

Edited by

Joseph Emonds
and Markéta Janebová

Language Use and Linguistic Structure

Proceedings of the Olomouc Linguistics Colloquium 2013

OLOMOUC
MODERN
LANGUAGE
SERIES

Vol. 3



OLOMOUC MODERN LANGUAGE SERIES
Vol. 3

Language Use and Linguistic Structure

Proceedings of the Olomouc Linguistics Colloquium 2013

organized by

Department of English and American Studies
Philosophical Faculty, Palacký University, Olomouc, Czech Republic
June 6–8, 2013



Edited by
Joseph Emonds and Markéta Janebová

Palacký University
Olomouc
2014

The Role of Information Structure in Czech Possessive Constructions

Jan Křivan

Charles University in Prague, Czech Republic
jan.krivan@ff.cuni.cz

Abstract: The paper deals with the functional properties of possessive constructions. It focuses on internal (adnominal) possession and external (affectedness) possession in Czech. The elements of both construction types (the possessor, the possessum, and the predicate) are first thoroughly investigated from a semantic and pragmatic perspective. In the main part of the article, it is claimed that the emergence of specialized possessive constructions can also be explained as a functional, speaker-oriented preference, based on different needs in terms of information structure. It concerns the hypothesis that the possessive construction types are correlated with different word order arrangements. A case study examining the relative frequency of word order patterns for each construction type corroborates this assumption. Specifically, for dative external possession, the emergence of the construction can be explained as a discourse preference of speakers to employ the possessor in the topic and the possessum in the focus.

Keywords: external possession; information structure; language use; animacy hierarchy; frequency

1. Introduction

Possession can be expressed by certain grammatical means in every language of the world. Three general concepts (i.e., construction types) conveying possession and ownership are distinguished cross-linguistically: internal (adnominal) possession (IP), predicative possession (PredP), and external (affectedness) possession (EP). Both IP (1a) and PredP (1b) are supposed to be encoded in all the languages of the world (Aikhenvald 2013 and Dixon 2010; cf. also Stassen 2009). On the other hand, EP (2) represents, for only some languages, a striking combination of the two basic types, in terms of both semantic and formal properties (see Payne and Barshi 1999).

- (1) (a) *My dog.*
(b) *I have a dog.*
- (2) CZECH
Zlomil mi nohu.
broke:3SG.M me:DAT leg(F):ACC
“He broke my leg,” lit. “He broke a leg to me.”¹

¹ The system of notation for the glosses and abbreviations used adheres to *The Leipzig Glossing Rules* (Comrie, Haspelmath, and Bickel 2008).

Not surprisingly, for each of the three general types we can observe formal and functional variation across languages, i.e., what possessive meanings are actually expressed and how they are specifically encoded in grammar. However, as has been shown in a number of cross-linguistic studies, the employment of language-specific constructions is not haphazard. As a starting point, one typological difference between IP and PredP is notable: whereas IP is conveyed in a specific language inherently as a presupposed relationship (frequently by a construction that also encodes attributive and more general associative meanings), PredP is commonly expressed by means of a dedicated possessive construction, or a special verbal form that explicitly establishes possession (see Seiler 1983). If we focus our attention on a specific language, intra-categorical distinctions within each construction type can often be observed. They are generally related to the functional properties of elements of the construction (the possessor, PR; the possessum, PM; and – if relevant – the predicate, Pred), and to the nature of the possessive relationship itself. Many cross-linguistic similarities have been revealed in this respect.

The focus language of the present study is Czech; to what extent and how cross-linguistic findings on possession apply to this specific language will be shown. In addition, I claim here that the actual usage of the IP and the EP construction type in Czech is not only determined by specific semantic features of constructional elements (as described, e.g., in Haspelmath 1999), but it is also related to distinct discourse properties, present in actual usage, mainly by means of word order (cf. Fried 2009, 221 and 236, who treats the role of information structure in a slightly different way). From a functional perspective, I will examine the properties of EP and IP constructions in terms of the information status of the PR and the PM. A case study, presented in the second part of the paper, works with one specific group of data (predicate *poškodit*, “damage” or “harm,” combined with both IP and EP constructions). I put forward the hypothesis that possessive constructions are correlated with different word order (i.e., information structure) arrangements. Subsequently, this could lead to diachronic explanations: the emergence of specialized possessive constructions in Czech can be explained as a functional, speaker-oriented preference based on different needs in terms of information structure. Therefore, the aim of this study is twofold: i) it should contribute to the detailed description of possession in Czech from a functional perspective and ii) it should extend our knowledge of how possessive constructions function in discourse.

The structure of the article is as follows: syntactic and semantic properties of possessive constructions in Czech, with respect to the subsequent survey, are presented in Section 2. The third section introduces the survey, a case study examining the role of information structure in Czech IP and EP. Section 4 briefly concludes the paper.

2. Properties of Possessive Constructions in Czech

As indicated above, possession in Czech may be conveyed by all three types of constructions. Internal and external possessive constructions will be described in more detail in § 2.1 and § 2.2; predicative possession will be shown only briefly in § 2.3, since for the case study, the most relevant point is the relationship between EP and IP. The relation of PredP and IP is parallel in some respects (see § 2.3), but it concerns a different set of language data.

2.1 Internal Possession

In general, Czech adnominal constructions can employ a wide range of possessive (and other comparable) relationships, similarly to English: i) central possessive meanings (ownership,

part-whole relations, kinship relations, e.g., [3a]–[3c]), ii) broader associational concepts (attribution, orientation and location, and association, e.g., [4a]–[4c]), iii) nominalizations, e.g., (5a)–(5b).²

- (3) (a) *Petrovo auto* “Peter’s car”
 (b) *moje ruka* “my hand”
 (c) *Evin manžel* “Eve’s husband”
- (4) (a) *prestíž pana prezidenta* “the prestige of Mr. President”
 (b) *předek auta* “the front of the car”
 (c) *jeho učitel* “his teacher”
- (5) (a) *Pavlovo bránění* “Paul’s defending”
 (b) *Pavlova obrana* “Paul’s defense”
 (c) *Obrana Manchesteru hrála dobře.* “The defense of Manchester played well.”

In line with Dixon (2010, 265), I deal with both i) and ii) as varieties of possession, in contrast to nominalizations, which I take as a distinct function of the same grammatical marker (unless they are reanalyzed with possessive meanings, as in [5c]). All the constructions in (3), (4), and (5c) assume a possessive relationship between the two constituents; in this sense they can be substituted by a *mít* (“have”) predicative construction; a substitutional schema for (3a) is illustrated in (6).³ Although the nature of possession in ii) is not without problems, I decided to take the whole group into consideration as possessive, since a substantial amount of information would get lost if such borderline instances were omitted.

- (6) *Petrovo auto* ← *Petr má auto.* “Peter’s car” ← “Peter has a car.”

The formal structure of IP in Czech varies. Two basic IP constructions can be distinguished: i) a prenominal construction, subsuming two subcategories: with adjectival possessive pronouns, as in (7a), or possessive adjectives, as in (7b)–(7c), in the role of PRs (henceforth Pro/Adj IP); ii) a post-nominal construction formed by a genitive noun phrase (7d)–(7f), which is the most general adnominal construction type (henceforth Genitive IP).

In addition, there is a third type of construction formed by a prepositional PR which can function in a fashion similar to the IP construction, see (7g). However, the prepositional phrase can furthermore be detached from the PM in some cases, and functions in a similar way to the dative PR in the EP construction, cf. (15). This will be discussed in § 2.2.

² Examples (3)–(5) briefly illustrate the semantic similarity of Czech and English internal possessive relations. Detailed examples with glosses follow.

³ A similar criterion for delimitation of possession was previously used in Zimek (1960). Possession in Czech was thoroughly surveyed in Piřha (1992; resumé in English on pages 139–45), yet his approach differs considerably from the one adopted in this study in its predominant focus on the clear-cut distinction between the “linguistic meaning” and the “cognitive content.”

- (7) (a) *moje ruka*
 my:NOM.SG.F hand(F):NOM
 “my hand”
- (b) *Petr-ov-a ruka*
 Peter(M)-POSS-NOM.SG.F hand(F):NOM
 “Peter’s hand”
- (c) *prezident-ov-a ruka*
 prezident(M)-POSS-NOM.SG.F hand(F):NOM
 “the President’s hand”
- (d) *neviditelná ruka trhu*
 invisible:NOM.SG.F hand(F):NOM market(M):GEN
 “the invisible hand of the market”
- (e) *ruka Petra Velikého*
 hand(F):NOM Peter(M):GEN Great:GEN.SG.M
 “the hand of Peter the Great”
- (f) *den matek*
 day(M):NOM mothers(F):GEN
 “Mothers’ Day,” lit. “the day of mothers”
- (g) *sklo u auta*
 windscreen(N):NOM at car(N):GEN
 “the windscreen of the car”

In terms of the systemic distribution of prenominal/postnominal IP constructions in Czech, there are obvious differences: i) Pro IP is the only choice if the PR is pronominal, as in (7a) (both singular and plural); ii) Adj IP can only be used with some proper names (7b) and human (and some animate) nouns (7c);⁴ (iii) Genitive IP applies to the remainder set of common nouns (7d), for which it is the only way of possessive expression.

Genitive IP must be also used if formal restrictions come into play: i.e., in all cases where the PR forms a plural (7f), or in a phrase consisting of more than one constituent, cf. (7e) and (8e), or where the PR is neuter. On the contrary, the genitive construction is not allowed to be combined with personal pronouns, cf. (7a) and (8a), and is stylistically disfavored (yet possible) with singular sole-constituent proper nouns (8b) and human animate nouns (8c). However, if the PR in (8c) is considered as referring to a non-specific “president” (e.g., in the context of the rules of conduct for presidents), the Genitive IP is appropriate.⁵

- (8) (a) **ruka mě*
 hand(F):NOM me:GEN
 “the hand of mine”

4 Adjectival PRs systematically only refer to a specific referent which is generally known from the context (i.e., similarly to proper nouns) or to a referent given in the context of the utterance. Some animals may be expressed as possessive adjectives if we refer to them in an anthropomorphic manner.

5 With respect to this fact, it should be mentioned that a special possessive adjectival form for generic animal referents also exists. However, its usage is more similar to the function of relational adjectives than to possession; for the sake of simplicity, these phenomena are omitted here.

- (b) ^{??}*ruka* *Petra*
 hand(F):NOM Peter(M):GEN
 “the hand of Peter”
- (c) [?]*ruka* *prezidenta*
 hand(F):NOM president(M):GEN
 “the hand of the President”
- (d) **trh-ov-a* *ruka*
 market(M)-POSS-NOM.SG hand(F):NOM
 “the hand of the market”
- (e) **Petr-ov-a* *Velik-?* / *Velikého* *ruka*
 Peter(M)-POSS-NOM.SG.F Great-POSS / Great:GEN.SG.M hand(F):NOM
 “Peter the Great’s hand”

If we compare the data with cross-linguistic observations, it should be remarked that the variable usage of IP constructions in Czech appears to reflect the different statuses of the PR in the typological Animacy (Nominal) hierarchy (see also § 2.2; cf. Silverstein 1976; Seiler 1983, 76):

- (9) 1/2 personal pronoun > 3 personal pronoun > proper name > other animate > inanimate
 (the Animacy hierarchy)

Aikhenvald (2013, 40), who has worked on possession in hundreds of different languages, states: “A possessor tends to occupy a relatively high position on the Nominal hierarchy: a prototypical possessor is animate, or human, and expressed with a personal pronoun or a proper name.” In the case of Czech IP, these properties are specifically mirrored in the fact that PRs occurring higher in the hierarchy tend to occupy a more specialized type (i.e., Pro/Adj IP) of construction (which is not possible for the PRs of a lower status, as, e.g., inanimate nouns in [8d]). Eventually, on this basis, we can construct a language-specific scale for PRs (10), in terms of their involvement in Czech IP constructions:⁶

- (10) A Pro/Adj IP construction is favored if the possessor is a
 pronoun > proper noun individual > human (animate) individual
 > inanimate, plural noun, neuter, and noun phrase
 (Animacy hierarchy of PRs in Czech IP constructions)

The statement in (10) means that if a Pro/Adj IP construction is possible for a position at any point of the scale, then it is also possible with all of the positions that are higher (further to the left) on the scale.

2.2 External Possession

The concept of external possession is defined as a construction “in which a semantic possessor-possessum relationship is expressed by coding the possessor as a core grammatical relation of the

⁶ The hierarchy in Czech is similar, in many respects, to the preferred distribution of *’s* IP over *of* IP in English. E.g., according to Dixon (2010, 295), the *’s* alternative in English is preferred if the PR is human, specific, and singular, has few words, and is familiar information.

verb, and in a constituent separate from that which contains the possessum” (Payne and Barshi 1999, 3), see (11a) and (2), here repeated for convenience as (12a).

(11) CZECH (simplified examples from the Czech National Corpus)

- (a) *pacientům může poškodit zdraví*
 patients(M):DAT may:3SG harm health(N):ACC
 “It may harm the health of (the) patients,” lit. “It may harm the health to (the) patients.”
- (b) *může poškodit zdraví pacientů*
 may:3SG harm health(N):ACC patients(M):GEN
 “It may harm the health of (the) patients.”

One striking property shows up here: although the PR is coded as a core dative relation, it is not a part of the argument frame of the verb. In addition, from a semantic viewpoint, the possessive relationship is not established by the predicate itself (cf. PredP in § 2.3).

Formal and functional properties of Czech EP constructions roughly conform (but see below) to the European prototype, whose characteristics are (Haspelmath 1999, 111): “i) the marking of the [external] PR by the dative case, ii) the strict affectedness condition, i.e., external possessors are only possible if the possessor is thought of as being mentally affected by the described situation.”

The affectedness condition mainly discriminates the meaning of EP if compared with IP. However, the semantic relationship between IP and EP is far from being uniform. For instance, in (11a)–(11b) there is only a subtle difference in meaning: the external PR is perceived as (contextually) more affected than the internal PR. On the other hand, if we compare (12a) with (12b), the interpretation of the PR is radically different: in the case of EP, it carries a basic (neutral) body part meaning (the PR is affected via the broken body part), whereas the internal PM in (12b) might be interpreted either as somehow detached from the PR, or not as a body part at all.

- (12) (a) *Zlomil mi nohu.*
 broke:3SG.M me:DAT leg(F):ACC
 “He broke my leg,” lit. “He broke a leg to me.”
- (b) *Zlomil moji nohu.*
 broke:3SG.M my:ACC.F leg(F):ACC
 “He broke my leg.”

Thus, the substantial nature of the EP construction is based on mutual relationships among the properties of the PR, the PM, and the predicate. The interplay in the Czech EP works as follows:

- i) PRs are marked by the dative. Their prototypical semantic properties correspond to the Animacy hierarchy, presented in (9): PRs tend to refer to a human animate being, overtly expressed as a personal pronoun. However, in contrast to other European languages,⁷

⁷ In most languages, EP is possible only with possessors positioned further left in the Animacy hierarchy than in Czech. E.g., in French, PRs are restricted to pronouns.

reference to the dative external PR is not so restricted; the external PR can also marginally refer to other animates and inanimates (e.g., plants in (14), cars, sports teams).

- ii) PMs can occupy the object, subject, or oblique position in the clause. They prototypically refer to body parts and other nouns closely related to the PR, i.e., they correspond to the Inalienability hierarchy, depicted in (13) (König and Haspelmath 1998, qtd. in Haspelmath 1999). Once again, in Czech there is no clear semantic restriction on which nouns cannot appear as PMs. However, if PRs are inanimate, only part-whole relationships, as in (14), generally seem to be possible.

- (13) body part > garment > other contextually unique item
(the Inalienability Hierarchy)

- (14) *Balzáminec opadávají poupata.*
Garden.balsam(F):DAT fall.off:PRS:3PL buds(N):NOM
“The buds of the garden balsam are falling off,” lit. “Buds are falling off to the garden balsam.”

- iii) Both transitive and intransitive predicates can occupy the Czech EP construction. Transitivity corresponds to the syntactic role of the PM. If the PM is an object, the predicate must be transitive, as in (12a). If the PM is a subject, the predicate is intransitive, as in (14). If the PM is an oblique, both types are possible. Semantically, predicates tend to be dynamic and active; their inherent semantics is patient-affecting (somehow changing the quality of possession).

It should be emphasized that the semantic properties mentioned in i)–iii) are not absolute properties of the clausal elements, but need to be considered as relative with regard to the (pragmatic) context. For instance, the predicate itself need not be affective; it can only bear a potential context-dependent affective interpretation, i.e., the resulting semantics of the construction is affective. Similarly, in contrast to Haspelmath’s (1999) account of the European external PR prototype, I do not regard the mental affectedness of the PR in Czech as an absolute requirement; rather, the affectedness of the PR concerns the pragmatic context; it pertains to the speaker’s view of reality (cf. Fried 2009); therefore, it does not depend on the true mental or physical state of the PR, but rather on the discourse context reflected by the speaker. Very similar pragmatic conditions seem to hold for the external PR in terms of the information status. Since it concerns the basic hypothesis tested in this survey, this issue will be thoroughly explored in Section 3.

One additional note here: in § 2.1 I showed the prepositional *u* (“at”) phrase, which resembles both IP (7g) and EP (15) constructions, since the PR phrase can be detached from the PM. This (nearly) possessive construction has its origin in the locative meaning. In addition, the *u* preposition is the most salient member of a larger set of similar prepositions. All of them are more natural with inanimates and/or with less prototypical possessive relations; therefore, they form an expression functioning on the boundary between possession and other (spatial) relations. For the purpose of this study, I introduce a new concept of mixed possession (MP) which subsumes both IP-like and EP-like instances of this prepositional PR construction.

- (15) *u auta poškodí sklo*
 at car(N):GEN damages windscreen(N):ACC
 “It damages the windscreen of the car,” lit. “It damages the windscreen at the car.”

To sum up the typical functional properties of EP in Czech, the predicate of the construction has an effect on the possessum, and thereby the possessor is perceived by the speaker as being affected.

2.3 Predicative Possession

Predicative possession functions within an entire clause, in which it establishes a possessive relationship. In this respect, it differs from both IP and EP, where the relation is presupposed (inherent), and not established by the predicate (or by the predicative construction). Marking PredP in Czech involves the transitive verb of ownership *mít* (“have”), with the PR as the subject and the PM as the object. It can express all the possessive meanings mentioned in § 1.1, as previously illustrated in the scheme (6), here repeated for convenience: *Petrovo auto* ← *Petr má auto* (“Peter’s car” ← “Peter has a car”). Specific possessive meanings can also be conveyed, as in English, by the special possessive verbs *patřit* (“belong”); *náležet, příslušet* (“pertain”); and *vlastnit* (“own”).

One specific structure of “have” constructions resembles the relationship between EP and IP, cf. (11a)–(11b) and (16a)–(16b).

- (16) (a) *pacienti mají poškozené zdraví*
 patients(M):NOM have:3PL harmed:ACC.SG.N health(N):ACC
 “The health of the patients is harmed,” lit. “The patients have the harmed health.”
 (b) *zdraví pacientů je poškozené*
 health(N):NOM patients(M):GEN is harmed:NOM.SG.N
 “The health of the patients is harmed.”

The difference in the meaning between (11) and (16) is obvious. In (11) the predicate (*poškodit*, “harm”) has its basic verbal meaning; in (16) the adjective (*poškozené*, “harmed”) forms a part of the resultative construction. The PredP construction (16a) forms a transitive counterpart to the *být* (“be”) predicate, which subsumes the IP construction (16b). In the sense that both EP and PredP can replace IP in some clauses, the relationships EP-IP and PredP-IP are parallel: both constructions enable the PR to be detached from the PM, and change the information structure of the PR and PM (see Section 3). However, the only focus of the present study will be the EP-IP relationship; the resultative construction will be left aside.

3. Survey on Information Structure in Possessive Constructions

A general aim of this survey is to contribute to the question of how distinct possessive constructions function in common usage, and how they possibly emerged, i.e., by investigating language use, I strive to explain how different types of constructions could have come to be grammaticalized. I adopt the functional approach to the emergence of linguistic structure (e.g., Bybee 2007; Hopper 1998; Haspelmath 2008). Some ideas on information structure presented here were inspired by Du Bois’s (1987) seminal article, which worked on the emergence of ergativity. Basically, I claim that the actual usage of IP or EP in Czech is not only determined by specific

features of constructional elements (introduced in § 2), but it is also correlated with word order (i.e., information structure) arrangements.

Some basic notions concerning information structure need to be introduced. First, I work on the assumption of the “given”–“accessible”–“new” continuum in the sense of Chafe (1987). According to Chafe, the notion of “given” is defined as a cognitively active concept, whereas “new” concerns the concepts previously inactive in the speaker’s consciousness (“accessible” lies in the middle: semi-active concepts). As has been proven in different languages, this continuum shows an effect on language structure: e.g., according to Chafe, new concepts tend to be expressed by full lexical items, while given concepts are frequently expressed by pronouns or pronominal inflections. This indirectly implies that the Animacy hierarchy, mentioned above, also generally corresponds to the given-new continuum (see also Du Bois 1987, 830); in my study, the frequency of pronouns in the position of the PR will serve as an indicator of the amount of given information in each construction type.

The second assumption, based on word order, is closely related to the previous one. On the one hand, word order rules in a Czech clause are relatively free, and thus, the position of the external PR can be variable, whereas the position of the internal PR is fixed inside the nominal phrase (see Fried 2009). On the other hand, it has been shown that the Animacy hierarchy effects also apply cross-linguistically to topic-like (and agent-like) properties. This poses a more specific assumption that prototypical PRs in special (Pro/Adj IP and EP) possessive constructions may somehow behave along the Animacy hierarchy with respect to their information status. In my survey, word order serves as the main indicator of information structure, (since information flow in Czech is expressed mainly through word order [see, e.g., Daneš 1974]): elements closer to the front edge of the clause are expected to be “given” or “accessible” more frequently than the ones at the opposite edge. Consequently, I assume that the prototypicality of the PR (its “givenness”) may be somehow manifested, besides its expression as a pronoun, by its proximity to the front edge of the clause. More radically, the prototypical PR is supposed to function in the clause as a topic,⁸ and (since PMs are, on the contrary, typically inanimate) the prototypical PM as a focus.

Note: I base my definitions of “topic” and “focus” on Daneš’s (e.g., 1985) simple definitions of “theme” and “rheme” (theme/topic: what is being talked about, rheme/focus: what is being said about the topic). For the purpose of the present study, I delimit the terms “topic” and “focus” simply by reference to word order: “topic” concerns the positions in front of the predicate; while the “focus” concerns the positions after the predicate (I only marginally assessed the role of phenomena violating this correlation, such as stress, in my data).⁹

Therefore, to achieve the main goal, I investigate which possessive word order patterns are most commonly used by Czech speakers in different constructions. Following that, in § 3.1, I present the object of the case study, the key categories I have used to describe the material, and how I have extracted the data from the corpora. In § 3.2 I describe the frequency data; the last section, § 3.3, is devoted to the proposed explanations.

8 The topical status of the external PR has already been observed in different languages: Tz’utujil (Aissen 1999) and Creek (Martin 1999).

9 The predicate is regarded as a “switch” between the topic and the focus; I do not determine its particular information status.

3.1 Data

I analyzed the data from the *Czech National Corpus* (*SYN2000*, *SYN2005*, *SYN2010*). Each of these representative corpora consists of 100 million tokens of written Czech, i.e., the whole corpus for the survey comprises 300 million tokens in total. Spoken corpora would certainly provide more spontaneous data; however, for various reasons I decided to avoid them. I take written corpora as a good substitute, since I presuppose that the hypothesized phenomena may appear in spoken language more saliently than in written language, but not the other way round. The reason for the choice of written corpora was also practical: none of the existing corpora of spoken Czech could provide a reasonable amount of comparable data for my investigation. (I decided to work on a strictly delimited set of constructions relating to just one predicate lemma, distinguished solely by grammatical means.)

The objects of the survey are all clauses comprising: i) the predicate *poškodit* (“damage”), and ii) one of the possessive constructions (Genitive IP, Adj/Pro IP, EP, MP). The reason for such a restriction was as follows: if I take just one specific predicate into account, there will be no unpredictable semantic and syntactic variation among the arguments of the clause (e.g., in the case of *poškodit*, the PM is always an accusative object,¹⁰ or a subject, if the verb is in the passive; the predicate is prototypically dynamic and affecting, i.e., attaching EP constructions). Hence, all different arrangements of word order (and that implies information structure) can be studied neatly, without major potential interferences.

The data from the corpus were mined using more than 60 queries to capture all the potential co-occurrences of all the grammatical constructions. The technique was based on identifying the boundaries of the clause, and finding the predicate (lemma *poškodit*, with all tokens in the active/passive voice, including negative forms) plus the PM and the PR in all potential possessive forms, and in different positions, with respect to word order (I distinguished constructions in terms of lexical and pronoun PRs; I used automatic lemmatization).

After that, I saved the concordances for each query and exported them to a spreadsheet editor, checked all the concordances, and excluded the inappropriate (non-possessive) data. Subsequently, the checked data were collected within individual spreadsheets, according to the type of the construction, and manually annotated.

For each line in the spreadsheet, I distinguished between the active and the passive construction, and identified the word order of the three constructional elements: the predicate, the PR, and the PM. Therefore, I did not take into account the word order position of the agent (mostly the subject, which can be dropped in Czech; this variation would concern a different set of issues). If the predicate consisted of the lexical element *poškodit* combined with a modal verb, I annotated both positions, but eventually used only the position of the lexical verb as relevant for the word order analysis. Finally, I counted the total number of occurrences for each construction, in all possible word order arrangements.

3.2 Description

For the three elements of possessive constructions with the predicate *poškodit* that were studied, six logically possible word order arrangements for the active voice and six arrangements for

¹⁰ Admittedly, the oblique PM (see § 2.2) might produce a slightly different distribution of word order positions than the object PM; however, I have no evidence that the results would be radically different.

the passive voice need to be taken into account. Since the word order of the constituents in the Czech clause is relatively free, all 12 combinations can theoretically be found among EP constructions.¹¹ IP constructions with a fixed constituent structure PM-PR (for genitive noun phrases) or PR-PM (for possessive pronouns and possessive adjectives) permit only four settings each.¹² In the charts below (Figures 1–3), I refer to each pattern on the x-axis by the numbers 1–12. The numbers 1–6 refer to the active voice (*poškodit*, “damage”; not surprisingly, 84% of all *poškodit* tokens within the possessive constructions are active):

- 1–3 comprise the PR in the topic: [1] PR-Pred-PM, [2] PR-PM-Pred, [3] PM-PR-Pred
- 4–6 comprise the PR in the focus: [4] Pred-PR-PM, [5] Pred-PM-PR, [6] PM-Pred-PR

The numbers 7–12 refer to the passive voice (*poškodit se*, reflexive passive or anticausative; or *být poškozen*, “be damaged”):

- 7–9 comprise the PR in the topic: [7] PR-Pred-PM, [8] PR-PM-Pred, [9] PM-PR-Pred
- 10–12 comprise the PR in the focus: [10] Pred-PR-PM, [11] Pred-PM-PR, [12] PM-Pred-PR

In the rest of this section, I will show the results of the word order analysis for Genitive IP, Pro/Adj IP, and EP, respectively, and make a remark on MP. The remaining paragraphs show the entire picture.

GENITIVE IP CONSTRUCTION, as stated above, shows the lowest status of the PR in the Animacy hierarchy: it can be used with non-human PRs, it does not comprise any pronoun PRs (see § 2.1), and the data subsume many instances of non-prototypical relations (attribution and association). Its most frequent word order pattern is [5] Pred-PM-PR (76%), illustrated in (17).¹³ It employs the PM in a position typical for objects; however, since the PR is firmly bound with the PM, the PR occupies the position in the focus, even to the right of the PM.

(17)	PRED	PM	PR
	<i>kdo</i>	<i>poškodí</i>	<i>zdraví</i>
	who	harms	health(N):ACC
			<i>pacientů</i>
			patients(M):GEN
	“who harms the health of (the) patients”		

A comparison of all four patterns is given in Figure 1 (note: the vertical lines at the top of the columns in Figures 1–3 show the 95% confidence interval for multinomial distribution, and therefore illustrate which differences between patterns should be considered as significant). The frequency counts are presented in Table 1 at the end of this section (note: the table shows absolute and relative counts of word order patterns for each of the constructions that were studied).

11 As an anonymous reviewer correctly noted, the word order of the Czech clause is not free in the case of clitics. Pronominal clitics (in the data, referring to external PRs) are fixed to the clausal second position. However, this does not affect the potential word order arrangements of the PR, the PM and the Pred in the EP construction. First, pronominal external PRs can also be expressed as non-clitics; second, the initial position of the clause can be occupied by any kind of constituent. Rather, clitics can be considered as another indicator of givenness (pronouns which are close to the front edge of the clause).

12 For the sake of completeness: in marginal (extremely low frequent) contrastive contexts, the reverse order of PR-PM elements in the Pro/Adj IP construction is possible. In addition, concerning the Genitive IP construction, there is a category of human nouns forming pronominal genitives which function just like possessive adjectives. Both phenomena were omitted here.

13 All examples in (17)–(19) from the Czech National Corpus.

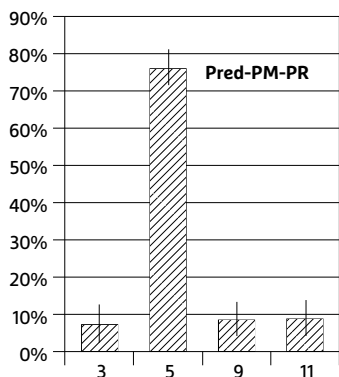


Figure 1. Relative frequency (in %) of the Genitive IP word order patterns for the *poškodit* predicate.

PRO/ADJ IP CONSTRUCTION has already been presented as a specialized sort of IP construction, and prototypical in terms of the semantic properties of the PR. Possessive pronouns (i.e., not possessive adjectives) are employed in a vast majority (91%) of all instances. I treat both adjectives and pronouns as one group, since the difference in distribution of word order patterns did not appear to be significant or relevant in terms of description.

- (18)
- | | | | | |
|----------|--------------|------------------|-------------|-----------------|
| | | PRED | PR | PM |
| <i>a</i> | <i>velmi</i> | <i>poškodily</i> | <i>naše</i> | <i>věci</i> |
| and | greatly | damaged:3PL | our:ACC.PL | stuff(F):ACC.PL |
- “and they greatly damaged our stuff”

The construction displays one major pattern [4] Pred-PR-PM (60%), as in (18). This pattern employs both PR and PM in the focus, similarly to [5] in the Genitive IP; however, the fixed PR-PM order is reversed, i.e., both elements are moderately approaching their presumed prototypes. In addition, if we compare Figures 1 and 2, we can observe a considerable increase in the relative usage of the pattern employing the PR in the topic; see [2] PR-PM-Pred in Figure 2.

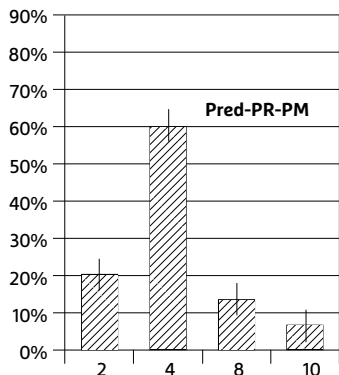


Figure 2. Relative frequency (in %) of the Pro/Adj IP word order patterns for the *poškodit* predicate.

EP CONSTRUCTION introduces an affected, dative PR. As demonstrated above, the PR is supposed to be very high in the Animacy hierarchy. This is corroborated by the data that were studied. Most external PRs (87%) are pronouns, and a vast majority of them are second position clitics. Again, there was no significant difference in the distribution of pronoun vs. adjectival word order patterns, and therefore no reason to split the category. Once again, there is only one major pattern [1] PR-Pred-PM (64%). It employs the PR in the topic, and the PM in the focus, illustrated in (19a)–(19b).

- (19) (a) PR PRED PM
že jí poškodili páteř
 that she:DAT harmed:3PL backbone(F):ACC
 “that they harmed her backbone,” lit. “that they harmed a backbone to her”
- (b) PR PRED PM
a policistům poškodila stejnokroj
 and policemen(M):DAT damaged:3SG.F uniform(M):ACC
 “and she damaged the uniforms of the policemen,” lit. “and she damaged a uniform to the policemen”

Additionally, the pattern [4] Pred-PR-PM, is relatively frequent (18%); see Figure 3. One note here: if we compare Figures 2 and 3, we can observe that both construction types, Pro/Adj IP and EP, make use of pattern [4]. My previous expectation was that if there were no tendency for external PRs to be used in a special word order position, the ratios of [4] in Figures 2 and 3 would appear to be comparable.

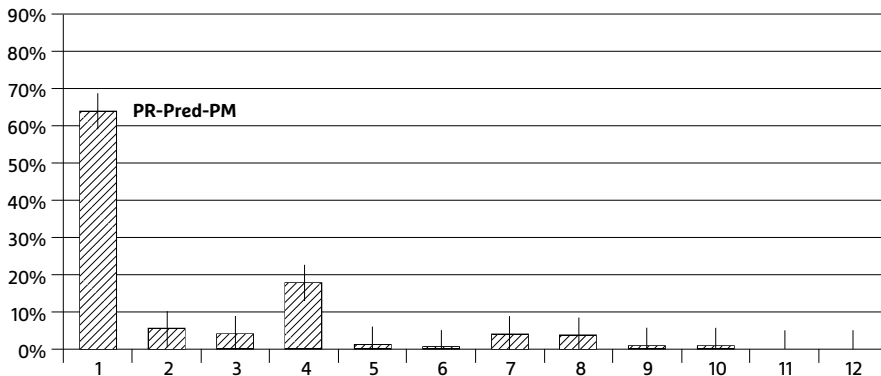


Figure 3. Relative frequency (in %) of the EP word order patterns for the *poškodit* predicate.

MP CONSTRUCTION, unfortunately, does not supply a sufficient amount of data for a plausible analysis. What should be noted is this: pattern [1], stated above as a major pattern for EP, shows significantly more occurrences than the rest of the patterns (not surprisingly, because of the mixed nature of the construction, except for the comparison with pattern [5], which turned out to be the major pattern for Genitive IP), see Table 1.

		EP		Pro/Adj IP		Genitive IP		MP		Total	
		f	%	f	%	f*	%	f	%	f	%
1	PR-Pred-PM	225	64%					16	37%	241	12%
2	PR-PM-Pred	19	5%	98	20%			0	0%	117	6%
3	PM-PR-Pred	13	4%			76	7%	5	12%	94	5%
4	Pred-PR-PM	62	18%	294	60%			1	2%	357	18%
5	Pred-PM-PR	3	1%			812	76%	12	28%	827	42%
6	PM-Pred-PR	1	0%					1	2%	2	0%
7	PR-Pred-PM	13	4%					6	14%	19	1%
8	PR-PM-Pred	12	3%	66	13%			0	0%	78	4%
9	PM-PR-Pred	2	1%			88	8%	1	2%	91	5%
10	Pred-PR-PM	2	1%	31	6%			0	0%	33	2%
11	Pred-PM-PR	0	0%			92	9%	1	2%	93	5%
12	PM-Pred-PR	0	0%					0	0%	0	0%
Total		352	100%	489	100%	1068	100%	43	100%	1952	100%

* The frequency counts for Genitive IP were approximated (based on a one-quarter random sample).

Table 1. Overall frequency overview of distinct word order patterns for the *poškodit* predicate. Frequency counts in the written corpora SYN2000 + SYN2005 + SYN2010.

THE OVERALL PICTURE is displayed in Table 1. We can imply three preferred word order patterns for each of the basic possessive constructions (note that 73% of all tokens are subsumed under these arrangements):

- Genitive IP construction: [5] Pred-PM-PR
- Pro/Adj IP construction: [4] Pred-PR-PM
- EP construction: [1] PR-Pred-PM

Additionally, seven other low-frequency patterns play a role in discourse. On the contrary, the patterns [6] and [12] PR-Pred-PM, which employ the PR in the topic and the PM in the focus, appeared not to be preferred at all.

3.3 Explanation

The frequency data in § 3.2 indicated the existence of certain functional motivations, which may have given rise to distinct possessive constructions in Czech. I will state them here more systematically.

Two specialized possessive constructions (Pro/Adj IP and EP) show a highly prevalent number of PR pronouns. This supports the “givenness” and the animacy status of the prototypical PR and it corroborates the hypothesis that the emergence of possessive constructions could be motivated by a speaker’s needs related to information structure.

First, the major word order pattern of the EP construction employs the PR as the topic and the PM as the focus. This supports the general idea of Aissen (1999, 189), based on findings in Tz’utujil, that one of the functions of the EP construction cross-linguistically may be “to present the possessor as the logical subject.” Similarly, in Creek, according to Martin (1999, 243), in all

of the natural examples in his data, the external PR “is old information and is generally omitted.” Consequently, the emergence of the EP construction appears to be functionally related not only to the semantic affectedness of the PR but also to its topical status.

Second, the Pro/Adj IP construction shows, in comparison with the Genitive IP construction, a reverse word order PR-PM; thus, the PR appears further to the left in the clause; this could be motivated by similar discourse needs. Simultaneously, pronouns are excluded from the Genitive IP construction, which appears to be the “least friendly” construction for PRs.

Third, as stated in the previous section, the pattern PM-Pred-PR is not used at all.

To sum up: my findings generally indicate that the Animacy hierarchy which corresponds to the typical semantic properties of the PRs is also operational in terms of the information status of the PR and the PM. It concerns a functional motivation for speakers to employ the possessor as given, typically in the topic, and to detach it from the PM, whose prototypical position is in the focus. Therefore, complementary needs to express the PM as new information and the PR as given could be the functional reasons for the emergence of the EP construction in Czech and, simultaneously, for the reversed PR-PM order of the Pro/Adj IP construction.

4. Conclusion

The data presented in this study suggest that information structure plays a crucial role in explaining the usage and emergence of distinct possessive structures. An overview of the Czech possessive construction types illustrates how different structures are related to the semantic properties of the constructional elements (the prototypical properties of the possessor can be widely explained by Animacy hierarchy effects); however, it cannot reveal the full story.

In the subsequent survey, the data on the *poškodit* construction brought remarkable results in terms of information structure and possessive constructions. It showed that each of the three IP and EP construction types has its own major word order pattern. Chiefly, the emergence of EP constructions supports the explanation in terms of different and complementary discourse needs: there is a tendency to express the “topical” information status of the possessor and the “focal” status of the possessum. These findings widely correspond to the Animacy hierarchy effects presented above.

Finally, the proposed hypothesis on the role of information structure has been corroborated by empirical findings; however, further research dealing with more variable data is needed to support it in a more general scope.

Funding Acknowledgement

Outcome of the project of Internal Grants at the Faculty of Arts, Charles University in Prague 2013.

Works Cited

- Aikhenvald, Alexandra Y. 2013. “Possession and Ownership: A Cross Linguistic Perspective.” In *Possession and Ownership: A Cross-Linguistic Typology*, edited by Alexandra Y. Aikhenvald and Robert M. W. Dixon, 1–64. Oxford: Oxford University Press.
- Aissen, Judith. 1999. “External Possessor and Logical Subject in Tz’utujil.” In *External Possession*, edited by Doris L. Payne and Immanuel Barshi, 167–93. Amsterdam: John Benjamins.
- Bybee, Joan L. 2007. *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

- Chafe, Wallace. 1987. "Cognitive Constraints on Information Flow." In *Coherence and Grounding in Discourse: Outcome of a Symposium; Eugene, Oregon, June 1984*, edited by Russell Tomlin, 21–51. Amsterdam: John Benjamins.
- Comrie, Bernard, Martin Haspelmath, and Balthasar Bickel. 2008. *The Leipzig Glossing Rules: Conventions for Interlinear Morpheme-by-Morpheme Glosses*. Last modified February 2008. <http://www.eva.mpg.de/lingua/resources/glossing-rules.php>.
- Daneš, František. 1974. "Functional Sentence Perspective and the Organization of the Text." In *Papers on Functional Sentence Perspective*, edited by František Daneš, 106–28. Prague: Academia.
- Daneš, František. 1985. *Věta a text: Studie ze syntaxe spisovné češtiny*. Prague: Academia.
- Dixon, Robert M. W. 2010. *Basic Linguistic Theory*. Vol. 2, *Grammatical Topics*. Oxford: Oxford University Press.
- Du Bois, John W. 1987. "The Discourse Basis of Ergativity." *Language* 63 (4): 805–55.
- Fried, Mirjam. 2009. "Plain vs. Situated Possession in Czech: A Constructional Account." In *The Expression of Possession*, edited by William McGregor. Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 1999. "External Possession in a European Areal Perspective." In *External Possession*, edited by Doris L. Payne and Immanuel Barshi, 109–35. Amsterdam: John Benjamins.
- Haspelmath, Martin. 2008. "Creating Economical Morphosyntactic Patterns in Language Change." In *Linguistic Universals and Language Change*, edited by Jeff Good, 185–214. Oxford: Oxford University Press.
- Hopper, Paul. 1998. "Emergent Grammar." In *The New Psychology of Language*, vol. 1, *Cognitive and Functional Approaches to Language Structure*, edited by Michael Tomasello, 155–75. Mahwah: Lawrence Erlbaum.
- König, Ekkehard, and Martin Haspelmath. 1998. "Les constructions à possesseur externe dans les langues d'Europe." In *Actance et valence dans les langues de l'Europe*, edited by Jack Feuillet, 525–606. Berlin: Mouton de Gruyter.
- Martin, Jack B. 1999. "External Possession in Creek." In *External Possession*, edited by Doris L. Payne and Immanuel Barshi, 229–50. Amsterdam: John Benjamins.
- Payne, Doris L., and Immanuel Barshi. 1999. "External Possession: What, Where, How, and Why." In *External Possession*, edited by Doris L. Payne and Immanuel Barshi, 3–29. Amsterdam: John Benjamins.
- Pitřha, Petr. 1992. *Posesivní vztah v češtině*. Prague: AVED.
- Seiler, Hansjakob. 1983. *Possession: As an Operational Dimension of Language*. Tübingen: Gunter Narr.
- Silverstein, Michael. 1976. *Hierarchy of Features and Ergativity*. Canberra: Australian National University.
- Stassen, Leon. 2009. *Predicative Possession*. Oxford: Oxford University Press.
- Zimek, Rudolf. 1960. "K chápání posesivnosti." In *Rusko-české studie*, edited by Ctirad Bosák et al., 131–56. Prague: SPN.

Corpora

Czech National Corpus – SYN2000. Institute of the Czech National Corpus. Available online at <http://www.korpus.cz>.

Czech National Corpus – SYN2005. Institute of the Czech National Corpus. Available online at <http://www.korpus.cz>.

Czech National Corpus – SYN2010. Institute of the Czech National Corpus. Available online at <http://www.korpus.cz>.

List of Abbreviations

The abbreviations used in the glosses are not listed here (see note 1).

EP – external possession

IP – internal possession

MP – mixed possession

PM – possessum, possessed

PR – possessor

Pred – predicate

PredP – predicative possession

Pro/Adj IP – pronoun/adjective internal possession