

Ambiguity of Prepositional Groups; Classification, Criteria and Method for Automatic Processing

Abstract

This paper is focused on the ambiguity of prepositional groups in Czech (Pg's). We specify two types of ambiguity relevant for these groups. A method of analysis based on reduction is introduced together with the proposal of the criteria suitable for the automatic processing. Finally, an evaluation of the criteria proposed is presented, with respect to the reliability of the analyses obtained.

1. Introduction

The problem of so-called “PP-attachment” is a notoriously well-known problem that has to be solved for any system of (linguistically based) machine translation or, in general, for any automatic parsing procedure as such. This problem seems to be interesting not only for the applied tasks, but also from the theoretical point of view. In the framework of a dependency based approach, the question of adverbial (a) or adnominal (b) dependency of the prepositional group (Pg, described in Section 2.) must be solved – the problem is often exemplified by the ambiguous sentence (1):

- (1) John saw a man *with a telescope*.
- (1a) John saw a man using a telescope. (= John had a telescope)
- (1b) John saw a man having a telescope. (= a man had a telescope)

It must be added that the same problem arises with nouns in non-prepositional cases, e.g. with the oblique case (dative) in (2), with two possible continuations:

- (2) Syn přinesl dárek *mamince*.
'The son brought a gift to (his) mother.' / 'The son brought a gift for (his) mother.'
- (2a) ... aby ho předala sestře.

- ‘... to deliver it (to) (his) sister.’
 (2b) ... a nechal ho u sestry.
 ‘...and left it with (his) sister.’

A Pg can also modify any adjective (c), deverbal as well as adjective not derived from verb. In example (3) the Pg *na ramínku*, ‘on (a) hanger’ is ambiguous, it can be treated both as a verbal and as an adjectival modifier:

- (3) Dívka rovná *na ramínku* vystavený kabát.
 ‘The girl arranges on a hanger displayed coat.’
 (3a) rovnat *na ramínku*, ‘to arrange on a hanger’
 (3c) vystavený *na ramínku*, ‘(st) displayed on a hanger’

In the following section we will demonstrate that the phenomenon just introduced (usually understood as a syntactic (or structural) ambiguity) is characteristic for Pg’s, but it is not the only type of ambiguity connected with Pg’s.

The following two points are relevant in this context.

1.1 The sources of Pg-ambiguity are different, they include (in addition to syntactic) also morphemic and phonemic facts (see Section 2.).

1.2 Though some pairs (or n-tuples) of sentences with these types of ambiguity exhibit the same content, having the same truth conditions (as in (4)), many of such ambiguous tuples differ in their truth conditions, thus, in general, the constructions have different linguistic meaning (see e.g. (5)). This is the main reason why the solution of this problem is crucial for any automatic analysis.

- (4) Stát neposkytne žádné slevy *na nákup bytů*.
 ‘The state will not provide any discounts for apartment purchase.’
 (4a) neposkytne (slevy) (na nákup bytů)
 ‘will not provide (discounts) (for purchase of apartments)’,
 i.e. adverbial dependency of Pg
 (4b) neposkytne (slevy (na nákup bytů))
 ‘will not provide (discounts (for purchase of apartments))’,
 i.e. adnominal dependency of Pg

(The simplified dependency structure is reflected by the brackets enclosing the dependents.)

- (5) Jednání o míru *na Blízkém Východě* splnila očekávání ...

‘The peace talks on the Middle East fulfilled the expectations.’
/‘The talks about peace on the Middle East fulfilled the expectations.’

(5a) na Blízkém Východě splnila, ‘fulfilled on (the) Middle East’

(5ba) jednání na Blízkém Východě, ‘talks on (the) Middle East’

(5bb) mír na Blízkém Východě, ‘peace on (the) Middle East’

Syntactically we have to deal with three different structures, which represents three different cognitive situations.

2. Types of ambiguity relevant for Pg’s

The linguistic framework of our research is constituted by the stratificational, dependency based Functional Generative Description of Czech (cf. SGALL ET AL. 1986) where four types of ambiguity are distinguished according to the levels of language description. This fact is important for the classification of ambiguity we have adopted. The ambiguity is understood there as an asymmetric relation between two adjacent levels of description (cf. PANEVOVÁ 1981). For our applied task of Pg-disambiguation, only two types of ambiguity are relevant. They are illustrated by example (6), where both the ambiguity of syntactic relations and the ambiguity connected with morphemic forms are present.

We use the term *prepositional group* as a parallel of the term *prepositional phrase* known from the phrase structure oriented grammars (we avoid using the latter term due to its connotations). In the simplest case, a Pg is constituted by a preposition and a noun in a relevant morphemic case (a “core” noun of the Pg). A more complex Pg may be simplified by the means of an analysis by reduction (which will be introduced later) in order to obtain the preceding case.

See the following example:

(6) Trenér Haber byl jako jediný schopen zabezpečit přípravu *na mistrovství světa*.

‘The coach Haber was the only one who was able to ensure the preparation for / on the World Championship.’

The Pg *na mistrovství (světa)*, ‘on World Championship’ modifies either the verb *zabezpečit*, ‘to ensure’ or the noun *příprava*, ‘preparation’, the sentence has two different syntactic structures.

In addition, the Pg *na mistrovství (světa)*, which means both ‘on World Championship’ and ‘for World Championship’, bears ambiguous morphemic information – the preposition *na* governs two morphemic cases, accusative and local, the (robustly ambiguous) word form *mistrovství* can be also analyzed as accusative or local. Thus the whole Pg *na mistrovství (světa)* can be treated in three ways:

(6a) as a verbal modifier (*zabezpečit na mistrovství (světa)*, ‘to ensure on World Championship’, free verbal modifier in local case, with local meaning – answering the question Where?) or

(6b) as a nominal modifier (*příprava na mistrovství (světa)*, ‘preparation on/for World Championship’, non-congruent attribute), namely

(ba) as a valency modifier of the noun *příprava* (Patient, with the form *na+Acc*)

(bb) as a free modifier of this noun, with local meaning (answering the question Where?, form *na+Loc*)

To resume the example: We have detected two types of ambiguity relevant for Pg’s according to the means “responsible” for the existence of two (or more) possible readings – ambiguity caused by general syntactic rules (2.1) and ambiguity originated in ambiguous morphemic information (2.2).

2.1 A structural (syntactic) ambiguity (or morphemic ambiguity, in the terms of PANEVOVÁ 1981) is an asymmetric relation between the level of the sentence structure and the adjacent morphemic level of description (i.e. the situation when two or more possible functions on the level of the syntactic representation of a sentence are expressed by the same morphemic form).

This type of ambiguity is characteristic for Pg’s. In principle, a Pg can depend on:

(a) any verb (autosemantic verb, modal verb as well as copula);

(b) any noun preceding the tested Pg in surface word order (prototypically any verb form is understood as the left-hand boundary, some atypical positions of Pg are mentioned below);

(c) any adjective (again, any verb form serves as a boundary in searching for a potential adjective governor of a Pg).

2.2 In addition, also morphemic and phonemic facts may be “responsible” for the ambiguity of Pg, as we have seen in example (6ba-6bb).

As some Czech prepositions govern two (exceptionally three) morphemic cases (as e.g. *na*, ‘on’ or *v*, ‘in’, both govern Accusative and Locative case, or *za* with Genitive, Accusative and Instrumental case) also phonemic ambiguity may appear within Pg’s. This type is connected with the concrete choice of lexemes and their morphemic characteristics.

We can summarise that the structural ambiguity is systematic for the sentences with Pg’s, it is determined by general syntactic rules. On the contrary, the occurrence of phonemic ambiguity is only accidental, inherent in the forms of concrete lexical items.

Whereas the potential structural ambiguity is resolved during syntactic analysis, the possibility of phonemic ambiguity is detected earlier, during morphological (pre)processing – the lexical item in consideration (a “core” noun in Pg, in our case) bears ambiguous morphemic information. In such cases, syntactic analysis splits into several branches in which all possible combinations of the input information are processed, i.e. each combination of possible morphemic information is treated separately. If two (or more) branches of syntactic analysis (with distinct Pg’s) are successful, then phonemic ambiguity is stated as relevant for the given sentence, i.e. as a source for its syntactic ambiguity.

This is the reason why we will focus on the structural ambiguity and on the proposal of criteria for its detection in the following sections. The evaluation of analyses obtained (concerning structural as well as phonemic ambiguity) is discussed in Section 5.

3. Method of analysis – analysis by reduction

Most of the reported approaches to disambiguation of Pg’s are statistical or corpus-based, and they usually try to solve a rather simplified problem: they consider only prepositions whose attachment is ambiguous between a preceding noun phrase and a verb phrase. Our approach is aimed at a formulation of linguistically-based rules for the detection of the respective sentence member (or members, in the case of ambiguity) on which Pg can depend. Such rules can be assigned with different priority (weight), which supports the evaluation of particular syntactic structures.

An analysis by reduction (RA) consists of a stepwise simplification of full sentence so that its syntactic correctness is preserved. In each step the simplification is realized by deleting one word of the

sentence and possibly rewriting other words. This process is non-deterministic, in each stage any of a set of mutually independent words can be deleted. The deletion is “justified” by linguistically based rules.

RA can be seen as a model coming close to human analysis¹. Then the condition on correctness preserving is OK. For an automatic procedure the criterion of correctness must be weakened – the fact, that the analysis is successful (that the sentence is reduced), “justifies” the particular branch of analysis.

The following example illustrates the analysis by reduction:

- (7) Tyto rodiny mají nárok *na státní vyrovnávací příspěvek*.
‘These families have a claim to a state compensatory allowance.’

We assume (for the sake of simplicity) that the congruent attributes *tyto*, ‘these’, *státní*, ‘state’ and *vyrovnávací*, ‘compensatory’ have been reduced in the first stages of an analysis (everyone is independent of each other). Thus, we obtain the following sentence (where only the phenomena important for the explanation are kept):

- (7’) Rodiny mají nárok *na příspěvek*
‘Families have (a) claim to (an) allowance.’

The analysis splits into several branches. The reduction of the Pg *na příspěvek*, ‘to allowance’ starts in several stages, with respect to the types of information that can be taken into account.

(a) If the verbonominal collocation *mít nárok*, ‘to have (a) claim’ is taken into account, the reduction of the Pg *na příspěvek*, ‘to allowance’ starts in two stages (marked by bold arrows, see Figure 1.). The dependency pair *nárok na příspěvek*, ‘claim to allowance’ is determined (in agreement with valency requirement of the noun *nárok*). This analysis has high preference (see evaluation of the criteria, Section 5.).

(b) If the verbonominal collocations are not considered, the analysis *mít na příspěvek* is also allowed. The reduction of Pg then starts in further two stages (marked by dotted arrows), the Pg is treated as a free verbal modifier, with low preference.

All these analyses satisfy the condition on preserving syntactic correctness, both are successful. Accidentally, the second structure *mít na*

1 Let us stress here that we do not aim at creating a real psychological model simulating human understanding. We describe only the possible view on a step-by-step reduction of a sentence which is used e.g. in grammar school.

příspěvek has another meaning than the original sentence has. Its meaning is connected with a frozen collocation *mít na něco (prostředky/peníze/...)*, ‘to have the wherewithals for st’. (The source of this coincidence is connected with the lexical analysis, it cannot be covered by general syntactic rules.) However, this analysis is marked as less probable.

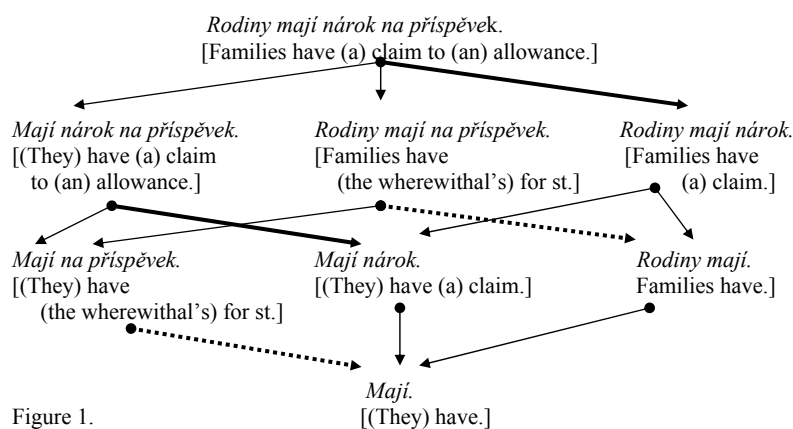


Figure 1.

Naturally, the more information can be used, the better analysis is obtained. However, despite of the incompleteness of the input information the automatic procedure can grant a satisfactory answer to our problem – the structures obtained are evaluated according to their reliability. This is very important with respect to the fact that the input information is always incomplete, the knowledge involved in natural language understanding (knowledge available for human) cannot be incorporated in its complexity in any automatic system.

4. Proposal of the criteria

We have proposed four types of criteria that seem to be relevant for the Pg-disambiguation: 1. Criteria based on surface word order, 2. criteria exploiting valency frames of verbs, nouns and adjectives, 3. criteria concerning the word order positions as combined with the relationship of a particular modifier to other modifiers; optionally, 4. rules based on semantic features can be used (cf. STRAŇÁKOVÁ 1999).

4.1 The position of a Pg in a sentence is the basic purely syntactic clue to the detection of its potential ambiguity. We have already specified the prototypical position of a Pg, which has been embodied in word order patterns.

4.1.1 The Pg-ambiguity is typically connected with special word order patterns, we can call them basic “suspicious” WOP’s.

(a) a sequence of nouns at the beginning of a clause immediately followed by Pg (“NNPg type”);

(b) a sequence of a verb, a noun and Pg (in this order, “VNPg type”);

(c) Pg between a noun and a verb (from the left to the right, “NPgV type”);

(d) Pg between a verb and an adjective (“VPgA type”);

(e) Pg being followed by an adjective and a verb (in this order, “PgAV type”).

There can be strings of nouns instead of a single N in all of these word order patterns, all of them either without a preposition or in a prepositional case.²

The first part of the sentence (5) satisfies type (a), NNPg, with two possible nominal governors of the Pg *na Blízkém Východě* (see (5')). The type (b), VNPg, is present e.g. in examples (4), (6) and (7). The particular Pg’s can be analyzed either as a verbal or as a nominal modifier there. In the sentence (3) the (d) type, VPgA, is met, again with two Pg-governors, a verbal and an adjectival one. The following examples (8) and (9) illustrate the remaining types:

(5') *Jednání o míru na Blízkém Východě.*
‘The peace talks on the Middle East.’ / ‘The talks about peace on the Middle East.’

(8) *Japonský lék na kanadské problémy nepomohl.*
‘The Japanese medicine failed with Canadian problems.’ / ‘The Japanese medicine for Canadian problems met with failure.’

(9) *Na severu pracující Kurdové vyhlásili stávkou.*
‘The Kurds working on the north declared a strike.’ / ‘The working Kurds declared a strike on the north.’

2 In addition, also basic “unambiguous” WOP’s and “WOP’s with a preferable reading” are stated, serving for Pg-disambiguation which are not commented here. (A WOP with preferable reading is used in example (13).)

In (8) the “suspicious” WOP, type (c), NPgV, is present. The sentence actually exhibits a syntactic ambiguity – the tested Pg can be treated either as a non-congruent attribute of the noun *lék*, ‘medicine’ or as a free verbal modifier of the verb *pomoci*, ‘to help’ (see also Subsection 4.2.2)

Sentence (9) satisfies WOP type PgAV, again with two readings:

- (9a) vyhlásit na severu, ‘to declare on (the) north’
- (9c) pracující na severu, ‘working on (the) north’

Both in (8) and (9) the sentences have different truth conditions.

4.1.2 Prototypically, Czech constructions with Pg’s meet the constraints defined above. However, there are some constructions in Czech where these conditions must be relaxed (for the sake of the adequacy of the resulting analysis).

Verbonominal collocations belong to constructions where it seems to be adequate to relax the word order constraints.³ The fact that such a Pg is treated as a nominal modifier (in agreement with the valency requirements, see Subsection 4.2.2) may lead to a non-projective structure (which seems to have the same degree of reliability as a prototypical structure with word order constraints satisfied).

In a dependency formalism, the non-projective construction can be characterised by the following subtree (Figure 2):

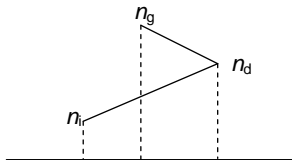


Figure 2.

- (10) *Na poslední místo v týmu měl největší šanci některý z mladíků.*
‘One of the youngsters had the greatest chance to get the last position on a team.’
mít, ‘to have’ ... Act Pat (Acc) (Orig (*od*+Gen / *z*+Gen))
šance, ‘chance’ ... Pat (*na*+Acc)

3 For our purposes, a verbonominal collocation can be described as a frozen collocation of a noun and a verb (often with very general meaning) where the noun part has (usually) a valency position filled in by a Pg.

The noun *šance*, ‘chance’ constitutes a verbonominal collocation together with the verb *mít*, ‘to have’ – it has a valency position asking for *na+Acc* modifier; the Pg *na (poslední) místo (v týmu)*, ‘on (the last) position (in the team)’ meets this requirement creating a non-projective construction (see Figure 3).

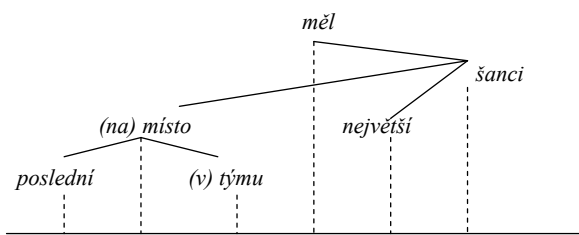


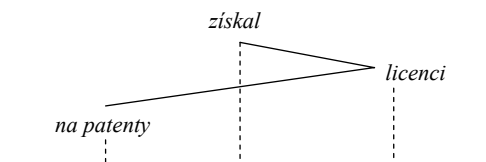
Figure 3.

Unfortunately, the field of verbonominal collocations has not been fully treated in a way available for automatic processing yet. However, such relaxation of word order constraints can be generalised for all nominal modifiers which satisfy valency requirements typical for verbs⁴, with the only additional condition that in the given sentence the tested Pg cannot be treated as a verbal valency modifier.

- (11) *Na některé z jejich patentů získal licenci.*
 ‘They obtained a license for some of their patents.’

licence, ‘license’ ... Pat (*na+Acc*)
získat, ‘to obtain’ ... Act Pat (*Acc*) (*Orig(od+Gen / z+Gen)*)

The noun *licence*, ‘license’ has a valency position for Patient, with the surface form *na+Acc*. The verb *získat*, ‘to obtain’ has no valency requirement which could be fulfilled by the tested Pg *na některé z jejich patentů*, ‘for some of their patents’. (The Pg can be simplified by the means of RA to its “core” – *na patenty*, ‘for patents’, *na+Acc*.)



⁴ We do not consider special adnominal modifiers here, such as Partitive or Identity (cf. PÍŤHA 1981).

Figure 4.

Again, it seems to be reasonable (in agreement with intuition) to permit a non-projective dependency pair *licence na patenty*, ‘licence for patents’ (Figure 4).

Another candidate for a relaxation of word order constraints is a non-projective Pg-modifier of an adjective, namely Pg-modifier which is separated from its adjective governor by a noun modified by this adjective, as in Figure 5 (“ $A_i N_i P_g$ type”, the i -indexies indicate congruency between noun and its adjectival modifier).

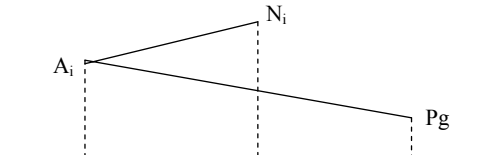


Figure 5.

- (3') *Dívka rovná vystavený kabát na ramínku.*
 ‘The girl arranges a coat displayed on a hanger.’
 (3a) *rovnat na ramínku*, ‘to arrange on a hanger’
 (3b) *kabát na ramínku*, ‘coat on a hanger’
 (3c) *vystavený na ramínku*, ‘st. displayed on a hanger’

Though the (c) case is stylistically not recommended, such formulations are frequently used. Their negative style markedness is reflected in lower preference of (c) dependency pair. (See also example (13) below.)

We have adapted the strategy of “careful” relaxation of the word order constraints, i.e. the relaxation is allowed only with well described phenomena in order to reach a linguistically appropriate solution. Each inadequate relaxation leads to an explosion of analyses obtained by parsing procedure, most of them inappropriate.

4.2 The valency frames of particular lexical items play a crucial role in the Pg-disambiguation. There is no doubt of the importance of valency information of verbs and nouns; we will discuss here also the usefulness of adjectival valency frames.

4.2.1 The notion of valency primarily pertains to the level of underlying representation of a sentence. However, the valency frames can be fruitfully interpreted also in what concerns the means of their expression in the surface structure. One feature of such interpretation is important from the point of view of automatic processing. In the surface (morphic) form

of the sentence (almost) any member of valency frame is deletable (at least in the specific contexts as e.g. a question-answer pair). Thus, the knowledge of the valency frame of a particular item can be used in ‘one direction’ only. If a member (e.g. a Pg) satisfying some valency requirement is present in the sentence, then it can be treated as such a valency modifier. Its absence, on the contrary, does not mean that the sentence is incorrect.

Originally, the valency theory was established for *v e r b s*. The valency frame of particular lexical item (in the narrower sense of “frame”)⁵ consists of the inner participants (i.e. Actor, Patient, Addressee, Origin and Effect), either obligatory or optional – with the list of one or more appropriate morphemic forms – and of the obligatory free modifiers (as local, temporal, manner, causal etc). Marked obligatory members of a valency frame can be omitted on the surface level. Each Czech verb, depending on the number of its valency frames, is represented by one or more lexical items in the lexicon (cf. PANEVOVÁ 1980).

4.2.2 Later valency theory was extended to *n o u n s* (cf. PŘÍHA 1981, PANEVOVÁ 2000). The following example shows a derivation of a deverbal noun and its frame, which is, in principle, inherited from the source verb.

- (12) Roční náklady *na provoz speciální školy J. A. Komenského* činí 6 miliónů korun.
‘Annual running expenses of special school of J. A. Komenský amount to 6 millions crowns’.

According to the “suspicious” WOP, type NPgV, which is met in the sentence, Pg potentially depends either on the noun *náklady*, ‘expenses’ or on the verb *činit*, ‘to amount’. However, the whole sentence is not ambiguous, and the valency requirement of noun allows to establish the most probable analysis.⁶

The Czech noun *náklad*, ‘load/expense/ printing’ is derived from the verb *nakládat*, ‘to load/to conserve/to treat’ which corresponds to three lexical units with their frames:

*nakládat*₁ ... Act Pat (Acc) Direction

5 Most of free modifiers are optional and belong only to a “valency frame” in a broader sense.

6 We will return to this problem and show why it is necessary to preserve all dependencies based on WOP.

The verb *nakládat*₁ is used e.g. in the sentence *Petr nakládá písek.Pat na auto.Dir*, ‘Peter loads a lorry with sand’. (In the sentence *Petr nakládá písek.Pat*, ‘Peter loads (a container) with sand’ the Direction position is fulfilled by the general modification, expressed as a ‘zero lexeme’ in the surface representation.)

This lexical unit *nakládat*₁ is often accompanied by a free modifier of Means (in instrumental case), *Petr nakládá písek lopatou*, ‘Peter loads up (a container) with sand using a shovel’.

The verb *nakládat*₁ may also be seen in the idiomatic usage *Petr nakládá okurky do octa*, ‘Peter – conserves – gherkins – to – pickle’, ‘Peter pickles gherkins’.

*nakládat*₂ ... Act Pat (Acc)

The sentence *Petr nakládá auto*, ‘Peter loads up a lorry’ illustrates this meaning. The lexical unit *nakládat*₂ is characterized by Patient prototypically expressed by a noun with the semantic feature “container”.

The verb is often accompanied by a free modifier of Means in its broad sense (in instrumental case), see *Petr nakládá auto lopatou*, ‘Peter loads up a lorry with a shovel’, but also *Petr nakládá auto pískem*, ‘Peter loads up a lorry with sand’.

*nakládat*₃ ... Act Pat (s+Ins) Manner

The sentence *Petr špatně nakládá s manželkou*, ‘Peter treats his wife badly’ shows this possibility’.

The noun *náklad* is primarily derived from the verb *nakládat*₁ or *nakládat*₂.

*náklad*_i ... Act Pat Direction

It can be illustrated by *náklad písku (na voze)*, ‘load of sand (on lorry)’, actor is not usually expressed.

*náklad*_j ... Act Pat

This meaning is exemplified by *náklad auta (pískem)*, ‘load of a lorry’ often with optional Means, again actor is not usually expressed.

Secondarily it can be used (usually in plural) with the meaning ‘expenses’, with Patient incorporated⁷. (This meaning has no counterpart in the lexical paradigm of the lexeme *nakládat*.)

*náklad*_k ... Act Eff (na+Acc)

⁷ The notion of incorporated inner participant is introduced e.g. in PANEVOVÁ 2000.

As in *něčí/někoho náklady na živobytí*, ‘somebody’s costs of living / the living costs of somebody’; the incorporated patient stands for ‘money, sum, energy, ...’.

Further, the noun *náklad* is used for ‘printing’:

*náklad*₁ ... Ø

As in *náklad knihy*, ‘printing of book’ where the modifier in genitive case must be classified as partitive, a special nominal modifier. Due to the abstraction, the verbal modifiers disappear with *náklad*₁.

In the examined sentence (12) the noun *náklad*_k must be identified. Its valency requirement for Effect (surface form *na+Acc*) is met, thus the dependency pair *náklady na provoz* is established.

A problem arose concerning preservation of dependencies based on WOP’s. Let us return to sentence (8) – we can demonstrate on this sentence that it is necessary to preserve all such readings of a sentence (as they can have different truth conditions).

- (8) Japonský lék *na kanadské problémy* nepomohl.
‘The Japanese medicine failed with Canadian problems.’ / ‘The Japanese medicine for Canadian problems met with failure.’

The noun *lék*, ‘medicine’, a primary noun, has a valency frame consisting of optional Patient modifier:

lék ... Pat (*na+Acc/proti+Dat*)⁸

As the Pg *na kanadské problémy*, ‘on Canadian problems’ fits into the valency frame of the noun *lék*, ‘medicine’ (Patient) it is treated as its nominal modifier.

As we stated above, the sentence is ambiguous – the tested Pg can be also treated as a free verbal modifier (Aim) of the verb *pomoci*, ‘to help’. Both structures will be established, the former with high preference (as satisfying

8 Compare this with the valency frames of the verb derived from this noun:

*léčit*₁, ‘to cure’ ... Act Pat (Acc)

(as in *lékař léčí angínu*, ‘a doctor cures an angina’)

*léčit*₂, ‘to cure’ ... Act Pat (Acc)

(as in *lékař léčí chlapce*, ‘a doctor cures a boy’).

The division into two units is motivated by the different prototypical requirements for semantic features of potential participants. It is supported also by the original noun *lék*, ‘medicine’ which requirements are identical with *léčit*₁.

valency requirements), and the latter with low preference (as a free modifier).

The fact that both structures are established is important because they have different truth conditions (and hence different meaning), as the possible continuations show:

- (8a) Japonský lék *na kanadské problémy* nepomohl (Kanadě, pomohl však na problémy Británie.)
'The Japanese medicine did not help Canada with Canadian problems, but it helped with the problems of Britain.'
- (8b) Japonský lék *na kanadské problémy nepomohl* (Británii, přestože Kanadě pomohl.)
'The Japanese medicine for Canadian problems did not help Britain, though it helped Canada.'

In (8a) the analysis (*ne*)*pomoci na (kanadské) problémy*, '(not) to help with Canadian problems' is adequate. On the other hand, in (8b) the different analysis *lék na (kanadské) problémy*, 'medicine for Canadian problems' seems to be proper.

The difference between the two analyses obtained seems to be clearer in the sentences (8a') and (8b') which have the same syntactic structures as (8a) and (8b) have:

- (8a') Lék *na kašel* nepomohl, ale teplotu srazil.
'The medicine did not help a cough, but the temperature lowered.'
- (8b') Lék *na kašel* nepomohl, spíš mu pomohlo, že zůstal v posteli.
'The medicine for a cough did not help, rather it helped that he stayed in bed.'

The preceding examples show that the problem of the derivation of nominal frames is rather complex, many different linguistic phenomena must be taken into account.

4.2.3 A *deverbal adjective* shares its valency frame with the original verb, with two regular differences (cf. PANEVOVÁ 1998):

- (i) None of the valency frame members is obligatorily expressed on the surface level, each of them can be omitted.
- (ii) One of the presupposed valency 'slots' is filled in by the word that is modified by the examined adjective (i.e. by the governor of the adjective).

A question emerges concerning the role of adjectival frames for the detection of Pg-ambiguity. First, we illustrate the relation between a verb and an adjective derived from it, then we focus on word order constraints.

(13) *Neschopnost způsobila skluz v realizaci omezené autonomie na území Golanských výšin.*

‘The incapability caused a delay in the realization of limited autonomy on the territory of the Golany Highs.’ / ‘The incapability caused a delay in the realization of an autonomy limited to the territory of the Golany Highs.’

We will focus on the Pg *na území (Golanských výšin)*, ‘on territory (of Golany Highs)’ in this example (reduced Pg *na území*, ‘on territory’).

1. Derivation of adjectival valency frame:

The adjective *omezený*, ‘limited’ is derived from the verb *omezit*, ‘to limit’:

omezit ... Act Pat (Acc) (Eff (na+Acc))

The verb *omezit* is often accompanied by free modifier of Means (in Instrumental case) and/or by free modifier of Regard (surface form *v+Loc*).

Deriving the adjective *omezený*, ‘limited’ the following valency frame is obtained (with “filled” Patient position):

omezený ... Act (Ins) Eff (na+Acc)

This derivation can be exemplified by the following examples:

Petr.Act omezuje kouření.Pat žvýkačkami.Means na minimum.Eff

‘Peter.Act limits (his) smoking.Pat to minimum.Eff with chewing gums.Means’

→ *kouření omezené Petrem.Act / žvýkačkami.Means na minimum.Eff*

‘smoking limited by Peter.Act / with chewing gums.Means to minimum.Eff’

Petr.Act omezuje Pavla.Pat v kouření.Regard na minimum.Eff žvýkačkami.Means

‘Peter.Act limits Paul.Pat in smoking.Regard to minimum.Eff with chewing gums.Means’

→ *Pavel omezený v kouření.Regard Petrem.Act / žvýkačkami.Means na minimum.Eff*

‘Paul limited in smoking.Regard by Peter.Act / with chewing gums.Means to minimum.Eff’)

2. Word order constraints:

Two “suspicious” WOP’s (specified in Subsection 4.1.1) are met in the sentence (13) – “VAPg type” and “VNPg type” (the possible WOP’s are selected non-deterministically):

The first one, “type VAPg”, is supported by the valency requirement of the adjective *omezený*, ‘limited’. In this case, the non-projective dependency pair is established *omezený na území*, ‘limited to the territory’ (with surface form *na+Acc*), which is possible, but has low preference, see Subsection 4.1.2.

On the contrary, four dependency pairs based on the second WOP, “type VNPg”, are more probable, all of them have the same degree of preference, the preference of free modifiers: *způsobit na území*, ‘caused on the territory’, *skluz na území*, ‘delay on the territory’, *realizace na území*, ‘realization on the territory’ and *autonomie na území*, ‘autonomy on the territory’ (with surface form *na+Loc*).

3. Summary:

The syntactic as well as phonemic ambiguity is detected in the sentence (13), five structures are obtained (and evaluated). Concerning deverbal adjectives – the word order position of a Pg seems to be more important than the valency requirements of adjective in consideration.

The same observations hold for free modifiers of deverbal adjectives (as in (3')) as well as for adjectives not derived from verb, primary adjectives.

5. Evaluation of the criteria

In the process of syntactic analysis of a particular preposition group, several dependency pairs may be obtained, more or less probable (with respect to the meaning of the sentence analyzed). We assume that the “reliability” of syntactic structures with prepositional groups is in correspondence with the type of linguistic phenomenon, which leads to their establishment.

We propose the evaluation of the dependency pairs according to the types of criteria the dependencies are based on⁹. The observations leading to this arrangement were briefly sketched in the examples in Section 4.

⁹ We abstract here from the “formal” criteria that have not been commented in this article.

Evaluation:

1. If a so called “WOP with preferable reading for adjective” is met in a sentence, then the highest preference is assigned to a structure with a Pg analyzed as an adjectival modifier, valency one as well as free one.
2. The structures satisfying valency requirements follow, in this order:
 - (i) a Pg as a valency modifier of a verb;
 - (ii) a Pg as a valency modifier of an adjective – if the word order constraints are met;
 - (iii) a Pg as a valency modifier of a noun – the possibility of the relaxation of word order constraints depends on valency requirements of a verb.
3. The dependency pairs based on “unambiguous” and “suspicious” WOP’s follow:
 - (i) rules using semantic features (if they are considered) may contribute for the evaluation of free modifiers;
 - (ii) otherwise all structures based on “unambiguous” and “suspicious” WOP’s have the same degree of preference.
4. The lowest preference is assigned to a non-projective structure with a Pg treated as a modifier of an adjective separated from it by the nominal governor of the whole nominal group.

This evaluation is proposed for all structures with Pg under consideration, independently of their morphemic characteristics. All dependency pairs obtained in all possible branches of RA (i.e. structures belonging to both types of ambiguity relevant for Pg’s, structural (syntactic) as well as phonological) are compared.

The evaluation of particular dependency pairs obtained (based on the type of criteria leading to their establishment) allows to order output analyses. Such ordering enables to introduce a boundary between acceptable and non-acceptable analyses of the prepositional group tested. This boundary is present in a particular sentence – considering e.g. the ambiguity of (8) versus unambiguity of (12)¹⁰, any generalization is disputable, it seems not to be sound enough.

10 In both cases, the concurrence of valency modifier and free modifier is detected. In (8) two analyses reflected in two syntactic structures are adequate (*lék na problémy*, ‘medicine for problems’, valency Aim, and *pomoci na problémy*, ‘to help a problems’, free Aim). On the other hand, in (12) only a valency requirement is relevant for the successful analysis *náklady na provoz*, ‘expenses on running, running expenses’.

The possibility of evaluation of syntactic structures obtained is very important from the point of view of our task – the possibility of the detection of Pg ambiguity. As we have already stated, it allows not to rely on the completeness of input information that, in general, cannot be granted within any automatic parsing system.

6. Conclusions

We have discussed here some linguistic phenomena that we understand as sources of the ambiguity of Pg's. Such phenomena, "responsible" for Pg-ambiguity, belong both to morphology and to syntax. We have introduced and exemplified the notion of analysis by reduction, which can naturally underlie an automatic processing of natural language. Two types of criteria aiming at the adequate solution of Pg's have been commented – syntactic rules based on the surface position of a Pg and criteria exploiting valency frames of particular lexical items. Their evaluation has been offered with respect to the reliability of the syntactic structures obtained.

Let us conclude that the surface position of a prepositional group is crucial for its potential ambiguity. For the detection of an adjectival governor of a particular Pg the surface word order is more important than valency requirements of the adjective. On the other hand, there are no constraints on the position of a verbal modifier expressed by a Pg, thus word order plays no role in this case. The valency requirement of a particular noun for a Pg-modifier is relevant despite the word order, if such requirement is not "blocked" by the same requirement of a verb.

Acknowledgement

The research for this article was supported by the project of the Czech Ministry of Education VS 96/151.

References

- ALLEN, J. 1987: *Natural Language Understanding*. The Benjamin/Cummings Publishing Company
- HAIČOVÁ, E.; PANEVOVÁ, J. 1984: Valency (Case) Frames of Verbs. In: Sgall, P. (ed.) *Contributions to Functional Syntax, Semantics, and*

- Language Comprehension*), Prague, Academia and Amsterdam, John Benjamins, 147-188
- HOLAN, T.; KUBOŇ, V.; OLIVA, K.; PLÁTEK, M. 2000: *On Complexity of Word Order*. ÚFAL Technical Report TR-2000-08, Prague
- KUBOŇ, V., HOLAN, T.; PLÁTEK, M. 1997: *A Grammar-Checker for Czech*. ÚFAL Technical Report TR-1997-02, Prague
- MEL'ČUK, I. 1988: *Dependency Syntax: Theory and Practise*. NY:SUNY Press, Albany
- PANEVOVÁ, J. 1980: *Formy a funkce ve stavbě české věty*. Academia, Praha
- PANEVOVÁ, J. 1981: K otázkám homonymie a neutralizace ve stavbě věty. In: *Jazykovedné štúdie* 16, Bratislava, 85-89
- PANEVOVÁ, J. 1998: Ještě k teorii valence. In: *Slovo a slovesnost* 59, Praha, 1-14
- PANEVOVÁ, J. 2000: Poznámky k valenci podstatných jmen. In: Hladká, Z.; Karlík, P. (eds.) *Čeština - univerzália a specifika 2. Sborník konference ve Šlapanicích u Brna, 17.-19.11.1999*. Brno, 173-180
- PÍŤHA, P. 1981: On the Case Frames of Nouns. In: *Prague Studies in Mathematical Linguistics* 7, Academia, Prague, 215-224
- RATNAPARKHI, A. 1998: Statistical Models for Unsupervised Prepositional Phrase Attachment. In: *COLING/ACL'1998, Proceedings*, 1079-1085
- SGALL, P.; HAJIČOVÁ, E.; PANEVOVÁ, J. 1986: *The meaning of the sentence in its semantic and pragmatic aspects*. Ed. by J. Mey, Dordrecht:Reidel and Prague:Academia
- STRAŇÁKOVÁ, M. 1999: Selected Types of Pg-ambiguity. In: *The Prague Bulletin of Mathematical Linguistics* 72, 29-58
- ŠMILAUER, VL. 1966: *Novočeská skladba* (2nd edition). SPN, Praha
- UHLÍŘOVÁ, L. 1972: On the Non-Projective Constructions in Czech. In: *Prague Studies in Mathematical Linguistics* 3, Academia, Prague, 171-181