call upon young Bohemists to research spoken Czech with the aim

of answering the question whether SC can be gradually made stylistically more neutral. They emphasize that researchers should: (1) carry out their investigations into the spoken language in the vein established by Kučera and later by Hammer (1985); (2) analyze the motivation for the choice of a particular variant and speakers' attitudes; (3) strive for a better understanding of the stratification of Czech, about actual language use in the Czech Republic and in other countries; and (4) campaign for a change in stance of schools and other institutions in order that children are objectively informed about the role of CC.

On the one hand, we could argue that some of these objectives have been met and research into spoken Czech has been begun with renewed vigour.[26] As Cvrček (2006: 89) points out, after 1989, investigation into CC was no longer a line of enquiry associated chiefly with foreign Bohemists, but it had become a central issue in Czech linguistics, with the publication of Sgall and Hronek's *Čeština bez příkras* [Czech as it is] in 1992 and Sgall et. al's *Variation in Language* in the same year. However, despite the important political changes, such research has not been as impressive as might have been anticipated or hoped for. Sgall (1994) comments that the post-1989 empirical research into CC is not at a level at which it can describe the language situation systematically and effectively; he adds (1996: 55) that there is still insufficient data with respect to actual language use both in Bohemia and in Moravia, and both in the towns and in the countryside. Čermák and Sgall express similar opinions about the lack of research into the vernacular; they argue that Czech

---

[26] Since 1989, there have been four conferences at which the stratification of Czech has been the focus of discussion. Three of these conferences took place between 1993 and 1995: *Spisovná čeština a jazyková kultura* in Olomouc in 1993 (conference proceedings published in 1995 (edited by Jančáková, J., Komárek, M. and O. Uličný)); *K diferenciaci současného mluveného jazyka* in Ostrava in 1994 (conference proceedings published in 1995 (edited by Davidová)); and *Spisovnost a nespisovnost dnes* in Brno in 1995 (conference proceedings published in 1996 (edited by Šrámek)). A further conference, also named *Spisovnost a nespisovnost dnes* (– *zdroje, proměny a perspektivy*), took place in Šlapanice near Brno in 2004 (conference proceedings published in 2004 edited by Minářová and Ondrášková).

linguistics has a long way to go before to a comprehensive account of the vernacular and a systematic description of the phonology, morphology, syntax and lexicon of spoken varieties has still not appeared (1997: 17).

This holds foremost for the implementation of Sgall's variationist approach to researching spoken Czech that has been largely neglected or given up at the expense of other less well-suited means of studying the vernacular. In the above-mentioned article by Čermák and Sgall, the earlier plans for empirical analysis seem to have been overshadowed by a corpus-based approach. This is unfortunate, insomuch as from the perspective of describing and defining the spoken language, existing spoken corpora have several limitations. Here I enumerate the limitations of the Prague Spoken Corpus (*Pražský mluvený korpus*) and the Brno Spoken Corpus (*Brněnský mluvený korpus*), which are frequently cited in modern studies of spoken Czech. First, both corpora are at present too small and thus do not offer representative data for many important linguistic variables. Second, the corpora do not present data in detailed phonetic transcription and consequently several important factors are potentially overlooked. For example, the corpora will not show whether a speaker realizes *s Jirkou* 'with Jirka' as ['sjirkou] or ['zjirkou], whether *z Indie* 'from India' is pronounced ['sʔɪndɪɛ], ['zɪndɪɛ] or ['sɪndɪɛ], whether *dobrý* is realized ['dobriː], ['dobriˑ] or ['dobrɪ], whether *v pořádku* 'okay' is pronounced ['fpoɾaːtku] or [fpo'ɾatku], or it will not show several other important phenomena. In this respect, we must also assume that the transcribers do not succumb to their native intuitions and that speech is presented graphically in the way that it was originally uttered. While few would fail to notice differences between, say, *dobrý* 'good' ['dobriː] and ['dobrɛj], differences in pronunciation between *bláto* 'mud' ['blaːto] and ['blato] or

tvůj 'your(s)' [tvuːj] and [tfuːj], or even between *jsme* 'we are' [smɛ] and [zmɛ] might go unnoticed. A further limitation is that spoken corpora are to date restricted to just two locations: Prague and Brno. This contradicts Sgall's principle that actual language use should be analyzed throughout Bohemia and Moravia in both rural and urban settings. In connection with this point, speakers in the two spoken corpora are not necessarily natives of either of these towns and this should be taken into consideration in interpreting the results. Additionally, there are major limitations in terms of independent variables. Speakers in both corpora are stratified according to only four criteria: sex, age, level of education and the type of communication. Moreover, for practical reasons the social variables are treated as dichotomous. Thus, for 'age' informants are defined as either 'old' or 'young' and for 'level of education' they are categorized into those with or without a university education. A more substantial problem concerns the treatment of 'type of communication', which is split between the categories 'formal' and 'informal'. [27] These categories are far too broad and as more data are gathered, more categories need to be introduced: communication with family or with friends, in-group or out-group communication, semi-official or official communication, and other types of communication. Nonetheless, the Czech National Corpus (*Český národní korpus*) is still in its infancy and, if it is developed and used correctly, it can undoubtedly make a valuable contribution to research into spoken Czech. It should be emphasized, however, that corpus-based studies alone are of little use in terms of the bigger picture of the stratification of the Czech national language and in terms of providing a reliable and accurate description of actual use.

---

[27] The term 'formal' is used to denote informants' answers to questions delivered by the interviewer, while the term 'informal' is used to describe communication between two individuals who know one another, in situations where the topic of conversation has not been designated by a third party.

Instead, the corpora should be used in combination with data elicited from empirical studies.

In sum, although taboos on various linguistic issues have been lifted and although there is greater freedom to pursue research into the various spoken varieties of Czech and especially into CC, the objectives advanced first by Sgall (1963) and Sgall and Trnková (1963) and 30 years later by E. Eckert (1993) and Hronek and Sgall (1999) have been fulfilled only in part.

3.5 The SC-versus-CC debate in the modern era

Although some linguists would assert that the SC – CC debate has waned since its heyday in the 1960s, there is no apparent ceasefire to the heated discussions and impassioned debates surrounding the relationship between the two codes in Bohemia and Moravia.[28] Bayerová-Nerlichová (2004: 174) believes that this issue is likely to remain at the forefront of Czech linguistics for the foreseeable future, Dickens (1994: 34) even argues that the debate has 'been resumed with renewed vigour in recent times' and Gammelgaard (1999: 32) states that SC – CC issue is 'the main subject of argument among Czech linguists today'. More recently, Nebeská (2003: 103) commented that the unsatisfactory state of affairs as regards the gap between SC and actual language use has in recent years instigated a number of scholarly debates with the aim of: (1) identifying why this situation came about and why it has lasted for so long; (2) describing the contemporary language situation; and (3) proposing an acceptable solution to this problem.

---

[28] Following the fall of Communism, discussions concerning spoken Czech and CC were rekindled in the early 1990s by a number of publications originating in the Sgall camp (see above). An ongoing debate is presently taking place in the Czech linguistic journal *Slova a slovesnost*, inspired by an article written by Čermák, Sgall and Vybíral (2005). The content of these later works is more or less identical to that of Sgall's works published in the 1960s.

The book *Čeština bez příkras* [Czech as it is] published in 1992 by Sgall and Hronek instigated a new vein of interest in the SC–CC debate and in the SC codification. In this monograph, which is not only aimed at linguists but is made accessible to the general public, Sgall and Hronek talk of the need to relax the rigid codification and argue that this is achievable by the admission of CC forms into SC. Sgall and Hronek assert that SC and CC have merged through natural language evolution, including the loss of certain marked items within CC and minor amendments to the still-rigid SC codification.[29] Sgall (1990: 62), for instance, says that Bohemians, on the one hand, are gradually dropping forms such as *ouvozů* (SC *úvozů* 'hollows' (Genitive Plural)), *vejletů* (SC *výletů* 'trips' (Genitive Plural)) and *voblastí* (SC *oblastí* 'districts' (Genitive Plural)),[30] and in another article he argues that even forms like *mlejn* (SC *mlýn* 'mill') and *dobrejm* (SC *dobrým* 'good' (Dative Plural)) are 'possibly' on the decline (2002: 9). Sgall also believes that Moravians are beginning to realize that forms of the type *ta města byla* 'those towns were', *bychom* 'we would', or *s novými stroji* 'with new machines' are archaic and stigmatized in non-formal communication.[31] Therefore, peripheral features of both SC and CC have

---

[29] Linguistic items that were previously marked such as *říct* 'to say', *psát* 'to write', *můžu* 'I can' have been codified alongside their now bookish counterparts *říci, psáti, mohu,* and later *beze mě* 'without me', *dej mně to* 'give it to me', *bez něj* 'without him' along side *bez mne, dej mi to, bez něho.* Within SC, forms of the type *přede dvěma…* 'in front of two…', *přese všecko* 'despite everything' have given way to *před dvěma, přes všecko* (Sgall and Hronek 1992: 107-108, Sgall 2002).

[30] From this list of individual words it is not entirely clear what Sgall has in mind when he talks of the features that Bohemians are giving up. This is especially important with regard to *v*-insertion and also *ý*-diphthongization in the prefix *vý*-, where some words appear to take /v/ or /ɛj/ more than others. Therefore, it is unclear whether we are to assume that by *voblastí* Sgall refers to the decline of *v*-insertion in all positions or just before (non-prefixed) lexical words where /v/ is least common. Hronek and Sgall (1999: 185) list similar forms (*ouvozy, vejstavy, vopory*), which are problematic for the same reasons.

[31] The forms listed by Sgall are generally not used in informal communication, anyway. A Moravian would be just as likely as Bohemian to say *ty města byly, bysme,* realized ['bɪzmɛ], and he or she would probably not say *s novými stroji,* as Sgall suggests, but more likely *s novýma stojama* (CC *s novejma strojema*).

been dropped and Sgall comments that there has been a 'strong shift' of CC in the direction of SC (Sgall 1963: 246), which is used in the way of justifying the addition of further CC forms into the codification.

In a scathing review of *Čeština bez přikras*, Daneš (1993) sums up Sgall and Hronek's aims with the slogan *Puste vobecnou češtinu do spisovný*.[32] He accuses the authors of having insufficient knowledge of the language situation in Moravia and he labels Sgall and Hronek's proposals 'purism inside out' (*brusičství naruby*), arguing that they are no less destructive than purist interventionism of previous generations. According to Daneš, allowing CC forms into the codification *en masse* would 'destroy' SC. In a further review of the controversial publication, Hausenblas (1993) also identifies several problems with the approach of accepting several CC forms into the codification. His principal argument is that admitting CC features into SC would result in a significant reduction in the boundaries between the spoken and written language, that the elimination of bookish forms in favour of CC forms would result in the 'degeneration' of the language, and that brining the two codes closer together would result in the loss of prestige of the standard (1993: 98-99). Hausenblas views the intention of the proposed modifications not as a measure of increasing variability within SC, but as the supplanting of one variety (SC) by another (CC), and this, according to Hausenblas, is comparable to purist interventionism.

In response, Sgall immediately in 1994 and later (2002; 2004) suggests that his aims have been misunderstood. He comments that bridging the gap between the prescribed codified standard and actual language use by admitting CC features into the standard language does not mean the complete replacement of SC by CC or the flooding of SC with CC forms. Instead, the admission of CC forms to the codification

---

[32] To emphasis his point, Daneš intentionally writes the slogan *Puste vobecnou češtinu do spisovný*, which translates as 'Let Common Czech into the standard language [into Standard Czech]', in CC. The SC equivalent is thus: *Pusťte obecnou češtinu do spisovné [češtiny]*).

is intended as a process whereby 'certain' CC forms are 'gradually' accorded the same status as their SC equivalents and then be introduced into the written language alongside them (2002: 4; 2004: 34). Such an approach would not 'destroy' SC but would 'enrich' and 'strengthen' it. Sgall believes that the terms 'spisovnost' and 'nespisovnost' more or less correspond to the purists' 'správnost' and 'nesprávnost',[33] that is, the opposition 'you must–you must not' (*musíš–nesmíš*), and that this precludes variation rather than allowing it (1996: 53). He talks of implementing a system of 'degrees of standardness' (*stupně spisovnosti*), which would allow greater variability and make SC functionally more flexible. If linguistic phenomena were viewed along a stylistic continuum as opposed to being categorized as either 'right' or 'wrong', then oppositions of the type *bychom – bysme* (*by jsme*) 'we would' would both be permissible. The form *bychom* would be considered bookish and *bysme* neutral; both, however, would be viewed as 'standard', as opposed to one form being perceived as 'right' and the other 'wrong'.[34] Sgall emphasizes the role of schools and other institutions in informing the general public of the actual language situation. At present, such institutions are still adhering to a prescriptive approach and the general public is still being told that it is right to write, say, *usnul* but 'wrong' to write *tisknul*; consequently, speakers tend to view bookish forms as 'more standard' than stylistically neutral ones (1996: 54). Sgall believes that the public should be informed that bookish forms are not 'more standard' and that *tisknul* is no less standard than

---

[33] The problem with the term 'nespisovnost' is that it is taken by the general public to denote uncultured or uncivilized speech, as opposed to being understood as a marker of informal communication.

[34] Other linguists also question the system of classifying linguistic phenomena as either 'standard' (*spisovný*) or 'non-standard' (*nespisovný*). Šticha (1995: 57), for example, argues that despite the vast literature on the standard language it is not explicitly clear which criteria a linguistic feature needs to meet in order to be considered 'standard'. He argues that not all forms can be accounted for by the traditional dichotomy 'spisovnost' – 'nespisovnost'; he considers the trichotomy 'spisovnost' – 'hovorovost' – 'expresívnost', which classifies linguistic features as 'standard', 'colloquial' and 'expressive', more appropriate.

tiskl. He believes a more liberal approach needs to be adopted whereby the transitional zone where SC and CC forms coexist should be made clear. Thus, instead, speakers should be instructed according to the principle 'you may–you need not' (*smíš–nemusíš*) (2004: 36-37).

On paper, Sgall's proposals appear to offer a realistic approach to increasing the variability and functionality of SC. First, enriching the standard by increasing its variability is very desirable. We can see the benefits of such variability in the lexicon where, say, *rov* and *hrob* (both meaning 'grave') are considered standard, the former being bookish and the latter stylistically neutral. Other examples include *stále* >< *pořád* 'incessantly', *jelikož* >< *poněvadž* 'since, insomuch as'. This works because the variant forms are common to all dialects. Conversely, most of the morphological items that Sgall and his associates have suggested should be upgraded are regionally restricted. Thus, although the opposition *bychom* >< *bysme* (or *by jsme*), realized ['bizme] in Moravia, is common to all dialects, other forms such as *dobrý sportovci* (SC *dobří sportovci* 'good sportsmen'), *von* (SC *on* 'he'), *velkýho* (SC *velkého* 'big' (Genitive Singular)), (*v*)*on řek* (SC *on řekl* 'he said') are not. Codifying forms that are not used in the whole of the Czech Republic is considered by some linguists to be 'out of the question' (Uličný 1995). Although I side with Daneš and Hausenblas that Sgall and his associates severely overlook the language situation in Moravia, which, we should not forget, constitutes a third of the Czech language territory, and that the codification of a large stock of what are essentially Bohemianisms is problematic, it is also true that certain non-standard forms that are used over a comparatively small area have been codified. Some Moravianisms or forms with a relatively small (or a smaller) areal distribution such as *dej mně* [*mi*] *to* 'give it to me' and the Moravian pronunciation ['zɦlɛdat] (SC / CC ['sxlɛdat]) for *shledat* 'to collect' have been

codified; therefore, Sgall is justified in expecting the codification of geographically widespread, stylistically neutral CC features as well. We might argue that if hypercorrect Moravianisms of the type *sází* 'they bet' have been codified, then surely CC equivalents (*sázej*) should also be codified.[35] Furthermore, if these new analogically levelled forms are accepted, then linguists are justified in their attempts to codify CC *prosej(í)* and *trpěj(í)* (Sgall 1996: 55). Nevertheless, we should also point out that the number of regionalisms admitted into the codification is very small and the principal objection seems to be against the amount of CC forms that would be upgraded.

There is a second problem with the CC proposal that we need to take into consideration. By simply re-classifying certain forms that are currently non-standard as standard, there will not necessarily be a change in lay speakers' perceptions of these forms: speakers will most likely continue to use them as non-standard, which means that little will have been achieved, at least in the short term. It is one thing to change the codification, but it is an entirely different matter to influence speakers' linguistic consciousness. We know this from the changes, say, in the use of capital letters, where speakers are often not aware of the latest conventions. We simply cannot expect that speakers will start to use forms such as *dobrý sportovci* or (*v*)*on řek* as standard. Additionally, such proposals to enrich SC via the codification of non-standard (CC) forms are flawed — at the present time, at least — since linguists' opinions differ dramatically about which variants form part of the hypothetical

---

[35] Third-conjugation verbs traditionally formed two sub-categories in SC, referred to in the present study as *trpět* and *sázet* verbs (§ 6.3.1), whose conjugation differs in the present-tense third-person plural, where a merger has occurred in most Czech dialects. In CC, both types conjugate according to the *sázet* paradigm (-*ej(í)*), while in Moravian dialects both types conjugate like *trpět* (-*í*) in the third-person plural. Forms such as *sází* first appeared in the *Pravidla českého pravopisu* [Rules of Czech Orthography] in 1993, although they had appeared in print and in some 'norm-creating' works before that date. This is an example of linguists' attempts to influence the standard language norm through codification (see footnote 36).

transitional belt of features that are suitable for codification. The problem is that if in a hypothetical situation the admittance of CC forms into SC was sanctioned, there is no reliable mechanism for delimiting which forms should be upgraded, since at the present time we rely solely on linguists' intuitions. This brings us back to the lack of empirical research into spoken Czech: large-scale amendments to the codification cannot be made until more research is carried out into the actual distribution and use of non-standard forms.

In sum, Nebeská (2003: 107) argues that all linguists agree that actual language use should be defined and described in its various communicative functions and they also agree that there should be some kind of regulation of the standard; they disagree, however, how and to what extent the standard should be regulated. Linguists are particularly divided with respect to their attitudes towards 'codification' and language 'norms'.[36] E. Eckert (1993: 9) describes the Czech linguistic community as being currently divided between 'defenders' and 'opponents' of SC as it is presently conceived – although this is somewhat of an oversimplification. The former 'traditional linguists' – here we might talk of linguists such as Bělič, Havránek and Jedlička and later Daneš, Chloupek, Uličný and Hausenblas (Cvrček 2006: 90) –

---

[36] The terms 'norm' and 'codification' were key concepts in the Prague School theoretical tradition. Generally, 'linguistic norms' denote 'linguistic practices which are typical or representative of a group' (Swann et al. 2004: 225). 'Codification' is the process employed by linguists to establish 'prescriptive norms of the linguistic code (i.e. the language system) through the publication of grammars, spelling rules, style manuals or dictionaries' (Swann et al. 2004: 41). Therefore, the codification reflects and represents the norm. With respect to the PLC, 'norm' is interpreted in two ways, both according to the definition listed above and in the sense of a 'standard language norm' (Nebeská 2003: 22-24). The latter is said to be intermediate between use and codification (Nebeská 2003: 26); it was a central element in the PLC's theory of language cultivation and its relationship with the codification is a complicated one. Although the standard language norm serves as the basis for codification, linguists also use the codification in an attempt to influence the standard language norm and as a means of bringing forms that they want to be established in the standard language norm to the public's attention. The problem with the current codification is that certain forms that are codified in dictionaries, style manuals and grammars are not part of the standard language norm and forms which are currently non-standard and that are reportedly used by speakers as if they were standard are not reflected in the codification. There is much disagreement in the Czech linguistic community over what criteria individual forms should meet in order to be codified, to what extent regional variation should be permitted, and so on. For a comprehensive discussion of the terms 'norm' and 'codification' in the Czech context see Nebeská (2003) and Cvrček (2006).

generally oppose rapprochement between SC and CC, although they are willing to accept less obtrusive CC forms into the codification. They continue to denigrate CC and consider it equal to other regionally and socially bound non-standard varieties. They reject the codification of several CC forms on the grounds that CC is not a national vernacular and that it is not used in all parts of the Czech Republic (see Bělič et. al. 1961, Skalička 1962 and Jelínek 1963; and also § 3.10), and that this would also result in the vulgarization of SC. The same linguists also concur that SC should be regulated and maintained strictly by linguists. It is suggested that the majority of (Czech) Bohemists belong in this camp (Nebeská 2003: 109). The latter, 'opponents of SC', as Eckert calls them – although it is probably more appropriate to label them as linguists with more liberal attitudes to the codification of SC (Hedin 2005: 16) – campaign for a relaxation in the level of regulation of SC. Here we might list linguists such as Čermák, Hronek, Sgall and Starý. Some of the linguists in this camp, proponents of CC (Sgall and his associates), campaign for the democratization of the standard language and an updated codification of SC; they envisage a thorough reevaluation of the codification and believe that CC should be described vis-à-vis SC in dictionaries and style manuals. They emphasize the role of CC as a variety with a higher communicative function that is both socially and geographically expansive and which in their opinions is assuming the role of a national vernacular, which is diffused into Moravia according to the 'urban hierarchical' model of geolinguistic diffusion, whereby innovative forms are spread in stages via the larger cities to the countryside. This, however, is not supported by empirical data (§ 3.10).

Owing to this important division between Czech linguists, it seems unlikely that SC will undergo any major changes, at least in the near future. Haugen, for example, argues that codification is unlikely to succeed 'unless the community can

agree on the selection of some kind of a model from which the norm can be derived' (1997: 349). In concluding this section, we can assert that any proposed amendments to SC must meet a number of criteria. A renewed codification must not result in the discontinuity of SC or impoverish its means of expression and it should be acceptable for speakers from all regions, of all ages and from all socioeconomic backgrounds (Nebeská 2003: 106). In addition, any changes that are made should be acceptable for all or most linguists and, most importantly, decisions should be made in view of their impact on the general public (Nebeská 2003: 114).

3.6 Varieties of Czech

3.6.1 Standard Czech

First, let us define what the term 'standard language' means. In Swann et al. (2004: 295) a 'standard' language (also called a 'literary' or 'official' language or dialect (see Chambers and Trudgill 1998: 3-12, Wilson 2003: 4-9 for a discussion of the terms 'language' and 'dialect') is succinctly described as a 'relatively uniform variety of a language which does not show regional variation, and which is used in a wide range of communicative functions'. As opposed to most non-standard varieties, the standard dialect tends to 'observe prescriptive written norms, which are codified in grammars and dictionaries', it is taught in schools and in most cases has a rich literary tradition. Thus, a standard dialect is more prestigious than other varieties of a national language, is functionally diversified and is made up of numerous registers and styles. As Haugen states, a fully-functional or complete standard language 'must answer the needs of a variety of communities, classes, occupations, and interest groups' (1997: 348), it needs to fulfil various cultural functions and it must be capable of meeting its users' communicative needs in all types of situations, both formal and informal.

Attempting to describe its inherent flexibility, Trávníček (1952: 31) likens a standard language to a large building, asserting that, on the one hand, the standard language constitutes a whole, overarching other forms of the national language, but it also has numerous floors and annexes that symbolize its various functional strata.

We have already established that SC due to its turbulent development does not fulfil all the functions that are typically associated with a standard language of the type described above. While SC is the only variety of the Czech national language that is used throughout the whole territory of the Czech Republic and is the unifying (written) code of all Czechs, it is functionally restricted by its lack of a colloquial variety or register (Sgall 1996: 53). Consequently, SC in non-formal communication is replaced by non-standard varieties. The chances of SC becoming the everyday spoken code in Bohemia are non-existent (Sgall 1996: 56) and the number of linguists holding the standpoint that SC is a fully-functional standard has dwindled considerably in recent years; even Daneš, who is often cited as one of the key figures in defence of SC, admits that SC 'does not function as the means of everyday spoken communication in the whole territory of the Czech Republic' (1997: 14; 2003: 11). SC, therefore, can been described as an archaic and semi-artificial (Short 1991:502; 1993: 455) or fully artificial (Čermák 1993: 27, Bílý 1999: 1), functionally restricted (Čermák 1997), socially stigmatized, at least in Bohemia, (Schmiedtová 1995), weak and primarily non-spoken standard that is not used as the everyday means of communication in any part of the Czech Republic and that has only bookish or lofty means of expression for some functions (Sgall et al. 1992: 49, Nebeská 2003: 98-99).

As in a diglossic community, nobody speaks SC as a mother-tongue variety, regardless of which part of the Czech Republic they are born and raised in, and, contrary to popular belief, there is no area of the Czech population where individuals

receive a better grounding in the use of standard from the home environment (Sgall et al. 1992: 30).[37] Instead, SC is acquired at school through formal education – though the school environment will not necessarily be their first encounter with the standard (Hauser 2004). Children are exposed to the frequent code-switching between SC and CC (or other non-standard forms) from a very early age, but it is later when they start their education that they learn to differentiate, with varying degrees of success, between these varieties. Although the rules of SC are for the most part adhered to rigidly in writing, there is a great deal of fluctuation in speakers' ability to use SC, in the way it is prescribed by linguists, in unprepared speech. Speakers' ability to use SC in spontaneous discourse is usually determined by their level of education and / or their profession. Daneš lists school instruction, public, political, scientific and similar discourse, newscasts on radio and television, various announcements by professional speakers as sociolinguistic domains where the standard is expected and observed in spoken discourse. He argues that in general politicians and people from 'higher administration, economy, and management' mostly use SC; 'scientists, technicians, and journalists' also use the standard, 'but to a lesser extent' (1997: 16; 2003: 14). However, with respect to the media the situation is, as Daneš puts it, 'very diversified' and the use / non-use of the standard depends on the type of television programme / radio show. The situation is also less than clear as regards school instruction. Although Hauser (2004) argues that SC should in theory function as the language of

---

[37] Many Moravians erroneously consider themselves to speak SC, primarily because certain morphological endings in some Moravian dialects are identical to those used in SC. Some of my informants said that they were brought up to speak SC and this is the variety they use as their everyday means of communication. Aleš, Alex, Luboš and Zdeňka from Frýdek-Místek in the Ostrava region all considered themselves to speak SC, although they admitted that many people from their region speak *ostravština* 'the Ostrava dialect'. Dan, a central Moravian, also considered SC his mother-tongue variety, despite using several CC forms. Some informants even named the regions in Moravia where they thought SC, or, as some put it, 'the best SC' is spoken. Šárka and Terezka, for example, stated that SC is the spoken code in their home town Zlín. Adélka considered SC to be spoken in the big towns and cities in Moravia. Denisa said she had been told at school that SC is spoken in Přerov (eastern Moravia). Linda thought that 'the best SC' is spoken around the town of Jeseník and in other areas along the north-Moravian border with Poland.

instruction at Czech schools and research has shown that pupils consider their teachers to use SC and see them as role models in this respect (see Svobodová 2004), the reality is that teachers in their unprepared monologues frequently use CC or other non-standard forms. Balkó (2004), Šindelářová (2004) and Zimová (2004) analyzed unprepared monologues of 22 students specializing in Czech language and literature at the Department of Czech and Slavonic Studies, Pedagogical Faculty, J. E. Purkyně University in Ústí nad Labem. The informants, future teachers of Czech language and literature, all used non-standard forms, in particular in phonology, although to a lesser extent in morphology and syntax.

Bílý (1999: 81) questions speakers' ability to use the standard in unprepared discourse, arguing that SC in speech is by and large 'restricted to reading predetermined written texts',[38] and that 'hardly anyone seems capable of adhering to this prescriptive code in its entirety in spontaneous speech'. Sgall et al. (1992: 22-24) list several factors that are linked to speakers' command of the standard which include 'characteristic features of the speaker's personality' and 'the situational motivation';[39] they argue that while some speakers are simply not accustomed to using SC, for others it may be an important part of their job to be proficient in using the standard. Chloupek (1974: 43) comments that, only 'linguistically talented foreign visitors' can speak SC of the type prescribed by linguists, since they have 'acquired this variety through reading grammars and text books'. That said, foreigners living in the Czech Republic who have undergone no linguistic training in Czech will undoubtedly speak the local variety, because they have no recourse to the standard. This is evident in,

---

[38] Bílý lists 'school instruction', 'public lectures' and 'church sermons' as occasions where speakers attempt to use the standard.

[39] For a discussion of the factors that can account for variation in speech see Dickens (1994: 21).

say, the Vietnamese community where shops signs are sometimes written in CC (Čmejrková 1996: 226).[40]

3.6.2 Colloquial Czech

The term *hovorová čeština* 'Colloquial Czech' (hereafter, ColC) is particularly difficult to pinpoint and its meaning has undoubtedly changed several times since it was introduced in the 1930s (Bělič, the most vigorous proponent of ColC believes it started to develop a century earlier in the 1830s and 1840s (see also Auty 1986: 85-86)). ColC was originally viewed by Havránek as a spoken form of SC, 'used in informal conversation but outside the intimacy of the family and close friends' (Dickens 1994: 30) or as a code used by 'relatively educated speakers in communications of a conversational nature, both within and outside the family circle' (Kučera 1955: 577), via which the 'upwardly mobile' forms of CC were admitted into SC, and which at one stage was anticipated to become the spoken norm. It occupies a

---

[40] Grygar-Rechziegel (1990: 18-19) considers that it would seem 'unnatural' for non-native speakers of Czech to use CC and that native speakers would perceive this as 'being out of place'. She also argues that a poor command of CC would be evaluated more negatively than a poor command of SC. Likewise, Short asserts that unless a speaker was fully competent in using CC the effect would be no less bizarre 'than a French attempt at Cockney' (1991: 503). From my own observations, I would suggest the converse. I have been told on numerous occasions (in Bohemia) not to use SC forms such as *to je dobré* 'it's good' or *velké pivo* 'a big beer', but to use the CC equivalents *to je dobrý* and *velký pivo*, which are more 'natural'. First, since native speakers frequently and sometimes randomly switch between SC and CC forms and nobody speaks 'pure' CC, it is difficult to imagine what a 'poor command' or 'competent usage' of CC actually is. Second, most foreigners living in the Czech Republic have not undergone any formal training in Czech before coming to the Czech Republic and have no recourse to SC – a predominantly unspoken code – but are familiar only with the regional variety. In terms of native speakers' perceptions, it obviously depends on the individual and on the communicative situation. While using CC during a class at the university may be perceived as strange and / or frowned upon, in informal communicative situations non-native speakers who use CC are, impressionistically speaking, not judged negatively for doing so – at least, this is my own experience. Short's comparison of CC to Cockney is a particularly bad example. While Cockney is a highly-stigmatized, highly-localized vernacular that is restricted to working-class speakers, CC is a socially unmarked variety over a large area of the Czech Republic and is the mother-tongue variety of up to 70 percent of all Czechs. It would be more appropriate – though still not a perfect match – to compare CC to Estuary English, a variety that over the last twenty years has become less geographically and socially bound (L. Milroy 2002: 8) and is perhaps becoming the regiolect of south-eastern England (see Britain 2005: 998-1000). This is important, since, again, impressionistically speaking, if a foreigner were to use highly localized forms of my native dialect, then I would probably find this amusing; if, however, he or she were to use supralocal features common to all or most dialects of northern England, then I would not consider this odd. I suspect that similar parallels can be drawn with respect to Czech.

special place in Czech linguistics, insofar as it was introduced by linguistic decree (Schmiedtová 1995: 84) as a means of bridging the gap between SC and CC. Bělič's original interpretation (1959: 437-438) of ColC is 'roughly the more economical spoken version of standard Czech, lacking all conspicuous high-style features and gradually absorbing some features previously considered sub-standard' (Short 1991: 503); he campaigned for the spread of ColC as the spoken form of SC to be used by speakers in informal communication, namely the one area where SC had failed to be implemented. Bělič believed that due to the growing status of ColC archaic and bookish forms such as the infinitive suffixes -ti were receding within SC and that ColC was bringing SC closer to actual language use (1960: 129). Kopečný (1949: 17) also views ColC as the spoken form of SC that contains neither formal means of expression associated with prepared speech nor regional or vulgar forms. Thus, ColC was initially viewed as an independent variety of the Czech national language and some linguists claimed that many Czech speakers are trilingual, using SC, CC and ColC. In this case, ColC forms a kind of intermediate zone (*mezistupeň*) between the two (Auty 1976). According to Auty, SC is reserved for the 'most formal of utterances', CC is used in 'completely relaxed', intimate and familiar situations or in utterances of an emotive character, and ColC is used in 'normal cultivated conversation'.

Empirical research has, however, more or less quashed this argument. The analyses carried out by Kučera in the 1950s and 1960s identified ColC as the result of a compromise or as the interference between SC and CC, and not as a clearly-demarcated distinguishable entity with its own phonological and morphological forms. Čermák, for instance, argues that ColC has no formal means of expression and does not form a complete code in itself (1993: 27-28). Sgall and his associates also

interpret ColC as term of convenience that 'describes a variety of styles of spoken Czech which fall between the phonological systems of both SC and CC' and which 'represents a compromise between the two varieties' (Dickens 1994: 31). Daneš (1988: 24) describes Bělič's formulation of ColC as chimerical, stating that it is 'wishful thinking' on the part of Czech linguists; but he does talk of 'a colloquial register of standard' (2003: 11) in delimiting the varieties of the national language. This holds for many other linguists who nowadays prefer to talk of a colloquial layer or stratum of SC. Chloupek, for instance, believes ColC to be more a 'style' of SC and he proposes the term *hovorový styl spisovné češtiny* as a more suitable alternative to the problematic *hovorová čeština* (1974: 40). The term *hovorový styl* is interpreted by Jelínek (1966: 108) as 'a set of tendencies for selecting the communicative means for an everyday, informal speech act'.

However, even in this second sense of a 'stratum' or 'transitional belt' of forms that are neither bookish nor non-standard or of a 'colloquial style', ColC is problematic, insomuch as linguists have not been able to successfully delimit the forms that belong in this colloquial stratum. Kopečný (1949) and Bělič (1958) have made subjective attempts to describe and define forms that they consider to be part of ColC (see, for example, Kopečný 1949:17-20 and also Cvrček 2006: 37), but these are incomplete (Jelínek 1963: 48) and seem to overlap in many cases with forms other linguists consider as belonging to CC. Čermák, for instance, states that 'it is far from clear' what should be understood by the term ColC (1987: 134) and nobody seems entirely sure which items can be labelled as belonging to it. Furthermore, the forms that Bělič has enumerated as belonging to ColC are considered by other linguistics as nothing more than deviations from SC in the direction of CC (Micklesen 1978: 445). In fact, many linguists have moved away from the term ColC and from the 1970s

onwards have leaned more towards the term *běžně mluvený jazyk* (or *běžná mluva*),[41] which literally translates as 'Ordinarily Spoken Czech' or 'Ordinary Speech', which more accurately describes the speech of everyday speech encounters that are made up of various elements, including features of SC, CC, interdialects, and traditional dialects. Chloupek (1974: 50; 1987: 15) points out that *běžně mluvený jazyk* is interpreted either as a mix of all elements that occur in speech (CC, interdialect, dialect, and items from the colloquial 'register' of SC), or as a mixture of purely non-standard items; the former definition is more frequent and corresponds to Daneš's oft-quoted interpretation of the term (see, for example, 1995b: 7-8; 1997: 15). Daneš stresses that this highly variable and heterogeneous mixture of varieties is observed in those situations where the standard is not expected.

3.6.3 Common Czech

SC fulfils a different role in Bohemia than it does in Moravia. In Bohemia and western Moravia, SC coexists with a regionally and socially expansive variety known as Common Czech (*obecná čeština*) and the language situation there has been compared to diglossia (see, for example, Grygar-Rechziegel 1990). Some linguists have gone as far as to say that the language situation in Bohemia is an almost perfect match for Ferguson's (1959) original formulation (Micklesen 1978), while others, more accurately, describe the language situation in Bohemia as manifesting certain features characteristic of a diglossic community, while pointing out that it differs from Ferguson's original formulation of diglossia on a number of levels (Sgall and Hronek

---

[41] Although not as vague as *hovorová čeština*, the term *běžně mluvený jazyk* (*běžná mluva*) is also interpreted in different ways (Dickens 1994: 39). Müllerová and Hoffmannová assert that the term *každodenní komunikace* 'Everyday Communication' would be a more adequate alternative to *běžně mluvený jazyk* in the way that it is used by Daneš (1995: 7-8).

1992, Bermel 2000).[42] In what I shall refer to hereafter as 'CC-speaking territories',

that is, the whole of Bohemia and western Moravia, SC is reserved for the utmost

formal situations and is generally avoided by native speakers in informal

communication, whereas CC serves as the everyday means of expression in all non-

formal communicative situations and is supplanting SC in certain (semi-)formal

sociolinguistic domains (Krčmová 1987: 129). Usage of SC – or rather certain SC

forms – in informal situations may be deemed 'stilted', 'pompous', 'arrogant', or

'inappropriate', or even 'amusing' (Sgall and Hronek 1992) and its use in Bohemia in

informal communication is no longer perceived as a marker of prestige as it was in

previous times (Schmiedtová 1995: 85).[43]

There is a great deal of overlap and borrowing between the two codes and CC

grammatical forms can be readily combined with lexical items from SC. In fact, it is

argued that CC is structurally closer to the standard than any of the traditional dialects

are (Sgall et al. 1992: 49-59), and differences between SC and CC forms are rarely

clear-cut (Čermák 1993: 27). Despite the apparent structural proximity between the

two codes, some linguists insist that CC is a language variety in its own right, rather

than a mere cluster of phonological and morphological items or a kind of allegro

---

[42] There are two problems with regarding Czech as a case of (classic) diglossia. First, the two codes are structurally very similar; they share grammatical and lexical items, there is extensive mutual borrowing between the two codes (Sgall 1963: 247) and speakers do not necessarily know what forms belong to SC or to CC, or which code they are using at a given time. According to Havránek (cited in Sgall et al. 1992: 71-72), even in morphemics where differences between the two codes are most pronounced their number does not exceed ten percent. However, later descriptions of diglossia do allow for this (see Paolillo 1997). Second, and more importantly, features of the low code are frequently observed in sociolinguistic domains of the high code and vice-versa, which under classic diglossia would be 'ludicrous' or 'outrageous' (Schiffman 1997: 206). Code-switching between SC and CC is frequently observed intra-sententially or even within the same word, thus some linguistics prefer to use the term 'code-mixing' to describe the rapid oscillation between forms in a single utterance (see Hammer 1985, Kraus 1993).

[43] This statement is not backed up by reliable evidence. As I assert at various points in this thesis, only individual SC forms 'appear' to be stigmatized in informal use in Bohemia and not SC as a whole. It would be very useful to research: (1) how Bohemians interpret the term spisovná čeština and what features they believe belong in this variety; and (2) what forms they consider inappropriate for informal communication.

pronunciation as is the case of the non-standard vernacular in French. Hronek (1972), Townsend (1990), Sgall and Hronek (1992), Sgall et al. (1992) are devoted to a comprehensive description of the differences between SC and CC phonology, morphology, syntax and lexicon, and the reader is referred to these works for further reference.

3.7 Interpretations of CC

The term *obecná čeština* 'Common Czech' has not been clearly defined and is generally subject to interpretation (see, for example, Krčmová 2000 and her definition of CC in *Encyklopedický slovník češtiny* 2002). The term was coined by Havránek (1924) in the sense of a 'variety' used by speakers in out-group or semi-official situations, in which they require a more prestigious code than their socially and regionally restricted dialects. Nowadays, definitions range from the view of its most active proponents, who consider it a second central variety of the Czech national language and / or a majority vernacular used in non-formal situations of a nationwide character (Sgall et al. 1992), to a mere cluster of phonological and morphological forms (Daneš 1995b: 8; 2003: 12-13; 2004: 29, Romportl 1996: 76, Müllerová and Hoffmannová 1997: 45), and even a 'chimera' (Uličný 1995: 23). This polar opinion is held by its staunchest rivals who reject its status as an independent language variety.[44] Daneš asserts that the adjective *obecná* 'common' in the term *obecná čeština* should be understood in its original meaning of 'vulgar' or 'demotic', not as 'widespread' or 'prevalent' (1995a: 92) – and this is how it is often interpreted by the general public. According to Townsend (1990: 12), many non-specialists use it interchangeably with other problematic and loosely-defined terms such as *hovorová*

---

[44] Those who question the existence of CC as an independent variety consistently adjoin the abbreviation 'tzv.' (*takzvaný* 'so-called') to the term *obecná čeština* (see comments in Müllerová and Hoffmannová 1997: 45, Gammelgaard 1999: 40).

čeština 'Colloquial Czech' *lidová čeština* 'Demotic [lit. the people's] Czech' and *běžná mluva (běžně mluvený jazyk)* 'Ordinary Speech'. CC was in fact formally used as a synonym of *lidový* or *obecný jazyk* (Daneš 1997: 15), both of which can be briefly and simply defined as any non-standard variety or a mixture of non-standard varieties used in non-formal spoken situations. CC has been interpreted in the following ways:

1. Any variety of non-standard (and substandard) speech, including traditional dialects, interdialects, slang, argot, and so on.

2. A mere cluster of morphological and phonological forms.

3. The interdialect spoken in central Bohemia.

4. A regionally and socially restricted interdialect used throughout Bohemia and western Moravia that is functionally equal to the interdialects used in Moravia.

5. A more or less homogeneous variety of language that is used with minimal variation and is socially unrestricted throughout Bohemia and western Moravia where it replaces SC in all non-formal communicative situations and is supplanting it in other 'semi-official' domains.

6. A second central variety of the Czech national language or a majority vernacular that 'constitutes the basis or core of the national language together with the standard' (Sgall et al. 1992).

To summarize, (1) is an out-dated view of CC or the erroneous interpretation of many non-specialists; (2) is the opinion of Daneš and his associates (see above); (3) describes CC from a historical perspective; (4) and (5) are probably the most common and least controversial definitions: many Moravian linguists often reject its functional superiority over other non-standard varieties (4), while others emphasize that it has exceeded its purely interdialectal status and has a higher communicative function in CC-speaking regions (5); finally, (6) is the view of some Praguian and foreign Bohemists who propagate CC as a spoken code that is used in all most of the Czech Republic. For the purpose of this study, CC is understood as (5). To elaborate, CC is considered here as an interdialect (§ 3.8) that resulted from intensive dialect levelling in the fourteenth and fifteenth centuries in central Bohemia as a result of increased mobility and urbanization and that has gradually spread throughout Bohemia and western Moravia, supplanting the more localized traditional dialects. Since it has been used for centuries as a nativized variety, it is relatively stable and homogeneous (Chloupek 1974: 55) and is used with little variation throughout the CC-speaking territories by approximately 65 percent of the Czech population.[45]

We should clarify what is meant by 'minimal regional variation'. Skalička, for instance, says that the regional origin of a speaker from Bohemia can only be

---

[45] According to the statistics presented in Englund (2004), Bohemia has an area of 52062 km^2 and a population of 6.27 million inhabitants, while Moravia and Silesia have an area of 26808 km^2 and 4.07 million habitants. In terms of area Bohemia is, therefore, twice the size of Moravia and Silesia and in view of the population over 60 percent of Czechs live in Bohemia. Since CC is the mother-tongue variety in western Moravia, the figure rises to between 65 and 70 percent.

identified by very occasional interference of more localized forms (1962: 202).[46]

Jančáková (1993: 181; 1995: 127), on the other hand, considers the popular belief that speakers in Bohemia use a variety of language that is by and large the same an oversimplification. She states that 'the widespread opinion that there are no longer any true dialects in Bohemia corresponds to the situation in the undifferentiated centre, rather than to the situation on the periphery'.[47] Dialects in the villages in Krkonoše mountain range in northern Bohemia and in the Chodsko and Doudlebsko regions in south and south-west Bohemia, for example, have maintained a number of regional forms (Bachmannová 1997; 1998, Janečková 1997). Historically, there were marked differences between East Bohemian dialects and South Bohemian dialects and Jančák (1997a: 202) lists Prague as a former transitional point, where the south-west and north-east Bohemian isoglosses converged and where the speech of some elder residents may still indicate which part of the city they are from. Nevertheless, as a result of intensive dialect levelling (towards CC) many regional features have been lost. This holds for enclaves in the core of Bohemia that had been resistant to CC forms (Jančáková 1987), for the peripheral areas (Janečková 1997), and also for the areas that were re-populated after the Second World War (Jančák 1997).

3.8 Interdialect

The term 'interdialect' is interpreted in different ways in the literature on dialect contact, where it is a relatively recent addition to the specialist terminology, and in

---

[46] Markéta (my insider, § 5.6) reported that when two speakers from different parts of Bohemia meet it is impossible for them to tell where the other is from. Sgall and Hronek (1992: 89) also suggest that for the most part speakers from Bohemia cannot be placed in a particular location and that an individual feature might be the only clue as to their region of origin.

[47] The view that Bohemia and western Moravia is a linguistically homogeneous region is generally taken as axiomatic and has yet to be investigated systematically.

Czech linguistics, where it has been established for many years. In Siegel (1993b: 114) Peter Trudgill is named as the coiner of the term 'interdialect', which he adapted from Selinker's 'interlanguage' (1972). In the literature on dialect contact, an interdialect is the result of 'situations where contact between two dialects leads to the development of forms that actually originally occurred in neither dialect' (Trudgill 1986: 62). However, in the literature on Czech linguistics the term 'interdialect' has a much longer tradition. It was first used by Havránek in 1932,[48] who described it as a variety that dialect speakers use in out-group but non-formal communication when they wish to use a variety that is not locally restricted (Hronek 1987: 82). This type of variety has also been described later as a 'macrodialect' (Dickens 1994: 40) and a 'supralect' (Kučera 1973: 500). The term is defined in the Czech-language literature (Nekvapil and Chloupek 1987: 14, Šipková 2002: 183-184, Čechová 2003: 50) as 'an unstable form of the national language which represents the highest evolutionary stage of the traditional territorial dialects – devoid of exclusively local features and yet containing innovations in accordance with the new evolutionary laws of the national language'.

This definition is somewhat vague and we can attempt here a more comprehensive description. An interdialect is the result of dialect contact and subsequent dialect mixing. In comparison with traditional dialects, it has simpler morphophonemics and is devoid of highly-localized forms. It consists of features from the source dialects that have a wide areal distribution; it borrows features from the standard language; it contains variants that have been diffused from one (inter)dialect region to another; and it also manifests some hybrid forms that are not observed in any of the source dialects but are the result of contact between them.

---

[48] Havránek had already talked of an intermediate variety between SC and traditional dialects in 1924 (see Sgall et al. 1992: 12).

Similar to other varieties that emerge as a result of dialect contact, an interdialect is highly unstable and inherently heterogeneous. Initially, it is used as a second language variety in out-group non-formal communication and, depending on extralinguistic factors, may undergo nativization. It is important to point out that an interdialect, although closer structurally to the standard dialect, remains socially and regionally restricted and is inappropriate outside informal communicative situations.

Let us take, for example, the Silesian interdialect. This variety is devoid of highly-localized and largely peripheral features such as the Opava instrumental ending *–um* or the phonological shift /o/ → /u/ in monosyllabic and disyllabic words *pozur, tchuř, druzd* >< SC *pozor* 'attention', *tchoř* 'polecat', *drozd* 'thrush (bird)', expect perhaps in lexicalized or idiomatic uses, for example, *ňema to ani ruk, ani nuh* (literally, 'it has neither hands nor feet'), which is the equivalent of SC *nemá to ani hlavu, ani patu* 'it makes no sense'. It maintains features such as penultimate stress and lack of vocalic length and progressive assimilation in the consonsant clusters <tv>, <kv>, <sv>, regional syntactic constructions like *do doktora* (SC *k doktorovi* 'to the doctor's), *přišla tady* (SC *přišla sem* 'she came here'), and *bo* (SC *protože* because'). There are also new innovations from the standard language such as *oni sou* 'they are' in place of regional *oni su* and *oni pišou* 'they write' (SC *oni pišou*) instead of regional *oni piš'u* (Šrámek 1979: 92), and from other Moravian interdialects *s těma hodinama* (SC *s těmi hodinami* 'with those clocks') instead of the traditional dialect form *s tyma hodž'ynoma*. In addition, there are forms common to all Moravian dialects that have spread from one dialect region to another / others: for example, old fourth-conjugation imperative suffixes *-ite* and *-ime* that were preserved in some Moravian dialects *myslite!* (SC *myslete!* 'think!'), *myslime!* (SC *mysleme!* 'let's think!'). These have been diffused throughout Moravia and by analogy are used with

first- and second-conjugation verbs: *otevřite!*, *otevřime!* (SC *otevřete!* 'open!',
otevřeme! 'let's open!') and *začnite!*, *začnime!* (SC *začněte!* 'start', *začněme!* 'let's
begin') (Šaur 1995a: 57-58). And, finally, there are other common Moravian variants
that did not occur in any of the source varieties but are the result of contact between
them: for instance, the common Moravian suffix *–ijou* in examples such as *mluvijou*
(SC *mluví* 'they speak') replaces regional forms like *mluvijó, mluvijú, mluviju,
mluvja, mluvija.*

From this description we can see that the principal difference between
Trudgill's use of the term 'interdialect' and its application in Czech linguistics lies in
how we interpret the term 'intermediate' in relation to the new mixed dialect. While
for Trudgill an interdialect is a 'structurally' intermediate variety between two or
more dialects in contact, an interdialect in the Czech sense is a 'functionally'
intermediate variety between a traditional dialect or dialects and the standard dialect.
Both varieties are the result of contact-induced change: Trudgill's 'interdialect' arises
out of a specific kind of mixing, while the formation of an interdialect in the Czech
sense encompasses all or several of the processes involved in dialect contact and
mixing.

3.9 Is Common Czech an Interdialect?

Undoubtedly, CC began to develop in a similar fashion to the Silesian interdialect
described above; it was in fact considered by Havránek in the 1930s equal to the
Moravian interdialects (Hronek 1987: 82). However, in its present state, CC differs
from an interdialect in the strict sense of the term on a number of levels. It is the
mother-tongue variety of all its speakers; it has been diffused (and continues to be
diffused) beyond its original heartland (Central Bohemia); and, most importantly, it is

not restricted to non-formal sociolinguistic domains, unlike the Moravian interdialects which are replaced by a higher code in such situations, but it fulfils a higher communicative function and has supplanted SC in a number of semi-formal communicative situations. CC's social and geographical expansion has been facilitated by a number of factors. CC is the variety spoken in Prague, the cultural, administrative and economic centre of the country. It has enjoyed centuries as nativized koine and it has had no competition from other supralocal non-standard varieties; therefore, CC has exceeded its purely interdialectal status and, as Nekvapil and Chloupek (1987: 14-15) and Čechová (2003: 51) assert, it also serves as 'a kind of standard' or 'a common language' with 'a higher communicative function'. CC is encountered on the television and on the radio, more so on private radio stations and television channels than on state-owned ones in programmes aimed towards younger viewers or listeners, on talk shows, reality shows, awards ceremonies and televised celebrity events, and it is used frequently in the theatre. [49] CC has gained currency in certain popular magazines, especially as a device for presenting interviews with celebrities (Hoffmannová 2004, Müllerová 2004), it is used in advertising and it is the norm for many speakers in means of electronic communication such as e-mails and SMS text messages. CC plays an important role in literary texts, not only as an aesthetic device in the speech of characters but also as the narrative language and it is observed in various types of public discourse. Furthermore, the social (and perhaps geographical) expansion of CC is inevitable as a direct result of the changing face of

---

[49] The use of CC in such semi-official situations generally differs from that in everyday spontaneous speech. I would argue, based solely on personal observations, that participants in talk shows and at awards ceremonies (and also in interviews) do generally make efforts to modify their speech, but as a consequence of the liberalization of society they feel less obliged to replace some features of their native dialects with SC equivalents. We observe, therefore, a kind of 'elevated' CC where certain non-standard forms that individuals for whatever reason consider appropriate are retained, while others are given up. Obviously, there is a great deal of idiolectal variation in this respect. The idea of a kind of elevated CC is supported by findings in Maglione (2003) and Hedin (2005), whose studies, which both analyze language variation in contemporary television discourse, show a considerably lower ratio of CC forms in comparison to the results highlighted in studies of spontaneous speech.

the Czech Republic and the democratization of Czech society. Bílý (1999: 7-8), for example, talks of a reversal in speakers' attitudes towards SC and CC in connection with the former regime: before the fall of Communism, proposals to bring SC and CC closer together were seen as a 'conspiracy' to 'take away' the mother-tongue variety, whereas nowadays SC is by some individuals deemed as the 'hypocritical language of the old regime'. Dickens (1994: 27) holds a similar opinion, believing that under the Communist regime 'the sterile debates, the self-important proclamations and the empty rhetoric and platitudes which characterized the Communist press' were exclusively SC domains, while nowadays CC has a firm position in the media, in particular in 'downmarket papers and journals'.

Consequently, CC is not an interdialect *per se*. It constitutes, in fact, a fourth variety of the Czech national language (Krčmová 1979: 69): a 'majority' or 'central' – though I emphasize not 'national' – variety (Hronek 1987: 83) that should not be treated as an interdialect of the Moravian type, and certainly not as a regionally and socially restricted traditional dialect. As a result, terms such as 'common language', 'regional standard', 'majority / national vernacular', 'second central variety of the Czech language' have been used to describe CC. One of the more frequent suggestions is the term 'koine'. Daneš labels both CC and the Moravian interdialects as 'regional koines with different degrees of common character and extent of use' (2003: 11); Swann et al. write that if an 'interdialect' (in Trudgill's sense) becomes stabilized and becomes the first language of a given speech community, then this is a koine (2004: 149). The term koine, however, (often written 'koiné', with an accented *é*) is problematic. Siegel (1993a: 5) argues that 'of all the imprecise terms used in sociolinguistics, "koine" may win the prize for the widest variety of interpretations' and Bubenik (1993: 9) views it as 'one of those useful but rather ill-defined linguistic

terms that has to be rediscovered and redefined because it has been around for twenty-three centuries or so'. The term comes from the Greek word 'koinē', which means 'common', used to describe Hellenistic koine, which emerged as a result of dialect contact and mixture and whose phonology and morphology were structurally less complex than that of the contributing dialects. It was predominantly used as a second dialect, although it eventually became standardized and used for writing. Many other varieties have been labelled as koines (or 'koinai') and linguists interpret the term differently. Siegel (1985: 360) lists six different definitions of the term, which include a lingua franca, a common language, a compromise variety, and a levelled dialect and it appears that some linguists emphasize the functional role of the koine, that is, the fact that it is shared as a common language by speakers of different vernaculars or that it is a language variety used over a large geographical area, not necessarily as a mother-tongue variety, while other linguists focus on standardization based on the original Greek koine.

Siegel (1985: 363) argues that defining a koine as a 'common language' or 'lingua franca' is too broad. Definitions, according to which a koine is the result of a spoken dialect being standardized, do not take into account mixing, which is the central element in the formation of a koine. Succinctly, a koine can be defined as the 'stabilized result of mixing of linguistic subsystems'. Trudgill (1986: 107) emphasizes that a koine is the result of the levelling out of minority and marked items and is stripped of, or has a significantly lower number of, irregular features that were present in the contributing varieties. Functionally, a koine is used as 'a lingua franca among speakers of the different contributing varieties'. Structurally, its morphophonemics are less complex and more regular than in the contributing varieties. There are various stages in the formation of a koine or in koineization. First, a 'prekoine' is formed

when various forms of the contact varieties are used concurrently and inconsistently, levelling, simplification and reduction have taken place, but few compromise forms have emerged (Siegel 1993a: 6-7). The interdialects that are developing in Moravia could be, therefore, considered as prekoines (prekoinai). The next stage is when stabilization and linguistic homogenization occur and phonological, morphological, syntactical and lexical forms of the emergent variety are used with minimal variation. Beyond the 'stabilization' stage the koine 'may be extended to other domains besides the intergroup communication' (Siegel 1993a: 6-7) and even become a standard language. A koine which undergoes social expansion is known as an 'extended koine'. Additionally, a koine may become nativized and may be spoken as a mother-tongue variety by its users ('nativized koine').

Therefore, CC does seem to be a reasonable match for Siegel's formulation of an extended koine – if the Central Bohemian interdialect, about which there is insufficient research (Sgall et al. 1992: 168), did in fact arise as a result of dialect contact and mixture. CC is used over a wide geographic area as a lingua franca and today almost exclusively as a mother-tongue variety; its phonology and morphology are, generally speaking, simpler and more regular than that of traditional dialects and it has extended into sociolinguistic domains beyond intergroup communication and is continually gaining currency in non-informal discourse. In view of the above, I would suggest the term 'extended interdialect' as an accurate description of contemporary CC.

3.10 Common Czech in Moravia

Problems regarding the areal distribution of CC were being discussed as early as the 1960s. Skalička (1962: 201) amidst the ongoing debate taking place in *Slovo a*

slovesnost expressed that at the time many things were still unclear about CC, one of which being its areal distribution. He did assert, nonetheless, that the vernacular in large Moravian towns such as Ostrava and Prostějov was 'completely uninfluenced' by CC. Today there is much debate with respect to whether or not CC is spreading eastwards into Moravia. It has been theorized that due to migration and movement of the population, increased contact and via the media (Hronek and Sgall 1999: 188)[50] CC forms are spreading beyond traditional CC-speaking territories and that the geolinguistic diffusion of CC forms into Moravia takes place via the large urban centres (Čermák and Sgall 1997). Cummins (1993:155) talks of CC items in the city speech of Brno and in the Silesian town of Ostrava, where it is encountered as a kind of hybrid form, mixed with regional Silesian items. Hronek and Sgall (1999: 188) claim that CC has made inroads in central Moravia, especially in Brno and Třebíč, though less so in towns in north-central Moravia such as Olomouc and Krnov (see the map 1.1 in the Appendices). Other linguists have also reported CC forms in Moravia, but they have stressed that they are used for effect in certain networks of individuals. For example, it has been claimed that CC forms are encountered in Silesia in intellectual circles (Chloupek, Lamprecht and Vašek 1962, Šrámek 1979; 1997, Hronek and Sgall 1999). It has also been suggested that some young speakers in Silesia are adopting certain CC intonation patterns, in particular the open realization

---

[50] Several Czech linguists believe that CC is being diffused via the mass media. However, it is disputed to what extent linguistic phenomena can be diffused from one group of speakers to another without face-to-face interaction. Hoffmannová (2001: 29) states, according to her own intuitions, that the spread of CC forms into Moravia is significantly influenced by the media, while Chambers (1998: 24) asserts that without face-to-face interaction the media can do little more than diffuse individual ephemeral lexical items; he rejects the influence of the media at the phonological or morphological level. This latter view has generally prevailed among sociolinguists. However, recent research carried out by Stuart-Smith et al. (2007) has uncovered a potentially meaningful relationship between television and language with respect to accent change. Under the influence of television, certain adolescents from Glasgow appear to be assimilating phonological forms characteristic of the dialects of southern England. Interestingly, a particularly strong and significant factor in the adoption of the linguistic variants was engagement with London-based television programmes, in particular with the soap opera *Eastenders*.

of short /ɪ/, which is similar to /ɛ/ (Höflerová 2004: 281), and that this odd linguistic behaviour is characteristic of adolescents who for whatever reason respect CC and use it as a way of indexing their identities.[51]

On the other hand, other linguists point out that the above opinions are not based on reliable evidence (Hausenblas 1993: 98) and they argue that there is a growing body of studies that does not support the view that CC is spreading eastwards. According to Skalička (1962: 203), CC has clearly expanded within Bohemia, becoming the dominant code in the repopulated Sudeten regions, but there is no reliable evidence to attest the eastward shift of its isoglosses into Moravia. Linguists, particularly those of Moravian origin,[52] have claimed that CC forms are not observed in any great quantity in Moravian towns and that CC is perceived as an alien variety in Silesia (Davidová 1993; 1997), in Uherský Brod and its immediate environs (Hlavsová 1988) and other regions in eastern Moravia (Krčmová 1993, Uličný 1995, Kolařík 1996), in the towns and villages of north Moravia (Kloferová 2000), and even in Brno (Krčmová 1987: 130), which is often used as a yardstick for measuring the diffusion of CC and where other linguists have claimed an increase in the use of CC features. Krčmová believes features such as raised /iː/, diphthongized /ɛj/ and prothetic /v/ were part of the original dialect and developed independently from CC; she rejects the notion that they have spread from CC-speaking territories (1977: 119). Uličný (1995: 23-24) argues that with very minor exceptions CC features have not

---

[51] This is very interesting, insofar as it suggests that speakers are picking up features of CC without engaging in face-to-face contact with native speakers. Šrámek (1997) believes that certain individuals in Silesia use CC forms in attempts to copy non-standard speech that is used on certain television shows. Krčmová (1977: 118) also states that certain speakers in Brno accommodate to the Bohemian intonation pattern typical of the Czech spoken on Czech television in official speeches.

[52] Gammelgaard ascribes the stance of Moravian linguists with respect to CC, which they frequently downgrade to a variety comparable with their own localized (inter)dialects, to 'extralinguistic motivations' and 'a political wish to reduce the influence of Prague' which 'may explain the opposition to any functional and territorial expansion of CC' (1999: 39).

made inroads beyond western Moravia and he states that CC has had no influence whatsoever in south-east Moravian or in Silesia. However, it is necessary to point out that most of the assertions that reject the geographical expansion of CC are also based on introspective evidence. A striking example of this is Šaur (2001: 344). He claims that if CC is heard in Opava, then it is definitely not spoken by a local resident. He also comments (1995b: 73) that the further east you go, that is, the further away from CC-speaking territories, the less accepted and the more marked CC becomes. This, of course, is Šaur's own opinion; it may indeed be representative of his immediate group of associates or the academics at the Department of Linguistics at the Silesian University in Opava where he works, but without systematic empirical research that analyzes language use in various sub-sections of the community it is impossible to claim that CC forms have not penetrated into local use.[53]

Probably the most accurate description of the geolinguistic diffusion of CC forms into Moravia is Krčmová (1979: 70): in her words, we can at best 'only guess and speculate' as to what extent CC is spreading into Moravia based on the available evidence. Krčmová believes that the only concrete example of the spread of CC is in western Moravia where contact between CC and local dialects is most intensive. Her various analyses of young speakers in Brno (1974; 1977; 1979) have shown that CC forms for paradigm unification (§ 6.3.1), *l*-truncation (§ 6.3.2) and gender neutralization (§ 6.3.3) have not made headway into the local dialect. Paradoxically, even Sgall admits that little is known about the actual spread of CC into Moravian towns or into the regions situated in direct proximity to CC-speaking territories (1996: 55), ascribing this lack of knowledge to the aversion of many linguists to study CC.

---

[53] Similarly, Cummins (1993: 156) comments that 'it is a fact of present-day society that nearly all native Brnoites [residents of Brno] (intellectuals and blue-collar workers alike) express strong dislike of *obecná čeština*'. However, once again, it should be noted that Cummins does not base this statement on the results of a survey completed by representative samples of intellectuals and blue-collar workers in Brno; thus, it cannot be taken as representative.

He calls for more research into the spread of CC forms in Moravian towns such as Jihlava and Svitavy, which are located in close proximity to CC-speaking territories.

3.11 The language situation in Moravia

First, it should be pointed out that the language situation in Moravia has not been researched in equal measure to CC and it has been neglected by many foreign and some domestic linguists at the expense of the quasi-diglossic situation in the western half of the Czech Republic. Traditionally, Moravian dialects are studied as three major (inter)dialect groups: Central Moravian (*Haná*), East Moravian (earlier known also as 'Moravian-Slovak dialects' (*moravskoslovenská nářečí*)) and Silesian ('Lach', 'Lachian' or 'Laskian') dialects, within which, despite intensive levelling towards supralocal regiolects (commonly referred to in the literature as 'interdialects' (§ 3.8)), there is further and in some cases extensive regional variation. For example, there are three specific subgroups of East Moravian dialects: Valašsko in the north of the region, Slovácko in the south, and Horňácko which is situated on the south Czech-Slovak border and consists of around ten villages, each of which are claimed to have their own distinct dialect that manifests several similarities with neighbouring Slovak (see Bělič 1972). Other groups of dialects in Moravia that should be mentioned are the transitional belt of Czech-Polish dialects (*Polsko-český smíšený pruh*), a series of distinct dialects spoken in the region where the Czech Republic meets Poland and Slovakia and that are heavily influenced by those neighbouring languages, in particular Polish, and to a lesser extent by Slovak and German, and the belt of transitional Bohemian-Moravian dialects (*nářečí přechodného pásu česko-moravského*) in the Vysočina region that manifest features of both CC to the west and Central Moravian dialects to the east. This latter transitional zone encompasses the

towns of Jihlava and Žd'ár nad Sázavou in the north of the region and Jemnice in the

south. For a more comprehensive account of Moravian dialects the reader is referred

to Bělič (1972). The various dialect regions are presented in Figure 3.1 (from

Davidová et al. 1997: 147).

A = Bohemian-Moravian transitional zone
B = Central Moravian dialects
C = East Moravian dialects
D = Silesian dialects
E = Czech-Polish mixed dialects
F = Area of mixed dialects

Figure 3.1 *Dialect map of Moravia*

Empirical research into language variation in Moravia has been largely carried

out in the traditional dialectological vein and looks to capture regional forms that have

survived despite the intensive dialect levelling that has been taking place there from

the latter half of the nineteenth century onwards; therefore, the emergence of

supralocal regiolects has been to a large degree overlooked. Wilson (2003), who sought to analyze Moravian dialects in terms of linguistic phenomena that were shared by all or most Moravian dialects, found substantial similarities between the three groups, in particular between East Moravian and Central Moravian dialects. Several forms were identified as 'Common Moravian', that is, as being common to all or most Moravian dialects. Importantly, this classification included several recent innovations that had emerged as a result of dialect contact and subsequent koineization or that had been diffused from one dialect area to other regions. Interdialectal innovations of the type described by Trudgill (1986) were also identified. Wilson believed that the differences between the three major interdialect groups were exaggerated and that since research had leaned heavily towards the mapping of peripheral localized forms used predominantly by older speakers, descriptions of language situation in Moravia are out-dated. Nevertheless, it was concluded that despite the intensive levelling of highly localized forms and the formation of supralocal interdialects that manifested several items of a 'Common Moravian' nature, a similar variety to CC is unlikely to develop in Moravia. While CC had the optimum social conditions to develop into a common language and supplant more localized varieties due to the fact that it was spoken in Prague, the economical, cultural and economic heart of the country, the formation of a parallel variety in Moravia is hindered because of the low prestige of Moravian dialects, the lack of a major cultural centre, and, possibly, due to competition from CC.

Consequently, the role, prestige and function of SC in Moravia, where the language situation resembles the classic 'standard – dialects'[54] situation (Chloupek 1995: 224), or more accurately the trichotomy 'standard – interdialects – dialects', are

---

[54] For a definition of a 'traditional territorial dialect' in the Czech sense, see Chloupek (1987: 13).

different than in Bohemia. In Moravia, the role of SC is more prominent in informal, especially out-group communication (Kučera 1955: 577). In a nutshell, while Bohemians strive to avoid SC (or at least some SC forms) in informal communication, Moravians strive to embrace it (Hammer 1993: 73), revering it as the sole prestige variety that they have at their disposal (Chloupek 1995: 224). Attitudes towards the standard are also positive (Nekvapil and Chloupek 1974: 13). Davidová (1997: 99) and Chloupek (1969: 148) point out that Moravian speakers acknowledge SC as the prestige code and strive to use it in non-informal contexts as much as possible. Praguian linguists often assert that, since Moravians neither have a common lingua franca parallel to CC, nor do they accept CC as 'a domestic (unmarked) variety of the national language' (Krčmová 1987: 123), they have a stronger desire or motivation towards the standard and they evaluate it more positively than speakers in Bohemia do (Sgall and Hronek 1992). Therefore, SC in non-formal communicative situations is in no way stigmatized. This does not mean, however, that Moravians are better at speaking SC than their Bohemian neighbours. Davidová (1997: 140) points out that speakers of Moravian dialects in attempting to use SC often make errors or slip back into their mother-tongue variety. In many cases, such errors are hypercorrect forms that are made under the influence of the speakers' native dialects. Such is the situation for speakers of Silesian dialects, which are structurally the most distant from SC and consequently impede acquisition of the standard (Balhar 1995, Bogoczová 2004). Balhar, for instance, states – again, not based on empirical data – that out of all Moravians, Silesians are the most linguistically insecure (1995: 253); they consider their dialects 'ugly' (škaredý) and view them as suitable only within the narrow family network. He argues that such linguistic insecurity motivates them to drop regional forms and as a result Silesian speakers are known to make hypercorrect

errors in an attempt to speak SC. Balhar (1995: 253-255) continues to say that such is the motivation to distance themselves from their highly localized dialect that Silesians in cases where SC permits more than one item as standard choose the one that is different from their dialectal form. For instance, the forms *z Jeseníku* and *z Jeseníka* are both accepted as standard in the genitive of *Jeseník* (a town in northern Moravia) and here Silesian speakers choose *z Jeseníku*, because, according to Balhar, the non-standard genitive ending *-a*, which in the standard is used mainly for animate objects, is abundant for inanimate objects in their regional dialects (*duba* (SC *dubu* 'oak'), *stroma* (SC *stromu* 'tree'), *nosa* (SC *nosu* 'nose')). Occasionally, this results in hypercorrect forms: *ze Zlínu* instead of SC *ze Zlína* 'from Zlín (a town in eastern Moravia)' and *do lesu* instead of SC *do lesa* 'into the woods', which, according to Balhar, have become permanent errors (1995: 255).

From the above, it is evident that the language situation in the Czech Republic is asymmetrical. On the one hand, in Bohemia, where industrialization and urbanization occurred almost two centuries ago and where this led to intensive dialect levelling, a stable and relatively homogeneous variety has taken shape. This variety is used over a wide geographical area, has, functionally, exceeded its role as socially restricted interdialect and now functions as a variety with a 'higher communicative function', which in view of the rapidly changing face of Czech society will inevitably replace SC in more sociolinguistic domains. On the other hand, industrialization and urbanization took place much later in Moravia and the intensive dialect levelling is a relatively recent phenomenon in the eastern part of the Czech Republic,[55] where feudalism caused linguistic diversity, a tendency to isolation, and the formation of

---

[55] According to Utěšený (1962: 579), dialect levelling was complete in Bohemia by the nineteenth century. On the other hand, in central Moravia intensive levelling started after the First World War and it did not begin in eastern Moravia or Silesia until after the Second World War, and then almost exclusively in urban areas.

distinct dialect boundaries (Chloupek 1969: 144). And this delayed levelling and formation of the interdialects.

Moravians are believed to evaluate CC negatively, as an alien code and 'an uncalled-for feature of "Pragocentrism"' (Daneš 2003: 12). Sgall et al. (1992: 252) assert that the reason for this is that Moravians are unaware of the markedly different language situation in Bohemia and of the fact that CC is functionally different from their own dialects. They state that most Moravians consider CC equal to their own regionally and socially restricted (inter)dialects[56] and that Moravians are often shocked or surprised when Bohemians use CC in out-group situations outside the circle of family or friends, considering such behaviour 'vulgar' or 'disrespectful'. Many Moravians argue that the Bohemians are murdering the language with their CC. Informants' attitudes towards CC are described in section 7.7.2.

3.12 Summary

The asymmetrical language situation described above, in particular the asymmetrical social relations of the varieties in contact, will undoubtedly influence the rate and intensity of accommodation in contact situations between speakers of regionally different varieties of Czech. We have established that Bohemians are at an evident advantage, insofar as their native variety is socially unmarked over a large geographical area and speakers have little need to diverge from CC in non-formal or out-group communication. Their native variety (CC) is used by over half of the Czech population and it is spoken as a mother-tongue variety in Prague, the economic, cultural and administrative centre of the country. Speakers of Moravian dialects, on the other hand, are likely to behave differently in situations of dialect contact, since

---

[56] Sgall et al. (1992: 252) assert that this is also the view of some linguists, particularly those of Moravian origin.

they constitute a linguistic minority in the Czech Republic and do not have a supralocal or socially unrestricted variety that is accepted outside the in-group. We can confidently say, therefore, that they will drop or reduce the number of regional forms and that they will accommodate towards a more prestigious variety. In situations where the contact is between speakers of different Moravian dialects we might expect accommodation in the direction of SC; however, in contact between speakers of CC and speakers of Moravian dialects the facts presented above suggest that accommodation might be, instead, towards CC. The optimal social conditions for the assimilation of CC forms are outlined below:

1. The students' native dialects are socially and regionally restricted and Moravian dialects are generally perceived negatively in Bohemia (as markers of provincialism).

2. SC does not fulfil the role of the everyday means of communication and its use in Bohemia is socially marked and disadvantageous in all but the most formal situations.

3. CC, although non-standard, is socially unmarked in non-formal communication in over half of the Czech Republic and it is accepted (by its users) in out-group and semi-formal communication.

4. Because of the social expansion of CC, Moravians are frequently exposed to its forms on television shows, radio broadcasts and in works of popular literature and in downmarket or 'trendy' newspapers and magazines, and since

over half of the Czech population speaks CC as a mother tongue, Moravians are likely to have engaged in face-to-face interaction with CC speakers before moving to Bohemia.

4 *The contact hypothesis*

4.1 General description

As we stated in the introduction, there is no systematic or reliable evidence with respect to the results of dialect contact between speakers of regionally different varieties of Czech. Nonetheless, certain linguists have speculated about the results of contact between speakers of CC and speakers of the more localized Moravian dialects, both in short-term interactions and more so during prolonged contact. Linguists have predicted the outcome of accommodation in situations where speakers of Moravian dialects have moved to Bohemia and vice-versa (Sgall 1990; 1996, Hronek 1968; 1992, Sgall and Hronek 1992; 1999, Sgall et al. 1992). I label this set of predictions the 'contact hypothesis'. Those who are most active in hypothesizing about the outcomes of dialect contact are typically the linguists who promulgate CC as a majority vernacular and / or a second central variety of the Czech national language. Consequently, their hypotheses reflect, and are influenced by, their stance with respect to the role of CC (and the role of SC). The contact hypothesis based on material presented in the above works is summarized below:

1. The fact that Moravia is divided into various dialect or interdialect regions and is linguistically more heterogeneous than Bohemia means that Moravians tend to avoid locally restricted variants of their native dialects in contact situations with speakers from other dialect regions (Sgall 1996: 55). The social stigma attached to their native dialects and their linguistic insecurity means that Moravians have a stronger motivation to use the standard in everyday conversations than speakers of CC do (Hronek 1992: 122, Sgall 1996: 55); the motivation is even stronger in out-group (contact situations).

Conversely, since Bohemians have at their disposal a socially unrestricted and supralocal 'koine', they do not need to use a variety other than their mother tongue in out-group (contact) situations.

2. Speakers of Moravian dialects or interdialects are at an 'evident disadvantage' in contact situations with speakers of CC (Hronek 1968: 34); and speakers of a highly-localized, non-prestigious variety are much more likely to converge towards the speech of speakers of a majority vernacular than vice-versa (Sgall et al. 1992: 22-23).

3. Speakers of Moravian dialects living in Bohemia quickly reduce and avoid features of their highly-localized vernaculars and assimilate features of CC (Sgall and Hronek 1992: 90, Sgall et al. 1992: 195-196).

4. Speakers of CC living in Moravia do not assimilate features of the highly-localized dialects or interdialects in equal measure: instead, they support the spread of CC into Moravia (Sgall and Hronek 1992: 90, Sgall et al. 1992: 195-196).

To summarize, according to Sgall and Hronek's contact hypothesis, speakers of CC, a majority though non-standard vernacular, do not face the same social pressures in interpersonal contact situations as speakers of the more localized Moravian dialects. Speakers of Moravian dialects are more likely to adjust their speech styles in order to gain the social approval of their interlocutors, although not necessarily with the aim of

improving communicative efficiency.[57] This holds not only for contact situations in which Moravians engage in face-to-face interaction with speakers of CC, but also in contact situations between speakers of the different Moravian dialects or interdialects (Balhar 1995).

4.2 Problems with the contact hypothesis

The contact hypothesis controversially points towards an asymmetrical pattern of accommodation, its formulators claiming that speakers of Moravian dialects who move to Bohemia quickly start to use CC, while accommodation in the opposite direction does not involve speakers of CC assimilating features from the localized vernaculars, but instead promotes the spread of CC beyond its heartland. Although the contact hypothesis is not supported by empirical evidence, it is easy to see why linguists such as Sgall and Hronek expect accommodation to follow such a pattern in view of the contact-specific factors that were summarized in the previous chapter. At least, this is the true for the first part of the hypothesis; the contact-specific factors do not, in my opinion, support the second part of the hypothesis that Bohemian migrants in Moravia help diffuse CC. The first part of the hypothesis is also problematic for several reasons. Sgall and Hronek do not base their claims on reliable data. Their assertions are based solely on introspection and their own intuitions. And, as we know, the intuitive assertions that linguists make about their native dialects must be treated with caution. As Labov (1972c: 113) argues, 'a linguist's idiolect is often an

---

[57] It is generally assumed that all dialects of Czech are mutually intelligible, although this has not been tested in any way. According to the results in Vybíral (2004), who conducted a survey to elicit informants' attitudes towards non-standard but frequently occurring items, Bohemians were 'one-hundred-percent sure' that they would be understood without any difficulty if they spoke CC in Moravia; the same informants were only eighty-percent sure that Moravians would be understood if they used their native dialects in Bohemia. From a personal observation, many Bohemians tend to have an incomplete or inaccurate knowledge of Moravian dialects, often listing features that are used only peripherally or talking of highly localized regionalisms as if they were commonly used by all Moravians.

accumulation of superposed varieties' and it is not representative of the local vernacular: linguists speech habits are often detached from the mainstream norms of the society in which they were born and raised. Since neither Sgall nor Hronek is Moravian, they are describing a situation of which they have no direct experience and we must take it to mean, therefore, that their argument is based on the linguistic behaviour of speakers of Moravian dialects who they know in some capacity, either as family members, friends, acquaintances, or work colleagues. This, again, is problematic in terms of representativeness. Our attitudes and perceptions are undoubtedly shaped by the communities of practice in which we engage and our commonly held beliefs are generated by the social behaviour of our immediate circle(s) or network(s) of associates. However, such communities of practice or networks constitute a very small sub-section of the community as a whole; thus, by relying solely on our intuitions we can only describe the processes that occur within those individual networks or groups.[58] While we should acknowledge that in recent years the situation has improved and, though fieldwork methods still need to be refined, more systematic and empirical research is being done, it should be emphasized that introspection as a methodological tool is still very much overused in

---

[58] To illustrate the above point, let us consider an example, which we looked at in the previous chapter. Šaur (2001: 344), a Moravian linguist, claims that people from Opava definitely do not use CC forms. We concluded that this statement reflects Šaur's own opinion and that it may be true of the attitudes of his immediate group of associates; we added, however, that this statement should not be taken as representative for the larger speech community in Opava, since without systematic empirical research that analyzes language use in various sub-sections of the community, it is impossible to claim that CC forms have not penetrated into the local vernacular.

Czech linguistics and the 'inductive' approach still outweighs the 'empirical'

approach.[59]

A second major problem in the proposed hypothesis regarding the linguistic

behaviour of Moravian migrants in Bohemia is the interference of language

ideologies. The contact hypothesis was formulated by linguists who actively support

the spread of CC. This is perhaps most evident in the prediction that Bohemian

migrants in Moravia help to diffuse CC eastwards. As we established in Chapter 3, the

name Petr Sgall is viewed by some scholars as being almost synonymous with the

campaign to have CC forms upgraded and codified alongside their SC equivalents.

Therefore, any suggestion about the functional expansion of CC or its functional

superiority over other non-standard varieties that originates in the Sgall camp is

sometimes quickly dismissed or not pursued by other linguists.[60] This is not to say

that the contact hypothesis, or for that matter other proposals made by Sgall and his

associates about the Czech language situation, is inaccurate or improbable; in fact, the

contact hypothesis does seem perfectly logical to the onlooker, given the information

---

[59] However, we should accept that the latter approach is not without its problems. Bermel, for instance, illustrates the problems associated with statistical analysis in the case of the Czech situation (see 2000: 40-41). For the linguistic variables that are analyzed in the present study, the statistical approach is problematic in *v*-insertion in lexical words and *ý*-diphthongization in word roots, where the use or non-use of the non-standard element may vary from word to word, rather than being constant throughout the whole lexical set.

[60] I put the above predictions to groups of undergraduate and graduate students of language and linguistics at Charles University. To my surprise, the contact hypothesis evoked strong criticism and was rejected almost unanimously. Many students considered that speakers would not accommodate to CC. Reporting on their own language use, students from Moravia claimed that they dropped regional forms that may preclude intelligibility or that were stylistically marked, but they claimed to use SC forms as 'compromise' or 'neutral' variants between their native dialects and CC. They were adamant that they would not assimilate CC forms and they expected my informants to behave in a similar manner. The students' comments can be interpreted (and criticized) in a similar way to the predictions that Sgall and Hronek make, since we can argue that both points of view are influenced by their attitudes towards the Czech language situation. Emphasizing the non-functionality of SC, the functional superiority of CC over other spoken varieties and the stigmatization and markedness of Moravian dialects, Sgall and Hronek logically conclude that CC will be used in this contact situation as the neutral variety. The students, on the other hand, while accepting that Moravian dialects are socially stigmatized in Bohemia and in out-group communication in general, viewed CC as a low-status territorially-based and socially-restricted dialect and assumed, again logically, that speakers would use SC, which in their opinion is the only neutral or appropriate variety for this contact situation.

that was considered in the preceding chapters. The problem is that to be taken seriously the hypothesis needs to be supported by empirical evidence.[61]

A further issue that the contact hypothesis makes no attempt to answer – and which the present study aims to clarify – concerns the statement that speakers of Moravian dialects who move to Bohemia 'start to use' or 'accommodate to' CC. To say that an individual uses a given variety or accommodates to it is vague and ambiguous. Both means of expression can be interpreted in several ways, ranging from the individual having fully mastered a given variety to simply using the odd feature from that variety every now and then. Presumably, Sgall and Hronek do not expect first-generation migrants to speak CC in the same way that native speakers use it. At least, this would go against the general literature on dialect contact. Speakers' assimilation of forms of a new variety is a complex process, shaped both by a combination of language-internal and extralinguistic factors that relate directly to individual linguistic variables and various other external constraints that are specific to the individual speaker. First-generation adult migrants seldom succeed in the full assimilation of a new language variety and we expect to observe partial or incomplete accommodation.

---

[61] The contribution of Petr Sgall to Czech linguistic merits our special attention. On the one hand, Sgall deserves much praise for bringing CC and non-standard use in general to the forefront of Czech linguistic enquiry. As we saw in Chapter 3, Sgall has fought consistently for a sociolinguistics of Czech and he has set the way for future research into spoken Czech, devising a methodology that young Bohemists can use to objectively describe the contemporary Czech language situation. Furthermore, many of Sgall's ideas have been developed by foreign Bohemists, who have carried out empirical studies of the semi-diglossic language situation in Bohemia. However, despite his efforts in promoting sociolinguistically-oriented research of the varieties of Czech, and his innovative ideas for the application of variationist methods in Czech linguistics, Sgall himself has relied almost exclusively on intuitions in his own descriptions the Czech language situation, propagating what are for the most part logical but nonetheless unsubstantiated hypotheses. Therefore, much of Sgall's own research goes against the methodology he advanced in 1963 and in later works. For this reason, Sgall's work is sometimes criticized by other Czech linguists. This is unfortunate, insomuch as Sgall's ideas, although undoubtedly influenced by his own views towards the Czech language situation, certainly merit empirical analysis.

4.3 Empirical studies of dialect contact in the Czech Republic

To my knowledge, there are only three studies that investigate the linguistic behaviour of speakers of Moravian dialects living in Bohemia (Jančák 1978, Bachmannová 1996, Jonášová 2001) and there are no accounts of accommodation in the opposite direction. In the first study, Jančák looks at the results of dialect contact in the second generation. Out of a larger sample of 107 informants, he studied the linguistic behaviour of five junior-school pupils who were born and had lived all their life in Prague but whose mothers were of Moravian origin. The mothers of three of the informants were from central Moravia, and the mothers of the other two informants were from Silesia. The study is a quantitative analysis of three phonological variables: *v*-insertion (§ 6.2.1), *é*-raising (§ 6.2.2) and *ý*-diphthongization (§ 6.2.3) and Jančák's findings showed that there were no discernable differences in the linguistic behaviour of his informants who had at least one parent of Moravian origin and their classmates whose parents were both born and raised in Bohemia. In his conclusion (1978: 205), Jančák attributes this to the strong influence of the Prague speech community on in-migrants.

However, for most sociolinguists, Jančák's findings are in no way revolutionary or surprising, since it is accepted that in cases where families move to another speech community, children acquire the local vernacular spoken by their classmates and friends rather than adopting the variety spoken by their parents or teachers (Labov 2001: 307; 423, Chambers 2003: 175). Chambers, for instance, states that in cases where families have migrated to a new speech community, children before the critical age under normal circumstances 'never' adopt the vernacular of

their parents[62] – although parents' region of origin has been shown to be important in the acquisition of complex rules[63] (Trudgill 1974, Payne 1980) and in new-dialect formation scenarios.[64] Therefore, the influence of CC in this particular case is insignificant and a similar pattern of accommodation in the second generation can be expected where, say, pre-adolescents of Bohemian parentage were born and lived in Moravia.

There are only two empirical studies that address the accommodation of speakers of Moravian dialects living in Bohemia in the first generation: an unpublished monograph (Jonášová 2001) and a short article (Bachmannová 1996). Jonášová provides a qualitative account of accommodation, recording informants from Moravia speaking amongst themselves and later with a native speaker of CC. She also asked them in a follow-up interview to listen to extracts of their own and their interlocutors' speech, which they were asked to evaluate, and to justify their selection of linguistic phenomena. Jonášová studied the speech of three individuals (two female, one male) who were born and raised in northern Moravia and who had moved to Prague and had all lived there for six years. The aim of Jonášová's study

---

[62] Newbrook found a striking counter-example in his West Wirral study (1982). Informant 143, a twenty-year-old male, used several Scottish features in spite of the fact that he had lived since the age of four in Merseyside and had not spent any periods of a month or over outside the area. Newbrook (1982: 80) describes informant 143's behaviour as typical of 'a young, educated Scot who had spent three or four years on Merseyside immediately preceding the interview'. Although his mother was Scottish, we would expect his linguistic behaviour to reflect more the norms of his peers. Newbrook attributes this anomalous behaviour to the family's involvement in an isolationist religious sect, which discouraged informant 143 from taking part in social activities outside school and denied him access to wider social networks.

[63] Trudgill (1974: 36) asked 20 informants aged between 30 and 40 to read out 'in a proper Norwich accent' the sentence *Norwich City scored an own goal*. The ten informants whose parents were born and raised locally produced the correct local pronunciation /ʌun guːl/, while the ten informants whose parents were born elsewhere did not. Similarly, the adolescents who participated in Payne's (1980) study failed to fully assimilate the Philadelphia short /a/ pattern if their parents were not born and raised locally.

[64] In such situations – in new towns or in colonial settings, for instance – there is no peer-group dialect for children to accommodate to. Instead they adopt 'individual accommodation strategies' and parents' dialects do play an important role (Trudgill 2004: 101).

was to: (1) analyse in what ways her informants speech habits had changed during the time they had lived in Prague; and (2) identify if and in what manner their attitudes towards the different varieties of the Czech national language, for example, their perception of their native dialect and also their opinions of CC, SC and any other variety had altered. The interviews were carried out on the same day at the flat of one of the participants. In the first interview, the three speakers from Moravia, one of whom was Jonášová, participated in a general conversation, the topic of which had been selected before the recording had begun. The second part of the interview was a continuation of the same topic, although this time a native speaker from CC was introduced into the conversation. Although Jonášová describes the interview as a natural, informal and spontaneous speech act, the topic of conversation had been decided upon before the recording began, thus presumably reducing the informality or naturalness of the proceedings. Jonášová predicted that in the second interview the Moravian informants would accommodate towards the informant from Bohemia and adjust their speech styles accordingly. However, the results show that the Moravian informants made no attempt to converge towards the speech patterns of the speaker from Bohemia and the non-accommodation is confirmed by the informants in the follow-up interview, which was recorded two days after the first interviews. In fact, the only traces of accommodation that were observed were on the part of the speaker from Bohemia, who admitted in the follow-up interview to accommodating towards the speech of the Moravians.

The linguistic behaviour of Jonášová's informants could have been, however, anticipated as a direct result of the methodology she employs. If it was her intention to identify to her informants' accommodation to CC, she should have looked to create an interview situation that reflects an everyday situation in which speakers of Moravian

dialects are usually outnumbered by speakers of CC. Furthermore, two of Jonášová's informants were students of Czech language and linguistics, who for the purpose of sociolinguistically-orientated research are particularly bad informants (§ 5.3).

Bachmannová (1996) uses Jančák's earlier study as her point of departure and investigates the same phonological variables. She uses a quantitative approach to compare the ratio of regional forms to CC forms for two women from the Břeclav region in southern Moravia, who had lived for a relatively short time in Prague. The first informant, known as 'informant A', had lived in Prague with her husband, a native resident of Prague, for two years and at the time of recording was on maternity leave. The second informant, labelled 'informant B', was a trained hairdresser; she had moved to Bohemia because she could not find work in her home town and she had lived there for four years with her aunt, who was born and raised in Prague. Informant A is a relative of Bachmannová and several conversations were recorded on family visits. Conversations with informant B were recorded in the hairdresser's salon where she worked and occasionally at her home – she was a neighbour of the interviewer. On all occasions, neither informant was aware that they were being recorded and the surreptitious recordings provided examples of spontaneous and naturalistic speech. Initially, Bachmannová states (1996: 208) that informant B, who comes into everyday contact with speakers of the host variety as part of her job, appears to show a higher frequency of CC forms than informant A, while informant A, who had a lower level of contact with speakers of the host variety, used more regional features, in particular localized lexical forms. This is most probably topic-related, since conversation with informant A centred almost exclusively around the family. Although the two informants differ in levels of exposure to the host variety and in a number of other social aspects, Bachmannová does not compare the levels of

accommodation for the two informants. Instead, she claims that the differences in the use of CC variants between the two informants were negligible and her tabulated data represent the combined score for both informants (adapted from Bachmannová 1996: 209-210):

Table 4.1 *The combined scores of Bachmannová's informants for v-insertion*

	/o/	/vo/
Personal pronouns (*on, ona, ono, oni*)	96.3%	3.7%
Prepositions (*o, od, ...*)	100%	–
Lexical words	100%	–

Table 4.2 *The combined scores of Bachmannová's informants for é-raising*

	/ɛː/	/iː/
Nominative / Accusative singular of hard neuter adjectives	10.3%	89.7%
Oblique cases of hard masculine / neuter adjectives / pronouns	14.3%	85.7%
Oblique cases of hard feminine adjectives / pronouns	33.4%	66.6%
Inanimate plurals	18.1%	81.9%
Word roots	25%	75%

Table 4.2 *The combined scores of Bachmannová's informants for ý-diphthongization*

	/iː/	/ɛj/
Word roots	88.8%	11.2%
Prefix *vý-*	79.6%	20.4%
Desinence-final position of hard adjectives / pronouns	69.2%	30.8%
Desinence-initial position of hard adjectives / pronouns	84.5%	15.5%

Bachmannová's informants use a very small proportion of CC forms for *v*-insertion and *ý*-diphthongization and the only instances of prothetic /v/ are in grammatical words – it is not clear how many tokens Bachmannová analyzed, since the data are presented only in percentages. Bachmannová lists *é*-raising as occurring in the dialects of both speakers and does not take this into consideration. She

concludes – rather succinctly and without further analysis – that the linguistic behaviour of both informants differs markedly from that of speakers from Prague and from Bohemia in general in terms of its proximity to SC, in particular in phonology (1996: 209). In addition, she describes occurrences of some CC phenomena as 'one-off accommodations', that is, the occasional reproduction of a feature or features of the host variety under the influence of the interlocutor's previous utterance. This was observed most frequently in informants' responses to questions posed by the interviewer (adapted from Bachmannová 1996: 210):

Interviewer: *To Blanka byla nemocná celej tejden?*

Informant A: *Jo, celej tejden, minulý týden byla nemocná, celý týden byla doma.*

(Interviewer: *Was Blanka ill all week?*

Informant A: *Yes, all week, she was ill last week and spent the whole week at home.*)

This example illustrates that Informant A uses the CC forms *celej tejden* immediately after Bachmannová's question, but in the other positions she maintains SC /i:/, using the SC forms *minulý týden* and *celý týden*. Therefore, it is argued that accommodation would be even lower if these anomalous cases were omitted from the study.

Bachmannová's conclusion – that the speech of her two informants from Moravia, in terms of the linguistic variables under study, differs markedly from that of native speakers of CC – is disappointing for a number of reasons. It goes without saying that we would not expect the two informants to use the three CC variants in the same way as native speakers do; however, the results show that some accommodation,

albeit minor, has occurred. Bachmannová could have analyzed the data more rigorously, suggesting why some CC forms were acquired and others rejected. For example, there are differences in the use of CC forms with diphthongized /ɛj/ in desinence-final position of hard adjectives, in desinence-initial position and in word roots, which correlates with the status and distribution of this variant in speech of native speakers of CC. A second feature which Bachmannová overlooks is the relatively high rate of acquisition of non-standard /iː/ in the oblique cases of hard feminine adjectives and pronouns, where it is 'not' encountered in the informants' native dialects. Additionally, Bachmannová could have suggested reasons why prothetic /v/ was categorically avoided by the informants by considering which factors could have inhibited its acquisition. We can argue that Bachmannová's study displays many of the methodological shortcomings identified in Jonášová's analysis. She can be criticized for her choice of informant, since, as Jonášová, she seems to have recorded the informants that were the easiest to find: in this case, a neighbour and a family member. Similarly, this micro-level study of the linguistic behaviour of two individuals from just one dialect region in Moravia is of little practical use at the macro level and in order to make generalizations about the linguistic behaviour of Moravians living in Prague a larger and more stratified sample of informants from various dialect regions is required.

In sum, it has been established that there is virtually no empirical evidence that we can use to confirm or disprove the contact hypothesis, insomuch as the very limited data that we have at our disposal are non-representative and have severe methodological shortcomings. This means that at the moment we can only speculate about the long-term linguistic accommodation of speakers of Moravian dialects living in Bohemia.

5 *Methodology: fieldwork strategies*

5.1 The research site

Fieldwork was carried out at the Kajetánka hall of residence in Prague from September 2004 to September 2005. The hall was originally built in the late 1960s as a dormitory for foreign students, although it now houses both home and international students enrolled at the Second Medical Faculty (*Druhá lékařská fakulta*) and the Faculty of Physical Education and Sport (*Fakulta tělesné výchovy a sportu*) at Charles University (pictures of the hall are included in section 2 of the Appendices). The hall consists of two autonomous buildings: the 'main' building, Kajetánka One, made up of two high-rise tower blocks that collectively accommodate up to 620 students – the total capacity is much larger, but part of the hall functions as a hotel and rooms are reserved for tourists – and Kajetánka Two, which also houses around 600 students and is situated approximately 100 metres away from Kajetánka 1. Most of the students who live at the hall are from the Czech Republic, although Kajetánka also houses a large number of students from Slovakia, who receive free education in the Czech Republic, and international students, either reading medicine, or on short-term Erasmus-Socrates programmes.

During the academic year 2004 – 2005 there were 792 students officially resident at the hall. This figure excludes students on short-term foreign exchange programmes. According to the hall records, 503 students were Czech (281 Bohemians and 222 Moravians), 108 Slovak, 112 were international students, and for 69 of the students on the list there were no details regarding their place of birth and / or permanent residence. The breakdown of students living at the hall is presented in Figure 5.1:
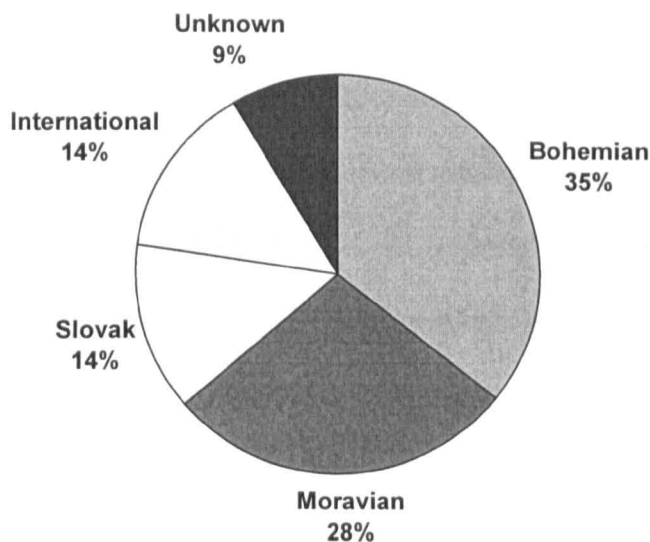
Figure 5.1 *Breakdown of the region (or country) of origin of the students officially registered at Kajetánka in the academic year 2004 — 2005*

We should also point out that 63 of the 222 Moravian students were from western Moravia, where CC is spoken as a mother-tongue variety. Therefore, of the students from the Czech Republic, a total of 344 (68.4%) were native speakers of CC – which corresponds to the national distribution of CC speakers in relation to speakers of other dialects. This is surprising, since students who live within 90 kilometres of Prague are not always assured of a place at Charles University halls and we might have expected the ratio of students from Moravia to be higher.[65] In terms of the aims of the study, this is a positive finding, insomuch as it is clear that students from Moravia are

---

[65] Because of the shortage of rooms, only students living outside the designated zone are guaranteed a place. Others enter a points-based system, where rooms are allocated based on academic achievement. Many students use the addresses of relatives living in other towns located further away from Prague in order to get a place, while others are forced to seek private accommodation. Since the price gap between university halls and private accommodation is very high, it is not common for Czech students to move out of the hall after their first year of study, although due to the prices for hall accommodation almost doubling in September / October 2005 the number of students seeking private accommodation has risen considerably.

unlikely to avoid contact with native speakers of CC.[66] There are, however, problems concerning the 'up-to-dateness' and the reliability of the given statistics. Working out the above distribution of students, it became apparent that names of students who had left the hall as long as two years ago had not been deleted, approximately a fifth of my informants were not on the list and international students were listed with the contact details of home students and vice-versa. The figures are further distorted by so-called 'hall-swaps': students living at different halls exchange rooms unofficially in the case that this works out to be mutually beneficial, for example, if the hall is closer to their faculty or department. In addition, some students (and non-students) were living at the hall illegally and others had moved into private accommodation and were renting out their rooms.

Kajetánka was selected as the research site for two important reasons. First, I already had a good knowledge of the local area. I had lived at Kajetánka on two occasions during previous study trips to Prague and many of my friends and acquaintances were still living at the hall at the time the fieldwork was carried out. They helped in the initial stages of recruiting informants, ensuring that an adequate number of informants were gathered in a relatively short time. Furthermore, the fact that I had already made contacts within the student community at the hall on previous visits facilitated the task of being accepted by the group of speakers whose social and linguistic behaviour I was observing. The second reason for selecting Kajetánka was that it houses predominantly students of medicine and sports science, who I considered ideal candidates for a sociolinguistically-orientated study. In planning the fieldwork, I decided that unless absolutely necessary I would avoid students of the

---

[66] According to informants' reports, the ratio of speakers of CC compared to Moravian dialect speakers at the university is even higher. In some cases, informants reported that over 80 percent of their classmates were from Bohemia.

arts and humanities, in particular those studying language or linguistics, who could have been recruited easily with the help of my host department at Charles University. Instead, I tried to recruit the informants myself, using a friend-of-a-friend approach. Impressionistically speaking, students of the arts and humanities are more aware of, or sensitive to, the complex nature of the Czech language situation and they might have deliberately attempted to monitor and / or better their speech habits during the recorded interview.[67]

5.2 Locating and selecting a representative sample of informants

In sociolinguistic studies, there has been a shift away from studies of the isolated individual to the group. Observing and recording speakers in communicative situations that they are used to provides more accurate and reliable data and the material elicited in the study of small groups can be made representative for a much larger community if certain selection criteria are met. This means that linguists can use data elicited in group studies to make generalizations about the linguistic behaviour within the larger speech community. In order to avoid group-based studies yielding similar results to studies of individuals and to generalize hypotheses beyond the communal setting of the individual study, a systematic selection process for recruiting informants needs to be employed. Because it is impractical to study the linguistic behaviour of every member of a speech community, sociolinguists must work with smaller groups taken from the larger community. The necessity to balance

---

[67] For readers unfamiliar with the Czech education system, it is important to point out that assessment of students of medicine and sports science (and several other disciplines) is based almost exclusively on oral examinations and students of medicine are not required to submit a written dissertation. Written examinations are much less common than at British universities. The greater linguistic awareness of students of the arts and humanities has been demonstrated by Bayerová-Nerlichová (2004), who in a questionnaire-based analysis asked informants to proof-read a short text containing CC forms (and other non-standard elements) and to replace the non-standard forms with their SC equivalents. Humanities students were the highest scorers, with an average score of 93 percent, while students of technical subjects scored only 79 percent.

a representative sample of informants with time restrictions in locating the required sample and subsequent data analysis can be problematic. Sociolinguistic studies generally do not require the vast numbers of informants typically used in sociological or scientific studies, and Sankoff (1980: 51-52) argues that 'samples of more than about 150 individuals tend to be redundant, bringing increasing data-handling problems with diminishing analytical returns'. Feagin (2002: 21) argues that 'a small amount of data is better than an unfinished grandiose project' and Tagliamonte also states that it is better to design a smaller and better circumscribed sample than 'to end up with lots of data but not enough funds (or energy) to use it' (2006: 33). To illustrate this point, Labov's New York study (1966) is based on data elicited from 88 informants (from a much larger random sample), Trudgill's Norwich study (1974) consisted of 60 informants, and P. Eckert in her study of Belten High (1989) had a sample of 69 informants, which she reduced from around 200 informants that had been interviewed during extensive participant observation. With regard to social network analyses (§ 5.9), Lippi-Green (1989) had a sample of 42 informants from a village in rural Austria, Edwards (1992) observed the linguistic behaviour of 66 black inner-city residents of Detroit, Michigan, Evans (2004) based her study of non-accommodation to an ongoing sound change in Michigan on a sample of 28 Appalachian migrants, and L. Milroy (1987a) based her study of language maintenance in Belfast on a sample of 46 informants.

To achieve representativeness and to make generalizations about the community as a whole, it is necessary to divide informants into stratified subgroups. Informants should not be selected at random but should be chosen in view of a number of predefined criteria. This stratified selection process is known as 'quasi-random', 'judgment' or 'quota' sampling. Informants, for example, can be categorized

according to age, sex, socioeconomic status, occupation, ethnic origin and other social parameters, depending on which factors the sociolinguist considers important to his or her study. These stratified groups of individuals are known as 'cells' and each cell contains a specific number of speakers. There are no concrete rules about how many individuals a cell should contain: Tagliamonte (2006: 31) comments that some statisticians quote three informants per cell as an adequate number, whereas others say that five is a reasonable objective. In theory, a cell that contains more than one informant is workable. Ultimately, the decision is heavily influenced by external factors such as the time a researcher has to do the study, the amount of money he or she has to spend on it and how much data needs to be elicited from each informant. L. Milroy (1987a), for example, stratified her sample in terms of sex, age and neighbourhood, having roughly an identical number of informants from Ballymacarrett, Clonard and Hammer who were further divided into cells that consisted of three or four individuals: male and female informants aged between 18 and 25 and male and female middle-aged (40-55) informants.

In this study, informants were categorized in view of four social parameters: sex, the time they had spent in Prague, their region of origin and their level of integration in the host community. My primary objective was to recruit at least ten informants from each of the three major interdialect regions in Moravia (Central Moravian, East Moravian and Silesian) and to have an even mix of male and female informants within each group. Another important factor was to have a sample of informants who had lived in the host community for varying periods of time. Obviously, it was impossible to know informants' level of network integration before the interviewing started. In view of subsequent data analysis and interpretation of the

results a maximum sample of 40 informants was considered both realistic and workable.

In gathering the informants, I employed established methods that had been successful in many other sociolinguistic studies. As an initial step, I established a contact within the student community: an undergraduate student, who in exchange for help with English assignments suggested friends and helped recruit informants to participate in the study. This is similar to the friend-of-a-friend approach, although, instead of approaching the subjects myself, the first informants were introduced to me by my helper. I also enlisted the help of my existing contacts at the hall, medical students who at the time of the field work were either in the final year of their degree programmes or already graduates either enrolled on postgraduate programmes and / or working as junior doctors to suggest their own friends. My aim was to take advantage of a network-sampling (snowball-sampling) approach, according to which the 'first-wave' informants suggest friends who are suitable candidates and who would be willing to participate in the study, who in turn then suggest their own friends, resulting in a snowball effect. Network sampling frequently results in the researcher having a surplus of informants that can be later reduced to a more manageable sample (see, for example, P. Eckert 1989). In this study, however, network sampling was unsuccessful and most first-wave informants did not suggest friends and said that they did not know anybody else from Moravia. In view of the number of students from Moravia who were living at the hall, a more plausible explanation for this lack of success can be sought in view of the loose-knit nature of the student community. I would suggest that informants did know other students from Moravia, but felt uneasy about disclosing their details to a third party, believing this to be an invasion of their privacy. Robinson (2007) encountered a similar problem in her study of semantic

change in South Yorkshire, finding network sampling to be unsuccessful among academics and white-collar workers. Thus, we could argue that network sampling is more effective in tight-knit, insular communities, where, generally speaking, there is a greater sense of group loyalty and a greater level of familiarity between members (see, for example, Bortoni-Ricardo 1985, Kerswill 1994).[68] In view of this setback, gathering a representative sample of informants took more time than anticipated and it was necessary to employ a range of different selection tactics.

A total of 39 informants (18 male / 21 female) participated in the study: 14 from eastern Moravia (7 male / 7 female), 13 from Silesia (5 male / 8 female), 10 from central Moravia (6 male / 4 female). Two informants from western Moravia (both female) were also recorded. With regard to length of residence, 11 had lived there for five years or more, 13 for between two and two-and-a-half years and 15 for up to two-and-a-half years. Three informants had lived in Prague for less than a year. Nineteen informants were students of medicine or physiotherapy and the rest were from other faculties. Most informants were on undergraduate programmes, although some postgraduates and trainee doctors also participated. Only one informant was not studying or had not studied at an institute of higher education. With regard to the methods used to gather the informants, 16 informants were recruited using the network-sampling approach described above, eight replied to posters I had placed around the hall, ten were my own or my insider's friends or acquaintances and five were recruited by chance during participant observation. The results for the two informants from western Moravia were used only to compare them to those of my insider (§ 5.6). In spite of the original plan to not interview students of the arts and humanities, Marek, a student of history from Silesia, was included in the analysis.

---

[68] In such societies, this is approach can be highly effective. Kerswill, for example, writes that he received 'enthusiastic co-operation and cordial hospitality' (1994: 75).

Therefore, despite the need to diverge from my initial plan in terms of finding a representative sample of informants, I was successful in recruiting an even mix of male and female informants, who had lived in Prague for various periods and who were from different regions of Moravia; in this respect, there was a large element of luck, since such methods of recruitment often produce an unbalanced sample where certain groups are often underrepresented. The necessity to depart from the initial plan provides a further potentially interesting subdivision: 'method of recruitment'. It is possible that differences in accommodation will be observed for informants who have been recruited in different ways. Kerswill, for example, states that a possible implication of the friend-of-a-friend approach is that it attracts an unbalanced sample of informants that are likely to be 'strongly ethnocentric' (1994: 76). He argues that his Stril informants, who were recruited using through a friend-of-a-friend approach, might have used less variants of the host variety than other more loose-knit migrants. It has also been argued that individuals who are quick to establish links with the outsider and those who are suggested or who put themselves forwards to take part in various kinds of research tend to be peripheral members of the community under study and their linguistic behaviour is, therefore, likely to differ from that of central members of the community (Saville-Troike 1997: 136).[69]

5.3 Students as informants in language-based research

Labov suggests (1972c: 113) that the college or university student provides a particularly bad source of data, since students' idiolects are likely to be an accumulation of superposed varieties distant from the vernacular of, say, working-class speakers or adolescents. We could argue, however, that Labov's comments are

---

[69] Labov (1972a: 81), for instance, asserts that in studies of non-standard language in schools researchers are often given peripheral, isolated members, so-called 'lames', who are the well-behaved members of the group, but whose linguistic behaviour often does not correspond to mainstream norms.

true only of sociolinguistic studies that are concerned with the study of language or dialect maintenance. Education is generally associated with a reduction in the use of highly-localized forms and convergence towards the standard; therefore, students may in fact be bad informants in this line of enquiry – depending, of course, on what the researcher wishes to prove. In the case of dialect accommodation, it is difficult to predict the influence of education on the linguistic behaviour of speakers moving to another speech community. In the case of English, we might assume, although we cannot be certain, that highly-educated speakers in situations of dialect contact would gravitate towards the standard more than less-educated speakers, who may retain more regional forms. The Czech situation, on the other hand, is different, since there are no social dialects and we do not expect to identify differences between students and blue-collar workers with respect to linguistic behaviour, where both are speakers of CC. This holds at least at for the distribution of CC forms, although differences are evident on a paralinguistic level. Although we cannot exclude the possibility that Moravian students living in Bohemia will adapt to the host variety differently than, say, manual labourers, we have no reliable means of predicting these differences.

Studying the linguistic behaviour of a group of working-class speakers would have been impractical. Recruiting an adequate sample of informants would have taken longer, arranging interviews would have been more complicated, and, most importantly, participant observation, an integral element to this study in supplying supplementary material to the data elicited in the recorded interviews, would have been impracticable. Bortoni-Ricardo (1985) tried a form of participation observation, whereby she made formal requests to visit informants at their homes in their recreation time. This, however, adds unwanted formality to the situation and is likely to create an artificial communicative situation. Spontaneity is key element in

participant observation, as P. Eckert (1989) demonstrates in her study at Belten High. In the present study, a similar approach was employed where, instead of arranging meetings with informants, I simply used to 'hang around' the hall, meeting individuals randomly. The underlying decision to study the speech of students was made primarily for practical reasons. As a non-Czech and a researcher from a different speech community, I was at a major disadvantage, insofar as I had few contacts within the group whose linguistic behaviour I was observing. My primary concerns were that: (1) I had to find a relatively large number of informants in a relatively short time; and (2) I needed to be accepted as soon as possible by the group under study in order to successfully and effectively carry out the fieldwork. From the point of view of ethnography, university halls of residence provide an ideal research site, since all informants are grouped together. In addition, I was roughly the same age as my informants and as students my informants and I shared several interests and activities. Thus, in theory this should have increased the probability that I would be accepted by the group whose linguistic behaviour I was observing.

A further factor that is important in terms of the specific aims of the present study is that a university hall of residence is a good locale from the perspective of dialect contact. As we described above, Moravians are outnumbered at the hall by students from Bohemia and they come into daily contact with speakers of CC. This ensured that I was analyzing the speech of individuals who are highly exposed to the host variety and not those who are relatively isolated from the larger community, which tends to be common in some migrant groups where new-comers seek out members from their native speech community. A final advantage is that students, generally speaking, have flexible schedules and have a lot of free time, during which I was able to approach them. Students are more accessible to the ethnographer than,

say, a worker with a regular nine-to-five job who is married with children and who might be less willing to give his or her free time to have a chat with an outsider. This combination of favourable factors meant that my informants were easily accessible in a range of informal situations and with respect to research-informant proximity there was a strong likelihood that I would observe the 'vernacular' (§ 5.4) and that I would obtain natural and unbiased data.

5.4. The vernacular

The aim was to elicit the vernacular (also termed in this thesis as 'actual (language) use') – the style of speech which sociolinguists strive most to observe. Like other ubiquitous terms in the specialist literature, the term 'vernacular' is variously defined. Milroy and Gordon argue that it is a 'fundamentally abstract object, rather like its counterpart, the standard language' (2003: 50), and L. Milroy (1987b: 66) believes that since the term is an abstract idealization, 'if by chance it were ever possible to locate this "vernacular" in the usage of our informants, we would have no criteria for demonstrating that we have located it'. It has been described as the variety acquired in pre-adolescent years and the style which is most regular in its structure (Labov 1972c: 112), more regular than the more formal 'superposed' styles that acquired later in life (L. Milroy 1987a: 23), and the variety that speakers use when paying the least attention to how they speak (Labov 1972c: 112). P. Eckert (2000: 17) describes the vernacular as the variety of 'locally based communities in opposition to the supralocal standard'. L. Milroy (1987a: 24) differentiates between an 'urban' vernacular and an 'individual's vernacular. The former denotes 'the kind of speech the majority of speakers of a city (usually low-status speakers) acquire in their adolescent years', she talks of the Belfast Vernacular or Black English Vernacular; the latter, the

unmonitored variety used in everyday, non-formal situations by the individual. L. Milroy (1987b: 66) uses the term 'vernacular' in a different way, 'roughly synonymous with "real language use"', and interprets it on a continuum 'of relative closeness to, or distance from, the idealized *norm* or (in some cases) the idealized *standard language*'. Tagliamonte (2006: 8) lists three reasons why the vernacular is considered 'the most systematic form of speech'. First, it is the variety speakers acquire first. Second, it has the least amount of hypercorrection and style-shifting. Third, it is 'the style from which every other style must be calibrated. For the purposes of the present study, the vernacular is to be understood as the variety that students use in informal, everyday communicative situations with their friends at the hall. This definition corresponds is closest to that of Swann et al. (2004: 327).

5.5 Interviews

The data for scoring: (1) informants' expected accommodation based on their integration within the host community; and (2) informants' quantitative score for the six linguistic variables under study were elicited in two recorded interviews. The first language or sociolinguistic interview (hereafter, I1) was in the format of a twenty-minute conversation with a native speaker of CC: Markéta, a twenty-six year old recent graduate of medicine from Ústí nad Labem in northern Bohemia. In the second interview (hereafter, I2), informants were asked questions concerning various aspects of their social life in Prague and they were requested to report on various language-related issues, including their attitudes towards the varieties of Czech and whether they thought they had started to speak differently since coming to Prague. I conducted

I2 myself[70] and the qualitative material was used to score informants' likelihood of accommodation. I2 was divided into two sections. First, it consisted of a series of 'life-style' questions which were used to calculate informants' level of integration in the host community and this is referred to as informants' score on the 'integration index'. In the second part of I2, informants answered a number of 'attitude' questions, where comparisons were drawn between Bohemia and Moravia for language, people and culture. In this section, informants were also asked questions about their own language use and to comment whether and in what manner their linguistic behaviour had changed in the time they had been living in Prague.

The interviews were recorded either at the hall or at the flat that I was renting, which was located approximately 150 metres from the hall. Before the recording began, the format of the interviews was clearly explained to the informants and they were asked to fill out a form of informed consent. Informants were told that they were participating in a sociological study of student life. The sociolinguistic content of the study, although possibly apparent in I2 which contained questions that directly addressed language use, was not disclosed to informants. Informing participants the exact aims of why they were being recorded might have influenced their linguistic behaviour and their choice of variant in the interview; therefore, they were told as little as possible before I1 was recorded. Markéta was also given only a very brief and non-technical description of the research, and although knowing the study sought to identify in what manner students from Moravia accommodated to the 'Prague dialect', she was unaware of the specific features under study.

Johnstone (2000: 48) states that this approach is ethically acceptable and widely used in sociolinguistic studies. Sociolinguists can take one of two routes in

---

[70] During I2, I spoke mainly SC, although I did use certain non-standard forms that are used both in Bohemia and in Moravia.

order to mask the sociolinguistic content of their studies. First, they can disguise the language aspect altogether. Second, as Johnstone suggests (2000: 48), sociolinguists may tell informants that the study concerns language use, but avoid going into any detail and revealing the specific features under study. The first option of inventing a pseudo-project is not necessarily the best approach and can be counter-productive. I did consider this before starting the fieldwork, but I found it difficult to openly lie about the nature of the research and I thought that informants may pick up on my hesitancy and uncertainty, and that this could have aroused unwanted suspicion. Consequently, I opted for the second approach. Informants were told as little as possible and they were given only a vague description of the aims of the research. Generally speaking, informants were indifferent to the specific aims of the research. The majority of them had a 'well-if-it-helps-you-out' attitude and rarely asked questions concerning the study. Douglas-Cowie experienced a similar response; her informants were willing to be recorded, although 'most did not ask what the nature of the experiment was' (1978: 40). Of course, this can only be achieved if informants are familiar with the researcher before the interviewing begins. Because informants had agreed to take part well in advance and because the interview was not our first meeting – in most cases we knew each other fairly well by this stage – any unwanted suspicion was eliminated.[71]

Both interviews, contrary to my original plan, were recorded on the same day. The order of the interviews was also switched: I1 was always recorded before I2. This decision was made primarily because in the final stages of planning the fieldwork new language-orientated questions were added to I2, initially intended to be recorded first. Therefore, once informants had been recorded speaking to Markéta, I was able to be

---

[71] This was one of the many advantages of carrying out participant observation. Baugh believes that trust is crucial in overcoming the observer's paradox. He argues that 'ethnographic familiarity with subjects is essential to successful fieldwork' (1993: 179).

more open in broaching the topic of language in I2, which afforded me the opportunity to elicit valuable attitudinal data on linguistic stereotypes and language use. The decision to record both interviews on the same day was considered to be mutually beneficial for Markéta, who had a tight work schedule, and for my informants, and, most importantly, it ruled out the possibility of informants withdrawing from the research between I1 and I2 or being for whatever reason unavailable for the second interview.

5.6 The sociolinguistic interview (I1)

In I1 (the sociolinguistic interview), informants were recorded during a twenty-minute chat with a native speaker of CC. Many sociolinguists in attempts to observe the vernacular give preference to the group interview with two or three informants and in the study of language variation and change more accurate representations of the vernacular have been gained recording interviews with more than one informant at a time (see, for example, Edwards 1992). In this study, however, all interviews were on a one-to-one basis. Although Labov (1972b: 116) argues that 'individual interviews give us only an approximation to the vernacular', this decision was taken for two reasons. First, speakers of Moravian dialects are in a linguistic minority in Bohemia and in the majority of communicative situations they are outnumbered by speakers of CC. Second, previous studies have shown that creating a situation where members of the host community (speakers of the host dialect) are outnumbered by members of the migrant group promotes the retention of regional forms, which leads to non-representative data (Jonášová 2001).

Besides having to decide whether to use individual and group interviews, I was also faced with the problem of who should record the interviews. Differences in

linguistic behaviour in view of the interlocutor are clearly highlighted in Douglas-Cowie's study of bidialectalism in Articlave in County Londonderry, Northern Ireland (1978). Ten adult informants were recorded speaking both to the investigator, who was born and lived in Articlave (Experiment One), and to an English outsider with an RP accent (Experiment Two). In Experiment One, informants were recorded in self-elected groups of two, while they were requested to come alone in Experiment Two. Douglas-Cowie wanted to highlight that in traditional studies of rural dialects, where informants are in the majority of cases interviewed in one-to-one interaction with a speaker from another dialect area who speaks with a different accent and where the interview is typically the first encounter between the two individuals, the speech that recorded is usually distant from the informants' vernacular. In both interviews, informants were recorded surreptitiously, although they had given their consent to be tape-recorded at an earlier date, for two hours in an identical situation. The results show that informants used more vernacular forms (for all the variables under study) when speaking to the local investigator, while they used more standardized forms when addressing the English outsider.[72]

Researchers have found factors such as age, social status and ethnicity to be a potential stumbling block in overcoming the observer's paradox. Bortoni-Ricardo (1985), for example, found that her social status caused informants to modify their speech habits. Labov's informants in his study of Puerto Rican speakers in New York (1972a) were wary of him because of his ethnicity and Bough (1993) found in his

---

[72] Although Douglas-Cowie clearly shows that informants better their speech habits in the interview with the English outsider, it would have been more objective if there were two informants present in both interviews, since it has been shown that informants use less standardized forms when interviewed in groups, regardless of whether the interlocutor is a native of their speech community or not (Edwards 1992). In addition, the informants were good friends of the investigator. It is thus difficult to conclude whether the informants used more standardized forms in Experiment Two in view of the interviewer or in view of the fact that they were recorded alone, or whether the informants would have reacted differently to an unknown speaker of their native dialect than to Douglas-Cowie.

study of African-American street speech that his own English was too standard, and this made his interlocutors increase the number of standard variants and speak differently from what they would do in everyday situations with speakers from their own community. In these studies, which all sought to capture the vernacular, an 'insider' or 'insiders' were used to conduct all or some of the recorded interviews, in order to overcome disruptive researcher-informant asymmetry. An insider was also required in the present study. In native – non-native discourse, it has been reported that native speakers often show a tendency towards standardization and simplification of various kinds, which Ellis terms 'foreigner talk' (1994: 251-256); therefore, it was considered essential that I1 be recorded by native speaker of CC.

The term 'insider' suggests a member of the group under study. In fact, insiders tend either to be from the in-group or they are in some way closer to the informants: of similar age, social status, of the same speech community, and so on. For example, Bortoni-Ricardo in her study of rural-to-urban transition in the speech of rural migrants used an insider of the informants' native region, rather than a member from the migrant community which she was investigating. Because informants in this study were taken from different groups or communities of practice, it was important to find an insider who matched several criteria. Markéta was selected to record the interviews for the following reasons:

1. She was in the same age cohort as my informants.

2. She lived for six years at the hall of residence where the research was carried out.

3. She was a medical student – the subject of most of my informants – and is a graduate of the Second Medical Faculty, where most of my informants were / are students.

4. She is a recent graduate who has maintained links with students still living at the hall, with whom she maintains regular contact; in effect, there has not been a sharp change in her social life despite the transition from student to working life.

5. She is a non-linguist.

6. She 'appeared', impressionistically speaking, to use all the CC variants under study as other speakers of CC do and her language 'seemed' to be typically 'Bohemian'.[73]

7. She is outgoing and sociable, without being overbearing.

In short, Markéta was a good choice for linguistic reasons and, more importantly because of her profession and her shared experiences with the informants. Markéta had shared experiences with the majority of my informants, either with regard to medicine or life at the hall (or both), which she used to good effect in the interviews. Common study-related topics were used to break down initial barriers and past experiences proved to be successful 'ice-breakers'. Informants who studied medicine sought her knowledge how to find work following the completion of their degrees and

---

[73] Obviously, at the time Markéta was chosen to record the interviews it was necessary to rely on intuitions in describing her linguistic behaviour.

they asked questions relating to the work she was doing, while others were interested in hall-related issues: former students, local pubs, living conditions, and so on. In addition, as a non-linguist with only a vague knowledge of the aims of the research, she was not in a position to direct the conversation in order to elicit particular variants; therefore, her performance in the interviews was in no way artificial.

Although the choice of insider has been accorded little research in sociolinguistic studies, most probably because most researchers carry out interviews themselves or record their informants speaking to each other, it would be worthwhile testing which factors are the most effective in capturing the vernacular. For instance, simply employing someone of the same age and social status as the speakers under study does not necessarily mean that the two speakers will be able to converse naturally. Topics of conversation that are unnatural for both interviewer and interviewee are likely to result in unnatural speech. On the other hand, common interests and shared experiences are evidently provocative and spur informants into positive or negative recollections of practices in which they have both engaged, though at different times and under different circumstances. Shared experiences lead to impassioned debates where the vernacular is most identifiable. This was clearly visible when informants became emotionally involved when they were talking about subjects, exams, lecturers, proposed rent increases and living conditions at the hall. Emotional involvement of any kind has been proved to be successful in allowing researchers to capture the vernacular. A second important consideration is the interviewer's position within a given network. As Labov (1972b: 115) argues, it may happen that informants cannot speak as openly or freely to a fellow member of their localized community as they would be able to speak with an outsider or stranger. Therefore, Markéta enjoyed the privilege of being close enough to the group of

speakers in order to elicit naturalistic data, but she was also distant enough not to make informants feel uncomfortable to the point that they might consciously monitor what they were saying.

The format of interview was left relatively unstructured, since the CC variants under study occur frequently enough without any special prompting. The aim was to elicit natural conversational narratives by creating what Wolfson (1997) calls a spontaneous interview: 'in which the subject is asked a few questions and then encouraged to develop any topic which seems to interest him'. Unless my insider and the informant had been discussing something before the recording equipment was switched on, the usual point of departure in the interview was life at the hall or study-related issues. Otherwise, Markéta or my informants were left to develop any topic that was of interest to them. Although sociolinguists often talk about 'good' or 'bad' interview questions (see Tagliamonte 2006: 37-45), asking Markéta to adhere to a rigid set of predetermined questions would have made her more conscious of her speech and in general would have made the interview more formal.

Predetermined questions of Labov's famous 'danger-of-death-question' type (1972a) – 'Have you ever been in a situation where you thought you were in serious danger of being killed – where you thought to yourself, "This is it?"' – which are typically used to spur informants into free-flowing, lengthy narratives, the effectiveness of which is restricted to certain communities (Trudgill 1974)[74], were not used. Instead, it was left to my insider's discretion to raise topical issues of her choice. In addition to the effectiveness of relying on shared experiences or practices to keep the conversation flowing and to avoid communicative disturbances and departures

---

[74] After Labov had used the 'danger-of-death' question with great success in his New York study, linguists tested its effectiveness in other studies. Trudgill tested it in his Norwich study (1974) study; his informants, instead of being spurred into a natural and emotionally charged narrative, struggled to come up with a response. They had evidently never been in a life-threatening situation.

from the vernacular, raising the issue of important events that are: (1) currently taking place; or (2) that have recently taken place in a given community is also an effective method of eliciting natural and prolonged narratives. Labov (1972b: 114) states that the fieldworker can make good use of local issues, humour and gossip identified during participant observation and Tagliamonte (2006: 38) says that the choice of interview question is shaped by the speaker's age and the type of community where the research is being conducted. As Baugh (1993) summarizes, unfamiliarity with the culture under analysis means that researchers are unaware of the topics that are potential sources of aggravation. Feagin (2002: 20) in a study in Alabama, for instance, asked informants to describe a snow storm that had recently happened there. As a freak occurrence in this part of the world, an exciting and unexpected phenomenon, it was big news at the time and everyone had their own recollection and version of the events. The desired effect of this type of questioning is for informants to become emotionally involved in the conversation, since in an impassioned debate speakers generally forget about delivering their utterance in a particular manner: 'they are more concerned with what they say than with how they say it' (Milroy and Gordon 2003: 65).[75]

5.7 The life-style questions

In contemporary sociolinguistic studies, researchers tend employ a 'triangulation' approach, using both quantitative and qualitative methods to elicit accurate and reliable data. Quantitative analyses are typically supported by qualitative data

---

[75] A benefit of participant observation is that it reveals interesting facts about the host community, which are otherwise not available to the researcher. During 2004 – 2005, several reality shows, a genre of television previously unknown in the Czech Republic, were launched and this was one of the most talked-about issues among the informants. Conversations involving the reality shows seemed to yield similar results to those engendered by the questions outlined above and informants become emotionally involved in their descriptions of the various contestants they either liked or disliked.

regarding the community under study, and researchers elicit various extralinguistic and attitudinal data about from their informants. In I2, I asked informants questions in Czech about various aspects of their life in Prague and I elicited material on their attitudes towards the varieties of Czech, as well as towards their own language use and whether they thought that during the time in Prague their speech habits had altered in any way. The second interview consisted of two sections. In the first part, informants answered questions concerning their life in Prague. I refer to these questions as the 'life-style' questions. In the second part, I elicited data on informants' opinions towards Bohemia and Moravia, and towards the varieties of Czech. This section is known as the 'attitudes' section. A total of 13 'life-style' and 15 'attitude' questions were used and 8 of the 15 'attitudes' questions were language-related. Since the second part of the interview was open-ended, the time it took to complete the interviews varied considerably: the longest interview lasted 49 minutes, the shortest, just eight, with an average of approximately 17 minutes.

Life-style questions

1. Do you have a roommate? Where does s/he come from?

2. Have you ever had a roommate from Bohemia/Moravia? / Have you always had a roommate from Bohemia/Moravia?

3. Have you always lived at this hall of residence?

4. How often do you go home?

5. What about your friends? Would you say that most of your friends come from Bohemia or Moravia?

6. Do you have many friends from Bohemia?

7. Are any of your relatives from Bohemia?

8. Did you ever live in Bohemia before coming to Prague?

9. What do you do in your free time?

10. Do you have a part-time job outside the university?

11. What do you do during the summer holidays? Do you stay in Prague or do you go home or go travelling?

12. Why did you decide to come to Prague?

13. What do you plan to do when you finish your studies? Have you already got an idea where will you look for work?

The life-style questions were intended to identify: (1) how often and in what capacity informants come into contact with native speakers of CC; (2) to what extent informants maintain contact with their native dialects; (3) what factors motivated informants to move to Prague and what are informants' long-term plans upon

completing their degrees. In theory, informants who are exposed the most to CC, who engage with speakers of CC in various practices, and who intend to set up home in Bohemia have a greater motivation to accommodate to the host dialect; therefore, it is predicted that these high-scorers would use the most variants. The scoring of the life-style questions is divided into sections. Informants are assigned scores along a scale from 0 (unintegrated) to 16 (highly integrated). A highly integrated individual is described as someone who lives and associates on a regular basis with members of the host community in various capacities, travels home infrequently, works and participates in some team-based activity in the host community and intends to stay in Bohemia upon completing his or her university degree. An unintegrated individual maintains close links with other individuals from his or her original speech community, leads an isolated way of life in the host community, goes home every week or every second week and intends to return to Moravia as soon as he or she has graduated. The scoring criteria are presented below.

1. Region of origin of informant's roommate(s) (questions 1, 2, 3):

0 – Informant lives (has lived) only with other Moravians; lives (and has always lived) alone.

1 – Informant lives (has lived) with both Bohemians and Moravians.

2 – Informant lives (has lived) only with Bohemians.

2. Level of contact with native community (question 4):

0 – Informant goes home every week.

1 – Informant goes home every second week.

2 – Informant goes home once a month.

3 – Informant goes home once every three months or so.

3. Immediate contacts within the host community (questions 5, 6):

0 – Informant's friends are mostly Moravian.

1 – Informant has a relatively even mix of friends from Bohemia and Moravia.

2 – Informant's friends are mostly Bohemian.

4. Exposure to CC from parents (question 7):

0 – Neither of the informant's parents is from Bohemia.

1 – One or more of the informant's parents is from Bohemia.

5. Exposure to CC on previous stays in the host community (question 8):

0 – Informant had not lived in a CC-speaking community for a prolonged period of time before coming to Prague.

1 – Informant had lived in a CC-speaking community.

6. Out-of-class activities / participation in local organizations (question 9):

0 – Informant's out-of-class activities involve spending a minimum of time with members of the host community.

1 – Informant pursues some activities that bring into contact with members of the host community.

2 – Informant participates regularly in some team-based activity in the host community.

7. Exposure to CC at the workplace (question 10):

0 – Informant does not work.

1 – Informant goes on the occasional 'brigáda' – jobs that typically last for a few days at a time.

2 – Informant has a regular part-time job in the host community.

8. Activities in summer holidays (question 11):

0 – Informant spends the summer holidays in Moravia or abroad.

1 – Informant stays in Prague and works.

9. Future plans (question 12):

0 – Informant intends to go home (or abroad) upon completing studies.

1 – Informant intends to stay in Prague.

10. Reasons for coming to the host community (question 13):

0 – Informant could not get a place at a Moravian university.

1 – Informant wanted to experience life and / or get away from home environment.

SCORE:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
N:	0	1	2	1	6	6	5	2	1	2	5	4	1	1	0	0	0

Although the above sub-variables are interlinked, it is likely that some will be more influential than others in shaping informants' linguistic behaviour. Kerswill (1994: 64) states that in his study of Stril migrants in Bergen, housemates' region of origin is probably the most influential sub-variable. In the present study, the region of origin of informants' immediate circle of friends is likely to override the influence of their roommates, since during participant observation it was evident that, besides sharing a room at the hall, roommates rarely socialize together or come into contact in any other capacity. This was confirmed by informants in I2. Similar findings were observed in the Heidelberger Forschungsprojekt "Pidgin-Deutsch" (HPD) project (1978) – a study in Germany of Spanish and Italian migrant workers' acquisition of syntactic forms – where the most important factor in acquiring the target language was contact with native speakers during leisure time. Contact with native speakers at work, on the other hand, was far less important. Informants' out-of-class activities and whether or not

they have a job in the local community are also expected to be very important factors in terms of informants' accommodation.

The criteria employed in this study are limited to those that are used to measure informants' integration in the host variety. Although in other studies, informants' life histories are taken into consideration with respect to their potential impact on accommodation, it was decided that such background information could not be objectively obtained in the time that was available. Kerswill, for instance, looks at informants' socioeconomic status, the age at which they moved to the host community, their level of education, and so forth (1994: 58-59); however, such considerations are more or less redundant here. First, all informants moved to the host community at roughly the same time and are all in the same age cohort. Second, all are at approximately the same stage in terms of their level of education. Third, since there are no social dialects as such in the Czech Republic – or at least no social dialects have yet been identified – the informants did not need to be stratified in terms of socioeconomic class. Undoubtedly, it would have been very useful to have at least some information about the informants' life histories and to find out whether they were core or peripheral members of their base community; however, because of time restrictions it would have been necessary to rely on speakers' self-reports.

5.8 Language attitudes

According to Kristiansen and Jørgensen (2005: 289), the study of language attitudes 'is commonly treated as not belonging to variationist sociolinguistics', but is 'usually grouped with the kind of sociolinguistics which has a practical interest in social issues, not with the kind of more theoretically and more linguistically oriented sociolinguistics which includes society in order to sharpen our understanding of

language'. The link between speakers' attitudes and their language use is a problematic one and data elicited in attitude studies do not 'necessarily carry a high degree of external validity' and 'cannot always be generalized to language use in "real life"' (Auer and Hinskens 1996: 21-22). Fasold (1987: 147-180) distinguishes between the 'mentalist' and 'behaviourist' views of attitudes. According to the latter view, 'attitudes are to be found simply in the responses people make to social situations'. Conversely, the mentalist view is 'an internal state of readiness' (1987: 147) and here individuals' attitudes are measured on the grounds of self-reported data that is elicited by a number of different methods. The majority of studies on language and attitudes are based on this latter approach and this inevitably causes problems regarding the validity of the data. We seem to, again, hit upon a paradox, insomuch as, according to Kerswill (1994: 67), it is probable that attitudes are linked in some way to language use, but for the time being there is no suitable methodology capable of demonstrating this fact. Thus, in sum, it is difficult to predict actual language use from attitudinal data: the relationship between individuals' attitudes towards a given variety and their actual use of that variety is complex and it is difficult to find any direct link between the two with the methods that we currently have at our disposal.

Methods in eliciting speaker attitudes can be classified as either 'direct' or 'indirect'. The direct approach, which is used in the present study, simply involves overt questioning about language varieties or forms. The indirect approach attempts to conceal the fact that informants' attitudes are the subject of investigation and it seeks to 'uncover tacit and (arguably) more deeply held beliefs and predispositions' (Coupland and Bishop 2007: 75). Obtaining unbiased attitudinal data is a hurdle that sociolinguists struggle to overcome and, since there is no universal methodology that can be used equally successfully in all studies, eliciting accurate and reliable data is

frequently problematic. Fasold (1987: 147) states, 'a great deal of effort in language-attitude research has gone into devising ingenious experiments designed to reveal attitudes without making subjects overtly conscious of the process', but no-one has developed a fool-proof methodology. Traditional methods such as the written questionnaire and multiple-choice questioning where individual answers are assigned a numerical value, although ideal from the point of view of data analysis and scoring, have several shortcomings. First, the written questionnaire adds unwanted formality. Informants avoid giving answers that in their opinion sound 'stupid' or 'out of place'. Multiple-choice questions used either in written or oral format tend to be biased and circular. Both methods generally call for categorical responses. Methods adopted from traditional dialectology and dialect geography can generally be criticized in view of the non-casual style of the interview situation, observer bias, and errors in data collection, where interviews were not recorded.

Underwood (1988), for example, in a bid to elicit informants' scale of 'Texasness', their allegiance towards the state of Texas, asked informants to whom they would give a job in a situation where the applicants were of equal ability and had an equally good education, but where only one was from Texas and the other(s) from another state. He also asked them whether in a local election they would vote for a candidate from Texas or a candidate from another state. Underwood's possible answers were: (1) person from Texas, (2) person from some other state, (3) it depends, don't know. This type of data elicitation is problematic. First, a closed-question format can potentially result in apathy as an 'easy exit', where informants predominantly use the 'don't know' or 'it depends' option and it also prevents informants from changing or expanding upon their initial answers or comments. Second, I would suggest that Underwood's questionnaire is written in such a way that

it is designed to elicit the answers that the interviewer wants to hear, insomuch as –
impressionistically speaking – I would argue that it is our natural reaction to defend
our home towns, local traditions, or fellow townsmen in the presence of a probing
outsider.

Because the relationship of language attitudes and accommodation was not a
major component of the present study, ultimately because of the inherent problems in
meaningfully describing the relationship the two, no elaborate methods were
employed. Informants were simply required to answer open-ended questions. Only
informants' attitudes towards the host variety were scored. All the questions are listed
below:

Attitude questions

1. What do you think of Prague (as a town)?

2. How do you find the people here, the 'Praguers'? Do you think there is
some kind of difference between people from Prague and other Bohemians?

3. Where do you like it the most? Here or in your home town in Moravia? If
you had to choose between the two, where would you live?

4. How would you describe Moravia or the part of Moravia where you live? In
what ways is it different from Bohemia?

5. Are the people different there? Are there some general differences between
Moravians and Bohemians?

6. In England, there is quite a big difference in the way people from the north and people from the south talk. Is there something like that here? Do Moravians speak differently than Bohemians?

7. What do you think about the language spoken in Bohemia? [Would you use it?][76]

8. What about your native dialect? How would you describe it? [Do you use it here?]

9. Whose variety is the most 'standard': Bohemians or Moravians?

10. Does the fact that you live and interact with students from all over the Czech Republic influence the way you speak? [Do you try and avoid using your native dialect? Do you try and use SC more? Or do you try to speak how the Bohemians speak?]

11. Do you feel you have changed the way you speak during the time you have lived in Prague? Can you give some examples? [Why do you think you have changed the way you speak: did it just happen? Did you make a conscious decision to do it? Did you feel under some sort of pressure to change?]

12. Do you speak differently with other Moravians than with people from Bohemia?

---

[76] Optional follow-up questions are given in square brackets.

13. Do you speak differently in Prague than when you are at home?

14. When you go home do people ever comment that you have picked up a Prague accent? Do they mean this in jest?

15. Has it ever happened that when you have been back in your home town someone has thought you are a Bohemian or from somewhere else?

The order in which the questions were given to the informants was flexible and often differed from the format given above. Not all the attitude questions were scored and some were therefore omitted. Likewise, some questions address roughly the same issue and it was occasionally appropriate to use only one of them. In all interviews, informants were asked the life-style questions first and there was a seamless transmission between the two sections. The more sensitive questions addressing differences in the behaviour and opinions of Bohemian and Moravian people, and informants' attitudes towards language use were intentionally left until the latter part of the interview, when informants felt comfortable and relaxed. Scoring the attitude questions was complicated, since open-ended questions are not easily quantifiable. Informants were graded – based my own judgements[77] – along a five-point continuum with regard to their attitudes towards CC (1 = very negative; 2 = negative; 3 = neither negative nor positive (indifferent); 4 = positive; 5 = very positive).

---

[77] I tried to be as objective as possible in assigning informants scores on the attitudinal index; however, I accept that the approach I employed is understandably open to criticism, insomuch as there is the possible effect of researcher bias. Should the data be used in further studies, in which attitudes are a principal component, the system for scoring informants will obviously be re-evaluated.

SCORE:	1	2	3	4	5
N:	6	13	14	4	0

The 'attitude' questions used in this study were developed to identify informants' attitudes towards CC, which forms informants find salient, which forms they believe they acquire the most, which forms they adopt / reject, and so on. The 'attitude' questions also highlight: (1) informants' attitudes towards their native dialects; (2) informants' views on cultural differences between Bohemia and Moravia; and (3) informants' opinions on what ways their language is influenced by living in a CC-speaking community. Only informants' attitudes towards the host variety were scored and quantified. The other data contribute to the fund of folk knowledge on Moravians' attitudes towards Bohemia and Bohemians vis-à-vis Moravia and Moravians, some of which is used in other sections of the present study (§ 7.7) and some of which will be developed further in other studies.

5.9 Social networks and scales of network integration

The present study draws heavily on the 'social network' framework. The social network was traditionally employed in social anthropology in the 1960s and 1970s and was introduced to sociolinguistics as a quantitative speaker variable by Lesley Milroy as part of her 1980 Belfast study (1987a) as an attempt to 'explain individual behaviour of various kinds which cannot be accounted for in terms of corporate group membership' (L. Milroy 1987a: 135). It constitutes a major advancement within the variationist paradigm, insofar as 'it is capable of revealing intra-community sociolinguistic patterns' that cannot be highlighted in terms of social class stratification or other social parameters (Edwards 1992: 95). That is, the social

network represents a shift away from a 'monadic view of the individual in isolation' and focuses upon 'the relationship among individuals' (Bortoni-Ricardo 1985: 70) and allows researchers to explain variation between speakers that is otherwise unaccountable by broader categories such as age, sex or socioeconomic class. The social network has been used most frequently and successfully in the study of the maintenance of non-standard forms within groups of speakers who are under pressure to adopt publicly legitimized varieties. Since its inception it has been used to good effect in the study of language or dialect maintenance in insular groups of non-mobile speakers in both urban (L. Milroy 1987a (1980), Edwards 1992) and rural (Lippi-Green 1989) communities that are exposed to currents of innovation, in bilingual communities where stable bilingualism is giving way to language shift as a result of changes in the social structure (Gumperz 1982), and in migrant communities that resist to varying degrees ongoing changes that are affecting the larger host speech community (Bortoni-Ricardo 1985, Kerswill 1994, Evans 2004).

One of the most important components of the social network framework is the 'network integration scale'. A network integration scale or index is a measure of an individual's involvement in a given community that is based upon a set of criteria believed to be important to the group (network) under study. Individuals are assigned a network score and are typically categorized as 'core' or 'peripheral' members of the network. Obviously, the criteria used to determine informants' network integration is different in individual studies. L. Milroy (1987a: 141-142) in her Belfast study, for example, ranks informants according to five criteria:

1. Membership of a high-density, territorially based cluster.

2. Having substantial ties of kinship in the neighbourhood.

3. Working at the same place as at least two others from the area.

4. The same place of work as at least two others of the same sex from the same area.

5. Voluntary association with workmates in leisure hours.

L. Milroy views the network in terms of density and plexity. Density is associated with the structures of an individual's contacts and a highly dense network would be where all members know each other. Plexity denotes the relationship between members of a group and the number of situations in which individuals in the group come into contact with each other. For instance, a uniplex tie is where network members meet in one capacity, for example, they are work colleagues. A multiplex tie means that the relationship between individuals is in more than one capacity: for instance, they may work and socialize together. The insular, non-mobile communities, that is, the type of community that has received the most scholarly attention, are typically dense and multiplex. In this type of network, most of the members know each other and contact between them takes place in a variety of situations: they work together, they go for a drink after work, they visit each other's families, they participate in the same sports team, and so forth.

Other researchers have also used network integration scales, both before and after the inception of the social network paradigm, and several elaborate models have been developed. Lippi-Green (1989: 218-219), for example, ranks informants

according to 16 differently weighted criteria in her study in Grossdorf, an isolated mountain village in the Bregenz Forest area of Austria. Here a more fine-grained approach was desirable because socioeconomic differences were irrelevant. She also looked at the familiar domains of kin and friendship, but in addition included several factors specific to Grossdorf, which allowed her to correlate informants' linguistic behaviour with their network integration as a whole and additionally with the specific criteria on the network integration scale. Gal (1978; 1979) in a study of language shift in the bilingual Hungarian-German town Oberwart in eastern Austria employed a scale of 'peasantness' – a scale of adherence to local village traditions, according to which informants were scored in terms of social indicators such as whether or not they baked bread or bought it in shops, whether they kept cattle and whether they had an inside toilet or an outside one. Cheshire (1982) in her study of the linguistic behaviour of working-class adolescents in Reading identified that informants who scored the most points on a 'vernacular culture' index that included six criteria such as 'swearing', 'fighting', and 'carrying weapons', used most forms of the localized vernacular. Possibly the most elaborate network integration indices have been employed in studies of language variation in migrant communities. Kerswill (1994), for example, includes 13 parameters and measures informants' integration and involvement in both their base speech communities and the host community, and Bortoni-Ricardo in her study of speakers' rural to urban transition in Brazil combined a network integration index with an 'urbanization index'; the latter consisted of indicators such as level of education, spatial mobility, level of exposure to the mass media, and political awareness.

A problem associated with scales of network integration is that the individual criteria are selected by the investigator on the basis of informed but ultimately

subjective judgements. Edwards (1992) shows a new application of the network integration scale by allowing informants to decide upon their own network integrations scores, believing that they are the best judges of their attitudes towards the community and posses the most in-depth knowledge with respect to their demographic characteristics, social history, and other cultural experiences. He considers indicators such as informants' 'desire to remain living in the neighbourhood, the level of their disapproval of the street culture, and their assessment of the suitability of the neighbourhood for raising children' (1992: 96). His vernacular culture index was divided into two sections that measured (1) 'the physical integration of respondents into the neighbourhood', and (2) 'their psychological integration into the neighbourhood and their racial isolation'. In his study that targeted the distribution of Standard English and African American Vernacular English (AAVE) variants, he identified striking differences between the speech of his oldest informants (over 60) and his youngest informants (18-25). The older informants had experienced a far greater level of racial segregation, which meant that ghettoized, dense and multiplex networks were formed promoting maintenance of the norms of the highly-localized vernacular. Due to societal changes with regard to racial segregation, Edwards's younger informants were far more socially and occupationally mobile; they worked and socialized in multicultural environments, and contact with speakers outside the group clearly influenced their linguistic behaviour. The younger informants showed a greater tendency towards Standard English variants. Qualitative data used to supplement the quantitative analysis showed that Edwards' oldest informants had lived and worked all their life in the same community, and their friends and family were almost all exclusively from that community, whereas the younger informants had extensive contacts outside the

community. Not only did Edwards identify that informants' scores on the vernacular culture index correlated significantly with their overall linguistic behaviour, he also yielded several other interesting sociolinguistic findings. One of Edwards' most important conclusions was that 'attitudinal characteristics are as important as objective social characteristics in influencing linguistic behaviour' (1992: 112).

The literature shows unanimously that the network acts as a norm-enforcement mechanism, which promotes the maintenance of low-status vernaculars and stigmatized localized forms in communities that are under pressure from diffusing standardized or supralocal norms. The persistence of stigmatized non-standard forms amidst strong currents of linguistic innovation can be explained by relationships within networks. As Gumperz states, long-term, interpersonal engagement of culturally and ethnically homogeneous individuals in carrying out regular tasks and the pursuit of shared goals results in the formation of specific social and linguistic behavioural routines (1982: 42) and this explains the persistence of various argots, trade languages and other vernaculars in spite of the social pressures to assimilate standardized or prestige norms. We can state that in close-knit or 'closed' networks, non-standard or minority forms stand the best chance of survival, while in loose-knit or 'open' networks culturally dominant or supralocal norms are preferred (Bortoni-Ricardo 1985: 84). It has also been observed that those individuals who are the highest scorers on network integration scales – the core or central members of a network – adhere most closely to the stigmatized localized norms, while the low scorers, the peripheral members, who have more extensive contacts outside the local community, are the leaders in linguistic changes and favour supralocal, prestige norms. Peripheral members are more innovative, since they have more contacts outside the network who are often speakers of other varieties and their linguistic

behaviour is shaped by outside influences; it becomes detached from the highly-localized vernacular and adopts features of other varieties, in particular of the supralocal norm. The social network framework has also highlighted striking differences in linguistic behaviour with regard to sex and age. For example, in L. Milroy's Belfast study, women were more geographically and socially mobile than men, who were more ghettoized – at least in the three working-class areas that were studied – and this was clear in their linguistic behaviour. Male informants were the heaviest users of the vernacular, whilst female informants showed a tendency to supralocal norms and were the leaders of language change. This gender differentiation has been consistent in other sociolinguistic studies.

The social network model has also been tested in migrant communities (Bortoni-Ricardo 1985, Kerswill 1994). Evans (2004) used a network analysis in her study of Appalachian migrants living in Ypsilanti, Michigan. The aim of the study was to identify whether the migrants were taking part in the Northern Cities Shift (NCS): an ongoing sound change in the north-eastern part of the United States. Evans concentrated on the fronting and raising of /æ/, which she describes as the 'oldest aspect of the NCS' (2004: 155). She hypothesized that speakers who associated predominantly with other members of the migrant community (high-scorers) and who had established the least ties outside the group would not participate in this sound change that was affecting the local area; she argues, 'migrant inhabitants ... whose social ties are dense and multiplex will not have accommodated or will have accommodated least to the local norm. The results show that informants with more dense and multiplex networks participate less in the sound change. In fact, only nine out of a total of 28 migrants fronted and raised /æ/.

However, the emphasis in studies that have used the social network model in migrant communities has been fixed on non-accommodation and the preservation of stigmatized norms of the base dialects in spite of the social pressures to accommodate to the host variety. Migration has often been accompanied with a sharp transition from rural to urban life and there are marked social differences between the two groups. Furthermore, in such communities in-migrants find it easy to maintain their dense and multiplex ties within a group of speakers from their native community and they can, if they so desire, participate minimally in the host community. In all the above studies, the networks are tight-knit, relatively homogeneous social clusters, and it has been argued that social network analysis is intrinsically better suited to highly insular, non-mobile communities. Chambers (2003: 76) asserts that 'the social groups under study must obviously be localized and close-knit ... they should be homogeneous with respect to social class, age, ethnicity and other independent variables'.

In the present migrant study, the situation is somewhat different for a number of reasons. First, informants are not from the same source region. Second, social differences and changes in the life-styles of the in-migrants are comparatively small in relation to the rural-to-urban transitions analyzed in Bortoni-Ricardo (1985) and Kerswill (1994). Third, it is very unlikely that the individuals in this study will maintain contact solely with members of their native community: by virtue of the situation, informants are all part of loose-knit, mobile groups and they are engaged in constant contact with speakers of the host variety, both at the university and at the hall. The emphasis here is firmly on the assimilation of a new variety as opposed to the maintenance of an old one and our aim is to find out whether network integration can be used to effectively predict speakers' accommodation to CC. That is, can the integration index, which consists of the criteria described earlier (§ 5.7), function as a

reliable predictor of innovative language use in the same way as other indices of network integration have been successful in predicting the maintenance of localized vernacular norms.

5.10 Participant observation

Participant observation, traditionally an approach used in ethnographic studies, involves spending a considerable amount of time in the community under study; it affords the researcher detailed local knowledge of the research site and of the informants 'by engagements in local affairs and / or developing personal associations with members' (Tagliamonte 2006: 20). It also affords researchers, in many cases outsiders from a different speech community, a better opportunity to understand both the linguistic and the social facets of the community and group of speakers they are observing, and this allows them to formulate an ethnographic description of the given community on the analysis of the informants' comments and behaviour and not in terms of preconceived categories (Saville-Troike 1997: 126). Unlike purely quantitative analyses, ethnographic research is based on the social dynamics of the community under study and interaction between members of that community (Auer and Hinskens 1996: 23). In fact, without ethnographically-based research it is difficult to successfully analyze informants' integration in a community or network. According to Milroy and Gordon, the two main advantages of participant observation are: '(1) the amount and quality of the data collected, and (2) the familiarity with community practices gained by the investigator' (2003: 68). In addition, researchers can break down cultural barriers between themselves and the informants, and this allows them access to more natural data.

The benefits of participant observation are explained in P. Eckert (1989), one of the most comprehensive studies to use the ethnographic approach. P. Eckert spent two years conducting fieldwork at Belten High, meeting informants outside class time, both within the grounds of the school and occasionally outside school after the school day was over. She was careful to distance herself from the formal aspects of the school, since she did not want to be seen in the same light as her informants' teachers. P. Eckert described her experience of participant observation and the goals she tried to fulfil (1989: 25-26) thus:

> ... the challenges and responsibilities of doing participant-observation in an American high school are not very different from those facing an ethnographer working in any other culture or age group. I was an outsider trying to get to know and understand a community. I needed to gain the confidence and trust of the members of the community so that they would allow me access to their activities and knowledge, and I needed to become sufficiently part of the local woodwork to be able to observe activities without producing a distraction. I needed to be sufficiently aware of my native cultural assumptions to monitor their effect on my observations and perceptions. And I need to overcome the mistrust of a subordinate community towards me as a member of a dominant power group.

At first, P. Eckert felt that the age gap between her and her informants – she was 38 at the time – could have been a problem and the fact that she in her own words was from an 'upper middle class' background and a 'Jock' meant that her own personal experiences leaned more towards the 'Jock' way of life. However, she found that the age gap did not form the expected barrier between her and her informants; instead, she concluded that 'adolescents can be extremely wary of field workers too close to their own age' (1989: 29). She also argues that 'there is no special way to deal with the interference of personal experience', and, if used carefully, personal experience can be introduced to help, rather than hinder, fieldworker observations (1989: 26). P.

Eckert's in-depth analysis of the Jocks and Burnouts[78] has been described as probably

providing 'the best evidence of social networks and social classes as independent (but

overlapping) social entities' (Chambers 2003: 78).

Participant observation for the present study was carried out at the hall and

other nearby places frequented by students from September 2004 to September 2005.

Although only the material that was elicited in I1 and I2 was analyzed, supplementary

data were gathered while observing the conversations of informants in a variety of

everyday situations. Being in the same age cohort as my informants and myself a

student, I was quickly accepted by the group under study. Joint activities with my

informants included going to football matches, watching films, listening to music,

going to the gym, going to the pub, eating out, and playing computer games. I also

occasionally helped informants with English-language assignments and translations.

Since there are no common rooms as such at the hall the main areas where participant

observation was carried out was around the refectory where students usually chat

before and after meals and in the hall, and in bars in the local area, where students

used to 'hang out'. I was in an ideal position to carry out participant observation for a

number of reasons. First, I am a student of roughly the same age as my informants and

therefore was not treated with caution or suspicion upon entering the community.

Second, I was living approximately 150 metres from the hall in a neighbouring street,

which meant that I had constant access to my research site and there was no fixed

time that I had to arrive or leave, and, using the same local amenities as the students, I

---

[78] The term 'Jock' is often used in reference to individuals, predominantly men, who are well-known for their athletic abilities and achievements, while 'Burnout' is usually synonymous with 'drug user'. In P. Eckert's work, however, the terms are used in a much broader sense. 'Jocks' are regarded as the students who accept the school and its institutions and look to continue their education at university, while the term 'Burnout' is not restricted to students who take drugs, but it encompasses all students who reject the school, are academic underachievers and who look more towards blue-collar work. Generally speaking, the distinction between these two social groups in Eckert's work can be seen as the opposition of working-class (Burnout) and middle-class (Jock) cultures.

frequently met informants in local shops and restaurants. Third, with regards to the interference of personal experience, I had only spent three months at a Czech hall of residence before fieldwork commenced; thus, I approached the research site 'from scratch', learning its social dynamics as I went along.

During the ethnographic part of the fieldwork, I had the opportunity to monitor my informants' social and linguistic behaviour and their rates of accommodation and acquisition in an array of communicative situations. Such communicative situations included conversations in which all the speakers were from Moravia, all were from Bohemia, or in which both Moravians and Bohemians were present, one-to-one conversations and group debates, conversations among students, conversations between students and locals, all female or all male discussions, and so on. This was particularly important and informative from the point of view of comparison with the recorded interviews, in which the format was always on a one-to-one basis, with one speaker from Bohemia and the other from Moravia. Participant observation allowed me to observe linguistic behaviour in different settings and in groups of different sizes from different regions conversing on a wide range of topics. I also devoted a proportion of my time to observing students' conversations without actually participating in them. I was conscious as a non-native speaker that speakers may accommodate to me in a different manner; therefore, I used to 'hang around' inside and next to the refectory area at lunch times and listen in on students' conversations. As Saville-Troike points out, observation without the participation can also be a useful form of data-collection (1997: 134).

There are problems with, or rather shortcomings of, the ethnographic study of informants and these are often overlooked. In some studies, particularly in those of adolescents and pre-adolescents, investigators must exercise caution when

establishing links among their group of informants, since establishing contacts with peripheral members of a community may hinder the researcher's chances of being accepted by other, more central, members of the community (Saville-Troike 1997: 136). A further important factor, which has not been documented in the literature, is that researchers, no matter how objective they are in attempting to distribute their time evenly among informants, can never achieve an even balance. Researchers will inevitably establish stronger links with some informants than with others, though not necessarily with those whose language is the most interesting to them, and some informants will be more or less able and / or willing to participate in informal chats than others. Some informants may become good friends, whilst other may remain distant acquaintances. This leads to a situation where researchers have a better knowledge of the linguistic behaviour of some informants then they do of others and the level of familiarity may influence informants' performance during recorded interviews. For example, I saw some of my informants on an almost daily basis, whereas I saw others, particular students studying for their final exams who had little free time, far less frequently. Nonetheless, I believe that participant observation was an essential component of the present study and this approach proved invaluable for several reasons in both I1 and I2.

6 *The linguistic variables*

6.1 Selecting the linguistic variables

The (socio)linguistic variable, which was introduced to sociolinguistics by William Labov in 1966 to help describe linguistic variation and its correlation with independent social variables, is a widely recognized construct in the variationist paradigm. Put very simply, a linguistic variable is a feature that consists of at least two variant forms (known as 'variants') that differ in form but not meaning (in the vast majority of cases) and ideally vary across various social groups. The type of linguistic variables that have been studied the most are phonological, since they are the easiest to elicit; however, linguistic variables have also been identified at the morphological, syntactic, lexical and suprasegmental levels. The linguistic variable brings the considerable benefit that it can be readily quantified, and quantification is a necessary tool if sociolinguists want to make valid and general claims about linguistic variation in a particular speech community. As L. Milroy and J. Milroy (1992: 1) argue, an important contribution of the Labovian quantitative paradigm is that it allows us 'to examine systematically and accountably the relationship between language variation and speaker variables such as sex, ethnicity, social network, and – most importantly perhaps – social class' and quantification allows us to 'make accurate statements about fine-grained differences between groups of speakers in a community' (J. Milroy and L. Milroy 1997: 49).

There are several procedural steps in delimiting linguistic variables, which are crucial to the aims of the research. Labov (1972a: 7-10) and Tagliamonte (2006: 82-84) list criteria that can help the researcher select 'good' linguistic variables. Besides the variable obviously having at least two variant forms, it is generally agreed that there should be adequate variation between the variants, that variants should be

distributed asymmetrically over a range of ordered strata of society, and that the variables should occur frequently without any special prompting in undirected natural conversation. With regard to well-studied varieties, sociolinguists may draw on existing material in selecting linguistic variables, whereas in societies that have received little systematic research, we must rely chiefly on our intuitions in selecting variables that we have identified as potentially interesting. Although the Czech language situation has received little systematic empirical study, and virtually no scholarly attention from a variationist perspective, the CC forms are well documented, though predominantly on the basis of linguists' intuitions or on value judgements of CC speakers (see, for example, Hronek 1972, Čermák 1987, Townsend 1990, Sgall and Hronek 1992, Sgall et al. 1992). Despite relying heavily on introspection, these works draw also on a number of earlier empirical studies, including Kravčišinová and Bednářová (1968), Brabcová (1973), Dejmek (1976; 1981; 1987), Krčmová (1981), Hammer (1985), and also on the first empirical studies of CC carried out by Vey (1946) and Kučera (1955; 1958; 1973). Thanks to the above literature, selecting the linguistic variables was a relatively straightforward task.

The dominant trend in the above works is to describe the geographical and social distribution of the CC variants and I shall use this data in order to make predictions about the extent to which the individual forms will be acquired. Sgall and Hronek categorize CC forms into hierarchies based on their areal distribution and social acceptance, while Čermák categorizes them according to their social acceptance and how often the CC forms are used in relation to their SC equivalents. Hronek, Sgall and Hronek and Sgall et al. classify CC features into the following geographical and functional categories (adapted from Hronek 1972: 19-22, Sgall and Hronek 1992: 28-29, Sgall et al. 1992: 77-78):

Geographical scale

1. Forms that are used throughout all or most of the Czech Republic.

2. Forms that are used both in Bohemia and in western Moravia, and occasionally in other parts of Moravia.

3. Forms that are by and large used only in Bohemia and are also encountered in the westernmost parts of Moravia.

4. Forms with a more restricted areal distribution that are used only in parts of CC-speaking territories.

Functional scale

A – Forms of a basically standard character, used in informal communicative situations without being considered (by most speakers) as non-standard.

B – Forms that are common in informal communicative situations in CC-speaking territories (Bohemia and many parts of western Moravia), when speakers do not attempt to use the standard.

C – Stylistically marked forms that are used in discourses or utterances of a specific nature.

Čermák (1987: 142-148), who draws heavily on Hronek's *Obecná čeština* (1972) and his own intuitions, categorizes CC forms into scales of frequency and acceptability (adapted from Čermák 1987: 142):

Frequency

1. Occurring always or in most instances.

2. Occurring often.

3. Occurring less often.

Acceptability

A – Accepted and used currently as normal.

B – Accepted only sometimes and in certain situations.

C – Accepted seldom or never.

Obviously, the criteria used for the selection of linguistic variables may differ depending on the aims and the nature of the individual study. The linguistic variables analyzed in this study were chosen in view of their frequency, social acceptance in informal conversation, areal distribution and categoricity. It was essential that all the selected forms have a CC variant that is restricted to CC-speaking territories, that the forms are stylistically neutral within CC, that is, they are perceived by native speakers as unmarked in all informal communicative situations, and that the SC to CC shift is

categorical or occurs in the majority of situations. It was also considered important that the selected forms should not be difficult to acquire (phonologically too complex) and that a twenty-minute interview should provide a representative amount of tokens without any special prompting on the part of the interviewer. Six linguistic variables were selected: three phonological and three grammatical. The phonological variables are word-initial (o); (é); and (ý); the processes that are connected with these variables are 'v-insertion', 'é-raising' and 'ý-diphthongization'.[79] The grammatical variables are third-person plural verbal inflection in third-, fourth- and fifth-conjugation verbs in the present tense ('paradigm unification'); verbal inflection in the masculine past tense of first- and second-conjugation verbs ('l-truncation'); and adjectival (and pronominal) inflection in the nominative plural ('gender neutralization').

6.2 The phonological variables

6.2.1 V-insertion

V-insertion is a process by which prothetic [v] is realized before word-initial /o/ and in some instances before word-internal /o/ that occurs after morpheme boundaries, particularly with non-vocalic or monosyllabic preposition-based prefixes:

1. Personal pronouns: SC *on* 'he' >< CC *von*; SC *ona* 'she' >< CC *vona*; SC *oni* 'they' >< CC *voni*; SC *ono* 'it' >< CC *vono*

2. Prepositions SC *o* 'about' >< CC *vo*; SC *od* 'from' >< CC *vod*

---

[79] In some cases, it is more appropriate to talk of the outcome of these processes: '(prothetic) /v/', 'raised /iː/' and 'diphthongized /ɛj/'.

3. Prefixed lexical words: SC *odejít* 'to leave' >< CC *vodejít*; SC *odkud* 'from where' >< CC *vodkad'* (*vodkud, vodkud'*)

4. Non-prefixed lexical words: SC *okno* 'window' >< CC *vokno*; SC *ocet* 'vinegar' >< CC *vocet*

5. Word-internally after a morpheme boundary: SC *naopak* 'on the contrary' >< CC *navopak*; SC *neočekávám* 'I don't expect' >< *nevočekávám*

Unlike *ý*-diphthongization and *é*-raising, for which the SC >< CC switch is, with very minor exceptions, categorical and constitutes a straightforward standard – non-standard opposition, *v*-insertion is governed by a number of external factors. Existing studies have highlighted a difference in the distribution of prothetic /v/ before grammatical words and before lexical words, and it seems that there are several words in which *v*-insertion is unlikely to take place. In grammatical words, *v*-insertion functions like *ý*-diphthongization and *é*-raising in adjectival desinences and some studies have identified an almost categorical use of /v/ in this position (Jančák 1974; 1978, Jančáková 1974). Before lexical words, however, *v*-insertion is more complicated and linguists believe that the insertion or non-insertion of prothetic /v/ should be treated on a word-by-word basis (Sgall and Hronek 1992: 33). Possible constraints include whether the word is a foreign borrowing, a borrowing from SC, an abstract term, or a technical term. Empirical research also shows that /v/ is more likely in prefixed lexical words than it is in non-prefixed lexical words (see, for example, Jančák 1974; 1978, Jančáková 1974, Dejmek 1986). Again, unlike /ɛj/ or

/iː/, which appear to be gaining currency in parts of Moravia (in certain positions) and have acquired a greater deal of social acceptance in non-formal communication and in various types of non-formal writing, /v/ seems to be going in the opposite direction, becoming more stylistically marked. A good example of the changing status of the phonological variables can be taken from the language of downmarket magazines: while transcriptions of interviews in popular magazines usually retain examples of /iː/ and /ɛj/, where these forms are uttered, editors generally correct /v/ and write, say, *otázka* although *votázka* was the form used in the interview (Hoffmannová 2004, Müllerová 2004). Likewise, many authors who write works or parts of works in CC tend to use prothetic /v/ considerably less than other CC forms or not at all.[80]

Townsend (1990: 36-40) gives an overview of prothetic /v/, which he considers 'by far the most important CC[81] phonetic feature and certainly the most difficult to analyze'. He argues that, unlike other phonological features, usage of prothetic /v/ is conditioned by extralinguistic factors such as speakers' social background and level of education, the nature and topic of the conversation, and social psychological factors such as the speakers' mood and their interlocutors at the time a conversation takes place, and that is generally more likely in emotive contexts or in expressive and emotionally coloured words. Townsend tries to categorize lexical items that are likely to take /v/ — he suggests among other things plants and animals, domestic things, common verbs and adjectives (37-38) — and he argues that abstract,

---

[80] Petra Hůlová writes her book *Paměť mojí babičce*, a recollection of her time spent in Mongolia, almost exclusively in CC, using all the other CC forms that are studied in the present study (and others); however, she does not use prothetic /v/, not even in pronouns forms such as *on, ona, od*, and so on.

[81] Although the subject of Townsend's study is 'Spoken Prague Czech' (SPC), his work can still be regarded as an overview of CC as a whole, since the two varieties are with very minor exceptions identical (Jančák: 1997: 200).

polite, bookish words and foreign borrowings are unlikely candidates to take /v/; Krčmová (1974) reveals similar patterns in the Brno dialect. Townsend also lists *otec* 'father', *ovoce* 'fruit' and *ovšem* 'of course' as lexical items in which the shift 'never' takes place. One of Townsend's more dubious claims is his assertion that *v*-insertion is more common in official or scientific words than *ý*-diphthongization or *é*-raising.

Townsend's method of classifying or categorizing words into those that are likely to take prothetic /v/ and those that are not has several shortcomings and his generalizations based on the judgements of 'non-ideal' native speakers[82] of CC cannot be taken as representative of how native speakers of CC actually use this feature. Sociolinguistic research has highlighted that there is often a considerable difference in the forms speakers think they use and the forms they actually do use. Speakers often over-report their use of standardized forms (or occasionally they over-report their use of non-standard regional forms). Only on the basis of analysis of thousands of tokens taken from naturally occurring speech is it possible to make reliable judgements about the use of a given feature. Townsend, for example, lists CC *votázka* (SC *otázka* 'question') as 'emotionally coloured' or 'unusual', and therefore unlikely; the form *votázka* however was recorded several times in my data in non-emotive contexts and it was often used in the same utterance as SC *otázka*, without any apparent change in emotiveness of the conversation (§ 7.2).

Other linguists have made similar attempts to classify words that do or do not take /v/ based on empirical material. Using data elicited in Hradec Králové, Dejmek also categorizes words into those that 'always' take /v/ (*votep* (SC *otep* 'bundle (of straw))', *votrok* (SC *otrok* 'slave'), those that 'never' take /v/ (*osoba* 'person', *okres*

---

[82] Most of Townsend's native speakers had lived in America for a long time (over ten years); therefore, it is conceivable that their value judgements would have been different from those of speakers living in the Czech Republic.

'district', *ovečka* 'sheep'), and those words in which the use of /v/ is variable ((*v*)*oběd*

'lunch', (*v*)*okno* 'window').

Some linguists have proposed that *v*-insertion is undergoing the same fate as

ú-diphthongization (/u:/ → /ou/) in word-initial position and *ý*-diphthongization in the

prefix *vý*- and that it is on the decline within CC (Müllerová 2004).[83] As early as the

1960s, Skalička asserted that prothetic /v/ is used only in lexicalized examples such as

grammatical words and sporadic lexical words like *vokno* and *vocet*; he specifically

lists *otrok* as a word which does not take /v/, although Dejmek later listed *votrok* as

categorical based on an empirical study. This clearly demonstrates that there is much

inconsistency and misunderstanding about the distribution of /v/ and claims to the

effect that it is on the decline are ill-founded. Townsend's assertion that *v*-insertion

may be constrained by sociolinguistic and sociopsychological factors is a valid point

and a possible consequence of emergence of the broader class stratification in the

Czech Republic is the emergence of social variation or social dialects. Sociolinguistic

studies carried out in many Western societies show significant class differentiation

with regard to the use of linguistic variables. Take, for example, the oft-studied

linguistic variables (h) or (ing) in varieties of English. Working-class speakers

generally have a higher percentage of zero realization for *h*-dropping than middle-

class speakers do (Milroy and Gordon 2003); and the same can be said for the

alternation of /n/ and /ŋ/ in unstressed /ɪŋ/ syllables: working-class speakers have a

---

[83] Müllerová looks at the transcripts of interviews published in magazines and she states that while raised /i:/ and diphthongized /ɛj/, together with some other non-standard forms, were 'let through' by the editing team, occurrences of prothetic /v/ were always corrected. Her conclusion is that prothetic /v/ is on the decline. This conclusion is, however, unfounded: the data show only that speakers' attitudes towards /v/ are less tolerant than they are towards some of the other CC forms, but there is no proof that speakers' use of prothetic /v/ is receding. Obviously, its growing markedness in terms of speaker perceptions may eventually trigger its decline in actual language use. Its use in public discourse does seem to be receding

higher percentage of /n/. It is conceivable that this could occur in Czech and if it were to do so, then, v-insertion would be a primary candidate, given that it is perceived more negatively (by the non-native speakers of CC in this study) than most other CC forms.

A further consideration is that v-insertion differs regionally (Krčmová 1974: 244). Utěšený reports that /v/ is gradually on the decline in central Moravia and is receding westwards (1962: 582); it is said to be on the decline in Brno under the influence of SC and other Moravian dialects (Sgall and Hronek 1992: 33). The literature shows that speakers in peripheral regions tend to have a lower distribution of /v/ than those from the core of Bohemia. In the repopulated Sudeten lands of northern Bohemia (Jančák 1997: 243) and in Brno (Krčmová 1974: 246), speakers have an almost even distribution of forms with and without prothesis. Regional differences may be reflected not only in the relative frequency of /v/, but also with respect to which words commonly take /v/. Material gathered for the *Český jazykový atlas* [The Czech Linguistic Atlas (hereafter, CLA)], for instance, highlights that the form *vovoce* (SC *ovoce* 'fruit'), considered impermissible by many linguists,[84] does in fact occur in parts of central Moravia, occasionally realized as ['voboˆtsɛ] with a bilabial plosive /b/ in place of fricative labiodental /v/, possibly to avoid cacophony. Interestingly, Šrámek (1997: 236) also identified the anomalous use of prothetic /v/ in the speech of certain individuals from Silesia: he recorded the forms *voltař* (SC *oltář* 'altar') and *voxid* (SC *oxid* 'oxyde') and he 'claims' that these forms would never be uttered by a native speaker of CC. In terms of its areal distribution, /v/ is observed

---

[84] Čermák (1987: 142), for example, argues that cacophonic *vovoce* (SC *ovoce* 'fruit') may sound 'funny' to native speakers of CC.

over a smaller area than the other CC phonological variants (see CLA Vol. 5, 2006: 370-371). In addition to being used throughout CC-speaking territories, it is also a (receding) feature of the local town dialect in Brno (Cummins 1993); its distribution in other parts of central Moravia is not well documented, though it is not thought to be part of the Olomouc dialect (Sgall and Hronek 1992: 33) and it is generally not encountered in East Moravian and Silesian dialects.

In summary, it appears that word-initial /o/ is currently undergoing change. Jančák (1974: 197) talks of two opposing trends in the development of *v*-insertion. On the one hand, SC forms without /v/, in particular technical and specialist terms like *orgán* 'organ' or *onkologie* 'oncology', have gained currency in CC and are causing a loss of prothesis in other words such as (*v*)*odešel* 'he left'. Jančáková believes that prothesis is on the decline under the influence of SC, especially in lexical words, while /v/ in grammatical words has proved more resistant. That said, Krčmová's study ranging across three generations did show frequent oscillation between grammatical forms with and without /v/ within the same utterance for middle-aged and younger speakers, though to a lesser extent in the speech of her older informants. Within lexical words important divisions are visible between prefixed and non-prefixed words, /v/ being considerably more common in prefixed words. The converse is also observed: technical additions to CC are undergoing prothesis, possibly under the influence of existing forms in which /v/ is especially resistant. Thus, we often encounter forms such as *vobdélník* (SC *obdélník* 'rectangle') and *voperace* (SC *operace* 'operation'). This trend also affects words that had previously not undergone prothesis like *vobrovskej* (SC *obrovský* 'huge'). In terms of its position on Sgall and Hronek's hierarchies of areal distribution and social acceptance, prothetic /v/ is listed

as **2/B** in grammatical words and some lexical words and **2/C** in other lexical words –
an explanation of what criteria are used to distinguish between the **2/B** and **2/C**
categories is not given. Čermák's interpretation is somewhat dubious: he evaluates
prothetic /v/ as **1/A-B**, that is, as occurring in all or most instances.

6.2.2 *É*-raising

É-raising denotes the raising of SC /ɛː/ to [iː] in adjectival desinences and word-
internally. It is observed in the following positions:

1. Nominative / accusative of neuter singular adjectives: SC *to je dobré*
'that's good' >< CC *to je dobrý*; SC *dobré pivo* 'good beer' >< CC *dobrý pivo*

2. Oblique cases of hard masculine / neuter adjectives and pronouns: SC
velkého 'big' (Genitive / Animate Accusative Singular) CC >< *velkýho*; SC
dobrému (Dative Singular) >< CC *dobrýmu*; SC *dobrém* (Locative Singular)
>< CC *dobrým*

3. Oblique cases of hard feminine adjectives and pronouns: SC *dobré*
(Genitive / Dative / Locative singular) >< CC *dobrý*

4. Inanimate plurals: SC *velká* (Neuter), SC *velké* (Feminine / Masculine
Inanimate) >< CC *velký*

5. In word roots: SC *mléko* 'milk' >< CC *mlíko*

6. In adjectival nouns: SC *na dovolené* 'on holiday' >< CC *na dovolený*; SC *z Karviné* 'from Karviná (a town in northeast Silesia)' >< CC *z Karviný*

7. In adverbial forms of the type: SC *poprvé* 'for the first time' >< CC *poprvý*; SC *za druhé* 'second(ly)' >< CC *za druhý*

The shift /ɛ:/ >< /i:/ occurs in almost all positions, with the exception of foreign borrowings like *šéf* 'boss', *šofér* 'chauffer', *fér* 'fair', *amatér* 'amateur', *foxteriér* 'fox terrier', *reportér* 'reporter', *aféra* 'affair'; when /ɛ:/ is shortened to [ɛ] in CC, as in *léto* 'summer' (CC *leto*) and *jméno* 'name' (CC *meno*); in the nominate masculine animate noun ending *-ové* (where /ɛ:/ is also frequently shortened to [ɛ]): *Rusové* 'Russians', *mnichové* 'monks' (also *mniši*); and in one or two isolated examples that often have an alternative form in CC: *lékařství* 'medicine' (CC *medicína*), *lék(y)* 'medicine(s), medication' (CC *prášek / prášky*), *lékař* 'doctor' (CC *doktor*) (see Townsend 1990: 32 for a more comprehensive list). Obviously, the phonetic environments that preclude variability are not included in the analysis. In all positions where *é*-raising takes place the shift is categorical and is not impeded by language-internal or extralinguistic constraints. That said, it has been suggested that /ɛ:/ is more resilient in adverbial forms of the type *poprvé* 'for the first time', *za druhé* 'second(ly)' and *pokaždé* 'every time', which Jančáková (1974: 182) calls 'borrowings from the standard language' and Jančák (1974: 194; 1978: 200) labels 'lexicalized adverbial constructions'; though, from a personal observation, in informal and undirected speech *é*-raising differs little if at all in the above forms from other positions. Dejmek (1986: 133) reports variable usage in proper names, stating that

speakers tend to realize place names of the type *u Železného Brodu* and

v Holakovickém lese more frequently with SC [ɛː].[85]

It is a fair assumption that /iː/ in all positions but the oblique cases of hard

feminine adjectives and pronouns is the most expansive CC form and thus is a

primary candidate to be acquired. It is used over the widest geographical area,

encompassing the whole of Bohemia, western Moravia and central Moravia, though

according to Hronek (1972), it has yet to gain currency in eastern Moravia or Silesia –

and this is supported by material presented in CLA (Vol. 4, 2002: 296-301; Vol. 5,

2006: 112-113). Non-standard /iː/ is also socially expansive: speakers use it in various

semi-formal linguistic situations and it has made inroads into various types of popular

literature and into popular television and radio programmes. Kraus et al. argue that

while other phonological forms such as /v/ and /ɛj/ are still clearly perceived as non-

standard, /iː/ is beginning to lose its markedness not only word-internally but also in

adjectival desinences (1981: 232). Hedin's data (2005: 83) also show that /iː/ is the

most socially acceptable of the CC phonological forms.[86] CC /iː/ is evaluated by Sgall

---

[85] CC /iː/ is, however, commonly observed in many place names, as in the recorded examples *do Mladý Boleslavi* (SC *do Mladé Boleslavi* 'to Mladá Boleslav'), *ze Starý Boleslavi* (SC *ze Staré Boleslavi* 'from Stará Boleslav'), *v Český Třebový* (SC *v České Třebové* 'in Česká Třebová').

[86] Hedin (2005) reported that /ɑ/ is used more frequently than /_ɔ/ and that the former appears to be less marked than the latter in formal discourse. Hedin studied *é*-raising in three positions: 'occurrence in endings', 'occurrence in endings before a consonant' and 'in word stems'. CC forms were observed the most in endings (30%) and only slightly less in endings before a consonant (28%) and in word stems (25%). No distinction is drawn between desinence-final position in attributive and predicative positions and presumably – although this is not clear – the 'occurrence in endings' category includes *é*-raising in the oblique cases of feminine adjectives (and pronouns). This is unfortunate, since, as we know, /ɑ/ in this position is regionally bound and socially more marked than in other positions; consequently, we would expect to observe it less outside informal communication. Similarly, as will be demonstrated in the present study, there is also likely to be variation in desinence-final position depending on whether adjectives are predicative (*to je dobré / dobrý*) or attributive (*velké / velký pivo*). I would suggest that Hedin would have observed an even higher incidence of /ɑ/ in desinence-final position, had she analyzed the oblique cases of feminine adjectives and pronouns separately.

and Hronek as 1/A on their geographical and functional scales in all positions except the oblique cases of hard feminine adjectives and pronouns, where it is restricted to traditional CC-speaking areas (3/A). It is listed by Čermák as a feature that occurs in all or most positions and is accepted as normal in some situations (1/A-B).

6.2.3 *Ý*-diphthongization

Ý-diphthongization is a process whereby the SC long vowel /iː/ is diphthongized to [ɛj]. The phonological shift takes place in the following positions:

1. Desinence-final position of hard adjectives: SC *velký* 'big' >< CC *velkej* (Nominative / Inanimate Accusative Singular)

2. Desinence-initial position of hard adjectives: SC *velkých* (Genitive / Locative Plural) >< CC *velkejch*; SC *velkým* >< CC *velkejm* (Dative Plural); SC *dobrými* >< CC *dobrejma* (Instrumental Plural)

3. In word roots: SC *týden* 'week' >< CC *tejden*; SC *být* 'to be' >< CC *bejt*

4. In the prefix *vý-*: SC *výlet* 'trip, outing' >< CC *vejlet*; SC *výplata* 'wage' >< CC *vejplata*

5. Occasionally in place of *í*:[87] SC *cítit se* 'to feel' >< CC *cejtit se*; SC *cizí* 'foreign' >< CC *cizej*

---

[87] As a rule, only *ý* (the reflex of Common Slavonic *y*) undergoes *ý*-diphthongization. There are, however, some 'new' cases where we find /ɛj/ in place of *í* from Common Slavonic *i* (Townsend 1990: 30). In contemporary Czech, both *ý* and *í* are pronounced identically as [iː].

The shift also takes places in SC *zítra* 'tomorrow' and *prý* 'they say; allegedly', which derive from the older forms *zajitra / za jitra* and *prají / praví* (Sgall and Hronek 1992: 33), yielding the forms *zejtra* and *prej*. The phonological shift /iː/ → /ɛj/ does not take place in foreign borrowings such as *tým* 'team' (**tejm*) and in some bookish words that sometimes (but not always) have a CC equivalent: *dým* 'smoke' (CC *kouř*), *nýbrž* 'but' (CC *ale*), *týž* 'the same' (CC *stejnej*) – for a more comprehensive list see Townsend (1990: 32). Interestingly, despite occurring in all other cases of adjectival morphology, *ý*-diphthongization does not take place in the instrumental singular of hard adjectives: *dobrým* 'good' (**dobrejm*). It is also not expected in words where /iː/ is frequently shortened to /ɪ/, as in *umývadlo* 'washbasin' (CC (and SC) *umyvadlo*), and in set phrases like *dobrý den*[88] 'hello' and *vážený pane* 'dear sir'.

Diphthongized /ɛj/ in desinence-final position of hard adjectives (but not in other positions) is encountered in central Moravian dialects and its areal distribution is roughly complementary to raised /iː/ in all but the oblique cases of feminine adjectives and pronouns (CLA Vol. 5, 2006: 80-81). However, although /ɛj/ in desinence-final position is part of the Brno dialect (and other Central Moravian dialects), it is not as categorical as it is in CC-speaking territories and is stylistically marked in non-informal communication (Krčmová 1979: 71). Parallels can also be drawn between desinence-final /ɛj/ and raised /iː/ in terms of social acceptance, since desinence-final /ɛj/ is also ubiquitous in various types of literature, on television and radio shows and it also has a high level of tolerance, although possibly not to the same

---

[88] Although *dobrý den* is generally preferred to *dobrej den*, speakers are much more likely to say *dobrej*, if the word *den* is omitted. The latter CC form is used either in response to *dobrý den* or as an initial greeting in place of it.

degree as raised /iː/. Although CC /ɛj/ is categorical in desinence-final position, it has been argued that speakers of CC do not consider forms with SC desinence-final /iː/ (*dobrý*) as unnatural as, say, forms with SC /ɛː/ in neuter singular adjectives or in the oblique cases of masculine or neuter adjectives and pronouns (*to je dobré*; *dobrého*). Some of Townsend's informants judged the SC form *starý* 'old (Nominative Singular)' as more appropriate in informal communication than *starého* (Genitive Singular); they said they would 'never' say *starého filmu* 'old film' but thought that they might use *starý film* (1990: 56). Nonetheless, Sgall and Hronek evaluate desinence-final /ɛj/ on a level footing as raised /iː/ in all but the oblique cases of hard feminine adjectives and pronouns as 1/A; surprisingly, Čermák considers /ɛj/ more socially acceptable than /iː/ (**1/A**).

Conversely, in desinence-initial position CC /ɛj/ is restricted both regionally and socially; it is used in CC-speaking territories, but not in central Moravia where it is perceived as 'unusual' (Krčmová 1979: 71). The areal distribution of desinence-initial /ɛj/ is therefore almost identical to /iː/ in the oblique cases of feminine adjectives and pronouns (CLA Vol. 4: 306-307; 312-313). Krčmová argues that even in Bohemia, /ɛj/ is not used categorically in desinence-initial position or root internally and that it is socially restricted (1979: 71), although empirical data suggest that in spontaneous and informal speech differences are minimal, at least between the use of /ɛj/ in desinence-initial and desinence-final position. In word roots, the situation is more complex and whether or not speakers diphthongize /iː/ to [ɛj] depends largely on the individual word and on the type of communicative situation (Sgall and Hronek 1992: 32). Thus, while words such as *tejden* (SC *týden* 'week') and

zejtra (SC *zítra* 'tomorrow') are almost categorical for native speakers of CC in all informal communication, *sejr(a)* (SC *sýr* 'cheese') and *rejže* (SC *rýže* 'rice') may be more variable, and we would be surprised if speakers realized, say, *nazývat* 'to call' with CC [εj]. The use of desinence-initial and word-internal /εj/ is sometimes considered inappropriate outside informal communication and the use of SC /iː/ therefore generally increases as the communicative situation becomes more formal. This is confirmed in studies that look at non-informal (spontaneous) spoken discourse (Kravčišinová and Bednářová 1968, Maglione 2003, Hedin 2005).[89] In both the above positions, /εj/ is regarded by Sgall and Hronek as **3/B**.

Additionally, *ý*-diphthongization is particularly problematic in the prefix *vý-*. Based on the value judgements of a group of native speakers of CC, Townsend (1990: 31-32) uses the traditional method of categorizing words prefixed with *vý-* into those that may undergo *ý*-diphthongization and those that probably do not. He goes on to assert that /εj/ in the prefix *vý-* is 'mostly emotive', but he lists the forms *vejlet* (SC *výlet* 'trip') and *vejmluva* (SC *výmluva* 'excuse') as 'so common that they approach neutral status' (1990: 31). However, the 'some-words-might-some-might-not' methodology typically employed in studies of the distribution of certain CC variants is not without its problems. Bermel (2000: 40-41) points out that *ý*-diphthongization in the prefix *vý-* does not constitute a straightforward choice between CC or SC variants, as it does in adjectival desinences. Bermel's argument is that while SC *výlet* and CC *vejlet* may be viewed as the opposition standard versus non-standard, SC *výbor* 'committee' and CC *vejbor* should be seen as neutral versus emotive; he also

---

[89] In desinence-final position, the CC forms were observed in 36 percent, 28 percent and 12 percent of cases in Kravčišinová and Bednářová (1968), Maglione (2003) and Hedin (2005), respectively; in the same studies, the CC forms were recorded in 45 percent, 28 percent and 23 percent of instances in word roots. It is interesting that in Kravčišinová and Bednářová (1968) and Hedin (2005) the CC forms were found more frequently in word roots than in desinence-final position – the converse would seem more likely from the introspective data presented in Sgall and Hronek (1992) and Sgall et al. (1992).

proposes that some words are restricted to SC domains and therefore do not have a CC equivalent: *výčitka* 'reproach' (**vejčitka*) is cited as an example. However, because of the striking lack of research into the vernacular it is necessary to consider that for the most part we are not basing our claims of whether a particular word takes or does not take the CC form on instances of spontaneous speech. Undoubtedly, it would be useful to approach the situation from more than one angle and perhaps a more comprehensive account of the distribution of CC forms could be achieved by investigating their correlation with various independent social variables such as age, sex, social class, level of education, and communicative situation. Dejmek, for instance, reports intergenerational differences in use of /ɛj/ in the prefix *vý-*, his youngest informants using only four out of 101 tokens with CC /ɛj/, while the older speakers used considerably more. Dejmek (1986: 134) also asserts that *ý*-diphthongization in the prefix *vý-* is on the decline under the influence of borrowings from SC (cf. *výčitka*) and he believes that speakers are beginning to shorten /viː/ in some words to [vɪ] and consequently diphthongization does not occur (cf. *umývadlo* >< *umyvadlo*).

6.3 The grammatical variables

To date, empirical studies have focused almost exclusively on the phonological variables and CC grammatical forms have not been researched in equal measure.[90]

---

[90] One of the exceptionally few works to investigate grammatical forms is Kravčišinová and Bednářová's study of Spoken Czech (1968). A total of 15 grammatical forms were studied, including the three CC variants analyzed in this study. The data are based on recordings of 79 informants of different ages, professions and levels of education, from different parts of Bohemia. The data are divided into three categories: (1) recordings of informal conversations on children's radio programmes in 1962; (2) similar recordings from 1963; and (3) 'recordings of everyday conversations'. The material is, however, somewhat vague. The authors do not mention how many informants are in each category, how the informants are stratified according to age, sex, education and profession, and it is unclear how the category 'recordings of everyday conversations' should be interpreted, since no further information

There is an important difference between the phonological and grammatical variants under study in terms of the area over which they are used. Unlike the phonological variants, which are to varying degrees encountered in parts of Moravia, either as a result of diffusion or independent development, none of the grammatical forms analyzed in this study have made inroads into Moravia, despite being socially unmarked in CC-speaking territories. Krčmová (1979) found no traces of the three grammatical features under study in her analysis of Brno city speech. This dialect is often said to be heavily influenced by CC; supposedly it is the Moravian dialect into which forms from CC are first diffused, even before those situated closer to CC-speaking territories.

6.3.1 Paradigm unification

The term 'paradigm unification' is used here to describe a process whereby in the present-tense third-person plural fourth-conjugation verbs (*prosit* 'to ask') and third-conjugation verbs of the *trpět* 'to suffer' type, which both have the verbal suffix -*í*, have merged with third-conjugation verbs of the *sázet* 'to bet' type,[91] which end in -*ejí*, yielding the CC form -*ej* where word-final /iː/ is truncated (*prosej, trpěj, sázej,*). Paradigm unification is also used here to denote the loss of word-final /iː/ in fifth-conjugation verbs (*dělat* 'to do') in the present-tense third-person plural.[92] This last

---

is given. Hedin (2005) also looks at morphological features in her study of contemporary Czech television discourse. She looks at *l*-truncation and gender neutralization, but not at paradigm unification.

[91] The classification of verbs into conjugation paradigms follows the guidelines in *Příruční mluvnice češtiny* [Handbook of Czech Grammar] (1995). Verbs of the third conjugation (which usually end in the suffix -*et* in the infinitive) are divided into two sub-categories according to their conjugation in the present-tense third-person plural. Verbs of the *sázet* type take the suffix -*ejí*, while those of the *trpět* type end in -*í*. The suffix can be deduced from the imperative form of the verb: *sázej* >< *trp*.

[92] Although the truncation of word-final /iː/ in the present-tense third-person plural verbal suffixes -*ejí* (-*ějí*) and -*ají* is a phonological process, it occurs only in the above morphological environment and is

case is not strictly the same phenomenon as for third- and fourth-conjugation verbs and I analyze it under paradigm unification more out of convenience; consequently, it is treated separately in terms of data analysis.

1. Third-conjugation: *trpí* 'they suffer' → *trpěj*; *sázejí* 'they bet' → *sázej*

2. Fourth-conjugation: *prosí* 'they request' → *prosej*

3. Fifth-conjugation: *dávají* 'they give' → *dávaj*

In SC, a distinction between verbs of the third conjugation, which Townsend (1990: 76) categorizes into 'ě'-verbs (*trpět* 'to suffer') and 'ěj'-verbs (*umět* 'to know how, be able'), is maintained in the third person plural present tense: 'ě'-verbs ('e'-verbs) take the suffix *-í* (*trpí*), whereas 'ěj'-verbs ('ej'-verbs) have the suffix *-ějí* (*umějí*). In spoken varieties, however, this distinction is lost. In Moravian dialects, alongside regional forms in *-ijou* (*-ijó*, *-ijú*, *-iju*), a supralocal trend has emerged whereby all third-conjugation verbs, both ě-verbs and 'ěj'-verbs, are levelled and take the suffix *-í* in the third-person plural by analogy to verbs of the fourth conjugation (*prosí* 'they request'). Therefore, the verbs *umět* and *sázet* have the supralocal forms *umí* and *sází* that are commonly encountered in Moravia. These new forms have been admitted into the codification. Conversely, in CC, levelling is in the opposite direction: all third-conjugation verbs are levelled to CC *-ej(i)*, for example, *sázej(i)*, *trpěj(i)* and by analogy fourth-conjugation verbs (or i-verbs as Townsend calls them) also take this suffix (*prosej(i)*). Word-final /í/ is usually truncated, although some traditional

---

therefore regarded as a morphological feature. In other positions, as in examples of the type *kolejí*, the instrumental singular form of the noun *kolej* 'university hall of residence', /iː/ is never truncated.

dialects in central Bohemia have resisted this feature until recent generations (see, for example, Jančáková 1987; 1993, Ireinová 2004). With respect to fifth-conjugation verbs, the SC suffix *-ají* is categorically replaced by CC *-aj*, with very minor regional exceptions. In Moravian dialects, several endings are observed, including *-ajó* (Central Moravian), *-ajú* (East Moravian), *-aju* (Silesian) and the innovative supralocal suffix *-ajou*. The regional variants for all the above verb categories are presented in CLA (Vol. 4. 448-459).

Paradigm unification is potentially very interesting in terms of speakers' accommodation. As is explained above, present-tense third-person plural verb forms are undergoing linguistic change and levelling is taking place. Furthermore, levelling is not in the direction of the standard but, instead, towards newly developed regiolects. Unlike the other two grammatical variables, informants' regional forms are not structurally the same as the SC equivalents and speakers are faced with the choice of adopting either the SC or the CC form – in the case that they drop their regional variant, that is. Interestingly, Krčmová (1979: 72) noted that regional forms in *-ijó* (*-ajó*) were receding in Brno – the city where CC has, in the opinions of some linguists, the strongest foothold outside its heartland – but were being replaced by (the new) SC rather than CC forms. This might give an indication to the route accommodation will take. The variant is relatively complex, insofar as it has a multivariant output: the verb *mluvit* 'to speak' has the following variant forms in the third-person plural: *mluví* (SC), *mluvěj* (CC), *mluví* / *mluvijou* (Common Moravian) *mluvijó, mluviju, mluvjá* (Moravian, traditional dialects). Surprisingly, paradigm unification is not included in Sgall and Hronek's geographical and functional tables; Čermák, however, categorizes the CC forms as **1/B**.

6.3.2 *L*-truncation

L-truncation is a process by which final syllabic /l/ is deleted in the masculine past

tense of certain first-conjugation verbs and certain second-conjugation '-nout' verbs:[93]

1. SC *nesl* 'he carried' >< *nes*; SC *četl* 'he read' >< *čet* (first conjugation)

2. SC *uběhl* 'he ran away' >< *uběh*; SC *tiskl* 'he printed' >< CC *tisk* (second-conjugation '-nout' verbs)[94]

---

[93] Although *l*-truncation is a phonological process, it is treated as a morphological variable because it occurs in a specific morphological environment: in the masculine past tense of certain first- and second-conjugation verbs. In other environments, the syllabic liquid /l̩/ is not truncated; thus, the nouns *rubl* 'rouble' and *smysl* 'sense' are never realized *rub or *smys.

[94] '-nout' verbs are those that either end in the suffix *-nou-* (or less frequently *-mou-*) before the infinitive suffix *-t*: *tisknout* 'to print', *stárnout* 'to grow old'; or verbs that conjugate identically to verbs that end in *-nout*: *říct* 'to say' (*řeknu, řekneš*, etc.), *začít* 'to begin' (*začnu, začneš*, etc.) and that have in some dialects alternative past-tense forms in *-nul* (*řeknul* (SC *řekl*) *začnul* (SC *začal*)).

In the case of '-nout' verbs, there is often an alternative masculine past-tense suffix -*nul*, which for some verbs is codified as standard, though usually considered less formal.[95] In such instances, the choice between *l*-truncation and *nul*-retention is said to be influenced, among other things, by whether or not the verb is prefixed and whether the root ends in a difficult consonant cluster (Townsend 1990: 86). Townsend tries to categorize verbs into those which are likely to undergo *l*-truncation in the past tense and those which are more likely to take *–nul*. He states (1990: 86) that for the verb *uběhnout* 'to cover a distance by running' the truncated form *uběh* is preferred by speakers of the Prague vernacular over *uběhnul* (and *uběhl*); for the prefixed verb *zestárnout* 'to get old)' the forms *zestár* and *zestárnul* are equally as likely (*zestárl* is unlikely); while for non-prefixed *stárnout* 'to get old' *stárnul* is used 'instead of *stár*' (and *stárl*). Sgall et al. (1992: 126) similarly assert that, while *příběh* is more common than *příběhnul* 'he came running', bare-stem forms like *uštk* or *zvlh* are less common than the 'nul'-forms *uštknul* and *zvlhnul* from the infinitives *uštknout* 'to sting, to bite' and *zvlhnout* 'to become damp', because of the subsequent consonant cluster. Hammer (1993: 65) argues that *l*-truncation may also be influenced by the rapidity of speech. The bare-stem forms were common in Czech poetry of the late nineteenth and early twentieth century and Bermel (2000: 19) points out that their use was 'probably closely tied to rhyme and meter, given that the final /l/ is syllabic'. *L*-truncation is

---

[95] '-nout' verbs are divided into three categories: (1) the 'tisknout' group, which consists of verbs whose stem ends in a consonant before the suffix *-nout* (*tisknout* 'to print', *táhnout* 'to pull', *kleknout* 'to kneel (down)'); (2) the 'minout' group, made up of verbs whose stem ends in a vowel, syllabic /l⌒/ or /_⌒, or have only one consonant before the suffix *-nout* (*minout* 'to pass', *zapomenout* 'to forget', *hnout* 'to move'); and (3) the 'tnout' (*Příruční mluvnice češtiny*) or 'začít' (Šaur 2004) group, including verbs such as *tnout* 'to strike', *začít* 'to start' and *vzít* 'to take'. In the first category, the alternative masculine past-tense suffix *-nul* is common, more so in some dialects than others. According to the guidelines published in the 1997 edition of *Pravidla českého pravopisu* [Rules of Czech Orthography], for some verbs, both the suffix *-l* and the suffix *-nul* suffix are considered standard, as in the examples *řízl* / *říznul* 'he cut' and *podvrtl* / *podvrtnul si nohu* 'he sprained his ankle'. For others, only the suffix *-l* is regarded as standard: *vyběhl* 'he ran away', *stiskl* 'he gripped', although non-standard *vyběhnul* and *stisknul* are frequently encountered and the *-nul* suffix is expansive (Šaur 2004: 90-91).

restricted to CC-speaking areas and has not been diffused beyond western Moravia.

Krčmová (1979: 73) found no traces of *l*-truncation in the language use of her

informants from Brno and in a more recent study Ondrášková (2004: 380) did not

record one instance of *l*-truncation in the speech of Brno school children. On the other

hand, bare-stem forms are encountered in certain traditional Silesian dialects, though

not in the regional interdialect. According to Šaur (1995b: 78), speakers in Opava use

bare-stem forms like *řek* or *utek* for verbs of the first conjugation but never for

second-conjugation '-nout' verbs; in this case, *nul*-retention is more widespread than

in other dialects: *spadnul* (SC *spadl* 'he fell'), *sednul* (SC *sedl* 'he sat down'),

vytahnul (SC *vytáhl* 'he removed') and for many other verbs of the same conjugation

(see CLA Vol. 4, 2002: 538-555; Vol. 5, 2006: 169).[96] Since there are only a few

verbs that undergo *l*-truncation, this variable is generally low-frequency. We would

expect to observe the bare-stem forms more in the speech of male informants in

common expressions such as *já bych řek(l)* or *řek(l) bych* 'I would say' or *moh(l)*

bych 'I could', and so on. The bare-stem forms are thought to be encountered

frequently in semi-official communication where native speakers view them as

normal (**1/A-B** according to Čermák's criteria and **3/A** according to Sgall and

Hronek). In Hedin's analysis of (semi-)formal television discourse (2005), the CC

bare-stem forms were recorded 39 times out of total number of 87 tokens (45%).

6.3.3 Gender neutralization

---

[96] The hitherto unknown author of *Denik Ostravaka* [An Ostravian's Diary] sticks to these conventions consistently in his (or her) representation of the Ostrava dialect. Marek, an informant from Opava, had a zero ratio of acquisition in I1, although in I2 he produced the common phrase *jak ji zobák narost*, truncating /l/. This might suggest that although the bare-stem forms are generally not encountered in the Silesian interdialect, certain lexicalized forms have been retained. Other informants from east and central Moravia also used this saying, but exclusively with SC *narostl*.

Gender neutralization denotes the process whereby the distinction between masculine

animate and inanimate hard adjectives and pronouns is lost in the nominative plural:

1. SC *dobří sportovci* 'good sportsmen' >< CC *dobrý sportovci*; SC *ti velcí*

   pavouci 'those big spiders' >< CC *ty velký pavouci*

2. SC *Oni jsou* [ˌ‿o‿] *bohatí* [-ciː] 'they are rich' >< CC *(v)oni jsou* [ˌ‿o‿]

   bohatý [-tiː]

SC maintains a distinction in the plural of hard adjectives and pronouns between the

masculine animate -*í* (*ti* [cɪ] *dobří*), the masculine inanimate and the feminine -*é* (*ty*

[tɪ] *dobré* for both), and the neuter -*á* (*ta dobrá*), with adjectives and pronouns in the

masculine animate often either undergoing a consonant mutation, as in the example

dobří (from the nominative singular form *dobrý*), or being palatalized *bohatí*

['bofiaciː] (from the nominative singular *bohatý* ['bofiatiː]).[97] CC, on the other hand,

has only one ending in the plural -*ý* (*ty dobrý*).[98] The loss of the gender distinction

means that masculine animate adjectives and pronouns in CC share the same form for

both the nominative and accusative plural adjectives and pronouns (*ty dobrý, ty velký*),

though masculine animate nouns have not merged: *sportovci* (Nominative Plural) ><

sportovce (Accusative Plural); *pavouci* (Nominative Plural) >< *pavouky* (Accusative

Plural). Čermák (1987) lists this feature as a form that occurs always or in most

---

[97] In the case of adjectives that do not undergo palatalization or a consonant mutation, the SC and CC forms are indistinguishable. Thus, the adjective *mrtvý* 'dead' in the masculine animate nominative plural is realized [mr̩tviː] in both SC and CC; it is represented orthographically in SC as *mrtví*.

[98] Moravian dialects also show a greater degree of syncretism in this respect. Although the masculine animate forms are retained in Moravian dialects, there has been a merger between masculine inanimate / feminine and neuter suffixes: Moravian *velké okna* >< SC *velká okna* 'big windows' (CC *velký* (*v*)*okna*).

instances within CC (1/A) and Krčmová (1979: 72) comments that gender neutralization is categorical in Bohemia and is used by speakers in speeches considered as standard.[99] Hronek and Sgall (1999: 184) and Sgall et al. (1992: 234-235) list SC forms such as *dobří sportovci* 'good sportsmen' alongside SC forms like *velká okna* 'big windows', *s těmi dobrými studenty* 'with those good students' and the first-person singular conditional auxiliary *bychom* as areas where SC possesses only bookish means of expression; it is argued by the authors that there is no stylistically neutral alternative to these bookish forms within SC (see Nebeská 2003: 98-99 for a fuller list of forms). In spoken discourse (of a non-formal character), they are replaced by their corresponding CC or interdialectal equivalents: *velký* *(v)okna* (CC) or *velké okna* (Moravian dialects); *s těma dobrejma studentama* (CC) or *s těma dobrýma studentama* (Moravian dialects); and *bysme* (CC ['bismε] and Moravian dialects ['bizmε]). However, the same does not hold for gender neutralization and forms of the type *dobří sportovci* should not be placed into the same category as the examples discussed above, since they are stylistically neutral throughout Moravia and are not in any way perceived as bookish or socially inappropriate for informal communication by their users. In fact, I suggest that it is just as unnatural for speakers of Moravian dialects to use forms like *dobrý sportovci* or *velký pavouci* as is it is for speakers of CC to use the SC equivalents.[100]

6.4 Scoring the linguistic variables

---

[99] Hedin recorded 22 occurrences of the CC forms out of 67 tokens. She expected, based on the sources considered above, that this figure would be much higher, and she concludes that the CC forms are in fact marked in (semi-)formal speech. It should be pointed, however, that Hedin has a non-representative sample of tokens for gender neutralization and this might distort the results.

[100] The SC forms were noted frequently during participant observation in the most informal settings and situations. Zdeněk used the forms *my zme dobří* 'we are good' and *zme druzi* 'we are second' while we were playing on a pub quiz machine and Radek used *velcí pavouci* 'big spiders' when talking to a group of classmates about insects that had got into his room. There were numerous other examples.

Variants of a given linguistic variable can be categorized as either 'dichotomous' ('binary'), where there are only two possible variants, generally discrete forms that are easy to code, or 'continuous', where there are several possible variants, in the case of vowels often identifiable across a phonetic continuum, where coding is more difficult. The variables in this study generally fall into the first category. Gender neutralization is the only discrete variant *per se*. *L*-truncation in some cases has a multivariant output in '-nout' verbs with three variant forms, while in the case of first-conjugation verbs the output is typically binary: *číst* 'to read', for example, has either *četl* (SC) or *čet* (CC). With regard to paradigm unification, all verbs have a multivariant output with distinct SC, CC and several regional forms. *É*-raising and *ý*-diphthongization have several variants; (ý), for example has the variant forms [iː] (SC and most Moravian dialects), [ɛj] (CC and Central Moravian dialects in some positions) and [ɛː] (Central Moravian dialects) and all the forms have variations in allophonic length. Word-initial (o) has the variant forms [o] (SC and most Moravian dialects), [vo][101] (CC and some Central Moravian dialects) and [ɦo] (some Central Moravian dialects). However, the level of detail required in distinguishing between variants of a variable depends on what the investigator wants to identify. Here, we are interested solely in whether or not a speaker adopts or rejects the CC variant. Therefore, categorizing and scoring the linguistic variables was a relatively straightforward task. Because our primary interest is whether a CC feature is adopted or rejected it was possible to categorize variables according to the criteria of CC versus non-CC (other) variants (SC, regional (interdialect), regional (traditional dialect)), which meant that all the variables could be treated as binary and it was not

---

[101] /v/ is said also to have distinct 'strong' and 'weak' realizations (Townsend 1990: 39).

necessary to calculate a mean (index) score for the variables that occur over a phonetic continuum or that have a multivariable output. In cases where the CC form is rejected, it is obviously interesting to see whether the speaker retains the form present in his or her regional dialect or whether he or she accommodates to the standard. However, the only cases where informants' regional forms differed from their SC equivalents were for paradigm unification and smaller distinctions such as central Moravian [ɛː] as a reflex of (é) as in *velké strom* (SC *velký strom* 'a big tree') or the East Moravian pronunciation [ɛj] ([ɛːj]) for (ý) in the oblique cases of hard feminine adjectives and pronouns and Central Moravian [ɦo] for (o). Moreover, as will be discussed in detail in the next chapter (§ 7.5), there were no examples of two of these localized regionalisms, only one incidence of East Moravian [ɛj] and exceptionally few regional forms for paradigm unification. Consequently, when I analyze the results, the data are scored and presented as either 'CC' or 'other' (SC / regional) forms.

Speakers of Central Moravian dialects were scored differently in view of the fact that some of the CC forms under study are identical to variants used in their native dialects. Scoring the grammatical variants is simple, insomuch as the CC variants under study are not found in any part of Moravia, other than in the westernmost parts where CC is considered the mother-tongue variety. The geographical distribution of the phonological variants is more complicated. All three variants are encountered in Central Moravian dialects to varying degrees, although not necessarily as a direct result of the spread of CC. All Central Moravian dialects have diphthongized /ɛj/ in desinence-final position (only) and raised /iː/ in the nominative and accusative plural of adjectives, the oblique cases of hard masculine adjectives and

pronouns and in the singular of neuter adjectives, although their use and social acceptance may differ from that of Bohemia and in some cases they are found on a stylistic continuum competing both with SC and regional forms. Additionally, in some parts of Central Moravia, in particular around Brno, prothetic /v/ is found in the local dialects. Informants were therefore scored according to their region of origin. For all central Moravians, ý-diphthongization was not scored in desinence-final position and é-raising was taken into consideration only in the oblique cases of hard feminine adjectives and pronouns. For those who come from areas where prothetic /v/ is a feature of the local dialect, v-insertion was omitted.

6.5 Tokens

The incidence of each variant was counted manually. In order to achieve representativeness and for quantitative analysis to be successful, a large number of tokens – instances of the occurrence of a particular variant of a linguistic variable – needs to be analyzed. There are no set rules with regard to the number of tokens a sociolinguist needs to make realistic generalizations about the distribution and use of linguistic variables, and hardly any advice is offered in the specialist literature. To use Czech as an example, studies working with many hundreds of tokens have identified that native speakers of CC almost categorically realize /ɛː/ as [iː] in adjectival desinences. Dejmek (1986), for example, collected a total of 9530 tokens for this variable and his data showed a ratio of CC 9432 >< SC 98 (99%) in favour of CC raised /iː/. If, however, Dejmek had based his analysis on, say, only five tokens and on the three occasions the variable was realized [iː], while on the other two [ɛː] –

which is quite possible – this would suggest the informant raises /ɛ:/ only 60 percent of the time, which obviously does not reflect his or her actual use of this variable.

Guy (1980: 26) suggests that 30 tokens of a given variable per speaker is a reasonable objective; he argues that a number of tokens above 10 'moves towards 90 percent conformity with the predicted norm, rising to 100 percent with 35 tokens', while a number below 10 cannot be taken as representative. This is a major handicap for quantitative sociolinguistics, since in some cases those variables that are considered interesting occur infrequently and it is difficult to yield representative results that can be used to make generalizations about the speech community under study. Methods have been advanced to elicit such problematic variables, such as specially selected word lists and reading passages, but this inevitably renders the interview situation less natural and non-casual. However, linguists can make reliable generalizations about low-frequency variables in the case that a regular pattern can be observed in the speech of a representative sample of informants. For example, if there is an average of only five tokens per interview but all or most informants (in a representative sample) either adopt or reject the feature on all or most occasions, then we can still talk of trends about how that feature is assimilated. Thus, with respect to low-frequency variables, inter-speaker similarities increase representativeness and allow us to generalize about the use of a particular linguistic feature.

6.6 Some general theories of dialect accommodation

Since the inception of the dialect contact framework in the 1980s (§ Chapter 2) there has been much theorizing about the outcomes of dialect contact. Linguists have tried to predict the results of short-term or prolonged encounters between speakers of different regional and / or social dialects and attempts have been made to predict

which linguistic features of the new host variety are likely to be acquired or rejected and / or which features of speakers' native dialect are likely to be retained or dropped. We earlier considered the social motivations for converging towards another variety and now we turn our attention to the result of the convergence. In this section, the theoretical concepts of dialect accommodation and second-dialect acquisition are considered and some general language-internal and extralinguistic constraints that are influential in determining to what extent individuals are able to assimilate the linguistic features of a second variety are discussed in detail. Here I pay particular attention to Trudgill's (1986) notion of 'salience', Schirmunski's (1928) theory of 'primary' and 'secondary' dialect features, Chambers's (1992) principles of dialect acquisition and Labov's (1972a) categorization of linguistic variables into 'markers', 'indicators' and 'stereotypes'.

First, though, let us consider the probable outcome of dialect contact between speakers of mutually intelligible but regionally different varieties. The direction and the extent of accommodation and convergence (or divergence) in situations where the individual varieties are mutually intelligible is inextricably linked to the social relations between the individual groups in contact. We should remember that non-standard varieties differ in terms of their prestige, stigmatization, areal distribution and functionality and that these factors can determine to what extent and in which direction accommodation takes place. If the migrants' dialects are in some way stigmatized or perceived negatively in the host community,[102] or if the dialects are localized or peripheral, then this might motivate the speakers to accommodate. Conversely, speakers of supralocalized regiolects that are devoid of highly-localized

---

[102] Some of the Stril migrants in Kerswill's study (1994) were overtly self-conscious about their localized dialects. Some of the older informants had been told by their employers that if they did not speak 'properly' they would have to leave (1994: 37) and they were under considerable pressure to converge to the speech habits of the host group.

and marked linguistic items and that are functionally intermediate between traditional dialects and the standard dialect may not consider it necessary to accommodate in out-group communication (Auer and Hinskens 2005: 335). In fact, we might expect speakers of traditional or localized varieties to accommodate in the direction of a more prestigious and supralocal (though not necessarily standard) variety (Sgall and Hronek 1992, Sgall et al. 1992).

Kerswill (1994: 4) argues that 'the outcome of language and dialect contact will depend on both the linguistic relationship between the varieties and, equally importantly, the social conditions underlying the contact'. In cases where the varieties in contact are linguistically (very) different – say, Czech and German – the possible results are code-switching, bilingualism and language shift. However, where the varieties are linguistically (very) similar, as in the present contact situation between CC and Moravian dialects, there are several possible outcomes. The migrants may accommodate minimally towards the host variety; they may switch between the two varieties (code-switching); they may emphasize features of their native dialects that do not occur in the host variety (divergence); or they may attempt to adopt the host variety at the expense of their native dialects (dialect shift). Their idiolects may also be augmented by items from the host variety (borrowing). Kerswill considers that extensive borrowing will prevail in situations where the varieties in contact are mutually intelligible and structurally similar, rather than code-switching and bilingualism (bidialectalism), although individual bidialectalism is also possible (Trudgill 1986: 1). The term 'borrowing' is usually interpreted as the transfer of some linguistic item from one language (variety) into another language (variety). Swann et al., for instance, list English *garage* as a borrowing from French and French *le weekend* as a borrowing from English (2004: 30). Borrowing is often associated

chiefly with lexemes, but it is much pervasive than that. Kerswill (1994: 21) argues

that in contact situations between varieties that are mutually intelligible, speakers will

borrow not only content words but also function words and morphological suffixes.

He calls this type of borrowing 'morpho-lexical mixing' (1994: 7) and he predicted

that his Stril migrants in Bergen would acquire morpho-syntactic features of the host

dialect more often than phonological or suprasegmental forms (1994: 79).

At this stage, therefore, we have already established: (1) what the social

motivations for accommodation are; (2) that speakers of Moravian dialects have a

strong motivation to accommodate; (3) what the direction of the accommodation is

likely to be; and (4) in what ways the accommodation is likely to be manifested. In

addition, we can also predict which features of the host variety are likely to be

acquired the most or the quickest and why. Some linguists, for example, assert that

accommodation does not occur arbitrarily but is rule-governed, both in terms of which

variants are acquired and in terms of the stages at which variants of the native variety

are dropped and forms of the host variety adopted. Trudgill (1986) believes that in

some cases accommodation follows a fixed route and that individual variants are

acquired in a structured order. Chambers argues that the elimination of old rules

comes easier than the acquisition of new ones (1992: 696) and it seems perfectly

natural to expect that certain features of the native dialect will be dropped before

features of the target variety are acquired. Labov (1972a: 178-180) states that variants

of linguistic features can be categorized as: (1) 'indicators', variants to which little or

no social significance is attached and that may only be perceived by observers with

special linguistic training; (2) 'markers', variants that are readily perceived by both

specialists and non-specialists and have social significance; and (3) 'stereotypes', 'a

variable which is popularly taken to represent the speech of a particular group,

whether or not it conforms to reality'. Schirmunski categorizes variants into 'primary' dialect features (*первичные признаки*) and 'secondary' dialect features (*вторичные признаки*). His features, which are discussed in more detail below, correspond loosely to Labov's 'markers' and 'indicators'. Generally speaking, it is expected that with respect to eliminating features of the native dialect, markers or primary features – that is, those forms that are high in speakers' consciousness – will be dropped, and in accommodating to the host variety, markers of that variety will be adopted, or rather will be the primary candidates to be adopted if there are no inhibiting factors. Conversely, features that are beneath the threshold of speakers' awareness (indicators or secondary dialect features) are not expected to be either dropped or acquired.

Therefore, this to an extent has clarified which features speakers retain or drop and adopt or reject in moving to another dialect region, or, rather, this is the expected pattern. Speakers are aware of markers and, therefore, can modify (drop or acquire) them in situations where they for whatever reason monitor their speech; indicators generally fall beneath speakers' level of consciousness and are, therefore, in many cases retained or not adopted. It is still unclear, nevertheless, how a particular variant becomes a marker (or a primary dialect feature) or what makes a particular dialect feature an indicator (or a secondary dialect feature). Trudgill (1986: 10-20), in his investigation of the accommodation rates of speakers of English English living in America, which includes a *post hoc* analysis of his own accommodation, believes that a key factor is 'salience'. Salience is described by Kerswill and Williams (2002: 81) as 'a property of a linguistic item or feature that makes it in some way perceptually or cognitively prominent'. According to Trudgill, if a variant is in some way well-known or 'salient' it becomes a 'marker' and is likely to be easily acquired or dropped, depending on other external factors. The notion of 'salience' was first employed by

Schirmunski in his investigation of dialect levelling in German language islands in Russia in the 1920s (see Auer, Barden and Grosskopf 1998:163-166) for which he developed a set of criteria that he believed made linguistic variants salient. In categorizing linguistic phenomena as either 'primary' (salient) or 'secondary' (non-salient) dialect features, he emphasises the following factors (adapted from Auer, Barden, and Grosskopf 1998: 163-164):

1. Articulatory and perceptual distinctness between the variants of a particular variable);

2. Lexicalization;

3. Phonetic dichotomy versus phonetic continuum;

4. Lay speakers' awareness;

5. Writing;

6. Mutual intelligibility.

According to Schirmunski, primary dialect features or salient dialect features tend to be phonetically distant as opposed to phonetically close, binary variants as opposed to continuous variants, and readily perceived by linguistically untrained speakers. Lexicalization increases the salience of a variant and consequently its learnability; therefore, lexicalized variants are also salient. With respect to writing, dialect forms

that are graphically distinct from standard orthography – those forms that cannot be transported into the writing system in a straightforward manner – are more salient than features which can be easily read into the orthography. And only primary dialect features may (but need not) preclude mutual intelligibility. Auer, Barden and Grosskopf (1998) list Schirmunski's first three criteria as objective factors and the last three as subjective factors. While Schirmunski uses the notion of salience in terms of dialect levelling within a speech community over a prolonged period, Trudgill tests its applicability in various studies of long-term accommodation of first-generation migrants. His framework for delimiting salient features of American English, for example, was to identify the forms of American English that British pop singers imitate. He lists the following factors that make linguistic features salient (adapted from Trudgill 1986: 11):

1. One of the variants of the variable is overtly stigmatized (stigmatization).

2. The variable is currently involved in linguistic change (linguistic change).

3. Variants of the variable are phonetically very different (phonetic distance).

4. Variants of the variable are involved in maintaining a phonological contrast (phonological contrast).

Trudgill argues that all salient features are, in theory, candidates to be dropped early, in the case of reducing or avoiding stigmatized features of the localized vernacular, or to be acquired early, in the case of assimilating forms of the host variety. However, in

practice adopting (though not dropping) a particular salient linguistic feature may be delayed, impeded, or even prevented by a number of factors. We can use Trudgill's study of speakers of English English in America as an example. Five linguistic variables were investigated: (1) the monophthongal realization of /aɪ/ as in the lexical set *life* as [aˑ]; (2) use of non-prevocalic /r/ as in *cart*; (3) acquisition of /ɑ/ in words like *hot* and *top* where *o* is realized in English English as [ɒ]; (4) /æ/ in words such as *dance* and *last*; and (5) the realization of intervocalic /t/ as [ɾ]. Trudgill identified a fixed route of accommodation whereby informants acquired the variants in the same order: (5), (4), (3), (2).[103] Interestingly, although Trudgill identifies all the forms as salient, certain forms were acquired earlier and to a greater extent than others. This can be explained by a number of inhibiting factors that may delay or even prevent the adoption of salient forms. Non-prevocalic /r/, though salient, was acquired only variably and only by informants who had lived in the host community for a relatively long time – Trudgill suggests informants would start to use this feature after ten years or so. The reason is that the dialect feature poses phonetic-articulatory problems for the acquirer – speakers struggle to or simply are unable to realize a particular phonetic feature. Trudgill calls this a 'phonotactic constraint' that inhibits accommodation: 'the phonotactic constraint present in their non-rhotic accents prevents them from accommodating to American English on this particular feature' (1986: 16).

If a feature is complex it may never be acquired, especially by adult migrants. It has been proved that phonologically complex variables such as the Philadelphia pattern of raising and tensing short /a/ (described in Roberts and Labov 1995: 102-

---

[103] Neither Trudgill, nor his informants used monophthongal /aˑ/. This feature is not used by Americans as a whole, but by speakers from the south and speakers of AAVE. Trudgill in his short stay had not been to the south and he had little contact with speakers of AAVE or speakers of southern dialects.

103), due to its extreme complexity, is sometimes not acquired even by pre-adolescents (Payne 1980). Payne identifies to what extent children who had moved from one dialect area to another are successful in acquiring the phonological system of the new host variety. She looks to answer two questions: (1) 'whether a child freely recognizes and / or restructures his grammar up to the age of 14'; (2) whether he 'will learn to speak like his peers or retain the system learned from his parents' (1980: 143). Her results show that, although the in-migrants assimilated other features of the host variety successfully, their acquisition of the Philadelphia short /ɑ/ pattern was generally unsuccessful. This holds also for children who were born and raised in King of Prussia, the middle-class suburb of Philadelphia where Payne carried out her research, but whose parents were not born and raised locally. Payne (1980: 156) argues that in order to successfully assimilate the short /ɑ/ pattern, in-migrants must learn not only the phonetic conditioning of its distribution, but also 'the grammatical conditioning and lexical exceptions'.

Conversely, acquisition of /ɑ/ does not involve phonotactic constraints. The relatively low level of accommodation is explained by a 'homonymic clash', since in English English /ɑ/ is used in the lexical set *heart, park*, and so forth. Therefore, homonymic clash is a second factor which might have a delaying effect on accommodation. The reason why speakers of English English used /æ/ only marginally in the lexical set *dance* is more complicated and Trudgill (1986: 18) offers a social psychological explanation. He bases his own non-assimilation of this variant on the fact that this feature is 'too' stereotypical of the host community; it is 'too' salient or 'too' American for him to use. He believes other speakers will follow suit and that the 'extra-strong salience' will inhibit accommodation. However, the idea of

extra-strong salience is problematic. If we were to compare Trudgill, who was in America on only a short-term basis, to, say, an individual who intends to stay in America forever, then the likelihood is that the latter speaker will have a stronger motivation to adopt this feature. Therefore, extra-strong salience can have different effects on accommodation, depending on extralinguistic factors that relate directly to the individual speaker.

With respect to the above argument, it is clear that in addition to the factors which, according to Trudgill, make linguistic phenomena salient, there are several other important factors that Trudgill overlooked. Kerswill and Williams (2002), for instance, challenge Trudgill's notion of salience, arguing that the factors he lists as making a feature salient are based solely on language-internal factors. They use Trudgill's notion of salience in their dialect levelling project in Milton Keynes, Reading and Hull. The conclusion is that by not emphasising the extralinguistic (sociopsychological and sociolinguistic) factors that contribute to making a linguistic feature salient, Trudgill fails to 'gain insight into the social patterning of linguistic features' (2001: 104). Trudgill's notion of extra-strong salience is given as an example. According to Trudgill's hypothesis, it appears that 'the very factors that lead speakers to notice and to adopt a feature ... are precisely those that also lead to a feature being avoided' (Kerswill and Williams 2002: 89). Therefore, in such circumstances acquisition or rejection will be determined by extralinguistic factors; in fact, it was the extralinguistic factors that were decisive in whether or not a feature was salient (in Kerswill and Williams' study). Auer, Barden and Grosskopf (1998) also test salience as a good predictor of the dropping and assimilation of dialect forms in a longitudinal study of the long-term accommodation of East German migrants in West Germany. They build on both Schirmunski's and Trudgill's criteria, using ten

factors (five objective, five subjective) to predict the salience of a linguistic feature (1998: 167). The results show that while there is a strong correlation between high salience and the loss of certain forms, the loss of other dialect features cannot be explained by their relative salience. On the one hand, weak intermediate forms of the native dialect were readily given up. The converse was observed for strong forms, in particular in the case of lexicalized, dichotomous variables which, although highly salient, were particularly resistant. Auer, Barden and Grosskopf also observed that subjective factors overrode objective factors for salience.

A further shortcoming in Trudgill's notion of salience, and also in Schirmunski's idea of primary and secondary dialect features, is that both models expect the dropping of regional forms and the assimilation of forms of the host variety to follow an identical route. Auer, Barden and Grosskopf (1998: 168) suggest that the two processes should be treated individually. They give as an example lexicalized items, which tend to be acquired in a different manner to which they are dropped. Although Schirmunski had identified lexicalized forms as primary features that are candidates to be dropped in dialect loss or assimilated in the acquisition of a new dialect, the East Germans in Auer, Barden and Grosskopf's study contradict this. They generally retained lexicalized forms of their native dialects, regardless of the fact that they were highly salient, and it is claimed that lexicalized variables 'may be particularly sheltered from loss' (1998: 184). However, the retention or loss of lexicalized forms is likely to be tightly linked to attitudinal factors and their relative prestige. We should not take the results from one study as axiomatic. In terms of reducing or avoiding highly salient lexicalized forms, speakers during accommodation are able to do this quickly and successfully. This, however, does not hold for dialect acquisition, since speakers must learn each example individually; they might decide to

reject the variable altogether or use it only variably, or they 'may turn this lexicalized variable into productive, non-lexicalized phonological rule and then overshoot the target' and produce what Trudgill calls a hyperdialectism (1986: 66-68), described by Britain (in press) as extending 'the local form to linguistic contexts where it was previously not used', or in this case where native speakers of CC do not use it.

This would certainly hold for my native dialect with regard to /aɪ/ in the lexical set *night, fight, light, sight*, and /iː/ in the lexical set *key, pea, bean, team, mean, seen*. In the first instance, many of words in this set are pronounced similar to the standard with [aɪ]: *sight, might, kite, height*; however, *night, light, frightened* (though not *fright*) are realized with [i] and *right* and *fight* with [eɪ]. Likewise, *prestige, reach, preach,* and many more words have standard [i], where as *keys, peas, meat, eat, teacher* have [eɪ] and *team, mean, season, cheap* are realized with [iˑə]. These stigmatized regional forms would be easy for me to drop in out-group situations; however, they would be far more difficult for a speaker from another speech community to acquire because he or she would need to learn each example individually in order to avoid hyperadaptation. In fact, the complexity of this variable may discourage acquisition altogether.

Chambers (1992), in his study of the linguistic behaviour of Canadian youngsters who had moved to southern England, focuses solely on dialect acquisition. Although the study deals with two standard dialects, not second-dialect acquisition in a contact situation between speakers of non-standard varieties, many of the points Chambers makes about second-dialect acquisition are directly relevant to the present study. Chambers summarizes that:

1. Lexical replacements are acquired faster than pronunciation and phonological variants.

2. Lexical replacements occur rapidly in the first stage of dialect acquisition and then slow down

3. Simple phonological rules progress faster than complex ones.

4. Acquisition of complex rules and new phonemes splits the population into early and later acquirers.

5. In the earliest stages of acquisition, both categorical rules and variable rules of the new dialect result in variability in the acquirers.

6. Phonological innovations are actuated as pronunciation variants.

7. Eliminating old rules occurs more rapidly than acquiring new ones.

8. Orthographically distinct variants are acquired faster than orthographically obscure ones.

Principles (3), (6), and (7) are especially interesting. First of all, it is necessary to define, and distinguish between, 'simple' and 'complex' phonological rules. According to Chambers, simple rules are automatic ones that admit no exceptions and complex rules 'have opaque outputs, that is, they have exceptions or variant forms' (1992: 682). In other words, we could talk of simple rules as categorical rules and we

could define complex rules as variable rules. Chambers lists (1992: 685-686), *t*-voicing in North American English whereby words such as *putting / pudding, hearty / hardy* are homophonic, or yod-insertion after velars and before low vowels in words like /kjat/ 'cat' and /gjas/ 'gas' in Jamaican Creole (adapted from Wells 1973) as simple rules. As an example of a complex rule, he lists the unmerging of Jamaican Creole /ie/ or Jamaican English /eː/ before tautosyllabic /r/ (1992: 686), whereby words such as 'steer' and 'stare' are realized identically: [stier] in Jamaican Creole or [steːr] in Jamaican English. The examples from my native dialect listed above can also be described as complex rules with respect to their opaque outputs.

With regard to principle (6), 'phonological innovations are actuated as pronunciation variants', Chambers (1992: 693-694) proposes that speakers who move to a new dialect area acquire new variants individually rather than identifying a phonological rule and adapting their phonological systems accordingly. Put another way, the adoption of new phonological variants is consistent with the theory of 'lexical diffusion', according to which 'phonological innovations are actuated by the acquisition of particular instances of the new rule or phoneme, and they only become rule-governed or systematic (if ever, in the first generation) after a critical mass of instances had been acquired' (Chambers 1992: 693). If this is true, my informants would be expected to acquire first individual words that undergo, say, *ý*-diphthongization, particularly in high-frequency lexical items such as *velkej* 'big' or *dobrej* 'good', rather than changing their phonological systems and consistently diphthongizing /iː/ to [ɛj] in desinence-final position of all hard masculine adjectives.

Principle (7) states that in second-dialect acquisition, eliminating old rules (features of the original variety) precedes acquiring new rules (features of the target

variety). As Chambers argues, second-dialect acquisition not only involves accommodating to the target variety through convergence, it also means sounding less like people from the old region through divergence from the native variety (1992: 695). Over time, therefore, speakers after the critical age often sound less like the people from the old region to the point that they are taken by speakers of that variety as 'non-natives', but rarely accommodate entirely to the target variety to the point where native speakers of that variety would accept them as being from the same speech community. Chambers also argues that it is easier for speakers to drop features of their native variety than it is for them to acquire features of the target variety. However, in considering word-final obstruent devoicing in German (1992: 696), he argues that although English students of German generally have no problems in devoicing word-final obstruents – this is an example of a 'simple rule' – German students of English have trouble realizing word-final obstruents as voiced consonants.[104] This seems to contradict the above and it also contradicts Chambers's earlier statement that simple rules are easy to both drop and acquire. A likely explanation why German students of English fail to produce word-final voiced consonants is because a phonotactic constraint is involved.[105]

6.7 Partial or incomplete accommodation

First-generation migrants are unlikely after the critical age to fully assimilate all features of the host variety successfully. Incomplete accommodation has a number of

---

[104] This also holds for Czech students of English. Many speakers whose English in terms of grammatical correctness, fluency and use of idiom is flawless, invariably pronounce the words *bed, bad, sad* as [bɛt], [bæt], [sæt]. However, contrary to Chambers's claim that English students of German easily master word-final obstruent devoicing, I would argue that English students of Czech, and of other Slavonic languages, struggle to overcome this, probably due to a lack of knowledge of the phonetic processes involved, especially within words and even more so across word boundaries.

[105] Interestingly, even knowing the rules, Czech students of English still realize words such as *bad* or *bed* with word-final [t] – although they are often convinced that they are pronouncing them with [d].

possible manifestations, including both the inability to drop features of the base variety and an unsuccessful command of the host variety. Trudgill (1986: 57-82) argues that the acquisition of features from the target dialect is frequently incomplete or imperfect, thus resulting in intermediate forms between the 'originating' and 'receiving' dialects. He goes on to discus the types of partial accommodation and concludes (1986: 62) that accommodation may be incomplete in three ways: (1) switching between (and mixing of) variants of the original and host variety; (2) inconsistently using features of the target variety, that is, using them in some words but not others; and (3) using intermediate (predominantly phonetic) forms that are not found either in the original or in the host variety. For the last type of partial accommodation, Trudgill coins the term 'interdialect' (§ 3.8), an outcome of contact between two dialects that leads to 'the development of forms that actually originally occurred in neither dialect' (1986: 62). Trudgill (1986: 66) also uses the term 'hyperdialectism' to explain another type of partial accommodation. The use of hyperdialectisms is identical to hypercorrection, although it involves making an error while striving to use a non-standard dialect, rather than the standard dialect. Such hypercorrections occur either because speakers believe the form to be correct (that is, they have an incomplete knowledge of the target variety) or as one-off slips, possibly under the influence of their native variety. A hyperdialectism works in the same way: speakers make generalizations about features of the target dialect and in which words they occur, and through overgeneralization they occasionally 'overshoot' and produce forms not found in the target variety.

6.8 Applying the general principles to the present contact situation

We considered above some general models about dialect accommodation and second-dialect acquisition, in particular about which factors influence the adoption, rejection, maintenance or dropping of linguistic phenomena in contact between speakers of regionally different varieties. Now that we have introduced the CC variants in terms of their areal distribution and social acceptance, we can try to apply these general models to the given contact situation and predict which forms of the host variety are the most likely to be adopted and why.

Let us begin by returning to the eight principles of second-dialect acquisition outlined by Chambers (1992). Chambers makes several interesting propositions and here I would like to address his notion of 'simple' and 'complex' phonological rules. According to Chambers, simple rules are acquired early and easily, while complex rules are acquired late or not at all. The notion that the adoption of new forms is facilitated or impeded in view of whether or not a feature constitutes a simple or complex rule is particularly relevant to the present study; however, the categories 'simple' and 'complex' are generally too vague and most variables fall somewhere between these polar categories. There are very few simple rules that admit 'no' exceptions whatsoever. Take, for instance, definite article reduction (DAR) in my native dialect, a process by which the definite article *the* is reduced and realized in most instances as a glottal stop [ʔ], or occasionally as a plosive [t], as in the lexicalized forms *told* 'the old' and *tother* 'the other'. DAR occurs in all positions, except before particular expletives, which in very colloquial usage replace a noun;[106] therefore, this is not strictly a categorical rule, because there are 'very minor' exceptions. The rule is simple, insomuch as whenever the switch takes place it is a

---

[106] For example, in a game of football a player might shout out *Pass us t'ball over here*; however, if he replaced the noun *ball* with an expletive like *bastard*, which is fairly common in such situations, the definite article cannot be reduced: *Pass us the bastard over here* (**Pass us t'bastard over here*).

discrete choice between the dialect and the non-dialect form and the dialect form has with very minor exceptions only one output; however, since the switch is not 'entirely' categorical, this may theoretically delay acquisition or result in hyperadaptation. Another example is *ý*-diphthongization in CC: although there is a discrete choice between the SC and the CC variant, the switch does not take place in certain foreign borrowings and in the instrumental singular of hard masculine and neuter adjectives; consequently, we might expect hyperadaptation.

Bearing this in mind, it may be preferable to categorize variables into at least three types of 'rule'. First, 'simple rules' in the sense of a rule which admits 'no' exceptions whatsoever: the switch takes place in 'all' positions and the variable has a binary output with no exceptions. Again, using an example from my native dialect, *h*-dropping in word-initial position could be considered a simple rule, since speakers have a choice between [h] and Ø and /h/ can be dropped in any word in word-initial position. A complex rule can be described as a rule that has an opaque output where speakers are faced with a choice of at least three variant forms. Again, a further distinction between complex rules could be made based on whether a switch occurs in all positions in all words or whether it occurs only variably. Finally, for the purpose of this study I use the term 'semi-simple rule' to describe those rules that have binary outputs, but where a switch is not entirely categorical. In terms of acquisition, semi-simple rules should be adopted in the same manner as simple rules and they are potential candidates for hyperadaptation, since they are all extremely frequent and easy to acquire, but non-natives may be unaware of the positions in which they do not occur.

All three phonological variables under study belong in this third category. They do not have multi-variant outputs – excluding variations in vocalic length which

are not considered important for the given analysis – but none of them occur in all positions: /ɛj/ is not used in the instrumental singular of hard masculine and neuter adjectives, /ɛː/ is not raised to [iː] in foreign words or in the nominal suffix -ové and several words beginning in /o/ are believed not to undergo prothesis. The grammatical variables are either simple or complex rules. L-truncation is a complex rule, since it has an opaque, multivariant output in second-conjugation verbs or is barred in certain positions due to its phonetic complexity in both first- and second-conjugation verbs. Despite having several forms in Moravian dialects and two standardized variants for third-conjugation verbs, paradigm unification purely in terms of acquisition constitutes a simple rule, insofar as the CC suffixes can replace their SC equivalents for all verbs and there is only one option in the host variety: -aj for fifth-conjugation verbs and -ej for third- or fourth-conjugation verbs. Gender neutralization is also a simple rule in this respect.

Another interesting point that Chambers makes concerns lexical diffusion. He proposes that linguistic phenomena are assimilated on a word-by-word basis, while he considers the wholesale adoption of a linguistic form throughout an entire lexical set unlikely. Chambers suggests that even simple rules are diffused predominantly on a word-by-word basis and migrants try to imitate and mimic speakers of the host dialect, focusing on high-frequency words that contain a salient phonological feature, rather than on the phonological feature itself. However, we might expect the converse in the present study, if informants' comments in I2 are accurate. The CC phonological variants are extremely high in non-native speakers' consciousness (though this is not necessarily true of the grammatical ones) and the migrant speakers may well adjust their phonological systems as opposed to assimilating features on a word-by-word basis. Such is the salience of the phonological forms that speakers often comment

directly on the individual feature rather than listing words in which the feature appears (Cummins 1993: 156); this was apparent in I2 when informants were asked to comment on CC forms that they would use or disliked.[107] Therefore, I would suggest that there is a strong possibility that the acquisition of the phonological forms will be rule-governed as opposed to lexicalized.

We should also take into consideration the factors that contribute to the salience of linguistic features and we should return to the notion of areal distribution, which, as we argued earlier, had been overlooked by linguists hypothesizing about the results of dialect contact. In contact-induced procedures such as dialect levelling, dialect supralocalization and the emergence of geographically expansive interdialects it has been demonstrated that features which unite speakers over a large area are generally retained, while those restricted to highly localized communities are more likely to undergo levelling and partially or totally disappear in situations of dialect contact and subsequent mixing. The language situation in Moravia supports this hypothesis. Take, for example, the levelling and koineization in Silesia, which was described in Chapter 3 (§ 3.8). Geographically widespread linguistic features such as lack of vocalic length and penultimate stress have remained, while more localized forms like instrumental plurals in –*um* in some parts of the Opava region or the phonological shift /o/ → /u/ in monosyllabic and disyllabic words of the type *pozur, tchuř, druzd* >< SC *pozor* 'attention', *tchoř* 'polecat', *drozd* 'thrush (bird)' are on the decline and now used exclusively by older speakers. We might expect, therefore, that geographically widespread features used by large groups of speakers are acquired earlier than those variants that have only a small areal distribution and / or are used only marginally by a small group of users.

---

[107] As well as listing CC forms in individual words such as *zejtra* or *tejden*, informants frequently talked about 'that *ý*' or 'that *v*'.

Linguists concur that dialect contact between adult speakers very rarely results in the complete assimilation of a new variety. First-generation contact is typically characterized by considerable levels of inter- and intra-speaker variation and various types of partial accommodation are usually observed. It is inevitable that informants in this study will show incomplete accommodation of Trudgill's type (1), which involves frequent switching between variants of the original and host variety, and (2), where in-migrants use features of the host variety inconsistently, applying them in some words but not in others (significantly more and in a more randomized manner than native speakers do). In view of the nature of the phonological variables, it is also possible, but somewhat less likely, that informants will produce interdialectal and / or hyperdialectal forms, especially if we find confirmation of the hypothesis that assimilation of phonological variants will not follow a lexicalized pattern but will be rule-governed. Misanalysis of the host variety and overgeneralization could result in anomalous use of most of the CC variants, although there are relatively few areas where informants can overshoot the target. The most likely way in which informants will perpetuate hyperdialectisms is by infringing the rules of acceptability established by Kučera (1955; 1958; 1973).[108] A possible outcome of second-dialect acquisition is

---

[108] Kučera sorted non-standard features into an acceptability hierarchy and he devised a set of rules according to which the absence of certain non-standard forms can exclude the presence of other non-standard features (Kučera 1955: 585). His main argument was that forms at the lower end of the hierarchy could be used only if the more acceptable non-standard features were present. Thus, while for 'every week' the combinations *každý týden* (SC + SC), *každej tejden* (CC + CC) and *každej týden* (CC + SC) are possible – or were deemed possible by the native speakers who were asked to evaluate the combinations – *každý tejden* is not. This is because /iː/ is less acceptable in word roots than it is in adjectival desinences and its absence in *každý* excludes its presence in *týden*. Kučera also states that in the case of two or more desinences of juxtaposed adjectives SC + CC or CC + SC combinations are not possible: **nový kamennéj dům* and **novej kamenný dům* (1955: 589). The same naturally holds for shifting between /ɛː/ and /iː/ in juxtaposed adjectives: **vysoké bílý čelo* and **vysoký bílé čelo* (1955: 588). Kučera asserts that the acceptability hierarchy of which speakers are not consciously cognizant is with the exception of very minor individual differences 'firmly established in the linguistic habits of every Czech speaker familiar with the CCL [CC] and the literary norm' (1955: 586). Another interesting claim was that CC morphology can be freely combined with SC phonology in the same word but not vice-versa; therefore, the combinations *mladými* (SC phonology + SC morphology), *mladejma* (CC phonology + CC morphology) and *mladýma* (SC phonology + CC morphology) are permissible in the instrumental plural of *mladý* 'young', whereas **mladejmi* (CC phonology + SC

'hyperdialectization'.[109] Based on a lack of knowledge or misanalysis of the host variety speakers may produce a local form in a position where it is not used by native speakers; take, for example, *é*-raising: SC /ɛː/ is raised to [iː] in many word roots, but not in the words *jméno* 'name' or *léto* 'summer'; therefore, the forms **jmýno* and **líto* are hyperdialectisms.[110] Another example would be *ý*-diphthongization in the instrumental singular of hard masculine and neuter adjectives: although /iː/ is diphthongized to /ɛj/ in other adjectival desinences, only *s velkým psem* 'with a big dog' is possible, while **s velkejm psem* is not. On the whole, continual oscillation between regional, SC and CC forms is expected.

6.9 Predictions with regard to the assimilation of CC forms

The contact-specific factors promoting accommodation towards CC were discussed in Chapter 3 on the basis of an analysis of the varieties in contact. Now, it is necessary for us to consider the variable-specific factors that may encourage, delay, or even prevent accommodation. We have analyzed the six individual linguistic variables in detail and it is now possible for us to make predictions that are based on the data presented so far. We can attempt to predict: (1) what factors are likely to increase the likelihood of accommodation (or what factors make a particular variant 'salient'); and (2) which forms of the host variety speakers are most likely to acquire. Although the

---

morphology) is not. However, in a later work (1973), Kučera appears to have retracted on some of his initial claims with respect to the combination of phonological and morphological items and lists forms where CC phonology + SC morphology as permissible, although, in his own words, 'unusual'. For a comprehensive description of these rules the reader is referred to Kučera (1955; 1958; 1973).

[109] The process by which speakers perpetuate what Trudgill calls 'hyperdialectisms'.

[110] Čermák (1987: 142) believes *é*-raising in the word *léto* does not take /iː/ because it would result in homonymic clash: *líto*, as in the phrase *je mi líto* 'I'm sorry'.

following list is by no means exhaustive, the following criteria have been selected with respect to their impact on the accommodation process:

1. Frequency

2. Areal distribution

3. Social acceptance

4. Speaker awareness

5. Complexity (including 'categoricity').

First, let us define the above set of criteria. The term 'frequency' seems to be ambiguous, insofar as it is sometimes used to denote the number of positions where a shift can take place, that is, in what positions /ɛː/ is raised to [iː] or /iː/ diphthongized to [ɛj], while it is used elsewhere more simply to signify the number of times a linguistic variable occurs in a given discourse, that is, how often (é) or (ý) is observed in positions that permit variability. Here it is used in this latter sense and the term 'categoricity' is used in the former. A variant's 'social acceptance' indicates to what extent the form has gained currency beyond informal communicative situations and in what sociolinguistic domains it is accepted (and tolerated). The geographical area over which the variants are used is referred to as the variant's 'areal distribution' and the extent to which speakers are aware of a given variant is termed 'speaker awareness'. The final factor labelled 'complexity' is a broader category that covers:

(1) the 'categoricity' of a given variant; and (2) the difficulty it poses to the acquirer. That is, whether the variant is used categorically or variably by native speakers and whether the variable has a binary or multivariant output or is in any other way phonetically complex. The five factors are closely linked and interrelated. This is very important not only with regard to the accommodation of speakers of Moravian dialects living in Bohemia, but also in view of the diffusion of CC forms into Moravia. We could argue that high-frequency CC forms that are categorical or semi-categorical and that are used over a large geographic area gradually become more tolerated by non-native speakers of CC. A higher level of tolerance in turn increases the social acceptance of the variants, which promotes their social and functional expansion into various semi-formal sociolinguistic domains. The result is that speakers of Moravian dialects are aware of the CC variants that have gained social acceptance, without necessarily engaging in face-to-face contact with native speakers of CC and these may prove to be the forms that are acquired the most.

Adhering to the traditional practice of grouping linguistic phenomena into predetermined categories, let us try to predict the 'probability of assimilation' for each of the CC variants under study. The following categories are posited:

A – The variant will be acquired in almost all situations

B – The variant will be acquired some of the time

C – The variant will be used only infrequently

D – It is very unlikely that the variant will be acquired

6.9.1 *V*-insertion

The variable (o) occurs frequently in spontaneous speech, with a recorded average of 74 tokens per twenty-minute interview (§ 7.1). Speakers of Moravian dialects have no problems in realizing the combination /v/ + /o/ in word-initial position, since this combination is present in SC words like *volant* 'steering wheel', *volby* 'elections' and *volavka* 'heron'; therefore, acquisition is not impeded by a phonetic constraint and *v*-insertion is a semi-simple rule. In terms of areal distribution, prothetic /v/ is found in dialects beyond CC-speaking territories in parts of central Moravia, but is reputedly on the decline under the influence of CC and other Moravian dialects. Despite these favourable circumstances, several factors hinder the adoption of prothetic /v/. First, although *v*-insertion is a semi-simple rule, the boundaries of where and where not /v/ can be inserted are highly variable and not clear-cut. Unlike the other CC phonological variants, the (non-)insertion of /v/ depends on a number of extralinguistic factors (§ 6.2.1) and besides regional and social variation there is also a significant amount of idiolectal variation involved. Speaker awareness is high and speakers' perception of prothetic /v/ is possibly the most important factor in terms of accommodation. All the three phonological forms are extremely salient and consequently high in speakers' consciousness, but whereas the functional range of non-standard /iː/ and /ɛj/ has expanded, resulting in the relaxing of attitudes towards these forms outside informal communication (more so in some positions than others),

the social status of prothetic /v/ appears to have dropped,[111] most probably in connection to the fact that this feature is rightly or wrongly considered to be on the decline within CC. Reactions from the informants in I2 also suggest that /v/, especially in lexical words, is among the CC forms that Moravians dislike the most. In sum, therefore, we might expect that due its high level of variability and its low level of social prestige /v/ will be acquired minimally (C), at least in lexical words. In grammatical words, where /v/ is categorical and perhaps has a higher level of social acceptance, we might expect acquisition to be much higher (A-B).

6.9.2 É-raising

Perhaps the most likely candidate for assimilation is CC raised /iː/. In terms of frequency, (é) occurs more often than any of the other variables (§ 7.1); raised /iː/ is encountered in all Central Moravian dialects, it is highly expansive, it is very common in various types of non-informal communication and informal writing and it is very high in non-native speakers' consciousness. In addition, é-raising is a semi-simple rule and there are no phonetic factors impeding the acquisition of CC /iː/ in any position. Therefore, it is a fair assumption that informants will use raised /iː/ more than any other CC feature. This hypothesis is also supported by informants' perceptions of /iː/ in I2. Not only did many of them tolerate /iː/ in adjectival desinences, some also preferred to use the CC form as an alternative to SC and their

---

[111] It appears that native speakers of CC do perceive prothetic /v/ as less appropriate outside informal communication than /ɑ/ or /ɔ/ and its use in public discourse seems to be receding. In her study of (semi-)formal television discourse, Hedin identified that of the phonological variants she studied, prothetic /v/ was observed the least (12%). In fact, Hedin describes the occurrence of prothetic /v/ in her data as 'remarkably low', and she concludes that this feature is not usual in public discourse (2005: 85).

native /ɛ:/, which they considered too 'official' or 'stilted' for everyday use. I suggest, therefore, that CC /i:/ will be acquired in almost all situations (A), although most probably to a lesser extent in the oblique cases of hard feminine adjectives and pronouns (B-C), where /i:/ is restricted both locally and socially.

6.9.3 Ý-diphthongization

As we have already seen (§ 6.2.3), an important distinction needs to be drawn between CC desinence-final /ɛj/ and /ɛj/ in other positions. Desinence-final /ɛj/ is one of the best known and most frequently referred to markers of Bohemian speech; it is expansive both socially and geographically and has gained currency in Moravia, either via the diffusion of CC forms beyond their original heartland or through natural independent evolution in Moravia, and it is also encountered outside informal sociolinguistic domains. Here we observe the opposite effect of extra-strong salience: while /v/ was consciously avoided, the functional range of desinence-final /ɛj/ has increased. In addition to these favourable factors, ý-diphthongization is a semi-simple rule and, on a phonetic level, there are no problems that should prevent informants from acquiring it. Several nouns end in the combination /ɛ/ + /j/ such as *kolej* 'university dorm' and *hokej* 'hockey' and *-ej* is the SC suffix for certain imperative forms *volej* 'ring', *počkej* 'wait', *dělej* 'get on with it', although in most traditional dialects in Moravia the regional suffix *-aj* is observed in these positions. Taking the above factors into consideration, desinence-final /ɛj/ is, therefore, a primary candidate for assimilation (A).

The situation with respect to ý-diphthongization in desinence-initial position of adjectives and in word roots is very different. Generally, (ý) in these positions

occurs less often than in desinence-final position (§ Tables 7.1 and 7.3). Root internally, as was mentioned earlier, the use or non-use of /ɛj/ may depend to an extent on the individual word and also on the communicative situation. In both of these positions, CC /ɛj/ is found only in CC-speaking territories and, unlike desinence-final /ɛj/, has not made inroads into central Moravia. Similarly, /ɛj/ has a lower level of social acceptance in non-informal domains, and native speakers of CC generally show a higher level of variation between SC /iː/ and CC /ɛj/ in desinence-final or word-internal position than they do in desinence-final position. Therefore, we might expect acquisition to be relatively low (C).

6.9.4 The grammatical forms

It is more difficult to predict the extent to which informants will adopt the grammatical variants. All three CC forms can be treated collectively, since all occur less frequently than the phonological variants, all three are restricted to CC-speaking territories and they are also much lower in speakers' consciousness[112] than the phonological forms. On the other hand, all the grammatical variants have a high level of social acceptance in CC-speaking territories, in particular *l*-truncation and gender neutralization, being used by speakers as standard outside informal communication. Complexity of acquisition is perhaps the only area where important differences between the grammatical variables need to be taken into account. The bare-stem *l*-truncated forms, for instance, cannot be used categorically and selecting the correct form is often more complicated in '-nout' verbs. There appears to be a number of language-internal constraints that can bar the bare-stem variant and this complexity might deter accommodation. The situation for gender neutralization is a relatively

---

[112] The only mention of the grammatical variants in I2 was a single description of paradigm unification.

straightforward choice between the CC and SC variant (regional forms are identical to SC); however, the elimination of a consonant cluster for gender neutralization may be perceived by speakers as more inappropriate than, say, the truncation of word-final /l/ or the replacing of *-í* by *-ej* for paradigm unification in fourth- and some third-conjugation verbs. Where a consonant mutation does takes place, structural differences between the SC and CC variants are relatively prominent and differences between, say, SC *čeští* and CC *český* 'Czech' (Nominative Plural) might be perceived differently from those between SC *řekl* and CC *řek* 'he said' or SC *dávají* and CC *dávaj* 'they give'. Whether or not informants will adopt the CC forms for paradigm unification is difficult to predict and the fact that third-person plural verb forms are undergoing change in Moravian dialects and SC could have a delaying influence on accommodation. This is the only variable for which informants' native forms, discounting the expansive *-í* forms, are different from the SC variants and this may also prove in some way significant. Besides these individual factors that appear to discourage accommodation there are also general considerations that are likely to have a negative impact. Chambers (2003: 57-58), for instance, argues that grammatical variables tend to mark differences in socioeconomic status or level of education more sharply than phonological variables. This means that informants may intentionally avoid adopting the non-standard CC grammatical forms, viewing them as more 'non-standard' than the CC phonological forms. In sum, all three grammatical variants are expected to be avoided or assimilated rarely (C-D), and by no means to the same extent as the phonological variants.

6.10 The impact of the social variables on accommodation

In the section above, the likelihood of acquisition was predicted in view of factors relating to the individual linguistic variables, and we were able to enumerate the linguistic forms which informants will potentially acquire the most. In this section, our attention turns to what effect a range of external social constraints can have on the accommodation process and to what extent factors related directly to the informant can be used to predict the level and intensity of accommodation. A wide range of social factors or a combination of them can influence migrants' levels of accommodation and their adoption of variants of the host variety. An important aspect of the present study is thus to investigate which constraints affect informants' linguistic behaviour the most. Correlations between the linguistic variables and the social parameters 'sex', 'region of origin', 'length of residence' and 'level of integration' are investigated in the present study. These four independent variables were selected because it is considered that any one of them is capable of influencing speakers' linguistic behaviour in a significant way. Importantly, the effects of these independent social variables on the assimilation of variants of the host variety can be quantified and statistically tested. In addition, the influences of speakers' 'attitudes' towards the host variety, their 'subject of study' and the way in which they were gathered for the study ('method of recruitment') are also taken into consideration.

6.10.1. Region of origin

As we have already established, linguists talk broadly of three major (inter)dialect groups within Moravia. Informants in this study were evenly stratified between the three (inter)dialect regions and it is considered that their region of origin will in some way influence their accommodation towards the host variety and that differences will

be observed between speakers of different varieties. L. Milroy (1987a: 113) comments that speakers' region of origin 'has not been an important variable in quantitative studies', probably since the majority of studies are directed towards the speech of individuals from a single dialect area or speech community. Payne (1980), however, did analyze the speech of informants from different dialect areas and she found that their assimilation of Philadelphia short /ɑ/ was influenced by their own or their parents' region of origin, although in most cases only subtly. To date, there is too little evidence from accommodation-based studies to suggest whether or not a speaker's region of origin is an influential factor in terms of accommodation. Nevertheless, we can hypothesize about the phenomena that may prove important. Potential differences in accommodation could arise due to geographical proximity between dialect groups, structural likeness between the source (base) and host variety and the prestige of the dialects in contact.

Let us examine informants' native dialects, beginning by examining lay speakers' perceptions of the individual dialects. Balhar (1995) argues that Silesian dialects are both the most distant from SC and the most stigmatized and that speakers of Silesian dialects are the most linguistically insecure. Conversely, East Moravian dialects are in many ways identical to the standard, especially in terms of adjectival and pronominal morphology, and this leads speakers to the assumption that their dialect is better or more standard than other dialects, including CC. Another important and perhaps more influential factor is the proximity between dialect areas and the amount of contact between them. Central Moravian dialects border CC-speaking territories, while East Moravian and Silesian dialects are more isolated from the influence of CC. We might assume that closer proximity between regions is likely to mean a greater amount of contact between the speakers from these regions, thus

speakers of CC are likely to come into more frequent contact with Central Moravians than with Silesians or East Moravians. Furthermore, Central Moravian dialects are closest to CC not only geographically, but also structurally. All three CC phonological variants are observed in parts of central Moravia, and since certain features that are identical to the CC forms are used here in some, but not all, the positions that were considered earlier, speakers may find it easier to acquire the CC forms in other (non-native) positions. *É*-raising can be used to illustrate this point. *É*-raising in Central Moravian dialects is used in the same positions as it is in CC, with the exception of the oblique cases of feminine adjectives and pronouns; therefore, in theory, speakers may find it easier to acquire /iː/ in this position, since raised /iː/ is a perfectly natural feature for them in other positions. The converse can be said for speakers of Silesian or East Moravian dialects, where *é*-raising is not part of the local dialects in any position. A final consideration is that of local loyalty. Chloupek (1974: 54) argues that East Moravians and Silesians are linguistically more conservative and more resistant to the influence of CC than Central Moravians, who, according to Chloupek – intuitive assertion – are indifferent to their local dialects and have a weaker linguistic consciousness. I would predict, therefore, that speakers of Central Moravian dialects will be the highest acquirers, since their dialects are geographically and structurally the closest to CC, their native dialects contain certain forms that are identical to CC features and that contact between speakers of CC and Central Moravians is likely to more intensive than with speakers from other parts of Moravia.

6.10.2. Sex

Many interesting findings have been identified in the variationist literature in view of sex-related differences in language use and we know that the linguistic behaviour of

men and women differs on a number of levels. Such sex-related differences range from men and women speaking separate dialects (Bradley 1998) to fine-grained differences in the distribution of linguistic variables. Based on data elicited in a wide range of sociolinguistic studies in a wide range of speech communities in rural and urban societies with different sociolinguistic profiles, sociolinguists have identified several important trends. Labov (2001: 261-319) analyzed a wide body of variationist studies and concluded that women: (1) in the case of stable sociolinguistic variables, 'show a lower rate of stigmatized variants and a higher rate of prestige variants than men' (2001: 266); (2) in view of linguistic change from above, 'adopt prestige forms at a higher rate than men' (2001: 274); and (3) with regard to linguistic change from below, 'use higher frequencies of innovative forms than men do' (2001: 292). Labov talks of a 'gender paradox', whereby women 'conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not' (2001: 93); or, as he states later, women 'deviate less than men from linguistic norms when the deviations are overtly proscribed, but more then men when the deviations are not proscribed' (2001: 367). Clearly then, women are more conservative than men in some situations, while they are more progressive in others.[113] Gal (1978: 2) also comments that women use more 'newer' or 'more advanced' forms than men. She backs this up with data from her study in Oberwart, a bilingual Hungarian-German village in eastern Austria where due to societal changes Hungarian is being supplanted by German in all communicative situations. Here young women are the leaders in language change and are more advanced than older people and young men. Gal also argues that women use more 'newly introduced

---

[113] Counterexamples to this well-established trend are rare and have been documented almost exclusively in societies where gender roles differ more markedly than in Western societies: for example, in certain Muslim-dominated and Arabic-speaking communities.

forms' than men, regardless of whether the forms are prestigious or non-prestigious, claiming that 'compared to men of the same social class' women use 'more of the *new* non-prestigious forms in casual speech, while moving further towards the prestige models in formal speech' (1978: 2). Labov's New York City data also back this up: in the raising of tense short /ɑ/ (eh) women use more of the innovative forms in casual speech than men do, but they are more conservative than men when reading words from reading lists (1972a: 301-302). Thus, women are often said to initiate linguistic change (L. Milroy 1987a: 112).

Although the trend for women to use more prestige variants than men has been identified in a number of studies, an adequate explanation or an explanation that is accepted by all or most linguists of what underlies this sex differentiation has yet to be formulated. Trudgill (1972: 182-183) hypothesizes that women are more self-conscious and that this might make them more linguistically aware and motivate them to use standardized norms, especially since women are more likely to be judged by the way they speak, while men are evaluated in view of other factors such as occupation and wealth. Trudgill also proposes that men might use fewer prestige variants, since a high use of vernacular norms is associated with 'masculinity' and the 'roughness and toughness' believed to be part of working-class life. This and other hypotheses (see, for example, Labov 1972a: 303-304) are rejected, however, by other linguists (Gal 1978: 2). The reader is referred to Chambers (2003: 143-153) for an in-depth review of linguists' opinions on the matter.

In view of sex-related differences in accommodation in contact situations between adult speakers of regionally different but mutually intelligible dialects, there is too little empirical evidence to suggest whether either men or women are more or less successful in acquiring a new dialect, and whether sex alone is an important

factor. Kerswill in his study of Stril migrants in Bergen, Norway (1994) found the differences in linguistic behaviour of his male and female informants to be negligible. A similar pattern was observed by Bortoni-Ricardo (1985) in Brazil. According to Labov's 'gender paradox', we might expect female informants to use more prestigious or standardized variants, while male informants might use more non-standard or vernacular forms. Here, however, we are faced with the problem of what constitutes a 'prestige' or a 'vernacular' form, since it is largely unknown how the migrants will perceive CC in relation to their own dialects. That is, we cannot say for certain whether CC variants will be regarded as prestige variants (evaluated higher than the informants' native forms). Alternatively, if we accept Gal's hypothesis that women are quicker to adopt innovative forms regardless of whether they are prestige variants or not, then we might expect the female informants to use more CC forms.

6.10.3. Length of residence

Without any background knowledge of the processes involved in acquiring a second dialect, it might naturally be assumed that the longer speakers live in a given community, the more they will accommodate to the variety used in that community. This seems perfectly logical. Nevertheless, existing studies do not support this assumption. Kerswill (1994: 74), for example, asserts that length of residence is only likely to be a relevant factor within the first few years of the accommodation process, citing as an example Nordenstam's (1979) study on the linguistic behaviour of Swedish women living in Bergen whose accommodation had stabilized after approximately five years. Data elicited in the HDP (1978) also show that length of residence does not play a major role in acquisition, especially after the first two years following migration, after which length of residence is 'overridden by other social

factors' (1978: 20). In fact, out of six social variables, length of residence had the least significant effect on informants' linguistic behaviour.[114] Therefore, according to the literature, length of residence will be important at the start of the accommodation process when migrants first arrive in the host community, but it will quickly be superseded by other external factors. It will be interesting to find out whether variants of the host variety are acquired intensively in the first year (or first couple of years) and to a lesser extent thereafter, whether variants are adopted gradually over a prolonged period, and whether accommodation has a 'cut-off point', after which informants stop assimilating features of the host variety. And it will be equally interesting to identify whether or not variants are acquired in a fixed order, as Trudgill (1986) has identified in a number of contact situations.

6.10.4. Network integration

The level to which a speaker is integrated into a community invariably has an important impact on his or her linguistic behaviour and analyzing the relationship between network integration and linguistic behaviour is one of the primary objectives of the present study. By quantifying speakers' network integration or participation in locally-based activities, sociolinguists have been able to demonstrate which individuals adhere the closest to localized vernacular norms and this is a tested method of predicting conservative linguistic behaviour and non-accommodation (§ 5.9). To date, the emphasis within the variationist paradigm has been to apply this methodology in studies that look to capture dialect or language maintenance in situations where speakers are under considerable social pressure to accommodate to a

---

[114] The social variables investigated, ranked here in descending order in terms of their impact on speakers' linguistic behaviour, are 'contact with Germans during leisure time', 'age at time of immigration', 'contact with Germans at work', 'professional training in the country of origin', 'level of education', and 'duration of stay'.

more standardized, supralocal or prestigious variety. On the other hand, scales of network integration or allegiance to certain criteria that are considered favourable in terms of divergence from stigmatized vernaculars have not been tested in terms of their ability to predict innovative language use. It is suggested here that the heaviest acquirers of CC forms will be those who score the most points on the integration index. In addition, I will try to identify which sub-variables are the most influential in shaping informants' linguistic behaviour.

6.10.5. Other factors

Besides the factors listed above, three other independent variables were taken into consideration. First, let us consider the relationship between speakers' attitudes towards the host variety and their accommodation towards that variety. It was established in Chapter 5 (§ 5.8) that although speakers' attitudes towards the host variety are probably linked in some way to shaping their speech habits, it is difficult to find a methodology than can meaningfully describe this link. This is compounded by data that have highlighted a substantial difference between speakers' perception of how they use certain linguistic phenomena and their actual language use elicited in sociolinguistic interviews. With respect to accommodation, there is no reliable evidence to suggest that speakers who strongly resent a particular variety will not accommodate to that variety, nor is there any evidence to suggest that a speaker who is keen to gain the social approval of his or her peers by making linguistic concessions, dropping stigmatized features of his or her native dialect or accommodating towards the host variety, will succeed in this endeavour. Nevertheless, informants were ranked according to their views towards CC and the relationship between these views and their actual accommodation compared. It is

predicted here that there will not be a strong positive correlation between informants' attitudes towards the host variety and their actual language use and that the impact of language attitudes on accommodation will be weaker than that of other social variables.

Differences in linguistic behaviour were also measured in view of students' subject of study and method of recruitment. Although the student community should not be taken as linguistically homogeneous, no major differences are expected in the accommodation of students from different schools and faculties, at least in the sense that we will not be able to predict whether an informant is a high or low acquirer based purely on his or her subject of study, except perhaps in the case of humanities students who were for this reason not included in the study (§ 5.1). The method by which informants were recruited creates a potentially interesting category, insomuch as some informants were enlisted during social events, while others replied to posters advertising the research, which perhaps indicates that they have a strong sense of 'Moravianness'. This could prove important – though, again, it is anticipated that method of recruitment is likely to be overridden by other extralinguistic factors.