APPENDIX A

CZECH

The Czech language is one of the West Slavic languages. It is spoken by slightly more than 10 million speakers, mostly in Czechia. In this section, we discuss properties of morphology and syntax of the language relevant to our work. For a more detailed discussion, see for example (Karlík et al. 1996; Petr 1987). Alas, there is no detailed grammar of Czech in English, but basic overviews can be found in (Fronek 1999; Harkins 1953; Janda and Townsend 2002; Naughton 2005; Short 1993).

For historical reasons, there are two variants of Czech: Official (Literary, Standard) Czech and Common (Colloquial) Czech. The official variant is a 19th-century resurrection of 16th-century Czech. Sometimes it is claimed, with some exaggeration, that it is the first foreign language Czechs learn. The differences are mainly in morphology and lexicon. The two variants are influencing each other, resulting in a significant amount of irregularity, especially in morphology.

form	lemma	gloss	category
měst-a	město	town	noun neut sg gen
			noun neut pl nom (voc)
			noun neut pl acc
tém-a	téma	theme	noun neut sg nom (voc)
			noun neut sg acc
žen-a	žena	woman	noun fem sg nom
pán-a	pán	man	noun masc-anim sg gen
			noun masc-anim sg acc
ostrov-a	ostrov	island	noun masc-inanim sg gen
předsed-a	předseda	president	noun masc-anim sg nom
viдě-l-a	vidět	see	verb past participle fem sg
			verb past participle neut pl
vidě-n-a			verb passive participle fem sg
			verb passive participle neut pl
vid-a			verb transgressive masc sg
dv-a	dv-a	two	numeral masc sg nom
			numeral masc sg acc

Table A.1: Homonymy of the *a* ending.

A.1 Morphology

Like other Slavic languages, Czech is a richly inflected language. The morphology is important in determining the grammatical functions of phrases. The inflectional morphemes are highly ambiguous, as Table A.1 shows. There are three genders: neuter, feminine and masculine. The masculine gender further distinguishes the subcategory of animacy. Sometimes, it is assumed that there are four genders: neuter (neut/n), feminine (fem/f), masc. animate (masc/m) and masc. inanimate (inam/i); we follow that practice. In addition to singular and plural, some dual number forms survive in body parts nouns and modifiers agreeing with them.[88] There are seven cases: nominative, genitive, dative, accusative, vocative, locative, instrumental. Only nouns, only in singular, and only about half of the paradigms have a special form for vocative, otherwise the vocative form is the same as nominative.

---

[88]In Common Czech, there is no dual. The colloquial plural forms are the same as the official dual forms. For example, official: *velkýma rukama* 'big$_{fem.dl.ins}$ hands$_{fem.dl.ins}$' vs. *velkými lžícemi* 'big$_{fem.pl.ins}$ spoons$_{fem.pl.ins}$' (there is no 'hands$_{fem.pl.ins}$' or spoons$_{fem.dl.ins}$'); colloquial: *velkejma rukama* 'big$_{fem.pl.ins}$ hands$_{fem.pl.ins}$' vs. *velkejma lžícema* 'big$_{fem.pl.ins}$ spoons$_{fem.pl.ins}$' (according to Oral2006, *ejma* ending is the most frequent accounting for 82% of 263 tokens, *ými* for 8% *ýma* for 10% and *ejmi* has no occurrences).

		N Monday	F song	F fly	M Jirka	M brother	I castle
nom.	sg.	pondělí	píseň	moucha	Jirka	bratr	hrad
gen.	sg.	pondělí	písně	mouchy	Jirky	bratra	hradu
dat.	sg.	pondělí	písni	mouše	Jirkovi	bratru/ovi	hradu
acc.	sg.	pondělí	píseň	mouchu	Jirku	bratra	hrad
voc.	sg.	pondělí	písni	moucho	Jirko	bratře	hrade
loc.	sg.	pondělí	písni	mouše	Jirkovi	bratru/ovi	hradu
ins.	sg.	pondělím	písní	mouchou	Jirkou	bratrem	hradem
nom.	pl.	pondělí	písně	mouchy	Jirkové	bratři/ové	hrady
gen.	pl.	pondělí	písní	much	Jirků	bratrů	hradů
dat.	pl.	pondělí	písním	mouchám	Jirkům	bratrům	hradům
acc.	pl.	pondělí	písně	mouchy	Jirky	bratry	hrady
voc.	pl.	pondělí	písně	mouchy	Jirkové	bratři	hrady
loc.	pl.	pondělích	písních	mouchách	Jircích*	bratřích*	hradech
ins.	pl.	pondělími*	písněmi*	mouchami*	Jirky*	bratry*	hrady*

* – Official Czech form

Table A.2: Examples of declined nouns.

A.1.1 Nouns

Traditionally, there are 13 basic noun paradigms distinguished – 4 neuter, 3 feminine, 4 animate and 2 inanimate; plus there are nouns with adjectival declension (other 2 paradigms). In addition, there many subparadigms, subsubparadigms. All of this involves a great amount of irregularity and variation. As an illustration, Table A.2 shows declension of few nouns.

A.1.2 Adjectives

Adjectives follow two paradigms: *hard* and *soft*. Both of them are highly ambiguous, filling the 60 (4 genders $\times$ (2 numbers $\times$ 7 cases + 1 dual form)) non-negated first grade categories with only 12 and 8 forms, respectively (10 and 8 in Common Czech). See Table A.3 for the hard paradigm and Table A.4 for the soft one.

Negation and comparison forms are expressed morphologically. Negation by the prefix *ne-*, comparative by the suffix *-(e)jší-* and superlative by adding the prefix *nej-* to the comparative. The comparative and superlative forms are declined as soft adjectives.

		Official Czech				Common Czech			
		M	I	N	F	M	I	N	F
nom.	sg.	mladý		mladé	mladá	mladej		mladý	mladá
gen.	sg.	mladého			mladé	mladýho			mladý
dat.	sg.	mladému			mladé	mladýmu			mladý
acc.	sg.	mladého	mladý	mladé	mladou	mladýho	mladej	mladý	mladou
voc.	sg.	mladý		mladé	mladá	mladej		mladý	mladá
loc.	sg.	mladém			mladé	mladým			mladý
ins.	sg.	mladým			mladou	mladým			mladou
nom.	pl.	mladí	mladé	mladá	mladé	mladý*			
gen.	pl.	mladých				mladých			
dat.	pl.	mladým				mladým			
acc.	pl.	mladé		mladá	mladé	mladý*			
voc.	pl.	mladí	mladý	mladá	mladé	mladý*			
loc.	pl.	mladých				mladých			
ins.	pl.	mladými				mladýma			
ins.	dl.	mladýma							

* – for neuter, and to some extent for feminine, can be also *mladé*

Table A.3: Hard adjectival paradigm.

		M	I	N	F
nom.	sg.	jarní			
gen.	sg.	jarního			jarní
dat.	sg.	jarnímu			jarní
acc.	sg.	jarního		jarní	
voc.	sg.	jarní			
loc.	sg.	jarním			jarní
ins.	sg.	jarním			jarní
nom.	pl.	jarní			
gen.	pl.	jarních			
dat.	pl.	jarním			
acc.	pl.	jarní			
voc.	pl.	jarní			
loc.	pl.	jarních			
ins.	pl.	jarními			
ins.	dl.	jarníma			

Table A.4: Soft adjectival paradigm.

A.1.3 Pronouns

Some pronouns have nominal declension, some have adjectival declension and some have their own (e.g., *já* 'I'). Some forms of personal pronouns are listed in Table 4.2.

A.1.4 Numerals

Only *jeden* '1', *dva* '2', *tři* '3', and *čtyři* '4' fully decline, all of them distinguishing case and *jeden* and *dva* also gender. The inflection of the other cardinal numerals is limited to distinguishing oblique and non-oblique forms. Numerals expressing hundreds and thousands have in certain categories a choice between an undeclined numeral form or a declined noun form (*sto dvaceti*, *sta dvaceti* '120.genitive'). Ordinal complex numerals have all parts in the ordinal form and fully declining (*dvacátý pátý* '25th')[89]. Similarly as in German, two-digit numerals may have an inverted one-word form (*pětadvacet* '25', lit: five-and-twenty, *pětadvacátý* '25th').

A.1.5 Verbs

As in all Slavic languages, verbs distinguish aspect – perfective and imperfective. Aspect is usually marked by prefixes, sometimes suffixes or by suppletion. Change of aspect is usually accompanied by a change, often subtle, in lexical meaning. For example, *psát* 'write$_{imp}$', *napsat* 'write$_{perf}$', *dopsat* 'finish writing$_{perf}$', *sepsat* 'write up$_{perf}$', *sepisovat* 'write up$_{imp}$', etc. For more information on Czech aspect see (Filip 1999).

There are three tenses – present, past and future. Present tense of imperfective verbs and future tense of perfective verbs is marked inflectionally, distinguishing number and person. Perfective verbs do not have a present tense. The conjugations of perfective future and imperfective present are the same; sometimes they are both called morphological present tense. Past tense and imperfective future is expressed periphrastically.[90] Sample conjugations are in Table A.5. In Common Czech, gender distinction in plural past participles is lost, all being pronounced as the official feminine plural form. Also Common Czech uses adjectives instead of passive participles. Modern Czech does not have a pluperfect or an aorist tense.

---

[89]Again, this is the case of the official language, complex numerals in Common Czech usually have only their tens and units in ordinal forms.

[90]Note however, that there is no auxiliary in 3rd person past tense. Although some linguists (Veselovská 1995), assume phonologically null auxiliary.

	'to be'	'lubricate$_{impf}$'	'say please$_{impf}$'	'do/make$_{impf}$'	'do/make$_{perf}$'
inf	být	mazat	prosit	dělat	udělat
present					
1.sg.	jsem	mažu	prosím	dělám	udělám
2.sg.	jsi	mažeš	prosíš	děláš	uděláš
3.sg.	je	maže	prosí	dělám	udělám
1.pl.	jsme	mažeme	prosíme	děláme	uděláme
2.pl.	jste	mažete	prosíte	děláte	uděláte
3.pl.	jsou	mažou	prosí	dělají	udělají
past prtcp					
M/I sg.	byl	mazal	prosil	dělal	udělal
F sg.	byla	mazala	prosila	dělala	udělala
N sg.	bylo	mazalo	prosilo	dělalo	udělalo
M pl.	byli	mazali	prosili	dělali	udělali
F/I pl.	byly	mazaly	prosily	dělaly	udělaly
N pl.	byla	mazala	prosila	dělala	udělala
pass prtcp					
M/I sg.	-	mazán	prosen	dělán	udělán
F sg.	-	mazána	prosena	dělána	udělána
N sg.	-	mazáno	proseno	děláno	uděláno
M pl.	-	mazáni	proseni	děláni	uděláni
F/I pl.	-	mazány	proseny	dělány	udělány
N pl.	-	mazána	prosena	dělána	udělána
imperative					
2.sg.	buď	maž	pros	dělej	udělej
1.pl.	buďme	mažme	proste	dělejme	udělejme
2.pl.	buďte	mažte	prosme	dělejte	udělejte

Table A.5: Sample Verbal Paradigms (Official Czech).

Five main conjugational types are recognized. Each class has several, quite similar, paradigms (6, 3, 2, 3, 1; 15 in total). Certain categories are expressed analytically; various forms of the verb *být* serve as the auxiliary. Some of the auxiliary forms are constant or inconstant clitics – see §4.3.4.

A.2   Syntax

A.2.1   Agreement

In Czech, there is agreement between subject and predicate and agreement within the NP. Below, we provide a basic overview; for a detailed description of Czech agreement see (Avgustinova et al. 1995).

A.2.1.1 Subject-predicate agreement

One can distinguish two types of agreement with subject:

- Subject – finite verb agreement.

  The finite verb agrees with the subject in person and number.

  (1) Střední Evropa je/*jsem/*jsou ve vzduchoprázdnu.
       Central Europe is$_{3sg}$/am/are$_{3pl}$ in vacuum

       'Central Europe is in vacuum.'   [syn6]

- Subject – participles/predicative adjectives agreement

  Predicative adjectives and participles in periphrastic constructions agree in number and gender
  with subject. In (2), the dropped 2nd person singular (and masculine since referring to *Oto*)
  subject agrees with the participle *byl* and adjective *zavřený* in number and gender. Similarly
  služba in (3) agrees with *povinná* in number and gender.

  (2) Oto,   za co   jsi   byl/*byla/*byli   zavřený/*zavřená/*zavření?
       Ota$_{m.sg}$, for what aux$_{2sg}$ was$_{m.sg}$/was$_{f.sg}$/was$_{m.pl}$ jailed$_{m.sg}$/jailed$_{f.sg}$/jailed$_{m.pl}$

       'Ota, what were you jailed for?'   [syn6]

  (3) Vojenská   služba   je ve Švédsku povinná.
       Military$_{fem.sg}$ service$_{fem.sg}$ is in Sweden obligatory$_{fem.sg}$

       'Military service is obligatory in Sweden.'   [syn6]

  Only Official Czech distinguishes gender for plural participles (see Table A.5). In spelling, there
  are three forms: *chrápali* [-lɪ] 'snored$_{m.pl}$', *chrápaly* [-lɪ] 'snored$_{f/i.pl}$', *chrápala* [-la] 'snored$_{n.pl}$'
  (note that *chrápali* and *chrápaly* have the same pronunciation). Common Czech uses the [-lɪ]
  form for all genders in plural (spelling is unclear). Plural adjectives pattern similarly (§A.1.2).

Non-nominative subjects   In case of non-nominative subjects (certain numeric expression (4a),[91]
(4b), etc.) and constructions that are traditionally analyzed as subject-less (4c or 4d), the predicate
is in 3rd person singular neuter form.

(4) a. Pět/Mnoho lodí   zmizelo.
        five/many   ships$_{fem.pl.gen}$ disappeared$_{n.sg}$

        'Five/Many ships disappeared'

---

[91]In similar phrases, the noun in genitive is traditionally seen as the head. Obviously we could also assume the
numeral to be the head. In such a case, it would be natural to assume the numeral is in the default form (neuter
singular).

b. Otevřít soubor je jednoduché.
open$_{inf}$ file$_{inam.sg}$ is$_{3sg}$ simple$_{n.sg}$

'To open a file is easy.'

c. Prší/Pršelo.
rains$_{3sg}$/rained$_{n.sg}$

'It is/was raining.'

d. Je mi příjemně.
is$_{3sg}$ me$_D$ fine$_{adverb}$

'I am feeling fine.'

Coordinated subjects   Agreement with coordinated subjects is rather complex. The gender of the predicate is the minimal gender of participants of coordination, computed under the following order: $m < \{i, f\} < n$. This covers also the trivial case when the gender of all participants is the same. However there is an exception: if all participants have neuter gender and at least one is in singular then the gender of the predicate is feminine. This complexity is absent in Common Czech because colloquial plural participles and adjectives do not distinguish gender. There is a similar hierarchy for determining person of subject with heterogenous persons. Under certain conditions (especially when the predicate precedes the subject, or the subject consists of abstract nouns), the predicate can agree only with the member of the coordinated subject it is closest – as (5c) and (5d) show.

(5)   a. *Two concrete nouns:*

Byl   jsem rád, že máma s   tátou byli/byly   v pořádku.
was$_{m.sg}$ aux$_{1sg}$ happy, that mom with dad   were$_{m.pl}$/were$_{f.pl}$ fine

'I was happy that mom and dad were fine.'   [syn6]

b. Hitler   a   Německo   už   měli   hotové plány na znovuzískání Horního
Hitler$_{m.sg}$ and Germany$_{n.sg}$ already had$_{m.pl}$ finished$_A$ plans$_A$ for reclaiming   Upper
Slezska ..
Silezia ..

'Hitler and Germany already had finished plans for reclaiming Upper Silesia ...'   [syn5]

c. *Two abstract nouns:*

Přesnost   a   srozumitelnost   je příznačná   / jsou příznačné   pro jeho
Accuracy$_{fem.sg}$ and comprehensibility$_{fem.sg}$ is typical$_{fem.sg}$ / are typical$_{fem.pl}$ for his
výklady.
explanations.

'Accuracy and comprehensibility are typical for his explanations.'   [Karlík et al. 1996]

d. *Verb preceding subject:*

Včera   přišla   / přišli   máma a   táta dumů brzo.
Yesterday came$_{fem.sg}$ / came$_{fem.pl}$ mom   and dad home early.
'Yesterday cama mom and dad home early.'

A.2.1.2   Agreement within the NP

So called *agreeing attributes* agree with the noun in gender, number and case. This includes

- normal adjective as *starý* 'old'. For example, in (3) the adjective *vojenská* 'military$_{fem.sg}$' agrees with the noun *služba* 'service$_{fem.sg}$'.

- possessive adjective as *otcův* 'father's'.[92]

- relative clauses. However the relative pronoun agrees with the modified noun only in gender and number; its case is dependent on its function in the relative clause. In Common Czech, relative clauses are often introduced by a universal nondeclined relative pronoun *co. jenž* 'that' is also often not declined.

- ordinal numerals

- possessive pronouns and various determiners

Note that there are some limited exceptions. For historical reasons, attributes modifying accusative or nominative pronouns like *nic* 'nothing' or *něco* iq'something' are in genitive as in (6a) In nominative or vocative, the gender can be feminine even when the noun is not, this gives the phrase an expressive flavor as in (6b).

(6) a. Nikdo   z   obou pánů nechtěl   říci   nic   konkrétního.
        Nobody from both men not-wanted say$_{inf}$ nothing$_{neut.sg.acc}$ concrete$_{neut.sg.gen}$
        'Neither gentleman wanted to say anything concrete'   [syn5]

     b. Kluku   líná!
        Boy$_{masc.sg.voc}$ lazy$_{fem.sg.voc}$
        'You lazy boy!'

---

[92]However, in the dialects of Southern and Western Bohemia, including my native dialect, the possessive adjectives do not decline. The form ending in *-ovo* (for masculine possessors) or *-ino* for feminine possessors is used regardless of case, number and gender of the possessed noun. In other dialects this form is used only for accusative singular However, the dialects of Southern and Western Bohemia also often use prenominal genitive to express possession instead, especially when the possessive adjective would involve a phonological change: *s Hanky kolem* 'with Hanka$_{gen}$ bike$_{i.sg.ins}$' for *s Hančino kolem* 'with Hanka's bike$_{i.sg.ins}$' for official *s Hančiným kolem$_{i.sg.ins}$* 'with Hanka's$_{i.sg.ins}$ bike$_{i.sg.ins}$'.

A.2.2 Numeral expressions

Numerals expressions with *jeden* '1', *dva* '2', *tři* '3', *čtyři*, '4', *oba* 'both' behave in a "normal" way: a numeral agrees with its noun in case; *jeden*, *dva* and *oba* also in gender. However, numerals *pět* and above in nominative or accusative positions are followed by nouns in genitive plural (see (4a)). Otherwise (other numerals or other cases), the noun is in the same case as the whole phrase.

A.2.3 Negation

Sentence negation in Czech is formed by the prefix *ne-* attached to the verb. As in the other Slavic languages, multiple negation is the rule, negative subject or object pronouns, adjectival pronouns and adverbs combine with negative verbs.

(7) Nikdy nikomu   nic   neslibuj.
     never nobody$_D$ nothing$_A$ not-promise$_{imper.2sg}$
     'Never promise anything to anybody.'