

Grammatical Functions in Case Languages: Subjecthood in Czech

Author(s): Mirjam Fried

Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: General Session Dedicated to the Contributions of Charles J. Fillmore (1994), pp. 184-193

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Grammatical Functions in Case Languages: Subjecthood in Czech*

Mirjam Fried

University of California, Berkeley

Case languages tend to resist straightforward treatment with respect to the standard inventory of grammatical functions (GF) in grammatical theories in which functional categories such as subject, object, etc. are taken as central to describing the relationship between a predicate and its arguments. One difficulty concerns the mapping between GFs and morphological cases, which often involves one-to-many relationships, thus making the occurrence of a significant number of case forms unpredictable. Moreover, certain constructions make it difficult to even identify GFs unambiguously. This problem is usually attributed to the discrepancy between a case form and the syntactic behavior of a particular argument, but often has more to do with theory-internal assumptions which may be in conflict with real data. Both problems raise the question of case assignment principles and of what role, if any, can be attributed to GFs in formulating those principles.

Assuming that a proper representation must include the level of morphological cases (formal expressions of arguments, at least in non-configurational languages) and the level of semantic roles (SR -- abstractions of conceptual structure needed for independent reasons), we can ask whether the additional level of abstract grammatical categories such as GFs is also necessary. With respect to case assignment, there are two options, both of which weaken the role of GFs:

- (i) Case is an expression of GF, as practiced for example in LFG. 'Unexpected' mappings either must be stipulated as unmotivated links between a GF and case, or the level of GFs is essentially left out in favor of linking the case directly to a particular SR (Neidle 1982). In either case, a large set of data must be treated as exceptions and in the latter approach, GFs become irrelevant for those data.
- (ii) Case is an expression of SR. Current approaches tend to either posit correlations between cases and SRs (Comrie 1981, Bhat 1991) or outright equate the two (Croft 1991); either way, GFs become superfluous in stating the rules of case assignment.

Schematically, the two kinds of representation could be summarized as follows:

- | | | | |
|--------|----------------------------|----|-----------------------|
| (1) a. | SR

GF

case | b. | SR (=case)

GF |
|--------|----------------------------|----|-----------------------|

Based on the empirical findings presented in this paper, I will propose a version of the latter approach. Specifically, I examine the issue of subject identification in Czech, focusing on predicates which are associated with NPs of dubious functional status.

Czech has been traditionally treated as a subject-prominent language in the sense of Li & Thompson (1976); most sentences contain a nominative NP which agrees in number and gender with the verb¹ and otherwise shows subject-like behavior (reflexive binding, target of equi NP deletion, etc.). But Czech is also known for a variety of truly subjectless sentences, which have two sources: syntactic and lexical. The syntactic subjectlessness is associated with certain modal

- (4) Petr po *jeho/svém vítězství uspořádal oslavu.
 Peter-N after *his/self's victory-L organize-Ps.3sg.m celebration-A.sg.f
 'Peter_i threw a big party after his_i victory.'

In (4), the non-reflexive pronoun is prohibited if the intended reading is for *Peter* and *his* to be coreferential. But if the possessor is not in the nominative, only the non-reflexive pronoun is possible, with the necessary ambiguity about its antecedent:

- (5) Petrovi po jeho/*svém nečekaném vítězství všichni blahopřáli.
 Peter-D after his/*self's unexpected-L victory-L all-N.pl.m congratulate-
 Ps.3pl.m
 'Everybody congratulated Peter_i after his_{i,j} unexpected victory.'

Here the antecedent is in the dative, which precludes the use of the reflexive. The sentence has two readings: the possessor is either *Peter* or some other entity.

This formal condition on the distribution of *svůj* vs. *jeho* predicts that the dative and the prepositional phrase with the verbs in (3) should also prohibit the use of the reflexive pronoun, and that indeed is the case:

- (6) Petrovi_i se po jeho_{i,j}/*svém nečekaném vítězství řádně ulevilo.
 Peter-D Rf after his/*self's unexpected-L victory-L very get.relieved-
 Ps.3sg.n
 'Peter_i felt a great relief after his_{i,j} unexpected victory.'⁴

3. Another behavioral test is *equi NP deletion*. Although Czech does have instances of both subject- and object-control equi constructions, this kind of embedding is difficult to use successfully to test the subjecthood of the arguments in question because of semantic incompatibility. Czech is one of the languages where the target of equi NP deletion is expected to have a certain amount of control, to be agentive (as Comrie 1984 has shown for Russian), which clashes with the spontaneous, uncontrolled nature of the predicates exemplified in (3). It is not possible to try/be able/be obliged/etc. to undergo the experiences or states of that kind. Similarly, it is not possible to order/allow/advise anybody to undergo them.

The only semantically plausible use of an equi-like verb would involve the verb *vidět* 'see', and with telling results with respect to the subjecthood of the target arguments. In general, the verbal complement of *vidět* 'see' can be either an equi-construction, illustrated in (7a), or a finite subordinate clause, shown in (7b):

- (7) a. Už jsem viděla Petra běžet.
 already Aux.1sg see-Ppl.sg.f Peter-A run-Inf
 'I have seen Peter run.'
- b. Už jsem viděla, jak Petr běží.
 already Aux.1sg see-Ppl.sg.f as Peter-N run-Pr.3sg
 'I have seen Peter run.'

By comparing these two sentences, we can see that the target of equi NP deletion (*Petr*) is the argument that is normally coded in the nominative case (subject). In contrast, when the embedded verb is one of the experiential (8) or locative (9) predicates, only the finite clause is acceptable:

- (8) a. *Viděla jsem Petrovi najednou přeskočit.
 see-Ppl.sg.f Aux.1sg Peter-D suddenly jump.over-Inf
 'I have seen Peter suddenly go nuts.'
- b. Viděla jsem, jak Petrovi najednou přeskočilo.
 see-Ppl.sg.f Aux.1sg as Peter-D suddenly jump.over-Ps.3sg.n
 'I have seen Peter suddenly go nuts.'
- (9) a. *Viděla jsem pářit se z lesů.
 see-Ppl.sg.f Aux.1sg vapor-Inf Rf from forest-G.pl
 'I saw vapor rise from the forest.'
- b. Viděla jsem, jak se páří z lesů.
 see-Ppl.sg.f Aux.1sg as Rf vapor-Pr.3sg from forest-G.pl
 'I saw vapor rise from the forest.'

Neither the dative NP *Petrovi* in (8), nor the prepositional phrase *z lesů* in (9) can be the target of subject-control equi NP deletion. We thus must conclude that as far as the equi operation is concerned (and to the extent that it can be used with these predicates at all), the arguments in question do not behave like subjects.

4. Quite interestingly, the most important behavioral test -- *raising* -- is not available because Czech does not allow raising of any kind. The only embedding strategy permitted with raising verbs is a finite subordinate clause, which is perhaps a telling typological property of the language. I would suggest that the very absence of raising constructions may be an important characteristic with respect to the (in)significance of subjecthood in the grammar of Czech.

To sum up, both the coding and behavioral tests clearly rule out the single argument of the above verbs as grammatical subjects.

Additional complications with subject identification arise with predicates that have two obligatory arguments neither of which is marked by the nominative.⁵ One class of such predicates can be described as a variation on the one-place predicates discussed above, reporting spontaneously occurring experiences located in a particular body part (sometimes expressed metaphorically); the experiencer is the body part's possessor:

- (10) a. Vyschlo mu v krku. b. Straší mu ve věži.
 dry.up-Ps.3sg.n he-D in throat-L haunt-Pr.3sg he-D in tower-L
 'His throat got all dry.' 'He's gone nuts.'

The semantic valence of these verbs (*znít* 'sound', *strašit* 'act like an apparition', *svítit* 'glow', etc.) can be described as <Exp, {Loc, So, Go}> and the case marking thus follows general case assignment principles. It is also important to note that the 'raising' of the possessor is obligatory. The sentences in (10) cannot be paraphrased with a possessive pronoun attached to the locative phrase:

- (11) a. *V jeho krku vyschlo. b. *V jeho věži straší.
 in his throat-L dry.up-Ps.3sg in his tower-L haunts-Pr.3sg

Again, according to the Subject Condition, one of these two arguments must be mapped on the subject function. The mapping would be presumably predicted by the thematic hierarchy, on which experiencers rank higher than locatives and therefore the experiencer argument should be the subject. However, when we apply

the available tests, we again find that the experiencer NP does not possess any characteristics associated with subjects in Czech. To the extent that we can construct a context in which it is possible to use a possessive phrase coreferential with the experiencer, the experiencer NP can be shown not to behave like a subject:

- (12) Při každém jeho/*svém slově mu zasvítilo v očích.
 with every his/*self's word-L.sg he-D light.up-Ps.3sg.n in eyes-L.pl
 'With every word he_{i,j} uttered, there was a flicker in his_i eyes.'

The patterning is the same as with the one-place predicates, namely the reflexive pronoun cannot be used with the dative NP. And to turn for a solution to the locative argument does not help either. Even if we somehow reconcile the blatant violation of the thematic hierarchy, the locative PP does not pass the subject tests any better than the dative.⁶

The application of the equi test also results in a failure to select either argument as the subject; only the subordinate clause (13c) is possible:

- (13) a. *Viděla jsem mu zasvítit v očích.
 see-Ppl.sg.f Aux.1sg he-D light.up-Inf in eyes-L
 'I saw his eyes flicker.'
- b. *Viděla jsem mu zasvítit oči.
 see-Ppl.sg.f Aux.1sg he-D light.up-Inf eyes-A
 'I saw his eyes flicker.'
- c. Viděla jsem, jak mu zasvítilo v očích.
 see-Ppl.sg.f Aux.1sg as he-D light.up-Ps.3sg.n in eyes-L
 'I saw his eyes flicker.'

And, of course, neither argument can be the subject in terms of coding: there is no subject-verb agreement. Thus the only conclusion we can draw from the facts is that neither argument can be assigned the subject function in any motivated way.

A different kind of situation is presented by a set of two-place predicates that permit three different case patterns (the corresponding case combinations are listed to the right of the data):

- (14) a. V domě páchla marihuana. Loc - Nom
 in house-L smell-Ps.3sg.f marihuana-N.sg.f
 'Marijuana smelled in the house.'
- b. V domě páchlo marihuanou. Loc - Ins
 in house-L smell-Ps.3sg.n marihuana-I.sg.f
 'There was the smell of marijuana in the house.'
- c. (Celý) dům páchl marihuanou. Nom - Ins
 whole house-N.sg.m smell-Ps.3sg.m marihuana-I.sg.f
 'The (whole) house smelled of marijuana.'

This set includes verbs such as *hemžit se* 'swarm', verbs of smelling, and verbs expressing intense coloring, equivalent to English expressions such as *The hill was yellow with daffodils*.

As the English translations show, the alternate patterns are not in free variation; the differences in case marking correspond to subtle differences in

meaning. Sentence (14a) emphasizes the second argument (marijuana) as the participant directly responsible for the resulting state. For this sentence to be felicitous, it must be true that somebody is currently smoking marijuana in the house. In contrast, when this argument is formally 'demoted' into the instrumental case (14b), its prominence is lost. In fact, the implication of (14b) is that there need not even be any marijuana in the house, let alone a smoker actively producing marijuana smoke. The version in (14c) is a rather subtle variation on (14b): while the latter has a certain existential quality, reporting a state (smell of marijuana) that exists in a particular location (house), the former presents the location as an entity to which the state is attributed as one of its characteristic properties. In fact, (14c) implies that the house itself exudes the smell, having been thoroughly steeped in it; that is also why the modifier *celý* 'whole' makes the sentence pragmatically more natural.

Assuming that the presence of a nominative NP in (14a) and (14c) makes subject identification unproblematic, let us suppose that these predicates subcategorize for a subject and an oblique function. We simply need to posit two different linking rules in order to account for the variation in case marking. But what would be the functional status of the arguments in (14b)? The case con-figuration in (14b) is semantically distinct from that in the other two sentences and this fact would necessarily be lost if we simply assigned subjecthood to one of the arguments. Moreover, such an assignment would have to be completely arbitrary since neither case form can have any subject properties. If they did, there would presumably be no reason to code them in non-nominative forms, since the nominative coding obviously is available.⁷ The only solution would be mapping them onto two oblique functions, which however does not really address the case marking. And since each case pattern is associated with a distinct meaning, some motivation for the case assignment should be established. I suggest to turn to the semantics of these predicates.

With some degree of imagination, their semantic valence could be perhaps specified as <Loc, Cause>, where 'Cause' should be understood as a generic substitute for initiators of states or actions.⁸ The case marking then becomes quite naturally motivated by the semantic roles. Czech expresses frame-internal locatives either by a locative PP or a nominative, with the subtle distinction in interpretation observed between (14a,b) on the one hand and (14c) on the other. Cause-like arguments also are associated with multiple forms: primary causes (the ones presented as being in charge, so to speak) are marked by the nominative, whereas secondary or demoted causes are marked most typically by the instrumental, less commonly by certain prepositional phrases. The differences in interpretation of the sentences in (14) are again consistent with the different coding of the cause argument. The information crucial for case assignment thus comes from the level of semantic roles rather than the abstract functional categories which after all are useful precisely in those cases where semantic distinctions have been neutralized.

Finally, compared to all these data, which involve sentences without any nominative NP, it is quite unproblematic to analyze sentences such as (15) below, which contain a dative experiencer but also a nominative NP:

- (15) Petrovi se nová pracovna líbila.
 Peter-D Rf new-N.sg.f study-N.sg.f appeal-Ps.3sg.f
 'Peter liked [his] new study.' (lit. '[his] new study appealed to Peter')

This configuration of case forms and semantic roles (usually described as <Exp, Stim>) is quite popular in the relevant literature concerning other languages (Sridhar 1979, Mohanan 1990, van Valin 1991), as an example of questionable subjects. It is a fairly straightforward issue in Czech, which uses this form with a handful of predicates of liking (*líbit se* 'appeal', *chutnat* 'appeal gustatorily', *vyhovovat* 'suit') and the verb *dařit se* 'be successful'.

Sentences of this kind are usually taken as examples of dative (or 'quirky') subject constructions, based on the behavioral properties of the dative NP, rather than on the formal subject properties of the nominative. However, the dative in this Czech sentence shows no subject behavior. For example, it cannot bind the reflexive possessive (see (16)), while the nominative must (see (17)):

- (16) Po celou dobu jeho_{i,j}/*svénemoci Petrovi_i nic nechutnalo.
for whole time-A his/*self's illness-G Peter-D nothing-N.n Ng-appeal-
Ps.3sg.n
'Throughout his_{i,j} illness, Peter_i didn't like [to eat] any food.'
- (17) Dana se ve svém/*jejím novém účesu všem líbila.
Dana-N.f Rf in self's/*her new hairdo-L all-D.pl appeal-
Ps.3sg.f
'Everybody liked Dana's new hairdo.' (lit. 'Dana in her new hairdo')

Also the equi NP deletion test works against the dative and in favor of the nominative with respect to subjecthood. In the following subject-control construction, the target of deletion is the argument that takes the nominative (*Dana*):

- (18) Dana se chtěla zalíbit každému.
Dana-N.f Rf want-Ps.3sg.f appeal-Inf everybody-D
'Dana wanted everybody to like her.' (lit. 'Dana wanted to be appealing to everybody')

In contrast, the experiencer dative cannot become the target, unless we use a different verb, such as *oblíbit si* '[come to] like', which implies more active, deliberate involvement on the part of the experiencer participant, and thus framing this participant as more of an agent. Consequently, the equivalent of the experiencer with *líbit se* 'be appealing' is marked by the nominative as shown in (19a) and it can become the target of equi NP deletion (19b):

- (19) a. Každý si Danu oblíbil.
everybody-N.sg.mRf Dana-A like-Ps.3sg.m
'Everybody liked Dana.'
- b. Každý se snažil si Danu oblíbit.
everybody-N.sg.mRf try-Ps.3sg.m Rf Dana-A like-Inf
'Everybody wanted to like Dana.' (...but she was a difficult person)

Once again, there seems to be no basis for analyzing the dative experiencer as the subject, regardless of our expectations based on English, where the experiencer maps onto the grammatical subject.

Given all these facts, what approach should one take to determining the functional status of all the questionable NPs? There are essentially two alternatives. By strict adherence to the Subject Condition as formulated in LFG, we would have to

assume that all these constituents are subjects after all. But such a solution would be very unattractive for several reasons:

(i) The notion of grammatical subject in Czech would become rather difficult to define. Since we would be adding constituents that do not behave like other subjects in the language, it would be practically impossible to find a property that would hold them all together as representing a single coherent notion. Needless to say, the diverse coding (nominative, dative, instrumental, and various prepositional phrases and adverbial expressions) would not be of much help either.

(ii) If we opted for a dative subject with the *libit se* type, we would be ignoring the fact that the valence contains a much better candidate for subjecthood, both in terms of coding and behavior. Furthermore, it would become inexplicable why some arguments marked by the nominative are subjects while others are not. The semantic valence alone could hardly be a factor, since there are other verbs with the same configuration of semantic roles which code the stimulus argument as the object and the experiencer as a nominative subject (e.g. *vidět* 'see', *slyšet* 'hear').

(iii) As a corollary to (ii), the dative-subject analysis of the *libit se* type would also complicate the linking rules responsible for the assignment of objects, since it is not a common property of the Czech objects that they appear in the nominative.

(iv) As a concern of more general nature, the introduction of oblique subjects would also have some undesirable consequences for the case assignment principles in the language. Czech exhibits quite consistent correlations between semantic roles and morphological cases, and particularly in oblique cases such as the dative, instrumental, and various prepositional phrases with locative meanings. The use of a specific case form in all of the oblique subjects is entirely predictable from the semantic roles associated with the arguments. By stipulating that the case is just a 'quirky' form of a subject, this semantic relationship would be severed and the form treated as unmotivated. This clearly is not in keeping with the nature of the language.

(v) Finally, it is unclear why we should insist on maintaining that the predicates in question have a subject in spite of all the evidence to the contrary, since Czech evidently tolerates the absence of grammatical subjects in other cases. Recall the subjectless sentences in (2).

This rather brute force solution could be softened by adding a link between the semantic roles and the cases, following Neidle's (1982) proposal for certain Russian data. But the very need to take such a step amounts to acknowledging that the abstract functional level is meaningless in these cases.

The alternative, then, to declaring subjecthood by *fiat* would be to leave the predicates that do not take any nominative subjectless, and in the *libit se* type to assign subjecthood to the stimulus argument. The arguments in question would thus map onto various oblique functions. But the oblique function in itself carries very little information of any value, unless it also specifies a particular semantic role it is linked to. And this very property of the obliques essentially reduces the representation to two levels only -- the semantic roles and the cases. The oblique function can thus be easily eliminated as an essentially superfluous mediator between two levels (semantic roles and case forms) that in fact do not need any mediation.

The minimum conclusion we can draw from all this is two-fold: (i) morphological cases in Czech cannot be just expressions of grammatical functions, as is assumed in representations of the type in (1a), and (ii) the Subject Condition cannot be a factor in evaluating the well-formedness of a large number of Czech sentences.⁹

The question is what else we can do with this finding. We can of course leave it at that -- this alone creates enough problems for stating well-formedness conditions for the language as a whole. But we can also take it as a basis for advancing the hypothesis that the grammar of Czech may not employ the notion of grammatical subject as a necessary category. There is certainly some evidence from other processes in the language that this might be so (agent-demoting constructions, equi NP deletion, general case assignment principles, etc. which are all sensitive to the information provided by the semantic valence), and it would not be merely a quirk of Czech, either; it has been suggested for other languages as well (LaPolla 1990, Bhat 1991). More research is required in order to really answer this question in all its complexity, but the data discussed in this paper justify at least a tentative suggestion that adequate representation of the Czech sentence might be structured along the following lines:

- (20) semantic roles
 |
 case
 |
 (GF)

Specifically, the level of grammatical functions may not be central to the organization of any Czech sentence, although it perhaps can be tagged on as an optional level of description (indicated by the dotted line and parentheses in (20) above) for formulating certain general typological statements.

Footnotes:

* I am grateful especially to Chuck Fillmore, Knud Lambrecht, and Alan Timberlake for helpful discussions and suggestions.

¹ The subject NP may not be always present, though, because Czech is a PRO-drop language.

² I will be using the following abbreviations: N-nominative, G-genitive, D-dative, A-accusative, L-locative, I-instrumental, sg-singular, pl-plural, m-masculine, f-feminine, n-neuter, Ps-past, Pr-present, Ppl-past participle, Aux-auxiliary, Inf-infinitive, Rf-reflexive, Ng-negative.

³ Czech uses the prepositionless dative to mark recipients, both in the literal and figurative sense.

⁴ It should be pointed out, however, that the use of any possessive pronoun (non-reflexive or reflexive) in these sentences is redundant and therefore somewhat artificial. The most neutral expression of coreferentiality within a clause is the absence of any pronoun. For example, the more neutral equivalent to (6) is *Petrovi se po š nečekaném vítězství řádně ulevilo*, without any pronoun, and it is unambiguously interpreted as 'Peter_i felt a great relief after his_{i,*j} unexpected victory'.

⁵ I will ignore expressions such as *Otce píchalo u srdce* 'Father had a chest pain' and *Helenu polilo horkem* 'Helen got all flushed', which involve case patterns Acc-PP and Acc-Ins, respectively. They both can be traced to deagentive constructions which demote the agent argument (subject) without promoting the patient.

⁶ It is in fact impossible to construct a sentence in which the reflexive could have the locative as its antecedent. The referent of this argument is always a body part which

is not easily cast in the role of a possessor of anything. Yet the possessive relationship is essential in this test.

⁷ The locative/nominative alternation in (14a,b) vs. (14c) clearly belongs to the type of alternation that is frequently associated with the difference between a partitive vs. holistic reading, respectively, as argued by Anderson (1971). And it is interesting to note that he explicitly correlates the oblique form with non-subjects.

⁸ If we wanted to use the standard inventory of semantic roles, Cause would correspond to agents (for verbs like *hemžit se*) and causes/sources/stimuli (for verbs like *vonět* 'give off a fragrance'). In either case, what I have in mind is a participant held somehow responsible for the action or state.

⁹ In fact, this conclusion also affects theories that do not work with the traditional notion of subject as such. The facts presented in this paper pose problems for any framework that operates on the assumption that every sentence must have a privileged argument, whether it is the 'external argument' in a GB-style analysis, i.e. the highest structural position subject to the Extended Projection Principle (Williams 1981), or RRG's 'syntactic pivot' as a privileged function (van Valin 1991). With respect to the Czech data, it is not clear which of the arguments in question could be the privileged or external arguments and on what basis these notions should be applied in the first place.

Bibliography:

- Anderson, Stephen R. 1971. On the Role of Deep Structure in Semantic Interpretation. *Foundations of Language* 6: 197-219.
- Bhat, Shankara D.N. 1991. Grammatical relations: the evidence against their necessity and universality. London: Routledge.
- Bresnan, Joan and J. Kanerva. 1989. Locative inversion in Chichewa: A case study in factorization in grammar. *LI* 20: 1-50.
- Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*. Chicago: The University of Chicago Press.
- Comrie, Bernard. 1984. Subject and object control: syntax, semantics, pragmatics. *BLS* 10: 450-464.
- Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Chicago: The University of Chicago Press.
- LaPolla, Randy J. 1990. Grammatical relations in Chinese: Synchronic and diachronic considerations. UC Berkeley dissertation.
- Li, Charles N. and S. A. Thompson. 1976. Subject and topic: A new typology of language. In Li & Thompson (eds.), *Subject and Topic*.
- Mohanan, Tara W. 1990. Arguments in Hindi. Unpublished dissertation. Stanford University.
- Neidle, Carol. 1982. Case agreement in Russian. In J. Bresnan (ed.), *The Mental Representation of Grammatical Relations*: 391-426. MIT Press.
- Sridhar, S. N. 1979. Dative Subjects and the Notion of Subject, *Lingua* 49: 99-125.
- van Valin, Robert D. 1991. Another Look at Icelandic Case Marking and Grammatical Relations. *Natural Language and Linguistic Theory* Vol. 9/1: 145-194.
- Williams, Edwin. 1981. Argument Structure and Morphology. *The Linguistic Review* 1: 81-114.