

Rise of canonical subjecthood

Ilja A. Seržant

University of Bergen

In this paper I have examined several instances in which an oblique constituent acquires canonical subject marking, i.e. nominative case and verbal agreement (in an accusative language). These instances show that an oblique constituent may acquire subject coding properties without being beforehand endowed with subject behavioural properties, if two requirements are met: (i) this oblique constituent must exhibit a considerable functional overlap with the prototypical subject in the given language and, (ii) there must be either no canonically case-marked subject in the construction at all, or the coding subject properties must be assigned to a constituent that has less functional-semantic overlap with the prototypical subject than the oblique constituent. Furthermore, I claim that there is often some minor semantic change concomitant with the acquisition of subject coding properties. I have also introduced the *control over the pre-stage property* (CoP) which is a weaker entailment than Dowty's (1991) *volitional involvement in event or state*. It only denotes whether or not the experiencer had the choice to resist the experience to come about. Differently from Dowty's (1991) approach, which presupposes that the proto-role entailments are lexical and provided primarily by the predicate, it is assumed that some of the proto-role entailments may also stem from the case frame. This becomes especially obvious with the labile predicates that allow for more than one case frame, each resulting in different sets of the proto-role entailments.

1. Introduction

The present paper aims at investigating the *canonicization* processes, that is the diachronic process by which an oblique, highest ranked argument acquires full-fledged subjecthood and becomes a canonical subject with canonical subject morphosyntax, i.e. (for the languages of concern) case-marked with nominative case and capable of triggering verbal agreement.

For this purpose I concentrate only on recently evolved nominative subjects. Moreover, only those instances will be treated where the original case frame is still present in the language, i.e. where the variation is motivated by the diachronic changes. An inquiry into such a near-to-synchronic kind of diachronic changes is the most reliable

tool for the diachronic typology (Croft 2003: 247). Specifically, it will allow testing the source case frame as to its semantics, function and morphosyntactic properties.

Extensive research on non-canonical subjects and their relation to canonical ones is taking place (cf., *inter alia*, Aikhenvald et al. 2001; Bhaskararao & Subbarao 2004; Barðdal & Eythórsson 2005), as well as research on synchronic alternations in the case marking of the subject, i.e. the *Differential Subject Marking* (Kittilä 2002; de Hoop & Swart 2008, *inter alia*). The latter typically includes (at least) one less-canonical subject marking option. However, much less attention has been paid to the diachronic dimension thereof, more specifically, to the question about how oblique constituents develop into canonical subjects.

The Germanic languages constitute here an important exception. Here, the development from a non-canonically marked, less prototypical subject into the canonical one has been thoroughly studied, e.g. in Old English (Allen 1995; Tousdale 2008) or Old Swedish (Falk 1997). An inherent difficulty of this type of data has always been that these languages were affected by a major restructuring of the morphological case, such as the loss of case in the history of both English and Swedish, which is why the development from an oblique case-marked subject-like argument into a canonically marked nominative subject might be biased by the general loss of case at some level.

At the same time, in languages with a rigid and stable case system, such as Lithuanian or Russian, such investigations have not been carried out yet. Exactly for this reason, the canonicization processes in these languages is highly relevant for historical morphosyntax, because canonicization of case marking in these languages is a much less frequent phenomenon pertaining to only a few predicates (not significantly more in number than those studied in the present paper). It is therefore not immediately interrelated with any major morphosyntactic process, as is the case with the Germanic languages. That is to say, the canonicization is not a part of a more complex change in these languages. This is why rather general explanations, e.g. in terms of productivity of certain case frames, that would have to apply to a class or a subclass of verbs, represent rather an insufficient motivation for the canonicization processes presented in this paper.

The process of the acquisition of the canonical subjecthood is a process that can be characterized as the progression along the grammaticalization cline, i.e. the progression towards *increase of internal dependencies* (Haspelmath 2004; Givón 1979: 208). The latter is found, e.g. in the establishment of obligatory verbal agreement. Additionally, other properties of a canonically case-marked subject in the languages of concern, such as the ability to control the subject reference of a conjoined clause, may be interpreted in these terms (except for the North Russian perfect (Section 5), whose non-canonical subject had this ability before the canonicization). Furthermore, the very change from lexical case assignment to the structural case assignment brings about what has been referred to as *semantic bleaching* in

the theory of grammaticalization, i.e. an increase in the degree of abstraction, since the structural cases generally contribute less information than do the lexical cases. Thus, as will be illustrated below, there are more contexts potentially compatible with the canonical subject than with its oblique counterpart in the same construction.¹ Another characteristic of this change is that “a grammatical unit assumes a more grammatical function” (cf. Heine et al. 1991:2; Traugott 2003:645).

In the following I will discuss six types predicates from Lithuanian, Russian, North Russian and German that have acquired canonical subjects. Sections from 2 to 6 are devoted to particular predicates, discussing the semantic differences between the canonical and the original non-canonical case frame, and giving a syntactic analysis of the respective oblique arguments before they acquire the canonical case marking and verbal agreement, in order to see how much syntactic subjecthood was present before the canonical subjecthood has been acquired. In Section 7, I summarize the empirical findings of the previous Sections and provide a theoretical account.

2. Lithuanian *su-/at-šalti* ‘to get cold’ and *su-šlapti* ‘to get wet’

2.1 Data

In Example (1) the Dat-Nom case frame of the Lithuanian verb *at-šalti* ‘to get cold’ (perf.) is illustrated. It consists of a dative case-marked experiencer (cf. *external possessor* König & Haspelmath 1997) and a nominative marked body part (Theme) with the 3rd person verb agreement with the Theme:²

- (1) *Man nuo lietaus atšalo rankos* (Lithuanian)
 I:DAT from rain froze:PST.3(PL) hand:NOM.PL
 ‘My hands got cold because of rain.’

1. In this respect these instances are different from similar cases known from other languages as, e.g. Icelandic or Faroese, where, as Eythórsson (2002:202) reports, there is no difference between the secondary nominative and the original oblique case-marked subject in terms of its semantic properties.

2. A finite verb in Baltic does not morphologically distinguish the number in the third person, both singular and plural having the same form. Nevertheless, the underlying grammatical number of the predicate can be determined on the basis of periphrastic tenses, in which the participle form of the main verb agrees with the subject in number and gender. I gloss all finite third person verb forms with a number in brackets (sg) in order to better demonstrate the agreement.

The cognate Latvian verb *sa-salt* ‘to get cold’ (perf.) also takes the Dat-Nom case frame:

- (2) *Man sasala rokas* (Latvian)
 I:DAT get.cold:PST.3(PL) hand:NOM.PL
 ‘My hands got cold.’

The same pattern is found with Lithuanian *su-šlapti* ‘to get wet’:

- (3) *Man sušlapo kojos* (Lithuanian)
 I:DAT get.wet:PST.3(PL) foot:NOM.PL
 ‘My feet got wet.’

The Latvian cognate corresponds here exactly to the Lithuanian one:

- (4) *Man saslapa kājas* (Latvian)
 I:DAT get.wet:PST.3(PL) foot:NOM.PL
 ‘I got my feet wet.’

In the following I will refer to Dat-Nom as to a case frame.

Before I continue with the discussion of this case frame, some remarks on the syntactic and semantic status of the Dat are in order. The syntactic status of the dative constituent is somewhat controversial. Thus, one could theoretically downplay the syntactic status of the dative argument to the one of an external possessor, i.e. an adjunct, just on the basis of the possessive semantics between the experiencer and the body part argument. Notably, the very possession relationship between the dative and the nominative referent does not principally entail anything about their syntactic status, exactly as the possession relationship between the agent and patient in (5) does not preclude the former from being analyzed as subject and the latter from being analyzed as object:

- (5) *Es mazgāju rokas* (Latvian)
 I:NOM wash:PRES.1SG hand:ACC.PL
 ‘I wash (my) hands.’

The dative argument is historically an external possessor indirectly affected participant that was not an argument of the verb to begin with. At the same time, synchronically, it is an obligatory constituent. Without it, the sentences (1)–(4) would have been either infelicitous or understood as ellipsis of the experiencer. Thus, it cannot be considered an adjunct at the present stage. I cannot go into a detailed discussion about which level of the syntactic organization makes the dative constituent obligatory here. I only note that it cannot be just the possessum (i.e. the body part) NP alone, as it would have been the case with the internal possessor. The difference between internal and external possessors is that the latter primarily enters a semantic relationship with the predicate denoting first and foremost the experiencer of the event, the possessive relation being rather a “by-product” motivated pragmatically (one cannot experience something at someone else’s body part). The former, in turn, establishes primarily a semantic and

syntactic relationship with its host NP only. It is thus natural for the original external possessors/affected participants to intrude the verb's valence and become a core argument thereof in the diachronic perspective.

Now, coming back to the discussion of the Dat-Nom case frame, one can state that given the obligatoriness and high status on the Referentiality Scale the Dat constituent became the most subject-worthy argument of the clause. This is why we find the development I am going to describe in this Subsection.

The Dat-Nom case frame as in (1)–(4) is the only available case frame to express the experiencer and the body part in Latvian. Lithuanian, however, allows these verbs also to take the Nom-Acc case frame, where the experiencer is encoded with nominative case and the body part with accusative case, cf. (6) and (7):

- (6) *Aš atšalau rankas, kol ģjau*
 I:NOM get.cold:PST.1SG hand:ACC.PL while went
į universitetą (Lithuanian)
 to university
 'While I was going to the university, I got my hands cold.'

- (7) *Aš sušlapau kojas, kol ģjau*
 I:NOM get.wet:PST.1SG hand:ACC.PL while went
į universitetą, (Lithuanian)
 to university
 'While I was going to the university, I got my hands wet.'

This Nom-Acc case frame in Lithuanian is a recent innovation with these verbs. It is not attested in either Old Lithuanian or closely related Latvian. Furthermore, the morphology of this verb, i.e. the nasal infix in the present stem (*šqla* 'it gets cold' from *ša-n-la, *šla-m-pa* 'it gets wet') with the past suffix -o-, signals that the nominative argument of this verb is semantically patient-like (cf. Arkadjevas 2008: 8–9). In the most cases, verbs with this middle-like morphology are formally intransitive in Baltic. This also suggests that there was no transitive alignment with this verb to begin with.

Superficially there appears to be no semantic difference between Dat-Nom and Nom-Acc case frames, since Nom-Acc in (6) can readily be replaced with Dat-Nom and vice versa in (3), with seemingly no consequence for the semantic interpretation (pace Holvoet, this volume). However, as cross-linguistic comparison shows, there is often a tiny difference between the dative and nominative experiencers even with one and the same predicate (cf. Croft 1994: 51f). Thus, Montaut (2004: 44–45) states that the semantic feature that discriminates the nominative experiencer from its dative counterpart in Hindi/Urdu is *reflexive consciousness*, entailing whether the experiencer is conscious about his experience or not (see also Montaut, this volume, for details; Mohanan 1994). In this language, contexts that explicitly state *reflexive consciousness*

allow nominative experiencers only. While *reflexive consciousness* is not specifically the distinctive feature that demarcates the semantic concepts encoded by both case-markings in Lithuanian, there is another tiny difference in terms of agentivity properties on the part of the experiencer with the Nom-Acc case frame. I shall demonstrate this in the following discussion.

2.2 Imperative

The fact that the verbs *su-šlapti* and *at-šalti* occur as imperatives only with the Nom-Acc case frame as in (8), whereas the imperative results in ungrammaticality with the Dat-Nom case frame, is semantically conditioned: if an addressee is requested not to admit a situation (*to get cold* in this case) to come about, it implies that the addressee is construed as having the control over whether or not the situation requested comes about (cf. similar argumentation in, *inter alia*, Rögnvaldsson 1996: 48 for Icelandic; Bickel 2004: 95 for Nepali; Montaut 2004: 44 for Hindi, Comrie 2004: 121 for Tsez):

- (8) *Ne-atšalk* *kojū* *basas* *vaikščiodamas!* (Lithuanian)
 not-get cold:IMPV.2.SG leg:GEN.PL barefoot walking
 ‘Don’t get your feet cold, while walking barefoot!’

The incompatibility of the Dat-Nom case frame with imperatives has alternatively been explained in purely morphosyntactic terms, e.g. by stipulating that only nominative subjects can be the addressees of an imperative in many languages. While this restriction is also valid for Lithuanian, I assume that this morphosyntactic explanation is not self-sufficient, and that there must be a functional-semantic motivation behind it.

The nominative case-marking always has the implication that there is a choice at the stage preceding to the event on whether or not to admit the situation to come about, while this is not necessarily true with datives. The compatibility of the Nom-Acc is motivated by the fact that the nominative case-marked addressees are endowed with control in the pre-stage of the event, whereas the incompatibility of Dat-Nom with imperative signals the lack of such a choice on the part of the Dat case-marked argument. Note that neither case frame entails control during the event the respective verbs refer to but rather – if at all – “the primary responsibility for the event” in DeLancey’s (DeLancey 1984, 1985: 6) terms. Indeed, there is more evidence that the Nom-Acc case frame presupposes that the Nom experiencer has the *control over the pre-stage* of the event (henceforth: CoP) referred to by the verb, which I now turn to.

2.3 Contexts with another participant having the CoP

The following examples independently illustrate the lack of the CoP entailment on the part of the experiencer with the Dat-Nom case frame. In both (9) and (10) it is the *wind* that logically has the CoP property, and this is why (9) is fully natural while (10) sounds less acceptable for some native speakers:

- (9) *Nuo vėjo man atšalo rankos* (Lithuanian)
 from wind I:dat get.cold:PST.3(PL) hands:NOM
 ‘My hands got cold because of the wind.’
- (10) *?Nuo vėjo aš atšalau rankas* (Lithuanian)
 from wind I:NOM get cold:PST.1SG hands:ACC
 [Intended] ‘My hands got cold because of the wind.’

The following context implies that there is another participant that is endowed with the CoP, namely, the *doctor*. In this context, only Dat-Nom of (11a) is grammatical, whereas Nom-Acc in (11b) results in ungrammaticality. This is because Nom-Acc would presuppose CoP at the nominative participant which is incompatible with the context provided by the preceding sentence:

- (11) *Gydytojas ant skaudančio piršto uždėjo ledų, ir po*
 doctor on aching finger put ice and after
dešimties minučių (Lithuanian)
 ten minutes
- (a) *man pirštas visai atšalo*
 I:DAT finger:NOM fully get cold:PST.3(SG)
- (b) **aš pirštą visai atšalau.*
 I:NOM finger:ACC fully get cold:PST.1SG.
- ‘The doctor put ice on [my] aching finger and after 10 minutes
my finger got cold.’

The Dat-Nom case frame encodes only the perception of the physical state by the experiencer without him having any responsibility for it, i.e. the lack of the CoP property on the part of the experiencer. In turn, with Nom-Acc, it is the experiencer that carries the responsibility for having let the state come about, even if it was not an intentional admission (i.e. non-volitionally).

Significantly, the CoP is close to Dowty’s (1991) Proto-Agent property *volitional involvement in event or state* which presupposes the presence of the CoP. However, inversely, the presence of the CoP does not always presuppose the *volitional involvement in event or state*. Thus, in order to endow the experiencer with the control over the whole event/result (*volitionality*), one has to use the causativized verb forms of the verbs under discussion:

- (12) *Skaudančią galvą*
 aching:ACC.SG.F head:ACC.SG.F
šal-dži-au ledu (Lithuanian)
 get.cold-CAUS-1.SG ice:INSTR.SG
 ‘I cooled down my aching head with ice.’

In (12), the Experiencer is construed as having the overall control over the event and acting volitionally during the whole event time. This is signalled by the causative verb stem formation derived by means of the causative suffix *-dy/-dži-*. In contrast, in (8) repeated here as (13) for convenience, the addressee is only requested to obey the speaker's volition. The context does not necessarily imply volitionality on the part of the addressee, but does always imply its endowment with the CoP:

- (13) *Ne-atšalk* *kojū* *basas* *vaikščiodamas!* (Lithuanian)
 not-get cold:IMPV.2SG leg:GEN.PL barefoot walking
 'Don't get your feet cold, while walking barefoot!'

I assume that Nom-Acc does always imply the CoP, whereas *volitionality* is something that can only (optionally) be inferred contextually, syntactically by an adverb, or by means of the causative formation. Crucially, the non-volitional contexts – which are also the most natural and frequent ones given the semantics of the verbs under investigation – do not result in ungrammaticality with the Nom-Acc case frame (cf. Masica 1991: 350 and Onishi 2001: 37–8 for parallels). This is because the Nom-Acc case frame presupposes only the CoP, and not volitionality, on the part of the experiencer, cf. (14), which – although non-volitional – is fully grammatical:

- (14) *Aš* *netyčia* *atšalau* *rankas* (Lithuanian)
 I:NOM accidentally:ADV get cold:PST.1SG hands:ACC
 'I've accidentally got my hands cold.'

2.4 Inchoative (change-of-state) *su-/at-šalti* 'to get cold' versus processual/stative³ *šalti* 'to be cold'

Another fact reflective of semantic differences between the Dat-Nom and Nom-Acc case frame is the difference in selectional restrictions on the predicate's input. Thus, the transitive Nom-Acc is only compatible with the inchoative/telic (change of state) stem of this verb, i.e. only with the verb "perfectivized" in the "Slavic manner" by means of a lexically-empty prefixation (bounder) *at-/su-šalti* 'to get cold' and *at-/su-šlapti* 'to get wet', while it is ungrammatical with their stative counterparts *šalti* 'to be cold' and *šlapti* 'to be wet':⁴

3. In principal, these verbs may also be regarded as statives in this specific (experiential) meaning of physical sensation as they do not imply any gradience in this meaning.

4. In fact, even though superficially reminiscent of the Slavic aspect systems, Lithuanian does not have a grammaticalized aspectual opposition but rather different kinds of actionality derivations (such as the one above), see Arkadiev (2011). The actional type of the verb *šlapti* can also be defined as processual (atelic 'to be getting gradually more and more wet') in its physical meaning, while the experiential meaning of this verb does not entail any increase in wetness but rather the very fact that a particular body part *is* wet and is hence in this meaning stative.

- (15) a. **Aš darbe nuolat*
 I:NOM work:LOC constantly
šqlu/šalau/šaldavau *kojas*
 be.cold:PRES/PAST/ITERAT.PST.1SG foot:ACC.PL
- b. *Man darbe nuolat*
 I:DAT work:LOC constantly
šqla/šalo/šaldavo *kojos*
 be.cold:PRES/PAST/ITERAT.PST.3 foot:NOM.PL
 ‘My feet get constantly cold in the office (because there is no heat there)’.

This distributional divergence in selectional restrictions is indicative of the denotational divergence, the differences in distribution being generally valid arguments for the respective functional differences in typology (Kibrik 1992). Thus, the compatibility of the Nom-Acc case frame with the inchoative/telic predicates only reveals that its semantics is divergent from the Dat-Nom frame. This is also expected, since the experiential change-of-state verbs generally outrank the experiential statives on transitivity scale, which is why the former but not the latter tend to the agent-like (here: nominative) case-marking of the experiencer (cf. Croft 1993 on crosslinguistic perspective).

To sum up, I argue that both Nom-Acc and Dat-Nom case frames do have small distinctions, in terms of symmetry with the latter and asymmetry with the former between the experiencer and the body part (*pace* Holvoet, this volume). Furthermore, while the stative *šalti* and *šlapti* exclude the endowment with the CoP of their experiencer argument, the respective telic (change-of-state) *su-šalti* and *su-šlapti* are labile. These verbs can equally well combine with the Dat-Nom and the Nom-Acc case frame and, hence, with the CoP entailment on their experiencer. Thus, the argument structure construction or, more specifically, the very case frame also has its own entailments on the arguments (cf. Goldberg 1995). I have claimed that only the Dat-Nom case frame entails lack of the CoP feature on the part of the experiencer, whereas only the Nom-Acc is compatible with its presence. Consequently, the inchoative/telic verbs at stake are underspecified or ambiguous with regard to the CoP entailment, and it is only the respective case frame that provides for disambiguation.

I assume, thus, that the CoP needs inherent prerequisites in the predicate, though becomes activated only if the respective case frame (i.e. Nom-Acc) is selected.

2.5 Subjecthood

The exact syntactic status of the dative experiencers of these verbs is difficult to determine. On the one hand, they have certain subject properties, such as their linear position in the unmarked word order that is typical (but not exclusive) of subjects in Lithuanian. Furthermore, the experiencer argument fits the notion of subject in

terms of its inherent lexical and referential properties: it is always animate, conscious, indispensable and referential (cf. Keenan 1976; cf. Dahl 2000 for the statistically significant correlation between animacy and subjecthood; de Hoop, this volume). On the other hand, it fails to control the co-referential subject NP in all kinds of clause linking strategies of Lithuanian, cf. control infinitives:

- (16) **Nenoriu sušlapti/sušalti kojoms*
 want:1.SG.PRES get.wet:INF/get.cold:INF foot:NOM.PL
 ‘I don’t want my feet to get wet/cold.’

Furthermore, the choice of the converb can be used as a subjecthood test in Lithuanian. This language has a kind of switch reference in the converb clause. There are two converbs: the one, ending in *-nt*, signals that the subject of its clause is *not* co-referential with the subject of the main clause, while the second converb formed with the suffix *-dam-* indicates that the subject of the matrix predicate *is* co-referential with the logical subject of the converb clause (Ambrazas 1997: 675). Thus, the choice of the converb can be used as a subjecthood test in Lithuanian. The next example shows that the dative experiencer is not the subject:

- (17) a. *Tiksliau, man vaikščioja-nt jos atšalo*
 surely I:DAT walk-OBL.CONV it:NOM.PL.F get-cold:PAST.3(PL)
 ‘More precisely, I got them [scil. the lips] cold surely while walking.’
 (<<http://www.supermama.lt/forumas/lofiversion/index.php/t530773.html>>)

The speaker has used here the *-nt* converb, which cannot be used with unequivocal matrix subjects in Lithuanian. At the same time, the use of the *-dam-*converb, which signals the co-referentiality with the subject of the matrix predicate, only makes the sentence ungrammatical:

- (17) b. **Tiksliau, man vaikščio-dama(s)*
 more precisely I:DAT walk-SUBJ.CONV:F/(M)
jos atšalo
 it:NOM.PL.F get-cold:PST.3(PL)
 Intended meaning: ‘More precisely, I got them [scil. the lips] cold while walking.’

The *-dam-*converb is half-inflected, i.e. it agrees with the matrix subject in number and gender but it cannot inflect for case. Thus, one could argue that this converb is morphologically nominative case-marked and, hence, cannot combine with the potential “dative subject” just because of case disagreement. However, crucially, the *-dam-*converb does not combine with every nominative, but only with those nominatives that are subjects. Thus, its logical subject cannot be co-referential with a nominative case-marked object or with a nominative case-marked time adjunct. Thus, the compatibility with the *-dam-*converb indicates the syntactic status of its controller

and is, hence, a reliable subjecthood test in Lithuanian. Furthermore, the very fact that the dative experiencer can be co-referential with the logical subject of the *-(a)nt* converb reveals its being not a subject, since the logical subject of the *-(a)nt* converb can be co-referential, in principle, with any mentioned and even not mentioned referent but the subject (Ambrazas, loc. cit.). Hence, it is not only the incompatibility of the *-dam-*converb but it is also the compatibility with the *-nt* converb that make the subject analysis highly improbable.

In addition, only the Nom-Acc case frame controls the reflexive anaphora, not the dative experiencer:

- (18) a. **Man atšalo kojoms dėl savo kaltės*
 I:DAT get.cold:PRES.3(PL) foot:NOM.PL due refl.GEN fault:GEN
 Intended meaning: ‘It is my own fault that I got my feet cold.’
- b. *Aš atšalau kojas dėl savo kaltės*
 I:NOM get.cold:PRES.1SG foot:ACC.PL due refl.GEN fault:GEN
 ‘It is my own fault that I got my feet cold.’

Summing up, I conclude that the dative experiencer with the verbs *su-/at-/nu-šalti* and *su-šlapti* lacks a number of syntactic subject properties (see also Holvoet, this volume), but reveals semantic subject properties, as well as appearing sentence initially in the unmarked word order. Despite lacking any syntactic properties of a subject, it has developed into a canonical, full-fledged nominative subject with the historically secondary Nom-Acc case frame. I will provide an account for this in detail below in Section 7.

3. Lithuanian (*pri*)*trūkti* ‘to be short of’, *už-tekti* ‘to have/be enough of’

3.1 Data

The Lithuanian verbs (*pri*)*trūkti* ‘to be short of’ and *už-tekti* ‘to have enough of’ occur in Standard Lithuanian and allow its experiencer argument to be marked either with dative (20) or with nominative (19) (Siližienė 1994–2000:II(2).350–1; LKŽ, *sub verbo*; Ambrazas 1997:663):

- (19) a. *Aš pritrūkau pinigų*
 I:NOM be.short:PAST.1.SG money:GEN.PL
 ‘I was short of money.’
- b. *Aš užtekau pinigų*
 I:NOM have.enough:PST.1.SG money:GEN.PL
 ‘I had enough money.’
- (20) a. *Jam pritrūko pinigų*
 he:DAT.SG be.short:PST.3.(SG) money:GEN.PL
 ‘He didn’t have money (for something).’

- b. *Jam užteko pinigų*
 he:DAT.SG have.enough:PST.3.(SG) money:GEN.PL
 ‘He had enough money (for something)’.

Parallel to the case of *su-šalti* ‘to be cold’, *su-šlapti* ‘to be wet’, the dative case-marking is the older one, while the nominative marking is secondary in Lithuanian with this verb as well. The nominative case marking is not attested in the closely related Latvian *pie-trūkt* ‘to be short of’ and *pie-tikt* ‘to have enough of’, which only allow Dat-Gen or Dat-Nom (cf. Berg-Olsen 2005: 186–7; 2009). Furthermore, I have not found any attestations of the Nom-Gen case frame in the Old Lithuanian texts.

In the following discussion I will discuss only *pri-trūkti* ‘to be short of’ as *už-tekti* ‘to have enough’ is exactly parallel to it.

3.2 Subjecthood

The syntactic status of the dative experiencer is the same as the one of the dative experiencer with the verbs *su-/at-šalti* ‘to get cold’ and *su-šlapti* ‘to get wet’. On the one hand, it occupies the linear position in an unmarked word order that is typical (but not exclusive) of subjects in Lithuanian.

The following example shows that the dative experiencer does not control the reference of the subject converb, instead the non-subject converb in *-nt* has to be used:

- (21) *Sudarinėjant programą, man pritrūko vienos grupės, kuri*
 compose:NON-S.CONV programme:ACC I:DAT was.short:3(SG) one
 group:GEN which

‘While composing the programme I need one group, which ...’.

(<http://www.scribd.com/doc/14759998/15min-Kaunas-20080424>)

The speaker has used here the *-nt* converb that cannot be used with subjects in Lithuanian. This converb indicates a switch in reference in the participial clause to a non-subject referent, be it coreferential with one of the arguments of the preceding clause or not. Similar to example (17b) above, the use of the *-dam*-converb, which indicates co-referentiality with the subject of the matrix predicate, would make the sentence in (21) ungrammatical.

Only the nominative case-marked experiencer in the Nom-Acc frame can control the reflexive anaphora (22b), while the dative experiencer fails to do so (22a):

- (22) a. **Man pritrūko pinigų dėl savo kaltės*
 I:DAT be:short:PST.3.(SG) money:GEN.PL due refl.GEN fault:GEN
 Intended meaning: ‘It is my fault that I was short of money.’

- b. *Aš pritrūkau pinigų dėl savo kaltės*
 I:NOM be:short:PST.1.SG money:GEN.PL due refl.GEN fault:GEN
 ‘It is my fault that I was short of money.’

Summing up, I conclude that the dative experiencer with the verb *pri-trūkti* lacks most of the syntactic subject properties, except, perhaps, the first position in the unmarked word order.

4. The Russian verb *vyrvat* ‘to vomit’

4.1 Data

This verb takes only an accusative-marked experiencer and is a one-place predicate in Standard Russian, as in (23a):

- (23) a. *Menja vyrvalo* (Standard Russian)
 I:ACC vomited:PST.NEUTR.SG
 ‘I vomited.’

However, in Colloquial Russian of some native speakers the nominative case-marking is equally possible, cf. (23b):

- (23) b. *Ja vyrval* (Colloquial Russian)
 I:NOM vomited:PST.M.SG
 ‘I vomited.’

The nominative case marking is more frequent in the speech of the younger generation, while the speakers of the older generation mostly do not accept the nominative case marking at all. It is difficult to establish an exact date when the nominative case-marking appeared, it cannot, however, be earlier than in the 20th century, since the Russian National Corpus (RNC) including a number of texts from the 19th century does not provide any example of the intransitive verb *vyrvat* ‘to vomit’ with the nominative case-marked experiencer.

The event in (23a–b) is expectedly non-volitional and uncontrolled. Though if one has to vomit, e.g. for the sake of recovery, then, for those speakers of Colloquial Russian who allow the nominative case marking, the nominative case-marking of the experiencer has to be used. Thus, *vz'al* in (24) below explicitly indicates a volitionally triggered event:

- (24) *Da i pogladite na životnyx: čto-to ne to sjel, libo volosy*
 just and look at animals: something not this eaten or hairs
v želudke komkom – vz'al i vyrval
 in stomach, – right away vomit:PST.M.SG
 ‘Just look at the animals: if an animal has eaten something not right or has
 hairs in its stomach, – [then] [it just] vomits right away.’
 <<http://forum.materinstvo.ru/index.php?showtopic=65103andst=10>>

This is, however, not true the other way around: if the event is non-volitional, then both Nom and Acc can be used in Colloquial Russian. The presence or absence of

volitionality on the part of the highest ranked argument is well known to be the conditioning factor behind canonical versus non-canonical case-marking alternations, in combination with additional verb morphology (cf. McCawley 1976; Klaiman 1980; Melis & Flores 2012). However, in some cases, as with Colloquial Russian *vyrvat*, there is no additional verb morphology involved, and the presence/absence of volitionality on the experiencer argument is entailed by the case frame only: only the Nom marked experiencer is compatible with volitionally on the part of the experiencer, while the Dat marked experiencer is not.

The Nom case-marking also invokes other effects that signal an increased degree in transitivity, as opposed to the original accusative case-marking: it allows the encoding of the theme as the direct object, which is not possible with the original accusative case-marking, cf. (25):

- (25) *Rebenok vyrval kašu* (Colloquial Russian)
 child:NOM vomit:PST.M.SG porridge:ACC.SG
 ‘The child vomited the porridge.’

In Standard Russian the theme would have to be encoded by the adjunct with the instrumental case, cf. (26):

- (26) *Rebenka vyrvalo kašej* (Standard Russian)
 child:ACC vomit:PST.NEUTR.SG porridge:INSTR.SG
 ‘The child vomited the porridge.’

Analogically to the case of the Lithuanian verbs *su-šalti* and *su-šlapti* (subsection 2.4 above), there is a clear preference for the perfective (telic) stem *vy-rvat* ‘to vomit once’ with the nominative experiencer, while its imperfective (atelic) counterpart *rvat* ‘to vomit (iteratively or duratively)’ is less natural with it. Here, I also assume that the correlation between the telicity feature with the nominative case on the experiencer argument, and the inverse, i.e. the correlation of stativity with the dative case-marking on the experiencer argument, indicate that there is a semantic difference between both experiencer markings. This difference lies in the degree of agentivity that the experiencer is endowed with in both cases: the nominative case-marking is more agentive or, in terms of Croft (1993, 1998, forthc.), encodes an asymmetrical force-dynamic relation between the experiencer and the stimulus/body part.

4.2 Subjecthood

The accusative case-marked experiencer exhibits lexical and functional properties of a subject: it is the only core argument (S in terms of Dixon 1994), it is always animate, conscious, indispensable and referential (cf. Keenan 1976). Apart from these properties, it shows up in the linear position typical of subjects in Russian, and it can control the reflexive anaphora:

- (27) *Amandu neožidanno vyrvalo*
 Amanda:ACC unexpected vomit:PST.3SG.NEUTR
na seb'ja (Russian)
 ON REFL.GEN
 'Amanda suddenly vomited on herself.'
 (<http://textik.ru/ca/films/name/16166/>)

At the same time, the accusative experiencer cannot control the reference of the subjects of participial and conjoined clauses, as can, e.g. be observed from (28), which is ungrammatical:

- (28) **Ego silno rvalo i voobše*
 he:ACC strongly vomit:PST.NEUTR and generally
 Ø *čuvstvoval seb'a očen' ploxo*
 pro:NOM.SG.M feel:PST.SG.M self:ACC very badly
 Intended meaning: 'He vomited and felt generally very badly.'

5. The North Russian perfect

In the North Russian perfect, the subject NP of an underlying transitive verb is always case-marked with an adessive-like PP (*u* 'at' with gen.). In some very rare cases, nominative subject marking is found. The nominative seems to substitute the adessive PP to indicate *strong topicalization* of the subject NP, as has been argued in Timberlake (1976: 562–3), cf. the regular subject case-marking in (29) and an instance of strong topicalization in (30):

- (29) *U nego armiju ne otsluženo*
 at he:GEN army:ACC.SG not serve.out:PART.PASS.PERF.NOM.NEUTR.SG
 'He hasn't served out [his] army term.'
- (30) *A on – eščo armiju*
 but he:NOM still army:ACC.SG
ne otsluženo
 not serve.out:PART.PASS.PERF.NOM.NEUTR.SG=INVAR
 'As for him – there hasn't been serving out [his] army term.'
 (*apud* Kuzmina & Nemčenko 1971: 38)

Strong topicalization or, in terms of Lambrecht (2000: 202), *accented topic expression*, serves to announce a new topic selected from the accessible participants of the discourse, or a topic shift, cf. (31)–(32):

(Vytegra, Vologda, North Russian)

- (31) *i teper' kto znato,*
 and now who:NOM.SG know:PART.PASS.PERF.NOM.NEUTR.SG = INVAR
dak zovut tak, a kto ne
 so call:3PL so but who:NOM.SG not
znato, dak ...
 know PART.PASS.PERF.NOM.NEUTR.SG = INVAR SO
 'And now, those who know call them so and those who don't
 know call them so.'
 (adopted from N. Markova, p.c.)

- (32) *Ja voda prineseno* (Il'men', North Russian)
 I:NOM water:NOM bring:PART.PASS.PERF.NOM.NEUTR.SG = INVAR
 'I have brought water.' (adopted from Ovčinnikova 1956: 82)

Lambrecht (2000:202–4) states that it is natural for *accented topic* NPs to occupy the first position (cf. examples above and those found in, *inter alia*, Ovčinnikova 1956: 82). This syntactic property is not the only way to encode such topics. Thus, the German anaphoric pronouns *der/die/das* are typically used when the referent is active but not yet an established topic, while in the latter case the anaphoric pronouns *er/sie/es* are used. In the North Russian perfect not only the first position but also the case-marking distinguish the accented topics: while (with transitives) the adessive-like PP subject marking is the regular, default option, the nominative subject marking signals the accented-topic status of the subject NP. Interestingly, the old nominative object marking is not necessarily lost, cf. (32) with two nominatives.

Notably, the adessive-like PP subject is endowed with all behavioural subject properties as has been extensively demonstrated in Timberlake (1976). The nominative subject of a transitive verb is not attested in the older language and must be historically secondary with transitive verbs here (cf. Seržant 2012). The rise of the nominative case-marking confirms the prediction made by the Behaviour-before-Coding-Principle (Haspelmath 2010) to be discussed below in subsection 7.1. Crucially, as in the examples above, there is a quite remarkable difference in denotation between the canonical nominative and the oblique (i.e. the adessive-like PP) encoding of the subject. In this case, the difference is rooted in the domain of information structure.

6. The German passive of 'helfen'

The next example discussed in a number of investigations (*inter alia*, Zaenen *et al.* 1985; Barðdal & Eythórsson 2005) comes from German.

- (33) *Unsere Mitarbeiter helfen Ihnen*
 our:NOM officer:NOM help:3PL you:DAT
 'Our officers will help you.'

The verb *helfen* ‘to help’, which takes a dative object cf. (33), retains the dative case-marking in the respective passive form in Standard German, cf. (34a). In non-standard varieties of German, however, the dative case-marking of the passive construction can be replaced with the nominative one, cf. (34b):

- (34) a. *Ihnen* *wird* *geholfen* (Standard German)
 YOU:DAT PASS.AUX.3SG help:PPP
 ‘There you will get assistance.’
- b. *Sie* *werden*
 YOU:NOM PASS.AUX.3PL
gleich *geholfen* (Non-standard German)
 immediately help:PPP
 ‘You will get assistance immediately.’

It has been repeatedly suggested by Barðdal and Eythórsson (*inter alia*, in Barðdal & Eythórsson 2005) that the dative case-marked constituent in (34a) is the syntactic subject of the clause. However, whatever syntactic status the dative argument in (34a) is given in a particular framework, on the descriptive level it remains indisputable that the dative constituent in (34a) does not exhibit *all* of the subject behavioural properties of German (cf., *inter alia*, Fischer 2004:196–7; Wunderlich 2009). In particular, it does not participate in conjunction reduction (Zaenen *et al.* 1985:477; Fischer 2004:196). Notwithstanding the lack of certain subject properties with the dative case-marking, there is a tendency in Colloquial German to replace the dative arguments of this passive with a nominative case marked subject.

7. Conclusions

I have examined six instances in which an oblique case-marked constituent secondarily acquires subject coding properties, i.e. nominative case and the ability to trigger verbal agreement. This can be a dative experiencer (Sections 2 and 3), an accusative experiencer (Section 4), a locational, aollesive-like adjunct (Section 5) or a dative underlying object, promoted via passive operation (Section 6). All of these arguments have recently acquired canonical subjecthood, i.e. nominative case-marking and access to verbal agreement are historically secondary with these predicates. Crucially, even though these constituents acquire typical subject coding properties, they had not been full-fledged behavioural subjects at the stage immediately before this acquisition, as evinced by their source constructions. Therefore, these cases contradict the predictions made by the *Behaviour-before-Coding-Principle* put forth in Haspelmath (2010), which assumes that the acquisition of the behavioural properties is the prerequisite for the acquisition of the coding properties. I discuss this principle in details below.

7.1 Haspelmath's Behaviour-before-Coding-Principle

The *Behaviour-before-Coding-Principle* was formulated by Haspelmath (2010), who bases his argumentation with regard to subjects on preceding work by Cole et al. (1980). Applied to the acquisition process of subject coding properties by oblique marked constituents, this principle predicts that first the behavioural properties of a subject must be acquired, and only after that, the coding properties may subsequently be acquired. Haspelmath (2010), following Cole et al. (1980), works with the following behavioural properties:

- i. control of reflexivization
- ii. target/control of coreferential NP deletion
- iii. compatibility with subject raising operations
- iv. initial position (in unmarked word order with SVO, SOV languages)⁵

The *Behaviour-before-Coding-Principle* (BbCP) correctly predicts, and attempts to coherently explain, diachronic cross-linguistic generalizations in language change. It indirectly implies that the development into a canonical subject is gradual and unidirectional, and that the acquisition of the coding and behavioural properties cannot proceed simultaneously. A number of case studies devoted to the acquisition of the canonical case marking in Germanic (Cole *et al.* 1980),⁶ in Latin (Fedriani 2009), as well as in some non-Indo-European languages (see Cole *et al.* 1980) and, finally, the acquisition of the canonical objecthood in the North Russian perfect (Seržant 2012) justify the aforementioned prediction made by this principle.

Nevertheless, it seems that this principle is not sufficient to account for the data presented in this paper. These data evidently violate this principle, because the constituent of concern is not endowed with (all) behavioural properties listed above before acquiring canonical subject marking (the North Russian perfect being the only exception). The question should then be raised as to whether all behavioural properties are necessary to trigger the acquisition of coding properties, or whether there are certain properties that would be sufficient for the principle to apply. In other words, the question about the internal ranking of the behavioural properties must be addressed (as, e.g. in Givón 1997). Additionally, it seems that certain behavioural properties might be language specific properties, or might be target to other (e.g. morphologically based) conditions, *cf.* Eythórsson & Barðdal (2005: 847ff) elaborating on this argument.

5. Note that Haspelmath regards the word order a behavioural, not a coding property (differently to, *inter alia*, Keenan 1976; Croft 1994: 31; Bickel 2010).

6. For criticism, see Barðdal & Eythórsson (2005).

Furthermore, it seems that it depends very much on the language type whether the prioritized status of the subject argument is revealed by the syntactic (behavioural) or rather morphosyntactic (coding) properties, or by a combination of both. There is no way to reasonably prioritize between these properties in the typological perspective, cf. the notion of *Grammatical Relation* in Bickel (forthc.), Witzlack-Makarevich (2010) that equally applies to both behavioral and coding properties. Thus, in one language (e.g. with no cases), it is the syntactic behaviour of a particular constituent that signals its prioritized status (or subjecthood); in another languages (with cases and/or verbal agreement), it might be primarily the dedicated case-marking and the verbal agreement that unequivocally encode this prioritized status, while a particular syntactic behaviour may just be a consequence of the morphological marking (e.g. due to matrix subject – converb agreement or the co-referential *pro* deletion in conjoined clauses).

Finally, this principle does not account for the fact that there is a semantic shift between the stage with no coding properties and the stage with coding properties. Thus, the presented data cannot be compared with the acquisition of the canonical subjecthood in Germanic languages, as, e.g. in the case of Middle English *me thinks* to Modern English *I think*, where no such semantic change has been detected so far.

In the following subsection, I will provide a modification of the aforementioned principle that would be able to capture the data with no or only few behavioural subject properties before acquiring the subject coding properties.

7.2 Discussion

As noted above, the BbCP leaves aside the fact that concomitantly to the change in the case marking, there is a change in the semantic and functional properties of the constituent at stake. The BbCP is formulated in syntactic and morphosyntactic terms with no involvement of the corresponding semantic and functional domains. Instead, I assume that the development towards canonical case marking is a functionally- or semantically-driven phenomenon, which is why the BbCP makes wrong predictions in the instances above.

The coding and behavioural properties do not exist unto themselves; their main task is to encode the *functional semantics* of the subject in a given language. A prototypical subject, i.e. endowed with all grammatical properties including nominative case and verbal agreement, represents a multi-faceted, cumulative category with notional correlates in different domains of grammar and semantics. Thus, typically, it is the most salient participant not only in the semantic structure of the event referred to, i.e. in terms of the proto-roles entailments (Dowty 1991) or in terms of the force dynamic relations among the participant (*antecedent vs. subsequent* in Croft 1998), but also in other domains that are not directly pre-determined by the predicate. The values of arguments in such functional domains as information structure (cf. the cross-linguistic

correlation with the *unmarked topic* in, *inter alia*, Andrews 1985; Lambrecht 2000:132), referentiality (correlation with *definite NPs*) and empathy (Kibrik 1997; Givón 1997; Croft 2001; Lambrecht 2000) also correlate with the prioritized argument of a clause.

The data provided in this paper illustrate that in each particular instance there has consistently been some sort of functional overlap between the respective properties of an oblique argument and the prototypical subject. An oblique constituent that shows consistent overlap with a prototypical subject within any of these domains (i.e. not exclusively in proto-role entailments or in terms of the role hierarchy) builds up, thereby, a functional relation to subjecthood (cf. Allen 1995:105ff; Gisborne 2011:177–8). This functional overlap enables, I propose, the diachronic change that consists in the assignment of the prototypical subject properties – both behavioural and coding – to the oblique constituent of concern. In Table 1, I summarize the instances of the functional-semantic overlap between the original oblique constituent and the prototypical subject in the respective language:

Table 1. Functional overlap between the oblique constituent and the prototypical subject

	Functional overlap between the oblique constituent and the prototypical subject in the following domains:			
	Proto-Agent properties	Inherent properties	Discourse	Definiteness/referentiality
Experiencer of <i>su-/at-/nu-šalti</i>	+ sentience/perception, exists independently (<i>only with Nom:+ CoP</i>)	+ always animate	+ default topic, default empathy focus	– no restrictions
Experiencer of <i>su-šlapti</i>	+ sentience/perception, exists independently (<i>only with Nom:+ CoP</i>)	+ always animate	+ default topic, default empathy focus	– no restrictions
Experiencer of <i>pri-trūkti</i>	+ (preferably) sentience, exists independently	± preferably animates	+ default topic	– no restrictions
Experiencer of <i>vy-rvat'</i>	+ sentience/perception, exists independently (<i>only with Nom:+ CoP</i>)	+ always animate	+ default topic, default empathy focus	– no restrictions
Subject of the <i>North Russian perfect</i> (from transitive verbs)	+ co-referential with the Agent of the preceding action	+ only animate	+ topic	+ presupposed information, definite
Beneficiary of the German passive of <i>helfen, assistieren</i>	+ sentience/perception, exists independently	± (mostly) animates	+ default topic	– no restrictions

I assume that the consistent endowment of a constituent with some functional properties of a prototypical subject is the main catalyst for the (re)assignment of subject coding and behavioural properties to that constituent; it is an adjustment of grammatical properties to function. I formulate the following constraint on the acquisition of subjecthood:

Functional-semantic constraint (FSC)

Consistent functional-semantic overlap of an oblique case-marked constituent with the prototypical subject may trigger the (re)assignment of the subject coding and behavioural properties to that constituent if there are no other constituents in the construction that would show even greater overlap.

However, it is not just this constraint that has to be satisfied for an oblique constituent to acquire canonical subjecthood. It seems that there are also other requirements for this to occur. Thus, it has often been noted that constructions lacking a nominative subject are generally dispreferred (in accusative languages). Thus, Tsunoda's "Unmarked-Case-Constraint" predicts that in a non-elliptical sentence at least one NP must be in the unmarked case, i.e. nominative or absolutive (Tsunoda 1981), cf. also the "Obligatory NOM Requirement" in Primus (1999) and the "Default Linking" in Wunderlich and Lakämper (2001), Malchukov (2005: 95). Hence, such a construction as in the case of Russian *vy-rvat'* that inherently lacks a nominative case is generally dispreferred and is, therefore, less stable. I assume that this nominative gap facilitates the acquisition of the nominative case by one of the constituents.

In turn, the Lithuanian verbs *su-/at-/nu-šalti* and *su-šlapti* assign nominative case to their body-part argument. This argument does not show, however, any functional-semantic overlap with the prototypical subject in Lithuanian: neither in terms of its discourse properties (default focus) nor in terms of its inherent lexical properties (inherently inanimate). Moreover, as Holvoet (this volume) shows, this nominative argument also fails in having syntactic subject properties, such as deletion of co-referential *pro* subject in coordinated clauses. In other words, there is a mismatch in such a construction between the function (and syntax) of the nominative constituent and its coding properties. Such mismatches may originally be rooted in the etymology of the predicates, but synchronically they represent a suitable environment for the FSC to operate. Hence, the following constraint:

Morphosyntactic constraint (MC)

In order for the FSC to apply, there must be either no canonically case-marked subject in the construction at all, or the coding subject properties must be assigned to a constituent that has less functional-semantic overlap with the prototypical subject than the non-canonically marked constituent.

I assume that both constraints, the FSC and the MC, must be satisfied in order for canonical subjecthood to be acquired. Additionally, if both FSC and MC are satisfied,

then both coding and behavioural properties can, in principal, be (re)assigned at once or, alternatively, along the path predicated by the BbCP. From this it follows that the BbCP can be reformulated in terms of an implicational “universal” (Croft 2003: 52ff):

Reformulation of the Behaviour-before-Coding-Principle (BbCP)

The coding properties cannot be acquired without the behavioural properties.

It has been pointed out that the core/structural cases, such as the nominative case, also bear semantic and pragmatic information, just as lexical or more specific cases do, and are not simply default case-markings of grammatical relations (Amberber and de Hoop 2005: 2). The replacement and subsequent generalization of the cognitively most common template, namely, the transitive template, necessarily involves certain shifts in the event structure that the predicate of concern evokes. This, in turn, leads to a change in the entailments on the arguments. The nominative case typically encodes the most active or the most salient participant that has been given a different notional background in different approaches, *cf., inter alia, controller* (Dixon 1994) or *antecedent in an asymmetric relation* in the causal approach (Croft 1993, 1998, forthc.), *trajector* in Langacker (2008). With the Lithuanian verbs *su-/at-/nu-šalti* ‘to get cold’ and *su-šlapti* ‘to get wet’, as well as the Russian verb *vyrvat*, the grammar makes use of the semantics of nominative, and infers the possibility to grammatically distinguish between an inactive and less inactive experiencing of the event in terms of the presence/absence of the CoP property. Recall that even when the experiencer is case-marked with nominative case, the event still does not imply full control over the experience (typically for such events that one would not deliberately work towards, e.g. ‘getting wet’ or ‘getting cold’).

The fact that the nominative case encodes saliency can be demonstrated independently from subjecthood, namely, with the time adjuncts in Latvian and Lithuanian. Typically, time adjuncts are encoded by the accusative case (alongside less frequent instrumental, locative or genitive) in these languages. Now, if the time duration should be emphasized, the nominative case, instead of the accusative case, can be selected (Roduner 2005) while the syntactic status of the adjunct does not change. The emphatic status of the nominative time adjuncts, as opposed to the same time adjuncts encoded by, e.g. accusative, follows also from the fact that only the former but not the latter are consistently used with emphasizing particles. Thus Ambrazas (2006: 135) writes that the nominative time adjuncts are only found with such particles as *tik* ‘just’, *nors* ‘though’, *jau* ‘already’ etc. Furthermore, an overview over the examples in Ambrazas (2006: 135–6) and Roduner (2005) with the nominative time adjuncts unequivocally makes evident that these adjuncts are most frequently found in the first part of the sentence, typically in the first position or immediately after the subject, evincing thereby their emphatic status. Independently from subjecthood, we observe here the semantics of the nominative case as a morphological marker in the languages of concern which is to encode (discursively or semantically) salient participants.

The acquisition of canonical subjecthood has been discussed in the literature mainly on the bases of such examples from Germanic as, e.g. Old Norse *lika* ‘to like’ with the dative experiencer and Norwegian *å like* ‘to like’ with the nominative experiencer. The instances that I have discussed in this paper are essentially different from the acquisition of the nominative subject by the Germanic oblique experiencers (see, *inter alia*, Cole *et al.* 1980; Eythórsson 2002, 2003 or Haspelmath 2010). The main difference is that in all cases discussed in this paper, there is a change in meaning concomitantly with the change from the original oblique case-marking into the nominative one, whereas there is no such a semantic change in Germanic. Thus, Eythórsson (2000, 2002, 2003) states that there is no difference in meaning with the secondary nominatives in Faroese and Icelandic that have replaced the former oblique subjects. This might be an indication that in Icelandic and Faroese such a coding property as the nominative case is an uninterpretable feature, i.e. it does not have any bearings on semantics. Indeed, such an account for the canonicization of oblique subjects in Middle English has been suggested in van Gelderen (2001). I hypothesize that the BbCP in its original formulation (Haspelmath 2010) accounts only for those languages where the nominative case became an uninterpretable feature. In languages such as Lithuanian, where the nominative case bears certain semantic properties, this principle fails to account.

7.3 Acquisition of the coding properties

Typically, the coding properties are acquired at once, and languages thereby often do not provide evidence on the order in the acquisition of the coding properties. Nevertheless, I have an unequivocal example in which the verbal agreement is acquired first, while there is still no nominative case-marking acquired. This is the case with the partitive genitive subjects in some North Russian subdialects. In these subdialects, the partitive genitive subject triggers verbal agreement along the number of the subject NP (Markova 2008: 153; Seržant, *forthc.*):

- (37) *Tut-to medvedej byvajut,*
 here-PRT bear:GEN.PL occur:3PL
tol’ko malo (Sujsar’, Onega North Russian)
 only few

‘There are bears, but only few.’

- (38) *A kto rabotal pokrepče, tak*
 But who worked stronger, so
ix byli (Sujsar’, Onega North Russian)
 they:GEN.PL were:3PL

‘As regards those who worked harder, there were (some) of them.’

- (39) *Zdes' vsjakix rastut* (Derevjannoe, Onega North Russian)
 here any-kind:GEN.PL grow:3PL
 'Here grow any kind (of plants).'

Also the partitive genitive in Ancient Greek has access to verbal agreement, though differently from the North Russian varieties, on semantic bases (Seržant 2012b).

From these two examples one may hypothesize that in the process of canonical subject acquisition, the ability to trigger verbal agreement is acquired prior to nominative case, if both are not acquired simultaneously.

8. Summary

I have analysed several instances of oblique case-marked constituents that have recently acquired the coding properties of a canonical subject. I have argued that the main driving force behind the acquisition of canonical subjecthood is not a particular syntactic behaviour of the constituent at issue, but rather its semantic and functional load, i.e. whether or not these constituents outrank other constituents of the clause with regard to the *functional* properties of a prototypical subject in the given language. The latter include not only the semantic properties, but also the information-structure properties of a prototypical subject. I claim that whenever there is a constituent that consistently scores higher with regard to the functional subject properties than the other constituents, it provides for instability of the pattern. In this case it is likely that the oblique constituent will acquire behavioural and coding subject properties. With regard to the *Behaviour-before-Coding-Principle* in Haspelmath (2010) I believe that the precise path of subject acquisition may vary cross-linguistically, and there is no universal restriction that the coding properties can only be acquired prior to the acquisition of the behavioural properties; I claim that the acquisition of the behavioural and coding properties may also occur simultaneously. I have slightly reformulated the *Behaviour-before-Coding* principle accordingly, weakening it so that the coding properties cannot be acquired without the acquisition of the behavioural ones.

Furthermore, as regards the exact acquisition of the coding properties, I have found two examples which suggest that, of the coding properties, verbal agreement is acquired first. However, in the majority of cases, it is acquired simultaneously with the nominative case-marking.

Acknowledgement

I am very much in debt to Peter Arkadiev (Moscow), Kristina Lenartaitė (Klaipėda/Vilnius), Nina Markova (Petrozavodsk) and Annie Montaut (Paris) for their help with this paper. All disclaimers apply.

Abbreviations

ACC	accusative	OBL.conv	subject of the converb is not co-referential with the matrix subject
ADV	adverb		
AUX	auxiliary		
CONV	converb	NP	Nominal phrase
DAT	dative	PART	participle
GEN	genitive	PASS	passive
IMPF	imperfect	PERF	perfective
INF	infinitive	PL	plural
INSTR	instrumental	PP	Prepositional phrase
INVAR	invariant	PPP	passive past participle
ITERAT	iterative	PRT	particle
LOC	locative	PRES	present
M	masculine	PST	past
NEG	negation	REFL	reflexive
NEUTR	neuter	SG	singular
NOM	nominative		

References

- Aikhenvald, A., Dixon, R.M.W. & Onishi, M. 2001. *Non-canonical Marking of Subjects and Objects* [Typological Studies in Language 46]. Amsterdam: John Benjamins.
- Allen, C.L. 1995. *Case Marking and Reanalysis: Grammatical Relations from Old to Early Modern English*. Oxford: OUP.
- Amberber, M. & de Hoop, H. 2005. Competition and variation in natural languages: The case for case. In *Competition and Variation in Natural Languages: The Case for Case*, M. Amberber & H. de Hoop (eds), 1–9. Amsterdam: Elsevier.
- Ambrazas, V. et al. 1997. V. Ambrazas, E. Geniušienė, A. Girdenis, N. Sližienė, D. Tekorienė, A. Valeckienė, E. Valiulytė, *Lithuanian Grammar*. Ed. by V. Ambrazas. Baltos lankos.
- Ambrazas, V. 2006. *Lietuvių kalbos istorinė sintaksė. Historische Syntax der Litauischen Sprache*. Vilnius: Lietuvių kalbos institutas/Institut für Litauische Sprache.
- Andrews, A. 1985. The major functions of the noun phrase. In *Language Typology and Syntactic Description*, Vol. I: *Clause Structure*, T. Shopen (ed.), 62–154. Cambridge: CUP.
- Arkadjevas, P. 2008. Lietuvių kalbos pirminių veiksmažodžių klasių semantika tipologinių duomenų kontekste, *Acta Linguistica Lithuanica* LIX: 1–27.
- Arkadiev, P.M. 2011. Aspect and Actionality in Lithuanian on a typological background In *Langues baltiques, langues slaves*, D. Petit, C. Le Feuvre & H. Menantaud (eds), 61–92. Paris: CNRS Editions.
- Berg-Olsen, S. 2005. *The Latvian Dative and Genitive: A Cognitive Grammar Account*. Oslo: Faculty of Humanities, University of Oslo.
- Berg-Olsen, S. 2009. Lacking in Latvian – Case variation from a construction grammar perspective. In *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case* [Studies in Language Companion Series 108], J. Barðdal & S.L. Chelliah (eds), 181–202. Amsterdam: John Benjamins.
- Bhaskararao, K. & Subbarao, S. 2004. *Non-nominative Subjects*, Vol. 1–2 [Typological Studies in Language 60–61]. Amsterdam: John Benjamins.

- Bickel, B. 2004. The syntax of experiencers in Himalayas. In Bhaskararao & Subbarao (eds), Vol. 1, 77–112.
- Bickel, B. Forthcoming. Grammatical relations typology. In *The Oxford Handbook of Language Typology*, J.J. Song (ed.). Oxford: OUP.
- Cole, P., Harbert, W., Hermon, G. & Sridhar, S.N. 1980. The acquisition of subjecthood. *Language* 56: 719–743.
- Comrie, B. 2004. Oblique-case subjects in Tsez'. In Bhaskararao & Subbarao (eds), Vol. 1, 113–127.
- Croft, W. 1993. The semantics of mental verbs. In *Semantics and the Lexicon*, J. Pustejovsky (ed.), 55–72. Dordrecht Kluwer.
- Croft, W. 1994. The semantics of subjecthood. In *Subjecthood and Subjectivity. The Status of the Subject in Linguistic Theory. Proceedings of the Colloquium "The Status of the Subject in Linguistic Theory", London, 19–20 March 1993*. M. Yaguello (ed.), 29–76. Paris: Ophrys.
- Croft, W. 1998. Event structure in argument linking. In *The Projection of Arguments. Lexical and Compositional Factors*, M. Butt & W. Geuder (eds). Stanford CA: CSLI.
- Croft, W. 2001. *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: OUP.
- Croft, W. Forthcoming. Force dynamics and directional change in event lexicalization and argument realization. (<http://www.unm.edu/~wcroft/WACpubs.html>)
- Croft, W. 2003[2009]. *Typology and Universals*. Cambridge: CUP.
- Dahl, Ö. 2000. Animacy and the notion of semantic gender. In *Gender in Grammar and Cognition, Vol.1: Approaches to Gender*, B. Unterbeck, M. Rissanen, T. Nevalainen & M. Saari (eds), 99–115. Berlin: De Gruyter.
- DeLancey, S. 1984. Notes on agentivity and causation. *Studies in Language* 8: 181–213.
- DeLancey, S. 1985. Agentivity and syntax. In *Papers from the Parasession on Causatives and Agentivity at the Twenty-First Regional Meeting [Chicago Linguistic Society 21, Part 2]*. W.H. Eilfort, P.D. Kroeber, K.L. Peterson (eds), 1–12. Chicago IL: CLS.
- Dowty, D.R. 1991. Thematic proto-roles and argument selection. *Language* 67: 547–619.
- Eythórsson, T. 2000. Dative vs. nominative: Changes in quirky subjects in Icelandic. *Leeds Working Papers in Linguistics* 8: 27–44.
- Eythórsson, T. 2002. Changes in subject case-marking in Icelandic. In *Syntactic effects of morphological change*, D. Lightfoot (ed.), 196–212. Oxford: OUP.
- Eythórsson, T. 2003. Change in progress in Faroese: Case marking with impersonal verbs. Ms, University of Manchester.
- Eythórsson, T. & Barðdal, J. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81(4): 824–881.
- Falk, C. 1997. *Fornsvenska upplevarverb* (Experiencer verbs in Old Swedish). Lund: Lund University Press.
- Fedriani, C. 2009. The 'Behaviour-before-Coding' principle: Further evidence from Latin. *Archivio Glottologico Italiano* 2: 156–184.
- Fischer, S. 2004. The diachronic relationship between quirky subjects and stylistic fronting. In Bhaskararao & Subbarao, Vol.1, 193–212.
- Givón, T. 1979. *On Understanding Grammar*. New York NY: Academic Press.
- Givón, T. 1997. Grammatical relations: An introduction. In *Grammatical Relations: A Functionalist Perspective [Typological Studies in Language 35]*, T. Givón (ed.), 1–84. Amsterdam: John Benjamins.

- Goldberg, A.E. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago IL: The University of Chicago Press.
- Haspelmath, M. 2004. On directionality in language change with particular reference to grammaticalization. In *Up and Down the Cline: The Nature of Grammaticalization* [Typological Studies in Language 59], O. Fischer, M. Norde & H. Perridon (eds), 17–44. Amsterdam: John Benjamins.
- Haspelmath, M. 2010. The Behaviour-before-Coding Principle in syntactic change. In *Mélanges Denis Creissels*, F. Floricic (ed.), 493–506. Paris: Presses de L'École Normale Supérieure.
- Heine, B., Claudi, U. & Hünnemeyer, F. 1991. *Grammaticalization: A Conceptual Framework*. Chicago IL: University of Chicago Press.
- Holvoet, A. & Semėnienė, L. 2005. Veiksnių ir tiesioginio papildinio sąvokos. In *Gramatinių funkcijų tyrimai* [Lietuvių kalbos gramatikos darbai 3], A. Holvoet & R. Mikulskas (eds), 39–64. Vilnius: Lietuvių kalbos institutas.
- Holvoet, A. 2009. Difuziniai subjektai ir objektai (Diffusional subjects and objects). In *Gramatinių funkcijų prigimtis ir raiška* (The nature and expression of grammatical functions), A. Holvoet & R. Mikulskas (eds), 37–68. Vilnius: Vilniaus universitetas and Asociacija 'Academia Salensis'.
- de Hoop, H. & Swart, P. 2008. *Differential Subject Marking* [Studies in Natural Language and Linguistic Theory 72]. Dordrecht: Springer.
- van Gelderen, E. 2001. Towards personal subjects in English: Variation in feature interpretability. Case and grammatical relations across languages. In *Grammatical Relations in Change* [Studies in Language Companion Series 56], J.T. Faarlund (ed.), 137–158. Amsterdam: John Benjamins.
- Keenan, E.L. 1976. Towards a universal definition of 'subject'. In *Subject and Topic*, C. Li (ed.), 303–333. New York NY: Academic Press.
- Kibrik, A.E. 1992. *Očerki po obščim i prikladnym voprosam jazykoznanija*. Moscow: Izdatel'stvo Moskovskogo Universiteta.
- Kibrik, A.E. 1997. Beyond subject and object: Towards a comprehensive relational typology. *Language Typology* 1: 279–346.
- Kittilä, S. 2002. *Transitivity: Toward a comprehensive typology*. Åbo: Åbo Akademiska Tryckeri.
- König, E. & Haspelmath, M. 1997. Les constructions à possesseur externe dans les langues d'Europe. In *Actance et valence dans les langues d'Europe*, J. Feuillet (ed.), 525–606. Berlin: Mouton de Gruyter.
- Kuzmina, I.B. & Nemčenko, Je.V. 1971. *Sintaksis pričastnyx form v russkich govorax* (Syntax of the participial forms in Russian dialects). Moscow.
- Lambrecht K. 2000. *Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Discourse Referents*. Cambridge: CUP.
- Langacker, R.W. 2008. *Cognitive Grammar. A Basic Introduction*. Oxford: OUP.
- LKŽ: *Lietuvių Kalbos Žodynas* (Dictionary of Lithuanian Language) (<http://www.lkz.lt/startas.htm>)
- Malchukov, A. 2005. Case pattern splits, verb types and construction competition. In *Competition and Variation in Natural Languages. The Case for Case*, M. Amberber & H. de Hoop (eds), 73–118. Amsterdam: Elsevier.
- Markova, N. 2008. Glagol'nye konstrukcii s roditel'nym padežom objekta i subjekta v onežskix govorax. *Severno-Russkie govory* 9: 146–155.
- Masica, C.P. 1991. *The Indo-Aryan Languages*. Cambridge: CUP.

- McCawley, N.A. 1976. From OE/ME 'impersonal' to 'personal' constructions: What is a 'subject-less' S? In *Papers from the Parasession on Diachronic Syntax*, S.B. Steever, C.A. Walker & S.S. Mufwene (eds), 192–204. Chicago IL: CLS.
- Melis, C. & Flores, M. 2012. Emergence and grammaticalization of constructions within the 'se me' network of Spanish. In *Grammaticalization and Language Change. New Reflections* [Studies in Language Companion Series 130], K. Davidse, T. Breban, L. Brems & T. Mortelmans (eds), 249–270. Amsterdam: John Benjamins.
- Mohanan, T. 1994. *Argument Structure in Hindi*. Stanford CA: CSLI.
- Montaut, A. 2004. Oqblique main arguments. In Bhaskararao & Subbarao, Vol.2, 33–57.
- Onishi, M. 2001. Non-canonically marked subjects and objects: Parameters and properties. In *Non-canonical Marking of Subjects and Objects* [Typological Studies in Language 46], A. Aikhenvald, R.M.W. Dixon & M. Onishi (eds), 1–52. Philadelphia PA: John Benjamins.
- Ovčinnikova, V.S. 1956. O nekotoryx tipax bezličnyx predloženj v poozerskom govore. *Učenyje zapiski XXX, Trudy istoriko-filologičeskogo fakul'teta Rostovskogo n/D gosudarstvennogo universiteta*, vyp. 4: 77–83.
- Primus, B. 1999. *Cases and Thematic Roles*. Tübingen. Niemeyer.
- RNC: Russian National Corpus (<http://www.ruscorpora.ru/search-main.html>)
- Rögnvaldsson, E. 1996. Frumlag og fall að fornu [Subject and case in Old Icelandic]. *Íslenskt mál* 18: 37–69.
- Roduner, M. 2005. Der Nominativ in Zeitadjuncten im Litauischen. *Acta Linguistica Lithuanica* 52: 41–58.
- Seržant, I.A. 2012a. The so-called possessive perfect in North Russian and the Circum-Baltic area. A diachronic and areal account. *Lingua* 122: 356–385.
- Seržant, I.A. 2012b. Morphosyntactic properties of the Ancient Greek partitive genitive in the subject position, *Indogermanische Forschungen* 117: 187–204.
- Siližienė, N. 1994–2000. *Lietuvių kalbos veiksmožodžių junglumo žodynas*. (Dictionary of Lithuanian Verb Valence), Vol. I (A–M), 1994; Vol. II(1) (N–R), 1998; Vol. II(2) (S–Ž), 2004.
- Timberlake, A. 1976. Subject properties in the North Russian passive. In *Subject and Topic*, C.N. Li (ed.), 545–594. New York NY: Academic Press.
- Trousdale, G. 2008. Words and constructions in grammaticalization: The end of the English impersonal construction. In *Studies in the History of the English Language, IV: Empirical and Analytical Advances in the Study of English Language Change*, S.M. Fitzmaurice & D. Minkova (eds), 301–326. Berlin: Mouton de Gruyter.
- Traugott, E.C. 2003. Constructions in grammaticalization. In *The Handbook of Historical Linguistics*, D. Joseph & R.D. Janda (eds), 624–46. Oxford: Blackwell.
- Tsunoda, T. 1981. Split case-marking in verb types and tense/aspect/mood. *Linguistics* 19: 389–438.
- Wunderlich, D. & Lakämper, R. 2001. On the interaction of structural and semantic case. In Special Issue *Effects of morphological case*, H. de Hoop, O. Koeneman, J. Mulders & F. Weerman (eds). *Lingua* 111: 377–417.
- Wunderlich, D. 2009. The force of lexical case: German and Icelandic compared. In *The Nature of the Word: Essays in Honor of Paul Kiparsky* [Current Studies in Linguistics 47], K. Hanson & S. Inkelas (eds). Cambridge MA: The MIT Press.