Adjectival and argumental Small Clauses vs. free adverbial Adjuncts – A phase-based approach within the Radical Minimalism with special criticism of the Agree, Case and Valuation notions

1. On definition of Secondary Predicates and Small Clauses: The Puzzle

In recent work on Russian and Slavic syntax, depictive secondary predicates are mostly considered to be “non-sentential adjuncts on the predicate layer of the clause” (cf. Schroeder, Hentschel, Boeder eds. 2008: preface, i). As opposed to adverbials, which modify the sentence or the VP, depictive secondary predicates modify the arguments (either the subjects or the objects). In fact, while adverbials are modifications of events (VP-adverbs) or propositions (sentence adverbs, cf. Kosta 2003ab), depictive predicates are modifications of arguments. There has been a highly controversial discussion in the nineties in generative framework regarding generative description of depictive secondary predicates or small clauses (henceforth, SC). To mention only a few studies, only few focused on their semantic properties (cf. e.g. Steube 1994, Hentschel 2008), most were concerned with their syntactic status (small clause vs. AP-, NP or VP-adjunction, cf. Stowel 1978, Stowel 1981, Williams 1984, Aarts 1992, Cardinaletti & Guasti eds. 1995, Staudinger 1997 vs. Wilder 1994, Emonds 2007, Hentschel 2008) and/or their Case assignment properties (cf. Bailyn 1995, Bailyn 2001, Bailyn & Citko 1999, Bowers 1997, Bowers 2001, Franks 1995, Strigin 2008, Bondaruk 2004, Bondaruk 2013ab). Only few studies mention also the diachrony, for Slavic to mention just a few Moser (1994), Hentschel (1993, 1994), Menzel (2008), Klemensiewicz (1926) and Timberlake (2014, in print).

Small Clauses (SCs) are structures, which have clausal characteristics in that they contain a subject phrase and a predicate phrase. They are, however, generally believed not to contain a complementizer position or an INFL-node (cf. Aarts 1992, but cf. section 2.5., 2.6. for discussion).

The bracketed sequences in the S-Structure (1)-(4) are examples of Small Clauses:

(1) Ja_i sčitaju [ego_j pjanyj / *pjanyj / pjanogo]

I consider him drunk

(2) Ja_i sčitaju [ego_j durakom]

I consider him a fool

(3) On_i vypil čaj / xolodnym

He drank the tea_Akk cold
Onivypil čaj, golyj; golymj/*golyj / golymj; xolodnyj.
He drank the tea_Akk naked / cold

The non trivial question, which has until now not being considered analyzing secondary predicates is their relation and division of labor between semantics and syntax (distribution of Theta-roles, Case assignment and Binding) which interests us most in our present talk. More specifically speaking, what is the difference between sentences like (1)-(4) w.r.t semantics, syntax and pragmatics? This is a part of a book I am working on right now (cf. the article Kosta in print, C). But for the time being, I shall try to demonstrate how the behavior of their Case Assignment and Agreement within the Phase oriented approach of Radical Minimalism can be explained (cf. Krivochen & Kosta 2013; Kosta & Krivochen 2014).

2. The Framework: Radical Minimalism

To begin with, we must outline the framework within which we will be working: Radical Minimalism. It is a theory that attempts to provide principled explanations for (linguistic) phenomena without ignoring the interactions between mental workspaces or biological-computational plausibility and, perhaps most importantly of all, seeking the elimination of all intra-theoretical stipulations via interface conditions and (we will claim, universal) economy principles in derivation and interpretation with basis on biology and mathematics. Let us review some of the basic tenets of Radical Minimalism, and then discuss the implications of this framework:

(5)

1. Language is part of the “natural world”; therefore, it is fundamentally a physical system.
2. As a consequence of 1, it shares the basic properties of physical systems and the same principles can be applied (Heisenberg’s uncertainty, the Conservation Principle, locality, etc.), the only difference being the properties of the elements that are manipulated in the relevant system.
3. The operations are taken to be very basic, simple and universal, as well as the constraints upon them, which are determined by the interaction with other systems, not by stipulative intra-theoretical filters.
4. 2 and 3 can be summarized as follows (2):

(6) Strong Radically Minimalist thesis (SRMT):
All differences between physical systems are “superficial” and rely only on the characteristics of their basic units [i.e., the elements that are manipulated], which require minimal adjustments in the formulation of operations and constraints [that is, only notational issues]. At a principled level, all physical systems are identical, make use of the same operations and respond to the same principles\(^1\).

We claim that there is only one generative operation (in the physical world in general, in the mind-brain in particular), call it Merge, which is free, “blind” (that is, insensitive to the inner characteristics of the objects it manipulates, we follow and extend the thesis of Boeckx 2010a that only format is relevant\(^2\)) and unbounded, and an operation Transfer that provides us with a way of delivering structured information across modules. Merge is an inherently (derivationally) diachronic operation that generates structures, endocentricity being merely a C-I interface requirement for interpretation purposes (i.e., take a referential variable from a sortal or extending-into-time perspective: NP / VP. See Borer 2005; Panagiotidis 2010 for details). Transfer, in a sense that will be clarified below, takes place as soon as it can, this timing being determined by the formation of a fully interpretable configuration in terms of “bare output conditions”, what we call, adopting the Chomskyan term (though not the whole theory associated with it), a phase\(^3\).

In a “dumb (i.e., blind and free) syntax” there are no syntactic constraints at all, so there is no point in positing feature-driven operations (e.g., Chomsky 1998; Pesetsky and Torrego 2004, 2007; Müller 2011a) as they represent a substantive complication of the theory, rather than being the null hypothesis. Merge applies because two whichever objects share a common format, and the relevant interfaces take the results of structure-building operations as soon as those objects are fully interpretable: we have a free Generator and Invasive Interfaces, capable of accessing the syntactic workspace at every point in the derivation (so-called “extremely local evaluation”, cf. Müller 2011b) and taking the minimal object they can fully read. The role of features, which has been one

\(^1\) Such a pretention of universality regarding physical explanation is somehow reinforced

\(^2\) From our proposal (which stems partly from Tegmark, 2007), it follows that the same generative mechanism underlies the mathematical structure of a sentence, the Fibonacci sequence, Theodorus’ Spiral and DNA, to mention some clear examples. If we are on the right track, the explanatory and justificative power of the theory would not only limit itself to language, which would be the optimal scenario. To give an example, Theodorus’ Spiral is generated by merging, in a non-trivial way, sides of right triangles and their hypotenuses (a = 1, 1, \(\sqrt{2}\); b = 1, \(\sqrt{2}\), \(\sqrt{3}\); c = 1, \(\sqrt{3}\), \(\sqrt{4}\)...).

of extreme importance in Minimalist syntax⁴, is questioned, and the very notion of distinctive feature as a primitive of syntactic theory is put to test. Our argument goes as follows: let us assume that we start with a fully interpretable Relational Semantic Structure (see Mateu Fontanals 2000a, b for the original presentation, Krivochen 2010b, d for developments within this framework⁵), built by structuring primitive, very simple and atomic generic semantic elements (à la Jackendoff 1987, for the earliest presentation, but see also Jackendoff 2002; Levin and Rappaport 2011, Mateu Fontanals 2000a, b. Hale and Keyser’s 1993 “l-syntax” and Pustejovsky’s 1995 “Generative Lexicon” which have been related to those positions as well, although their abstract pre-syntactic structures are lexical, not purely semantic, and the presence of syntactic principles like the ECP is strong, constraining processes like conflation / incorporation). The reason why such a primitive semantic structure is fully interpretable is simple: the atomic units of which such a structure is built are very few and underspecified and, moreover, there are no superfluous elements: a conceptual structure “embodies” a global semantic-pragmatic intention following strict faithfulness constraints (nothing is either added or taken away). These conceptual templates are built with the input from experience, as the actants in events in the phenomenological world are assigned an interpretation. Information enters the mind chaotically, but it is organized in very much a Kantian way, into categories that exist as a priori forms: mainly, space (if we accept that time is conceptualized as a metaphor of space, in favor of which there is both linguistic and neurological evidence: see Talmy 2000 and Dehaene, 2011). However, the existence of more complex conceptual templates must be posited, as our awareness and understanding of the phenomenological world around us does not limit to things in a location (be it concrete or abstract). In previous works, we have depicted a theory of semantic primitives, which we will sum up here. Ontogenetically (and, perhaps, phylogenetically), the most primitive category is the noun, denoting things (i.e., sortal entities in the sense of

⁴ For an exhaustive analysis and references, see Adger (2011), Adger and Svenonius (2011).

⁵ We will not analyze initiatives like the Lexicon Project or Pustejovsky’s (1995) “Generative Lexicon” simply because our interest is precisely to separate language from semantic structures, not to offer univocal mapping rules. A syntactic theory of semantic structures (in the broad sense of “how semantic structures are generated”) is both possible and desirable, without the need to resort to strictly linguistic concepts. Therefore, the mere mention of a LEX (for example) is out of question. Contrarily to the aforementioned approaches, we impoverish the so-called “syntactic component” to a single algorithm which applies without faculty restrictions, as it would be the optimal scenario.
Things, however, are not isolated, but related in various ways, the most basic of which is a purely spatial relation in terms of central or terminal coincidence (Hale and Keyser 1993, 1997; Mateu Fontanals 2000a). We have, then, two categories so far, one conceptual, the other, procedural: N and P respectively. Further up on the structure, a spatial relation between two entities is a static event, and we have thus derived uncaused verbs (i.e., Unaccusative). Different kinds of Unaccusative Vs arise when varying the nature of the P node: telic and atelic, static and dynamic Unaccusative Vs. The most complex structures appear when the event has an external initiator, which requires the presence of a cause primitive. We have now caused events, which may or may not include a spatial (i.e., prepositional) relation between entities: their linguistic instantiation is either a (Di)Transitive or Unergative verb respectively. Having this conceptual (pre-linguistic) skeleton, we can fill the places with the available information, participants, time, place, etc. Such an approach to pre-linguistic semantic structures has the following advantage: if these structures are built via the same algorithm that applies in other faculties (and in other systems within the physical world, as Tegmark 2007 suggests), the theory is substantially simplified. Moreover, as we will see below, if the driven force of these structures is the speaker’s intention (i.e., the I part of C-I, which has been systematically swept under the rug in Minimalist accounts), there is a principle driving the selection of the Array: choose only those elements that minimally instantiate the semantic structure via linguistic means. Under the simplest assumptions, these generic semantic elements include non-relational elements (i.e., logical arguments) and relational primitives, which link those elements. We will assume the following typology of primitives, based on Mateu Fontanals (2000a):

- **Non-relational element**: X (Hale and Keyser’s N, Jackendoff’s [THING]), a logical argument.
- **Relational predicative primitives**:
  - Cause (CAUSE / HAVE)
  - Event (BE –static- / GO –dynamic-)
  - Location (TO –terminal coincidence / WITH –central coincidence)

Those primitives, which we have adopted and adapted from Mateu Fontanals’s work (in turn, based heavily on Jackendoff 1987, 2002 and Hale and Keyser 2005). At this point, we are talking about “nouns”, but, as the reader will immediately find out, so-called “nouns” are actually an interface reading of a local relation between a root and a distributionally specified procedural element D, such that N = {D, √}. Moreover, when we refer to “entities”, we do so in a very underspecified way, so that we will talk about “events” as “entities”, in a sense that will be clear by Chapter 4. See also Krivochen (2012c) for details.
1993, 1997a, b, 2002) can be linguistically represented (e.g., in a sentence like [John made Mary cry], [made] is Spelling-Out [cause]), but this does not mean that they are necessarily linguistic but rather that they can be linguistically instantiated and Spelled Out if the language L in question has Vocabulary Item availability. On the other hand, other categories like Tense, Aspect and Modality require a linguistic structure to modify. Let us take Tense as an example: it would be a mistake to confuse linguistic Tense with either physical Time (external and objective, defined as the direction towards which entropy increases) or conceptual temporal notions, what we have called “Zeit” in earlier works (Krivochen 2012a). Moreover, the conceptual Zeit is expressible in locative terms (a proposal that is based on Talmy 2000 and related work): an event, regardless its linguistic instantiation (i.e., a V, a gerundive nominal or a derived nominal) is expressible as an ordered pair \((e, t)\), where \(e\) is the generic event denoted by a bare root and \(t\) is its anchor in the mental timeline, clearly spatial\(^7\). Tense, then, is not a viable pre-linguistic primitive, and Zeit can be subsumed to Location (a position that has its roots in Einstein’s works on special relativity and its philosophical implications). Aspect and Modality, other possible candidates, as they are commonly defined, need to have scope over a defined event (i.e., a \(\{T, V\}\) relation) and a fully-fledged dictum respectively, both eminently linguistic entities. As such, in our opinion, they are not candidates for semantic non-linguistic primitives.

With these tools, provided that the non-relational element “X” comprises the whole set of generic entities (whose identity is irrelevant for the generative component), an “Unaccusative” Relational Semantic Structure (hereafter, RSS) would look as follows:

\[
(7) \quad \text{[event GO [location BOYfigure [[TO] HOUSEground]]]}
\]

The elements of the RSS are not linguistic, but, at most, pre-linguistic (or, in the strongest interpretation, completely extra-linguistic). This means that a RSS is not necessarily instantiated as a sentence (i.e., a linguistic unit), but it conveys pure semantic substance with the potentiality (but, crucially, not necessity) of being instantiated via language. As such, the English words we have used serve only expository purposes, since RSSs do not belong to any particular natural language. They can be seen as semantic genotypes (this time, it is a metaphor), conveying such underspecified semantic substance its instantiation is by no means fixed beforehand.

\(^7\) See Talmy (2000) for a recent review of the foundations of Localism, and Dehaene (2011) for further evidence of the localist nature of human cognition, particularly regarding the “number sense”. See also D’Espósito (2007) for an overview of cognitive / neural models of working memory, which are essential for the localist theory.
According to the Conservation Principle (Krivochen 2011b; Lasnik, Uriagereka and Boeckx 2005 for an earlier and somewhat different presentation under different assumptions), this information must be carried along the whole derivational path (i.e., information cannot be erased, but instantiated in a way so that it can be manipulated by the relevant module: we see here the first important departure from the concept of deletion), which implies that the aforementioned concepts will have to be instantiated in such a way that they can be manipulated by the syntax: in our model (and Distributed Morphology and some versions of Exo Skeletal Models – XSM from now on – like De Belder’s 2011 or Borer’s 2009) those concepts take the form of roots. So far, we have no features or procedural instructions, only semantic primitives, either relational (predicates, interpretable only as determining how the relation between other conceptual objects is to be read) or not (arguments). Apparently, features should be added at this point in the derivation, when a semantic object is transformed into a syntactic object, if we accept that syntactic operations are always feature-driven as in the strongest constructivist models (e.g., Chomsky 1998; Pesetsky and Torrego 2007; Lasnik, Uriagereka and Boeckx 2005). (Un-) Interpretability depends on valuation (Chomsky 1999) and, in turn, valuation depends on the category on which those features appear. Those features that enter to the derivation unvalued in a category must be eliminated for the derivation to converge. Our objection here is: assuming the input of language is already structured and syntactic (i.e., built via Merge) which is the point of adding features in the first place if the system will then eliminate (some of) them? This, without taking into account the stipulation that underlies the whole system regarding the fact that a feature [F] enters the derivation valued in category P but not in category Q. Even if the reader does not accept our use of the Conservation Principle, this second objection is valid within an orthodox Minimalist framework. Feature valuation-deletion also entails the following problem, first noticed by Epstein and Seely (2002): the timing of Spell-Out. If we accept the orthodox view that Spell-Out deletes the features that are uninterpretable by LF (i.e., those which have entered the derivation unvalued, and have therefore acted as probes, copying the value of a c-commanded goal), then we have to indicate to the system which of all the features that we have in a determined derivational point had entered the derivation unvalued. But, in order to do so, we would have to look back and see the derivational point immediately

---

8 Even if the reader rejects the so-called interpretability-valuation correlation posited in Chomsky (1999), the objection regarding the unprincipled character of feature and value assignment holds. Moreover, the advantage that Chomsky’s system represented by unifying two characteristics into one is lost, thus complicating the theoretical apparatus with no (further) empirical support.
before valuation, which is impossible in a derivational (even if it is not as strong as Epstein’s 1999) approach as the derivation is a diachronic process, and past states of the system are no longer accessible. The situation can be summed up like this:

(8) Spell-Out timing:
   b. *After valuation.* Result: crash. There is no way of knowing which features entered the derivation unvalued (and were, therefore, uninterpretable by LF).

Chomsky (1999) attempted to solve the problem by stipulating that Spell-Out (i.e., Transfer to PF) takes place “shortly after” valuation, but we do not see how this could help solving the problem. Epstein and Seely also tried to provide an explanation by saying that the transference took place within a transformational cycle, that is, not before, not after valuation, but during the process. This is:

(9) \[ a \ldots u-F \ldots \] → \[ a \ldots i-F \ldots \] → Interface

Interestingly enough, their concept of “trace” is that they are “remnants” of previous derivational stages. They have no real entity, but are “diachronic” occurrences of the same element (we will return to the conception of traces as tokens below), thus strongly considering the temporal dimension of Merge, its inherent diachrony in an on-line derivational system, like the one we argue in favour of. The system makes a derivational backtracking to the immediate previous derivational step, i.e. to the input of the transformational rule to see which feature/s was/were unvalued and, therefore, uninterpretable. For us, that is not a satisfactory answer either, since it is simply not principled, but stated, and it does not follow from any interface condition or from conceptual necessity that the application of a rule R will always result in a legible object. If it does not, and the object gets transferred (a strong version of the “every phrase is a phase” desideratum), the model is clearly crash-rife, which is not a desirable state of affairs as it implies a considerable computational overload at the interfaces, which receive much more than they can actually interpret. Moreover, it relies on a computational system that can have access to inner characteristics.
of the manipulated elements, a substantive complication we will avoid with our “blind” generator. Our solution is quite more radical: we just eliminate features from the picture. If syntax (in the wide sense of “generative engine”, as in an OT-like architecture) only cares about putting things together, why should it bother about valuation? After all, feature valuation is an operation that only makes sense taking convergence at the interface levels into account, but in the syntax proper (or “narrow syntax”, using traditional terminology) it is perfectly superfluous, since nothing “converges” or “crashes” in the syntax: there is no way to determine well-/ill-formedness in a purely generative working space.

What we propose regarding all so-called "(un-)interpretable features" is that they do not exist at all, particularly (but not necessarily) considering the proposal made in Chomsky (1999) that uninterpretability is concomitant to unvaluation. That would be the strong (and optimal) thesis. Instead of a number of features (number that has increased over the years) which enter the derivation valued or unvalued (with an arbitrary number of possible outcome states) depending on the category they compose, we have a minimal number of (interface-required) interpretable quantum dimensions conveying, in abstracto, all possible outcomes licensed by the interface systems, but adopting one value or another in a local relation with another element whose distribution is specific enough to provide the interfaces with an unambiguous object. We will express it by using this notation: for any dimension \( D \), \([D_X]\) expresses its quantum state (that is, comprising all possible outcomes). Take, for example, Case. As we have proposed in earlier works, there are only three Case Spheres, any further refinement being a morphological epiphenomenon (on the PF side) or an inference (on the LF side). Should we accept this claim, the Case dimension would look as follows:

\[
(10) \quad \text{Case}_X = \text{NOM}+\text{ACC}+\text{DAT}
\]

We will come back to this below, when deriving a sentence following RM claims.

2.1 The Architecture of the System

In this section we will compare our version of the architecture of the mental grammar with other proposals, mainly Optimality Theory (Prince and Smolensky 2004; Smolensky and Legendre 2006) and the orthodox Minimalist Program (Chomsky 1995 et. seq.)\(^9\). Our goal will be to point out problems we find in those approaches that Radical Minimalism can solve without resorting to extra stipulations, but only to general interface conditions. Our architecture will...

\(^9\) For a presentation of OT-syntax, see Müller (2011b); for a Radically Minimalist version of OT, and detailed discussion of the present state of OT-syntax, see Krivochen (2012b)
share some components with OT and also the traditional Chomskyan version of
the MP, but we will see that both the local (the derivational engine) and global
(the architecture of the cognitive system we assume) characteristics of the
system differentiate RM from its antecessors.

The model of the “faculty of language” as a dynamic workspace we will
assume from now on is as follows, without entering into details (which will be
provided further below):

Table (1): The architecture of the RM-system

Thus, if we want to analyze secondary predicates we have to decide which status
they have on the level of syntax and semantics (e.g. within the C-I and S-M
interfaces system) and which role play the different levels before Spell-Out.

For me, at present, it is l important to stress that I do not come from the
theoretical background of the Mainstream Generative Framework (henceforth,
MGF) but my recent studies together with my colleague Diego Krivochen (cf.
Kosta/Krivochen 2012, Krivochen/Kosta 2013, Kosta/Krivochen 2014) would
rather allow to place myself in the more Minimalist theoretical approach, called
Radical Minimalism as it has briefly – for lack of time – been outlined here.

Within this framework, thoroughly described in Krivochen/Kosta (2013), a
theory of language, which not only wants to capture (observe) and describe the
data but seeks to explain how a child of any language in the world can learn these structures based on a relatively short time, poor evidence and the lack of negative evidence must be able to observe the variety of structures, describe and classify their intra-linguistic and cross-linguistic differences (presumably on the basis of a theory of language), and to explain their learnability from the standpoint of universal and specific properties of the structures themselves.

2.2. Agree and Case in Russian secondary predicates

While saying only a few words on the semantic status of the so-called SC in this section of my talk, I shall concentrate on the latter question, namely the behavior of two syntactically crucial properties – the Agree and the Case assignment properties in Russian secondary predicate construction. My talk will include not only adjectival SC of the type (11) but also nominal SC (12) and adverbial SC (13)-(14):

(11) Ja, sčitaju ego j pjanymj / *pjanyj / pjanogo
I consider him drunk

(12) Ja, sčitaju ego j durakom
I consider him a fool

(13) Oni vypil čajj xolodnymij
He drank the teaAkk cold

(14) Oni vypil čajj golyj / golymj / *golyj / *xolodnyj
He drank the teaAkk naked

In (11), an ECM matrix verb heads the embedded SC, the NP ego is marked Acc case and the argument position of the secondary predicate – an adjective – is assigned Instrumental case. Also, in (11) only the direct object of the matrix clause can be the antecedent of the secondary predicate, but not the subject.

Note, however, that both individual level predicates such as byt’ (and the zero copula), and stage level predicates such as stat’ ‘to become’ can take Instrumental as depictive predicate modifying the subject (external argument) in Russian. But with an individual level predicate the situation changes with the change of the event or tense: then the nominative is preferred in individual use in present and the instrumental can be used only with the individual level predicate byt’ if the situation has already passed:

10
It seems to be the case that NOMINATIVE IN PAST with the verb byt’ has also a qualifying meaning of the individual’s quality in Past which in a way rests in the memory of the speaker, whereas the INSTRUMENTAL IN PAST TENSE has the meaning of a characterization of a person during the time span we are talking about, but all the rest being inferred by conversational implicatures (cf. Kosta, in print), by assuming that person X is either not alive anymore (15b), or a student who is not student anymore (15f) (cf. on conversational implicature and the three levels of meaning in Kosta 2012 and Kosta, in print).

Another peculiarity, which deserves attention, is the possibility to scramble (move) a part of the SC leaving a remnant in situ which leads to a contrastive focus reading:

(16) Студентом я был "бо́ГАТЫМ" — повышенная стипендия да ещё подрабатывал в институтской лаборатории и сантехником в общежитии. (NKRJa

There is also possible to pied-piep a SC which resembles very much on adverbials, cf. Kosta (1998:148). So I assume that Small Clauses behave in a similar way like adverbial VP-adjuncts behave, which generally remain within the lexical VP and or rightadjoin or leftadjoin to VP - depending on the language. In contrast, adverbial adjuncts of the sentential class (cf. Kosta 2013ab) are hierarchically ordered in syntax, according to their modal or temporal affiliation either by moving to SpecTP (temporal adverbs such as often), or to SpecCP position (epistemic, faktive, verificative or evidential adverbs, cf. 1998:148 Kosta, Kosta 2013ab).

Overt movement of adjuncts does not follow for the sake of checking and Valuing strong features - such as it is usually in theories such as Pesetsky / Torego (1994, 1997) or Biskup().

11 As Petr Biskup (2011) tries to demonstrate, the reason why manner adverbs move is not because of the need to check F-features in syntax, but because of information structural reasons. Thus, if a manner adverb moves out of its base generated position right or left adjoined to the lower VP, it is because it has to check a EPP feature above the domain of the left periphery (CP), thus it gets either focused or back-grounded, just like in (b) as opposed to the unmarked reading in (a): (a) Zatím to vypadá dobře (b) Zatím to dobře vypadá. Note that Czech is
processes such as for pied-piping (Haj Ross 1967 mentioned pied piping first, see also Ruzicka in 1998 and Chomsky 1995).


The National Corpus of Russian Language shows that ECM verbs like to consider assign the Instrumental case to the secondary predicate (be it adjective or argument/noun) in object, but never in subject position, whereas subject control verbs such as to be, to become assign Instrumental case to the secondary predicate controlled by the subject position which is the controller of the secondary predicate. We give some examples for both:

Subject control verbs:


(19) — Ничего лучшего ты бы не смог придумать! Папаша, да ты, мы видим, просто молодец! Вечер провели в квартире у Чистых прудов, которую московский ФК снимал для своего председателя у бывшего советского министра хлебозаготовок. Последний махал метлой у ворот, придураиваясь под простого мужика; вскоре он стал хорошим капиталистом. [Василий Аксенов. Новый сладостный стиль (2005)]

(20) За коммунизм воюют «добрым оружием» — молоком, маслом, насущным хлебом справедливости! Но и этим добрьм оружием должны отважно воевать настоящие бойцы. — Он повернулся к Бахирову, склонив голову набок, посмотрел на него. — Вот тут говорили о товарище Бахиреве, что он стал хорошим бойцом. [Г. Е. Николаева. Битва в пути (1959)]

Out of all documents we have found in the National corpus, there was no single example in which the subject control verbs like byt’ or stat’ would assign a Nominative instead of Instrument. Interestingly, even if controversially

a VO language so that the reverse relative order between the verb and the adverbial makes the sentence marked as in (b). Given the correlation between the notion background and CP phase, it means that a back-grounded manner adverbial moves across the vP and targets the CP phase. Cf. Biskup (164).
discussed cases with verbs of ‘expected’ object control with subject control exercised by the same object control verb like *poprosit* ‘to ask’ come into play, the secondary predicate needs to have assigned the instrumental case (and not any other structural case), irrespective whether it is co-indexed with the object of the matrix clause, cf. (21) or with the subject (22):

(21) Ja poprosila ego ne byt’ žestokim
(22) Ja poprosila ego ne byt’ isključennoj iz školy
(Both examples are from Růžička 1999: 29, ex. 55 and 56).

The broad spectrum of functions and uses of Russian instrumental in different syntactic positions as described with the Jakobsonian notion of peripheral case such as subject and object can typically be detested in context of secondary predication. Strigin (2008:383seq.) mentions the following uses of Instrumental in Russian taken from the seminal work Nichols (1978), where even intransitive verbs like *rabotat* ‘to work’, *žit* ‘to stay’, simple transitive verbs like *vybrat* ‘to elect’ or *služit* and unaccusatives like *vernuts’ija* assign Instr to their complement. It has, however, to be mentioned, that the instr phrase is very closely attached to the verb and that it is not just a simple adjunct but more probably a complement, thus argumental position (A-position) to which the case Instr is assigned. Under standard analysis this position must then also be by definition a Theta-position visible to case assignment. We are left with the explanation how a verb like *rabotat’* or *žit’* can be assigned instr case in the first place.

Type 1: (23) a. On rabotaet inženerom
   He works engineer: INS
   ’He works as an engineer’

   b. Ego vybrali prezidentom
   he:Acc elected:3Pl president:INS
   they elected him president

12 In Polish, predicative adjectiv es in subject control constructions usually agree with the subject of the matrix clause and assign Nominative except the subject of the matrix clause is assigned other case than nominative, then the predicative adjective is assigned the instrumental, cf. (a) Marek, pragnął [PRO₁ być najlepszy w czytaniu], (b) Marek, każął Marii [PRO₁ być bardziej pewną siebie /* pewną siebie], (c) Marek, twierdzi, że ważne jest [PRO₁/arb być pewnym siebie /* pewny siebie], (d) Jest mu₁ żłe [PRO₁ być starym], cf. Bondaruk 2004:229seq. All examples are taken from Anna Bondaruk’s work (Bondaruk 2004, 20013ab) who works in Minimalism and I very much appreciate her work.
c. Kamni im služat oporoj
rocks they:Dat serve support:Instr
Rocks serve them as support
d. On igral vratarem
he:Nom palyed goalkeeper:Ins
He played goalkeeper
Type 2 (24) a. On sidel grustnyj
    He:Nom sat sad:Nom
b. On vernulsja geroem
    he returned hero:Ins
    He returned a hero
Type 3 (25) a. Snačala mašinu vzvešivajut pustuju
    First truck:Acc weigh empty:Acc
b. On vypil čaj xolodnym
    he drank.uo the tea:Acc cold:Ins
Type 4 (26) a. Rebenkom on žil v Pariže
    Child:Ins he lived in Paris
b. Xolodnym etot čaj ne vkusnyj
    cold:Ins this tea:Nom not tasty:Nom

It seems that the case assignment property, especially the contrast between Instr and Nom case, has a reason in event semantics, as has often been stressed in the literature but never really explained (cf. Steube 1994). Strigin (2008) tries to derive all uses of secondary predicates (depictives) from one general meaning, leaving the rest to the inference and context. I state that the major difference between Instr and Nom as case which modifies the subject of the matrix verb is that between stage level and individual level predicates. Other uses such as (23c) are conditioned by the proper meaning of the Instr case as Instrument. In (23a), it is a temporary profession as engineer, which predicates over the subject (on). Similarly, in (23d) it is not the quality of the individual which makes him a goalkeeper but the profession or activity, which itself cannot be permanent. In (24b), it is the result of a change of state which made the person a hero (no one is born as hero, as we all know). In (25a), the tea may have changed the quality while drinking, so in the beginning it might have been still
warm and then, maybe because the person did not pay attention and waited too long before drinking the tea, it got cold, so that in the final stage the tea was cold and the person drunk it cold. Both examples under (26) must be interpreted as events that have taken place in a certain time span (period), which is not true in the present state. So if we say that a child will not grow we might also say that this is not change of state. Any kind of change of state belongs to the class of stage level predicates, this is also known in languages such as Spanish or Italian, which have two different verbs of ‘to be’: stage level predicates (estar, stare) or individual level (ser, essere). This differentiation one can also see when one asks for how are you, one has to say ¿cómo estás? and in Italian come stai, not cómo eres, come sei, meaning how are you just now, not for ever.

In (24), it is rather the individual property (quality) of the person who is sad than his momentary disposition and maybe the fact that sidet’ is a state and not a change of state predicate qualifies and even forces the selection of Nom. But as far as I know, even the Instr would not be excluded in this context if the meaning would be something like in the present situation I saw him sitting sad in his chair.

Thus, we are left to explain the use of Instr in depictives where the Instr stands as second internal argument or adjunct after ECM matrix verbs such as sčitat’ ‘to consider’ which assigns the accusative case to the direct object and the secondary predicate – a noun – is assigned Instrumental case. This Instrumental case comes as default case, maybe because there is no other possibility to assign another (structural) case in adjunct position. Adjuncts in Russian are mostly inherent (or lexical) cases. This can be also confirmed by the fact that in passives, in which the object (Patient) of the active clause moves to the subject position and is assigned case, the agent is either suppressed (impersonal passive form) or it is expressed with the Instr case:

(27) a. Dom stroitsja
b. Dom postroen inženerom (cf. Kosta 1992, chapter 5)

Bailyn (1999:20) assumes that the head of a Predicate Phrase is an inherent Instrumental case assigner. But how can an empty head such as zero copula assign case? Thus, it must either be the matrix verb which assigns two cases, the case Accusative to the direct object and the case Instrumental to the argument of the secondary predicate as it would be the case with ditransitive verbs such as to give which assign two cases, namely Dative to the first internal argument (Addressee or Goal) and Accusative to the second internal argument (Theme) or with Un accusatives after having incorporated in a light vP or a VoiceP (cf. Kosta 2011:276, and Kosta in

Or we have to do with an Incorporation of a lexical verb into the head of a light verb (like in the case of causative constructions, cf. Kosta 2010, 2011, 2014), which allow then for double case assignment.
In (3), as it would seem on first sight, the adjective as the overt lexical part of the secondary predicate has a depictive and not an attributive reading because it is not the cold tea he drank but the adjective being a part of the predicate ‘cold’ which has a depictive reading and refers to the object.

The variation between Case assignment, argument position (external or internal argument), control and semantics of the predicate seem to display an important role w.r.t. Case assignment between predicate and subject and/or object.

Let us first consider those examples in the National Corpus of Russian Language where the ECM verb считать to consider heads a SC. In (5), we can see that a SC itself can introduce an adjunct clause containing a degree phrase, DegP and embedding a restrictive DP clause.

**ECM verbs assign two cases**

(28) Можно сказать, что я считаю его работающим неравнонечно лучше и гораздо более актуальнее и приспособленнее к современной жизни, чем, скажем, христианство, которое, на мой личный взгляд, изжило себя полностью и превратилось …. Александр Клейн. Контракт с самим собой // «Пятое измерение», 2003
which is v. This ability to my mind allows a light verb to be merged with another, lexical predicate which then bears the lexical or idiomatic meaning, forming a complex predicate structure. But contrary to Bailyn (1999, tree below under 2.3), we do not assume a PredP in its own right because we do not believe that an empty node can assign case but that the incorporation of a lexical verb into a light verb allows for assignment of two cases: accusative and instrumental (cf. section 2.5, 2.6).

2.3 Argument-Small Clauses: Subjects of Causatives (on SC and pro, PRO)


Under MGF, subjects of causative verbs entailing argument-Small Clauses that are selected by the V must be represented structurally. Otherwise, the sentence - regardless of the theory of SC – renders ungrammatical, by the predication condition of Williams (1980). The contrast in (29) vs. (30) can be explained by the presence or absence of a light causative verb, strictly governing the empty category pro in surface object position of the matrix clause and its binding into a SC, coindexed with the PRO (subject of the embedded small or better lexical VP clause).

(29) Questa musica rende pro_{arb} [SC PRO_{arb} allegri ]

(30) *This music renders pro_{arb} [SC PRO_{arb} happy ]

In Russian, the presence or absence of an empty category pro in object position is highly problematic form the viewpoint of RM (cf. Kosta 1992 vs. Krivochen, Kosta 2013). It seems, however, that there is new evidence that in certain contexts, Russian, being itself only a semi-pro-drop-language, replaces such generic sentences with an arbitrary object pro with the generic lexeme people:

(31) Эта музыка делает [pro_{arb}/людь arb]^{13} [SC счастливыми ]

Just like in examples (1)-(4), causative matrix verbs assign the direct object the Accusative case and the lower part of the predication, the adjoined Adjective Phrase of the lower lexical VP, the Instrumental case.

In Russian, we find that Russian secondary predicates, whether arguments (6) or adjuncts (1-5, 7, 10), are marked with Instrumental case. Thus, Instrumental

---

^{13} In our book (Krivochen & Kosta 2013: p. 96, ff., section 4.2.2; also when dealing with object pro, section 4.1.3, p. 87), we treated these kind of sentences as containing a Universal Quantifier at LF, which might or might not be materialized (apparently, in Russian, it does materialize), and which serves to license a predicate in the embedded clause without resorting to PRO. Because there is no other element to fulfill that role, therefore, to prevent crash at the interfaces, the predicate takes that UQ as its subject.

2.4 PredP (Bowers 1993 and Bailyn & Citko 1999)

The theoretical framework, which usually serves as starting point of analysis, is Bowers’ (1993) PredP view of predication (adapted for Russian in Bailyn and Rubin 1991, Bailyn 1995 and then revised in Bailyn and Citko 1999). In this approach, the predicational structure are headed by the functional projection PredP. Thus, under PredP analysis the structure of a small clause

(32) a. Ja, sčitaju ego, durakom

is as shown in (32b) (adapted from Bowers 1993 and Bailyn/Citko 1999:18):

(32) b. [TP I [PredP [t_i consider [VP him [PredP t_k Pred [NP a fool]]]]]]

(32) c.

Under Minimalist assumptions, the morphology of both adjectival and nominal predicates follows from two universal assumptions on well-formedness of linguistic structures:
A. UNIVERSAL: All NPs (including predicates) must have Case checked in an appropriate configuration.

B. UNIVERSAL: All APs (including secondary predicates) must be in an agreement relation with an appropriate head. (cf. Bailyn & Citko 1999:19)

In our view, the structures (32b) and (32c) are not minimalist at all, instead they assume empty nodes that cannot assign case in the first place if the are not Theta-marked or L-marked by a predicate (just like e.g. relational prepositions do not assign Case in their own right because they do not have a Theta-role of their own to assign, cf. Kosta 2011:276).

We already gave an alternative approach in section 2.3, but we give another alternative in 2.6., which seems to be even more minimalist in spirit, and that has already been outlined in our recent work (Krivochen & Kosta 2013: 80seq.), after we have clarified and rejected some further possibilities, namely the feature based approach in 2.5.

2.5. Against a Valuation feature based and in favor of a PHASE based approach

There are several points, which militate against a new projection PrP as proposed by Bowers (1993, 2001) and Bailyn & Citko (1999) and also against a feature based approach in syntax.

Early minimalism, ranging from Chomsky (1989) to Chomsky’s (1995) *Minimalist Program (MP)*, incorporated a weakly derivational approach. The computational system (narrow syntax, CHL₁) manipulates a selection of lexical items (LI) by means of a step-by-step application of the operations Merge and Move, until Spell-Out occurs. Then, PF and LF are created, the two levels of representation interfacing with the syntax-external modules A-P and C-I, respectively. Chomsky’s (2000) *Minimalist Inquiries (MI)* sought to reduce derivational complexity by chopping the lexical array (LA) up into sub-arrays², each feeding CHL to derive a particular phase – a derivational cycle. Phases are well-defined chunks of a derivation – vP, CP, and DP (more on the reasons for this choice below) –, each of which, upon completion, is transferred to the interfaces, and thus does no longer bothers the computation with its weight. This entails a theory of Multiple Spell-Out (cf. Uriagereka 1999). There is no PrP Phase in the Theory, no valuation.

We believe that a feature-based approach cannot be the solution. Anna Bondaruk (2013:82seq.) assumes that the contrast between the two sentences with secondary predicates, one with a non-defective Pred and with a φ-features containing head assigning Instr (34), the other with a defective Pred head with no φ-features and thus with no case to assign (35), can be explained by the presence of a non-defective head:
In order to account for the differences and to account for the nominative case marking in (35) Bondaruk (2013a) adopts the mechanism of case agreement along the lines proposed by Bondaruk (2013b). The proposal is based on two main mechanisms: 1) feature sharing, and 2) probing by a maximal projection. But if the same verb form to be (jesteś) selects two different Pred heads with different values, it would be highly problematic to assume that in one case it is the property of the Pred head to assign instrumental just because there is a full set of φ-features and in the other case we have no features for the probe to find at all. How does the head (call it probe) know how and where to find the appropriate goal if the same verb selects the same Pred head from the same
lexical array? There is another problem with the theoretical approach chosen by Bondaruk (2013), namely the blindness for features in narrow syntax. A rather nice conceptual argument for phases concerns the uninterpretability of features in narrow syntax: If Spell-Out has to remove uninterpretable features in syntax before Spell-Out to avoid a crash at the interfaces, it must know which features are interpretable, and which are uninterpretable. However, narrow syntax is supposed to be blind as a bat, and thus, would need to be able to look ahead (up the tree, to LF and PF) in order to determine the interpretability of a feature. A transfer of derivational chunks to the interfaces remedies this issue, search space being reduced to a local Domain (a Phase) and the rest being left for the interfaces (cf. my plea for an interface oriented approach in Kosta & Krivochen 2014).

If we want to work – just as an intellectual challenge and exercise for thought and for the sake of the argumentation within the Valuation Theory (Pesetsky & Torrego 2007) – within a feature driven Valuation approach, we would have to consider other features than just the Agree and Case Assignment f-features.

We would be forced to assume that there are features to be checked in a functional head which decide between the notion / function stage level vs. individual level predicates. Given there is only one lexical root √, which is highly underspecified and maybe even semantically empty (like Auxiliaries have and be and the expletive verb to do in English), there must be another Projection, call it Phase, which must be the holder domain of features. Since there are only three possible functional domains or workspaces/Phases where functional features can be hosted and checked (namely CP which selects TP, vP which selects VP, and DP which selects NP), we must rely on these categories. One possible solution would be that stage level predicates are events and must check strong eventive features against vP (since vP is aspectual, it is the locus of aspectual features). So first, the lexical item (presumably merged under lexical VP first) as probe has to search the first accessible Phase domain which is not VP but vP. The lexical item itself has unvalued and uninterpretable event features, but interpretable lexical semantics. It has to be checked against the head of v which itself entails not only features of event structure, but maybe (per inheritance) also modal and tense features of the CP Phase. The prediction would then be that only events, change-of-state verbs and actions probe within
this working space and value their event-features. As soon as they are valued they are deleted and the head of $v$ allows the lexical head $V$ to assign case $\text{Instr}$ to its specifier (argument) or complement (adjunct). The case $\text{Nom}$ (if there is no stage level feature available) is checked against the Spec of the $vP$ (just like in case of unaccusative verbs, cf. Kosta 2010, 2011, and Kosta in print, B). No agree relation via long distance agreement is needed. cf. (36):

(36)

Empirical support for phases comes from the EPP-feature on T: how does $C_{HL}$ decide between attracting a subject DP and merging an expletive there? Given the economic preference of Merge over Move (Move is more costly than Merge since $\text{Move} \approx \text{Copy} + \text{Merge}$), an insertion of there should be expected in every instance, and raising should never be possible. Phases circumvent this issue since lexical sub-arrays can be defined for every cycle. To get an idea of the technical side of this argument, first consider the following two examples, which illustrate $\text{Merge-over-Move}$. They share one and the same numeration, but one derivation yields the ungrammatical (37b).

- (37) Num = \{there1, T2, seem1, to1, be1, someone1, here1\}
  
  a. There seems to be someone here.

- b. *There seems someone to be here.
  
  Let’s take a closer look at the steps of the derivation of (37a).
A. \([\text{TP} \text{[EPP]} \text{to be someone here}}] \rightarrow \text{Merge T}

B. \([\text{TP} \text{there T[EPP]} \text{to be someone here}}] \rightarrow \text{Merge there & check EPP}

D. \([\text{TP} \text{T[EPP]} \text{[VP seems [TP there T[EPP]} \text{to be someone here}]]] \rightarrow \text{Merge T}

E. \([\text{TP} \text{there T[EPP]} \text{[VP seems [TP there T[EPP]} \text{to be someone here}]]] \rightarrow \text{Move there & check EPP}

Now consider the derivation of the ungrammatical (37b), based on the same numeration, taking another option at step B.

A. \([\text{TP} \text{T[EPP]} \text{to be someone here}}] \rightarrow \text{Merge T}

B. \([\text{TP} \text{someone T[EPP]} \text{to be someone here}}] \rightarrow \text{Move someone & check EPP}

D. \([\text{TP} \text{T[EPP]} \text{[VP seems [TP someone T[EPP]} \text{to be someone here}]]] \rightarrow \text{Merge T}

E. \([\text{TP} \text{there T[EPP]} \text{[VP seems [TP someone T[EPP]} \text{to be someone here}]]] \rightarrow \text{Move there & check EPP}

The derivational step B of (37b) violates \text{Merge-over-Move}, moving someone instead of merging the expletive there available in the Num, which is why the derivation produces an ill-formed sentence. Defining different sub-arrays for (37b), provided the phasehood of \(vP\) and CP, can capture this issue derivationally:

(38) a. \{"C, T\}3 \{\text{seem, there, T, to}\}2 \{\text{be, someone, here}\}1\}

There seems to be someone here. – (2)

b. \{"C, there, T\}3 \{\text{seem, T, to}\}2 \{\text{be, someone, here}\}1\}. *There seems someone to be here. – (2)

c. \{"C, T\}3 \{\text{seem, T, to}\}2 \{\text{be, someone, here}\}1\}. Someone seems to be here. – \text{w/o expletive}

The only way how to capture this problem avoiding the look ahead problem within a highly non-minimalist valuation approach is to define two different lexical arrays, one in which the verb \text{to be} is classified as a stage level predicate
and is merged (or adjoined) at the lowest possible domain or incorporated as in our theory which we will adopt from Radical Minimalism.

If Bondaruk (2013) has to have the same lexical array (and not two different Pred heads in maybe two different positions), she needs to decide which of two competing constructions to select, but this is not possible because in her model both lexical arrays would be identical for narrow syntax. Phases consist of three parts: the phase head H (v/C), its complement ZP (the internal domain), and its edge YP (the specifier domain). After completion, the phase is inaccessible to further operations, as formally captured by the Phase-Impenetrability Condition (PIC):

(38) **Phase-Impenetrability Condition (PIC):**

“In phase α with the head H, the domain of H is not accessible to operations outside α, only H and its edge are accessible to such operations.” (Chomsky 2000:108)

Our view of case assignment is different: it is based on accessibility, not on feature-related assumptions. Under the theory explained in Krivochen (2012: 78, ff) and Krivochen & Kosta (2013: 93, ff.), case is not strictly assigned, but read off a local syntactic configuration. We have already pointed out that an empty zero Pred head should not be able to assign case, except under special stipulations, and argued that V-to-ν incorporation might help solving the problem. Defectivity in a certain head is a stipulation that should be avoided if possible, therefore, we will stick to fully LF-interpretable nodes. Let us consider a Hale & Keyser (2002)-like structure for transitives:

(39) [TP [vP [VP [PP]]]]

With vP comprising causativity, VP eventivity, and PP introducing two arguments: theme and location, If an argument is accessible by T, this local configuration will be read off at LF and PF as Nominative, as argued in past works. Accessibility has to do with the absence of an intervenient head and with

---

14 It is possible that each of these nodes configures a cycle, perhaps v-V being both a single cycle (i.e., a single phase) if the definition of phase is sensitive to the kind of information conveyed by each projection (locative, causative-eventive, and temporal respectively).
the interfaces filtering out sub-optimal candidates, such that a relation that would be forced by orthodox Minimality could be overseen if necessary (as in the case of Latin Long Distance Anaphors, which are bound outside their governing category, thus ignoring their closest possible governor). This means that, if there is no initiator, the outer argument of P can receive Nom case, as the more local relation $\nu$-Spec PP could be ignored because it would lead to crash at the interfaces. If, however, there is an argument between T and Spec-PP, then this argument will be Nominative (the subject / initiator of transitive caused events) and the theme will receive an Acc-related case, in turn related to its semantic contribution. At this respect, P is essential not only for Dative interpretation (a locative case *par excellence*), but also other locative-related cases, like Genitive (e.g., Source Genitive in ancient Greek) and Ablative (e.g., Latin’s *ubi* adjuncts). Instrumental case seems to us to be a variant of non-Nominative which, as any other case, is interpreted within the $\nu$P domain. The contrast Bondaruk (2013) notices can thus be subsumed to the two variants of P (Hale & Keyser, 2002):

Theme [TO] Location (e.g., Prepositional Indirect Object Constructions, like ‘Give the book to mary’)

Location [WITH] Theme (e.g., Double Object Construction, like ‘Give Mary the book’)

Semantically interpretable variants of P head can account for variants in Case ‘assignment’ without the need to resort to feature processes, if the definition of Locality is semantically sensitive (as we proposed in Krivochen & Kosta, 2013: 179), thus derivable as a third-factor (i.e., interface-based) constraint over syntactic representations. Maybe, this needs to be explained in terms of two competing constructions of Small Clauses, one assigning the instrumental, the other assigning the accusative. I shall now proceed with section 2.6 in which our own theory of Radical Minimalism and case will be presented.

2.6. Expletive pro in small clauses: an incorporationist perspective

In this section we will focus on the derivation of so-called “small clauses” (SC), in which the null element *pro* plays a very important role\(^\text{15}\). To account for the

\(^{15}\) On pro-drop cf. Franks (1995, chapter 5), Kosta (1992), but above all the recent study Krivochen/Kosta (2013) and Corbara (2012, in print) and from diachronical point of view Meyer (2012) and Timberlake (2014).
derivation of SC in an elegant way, we will retrieve an intuition of Abney at the
beginning of his thesis, namely, that there is an Agr-like node in nominal
structures. That very same Agr node would be present, for example, in SC.
Bosque and Gutierrez Rexach (2008: 423) define SC as “quasipropositional
units without verbal inflection” (our translation). The absence of morphological
verbal inflection is often confused with the absence of agreement, especially
within the Anglo-Saxon linguistic literature. In English, SC have no overt
agreement between the predicate and argument. Thus, a sentence like

(40) I consider [him intelligent]

has been analyzed as a transitive sentence with an “ECM” verb and a SC as a
complement, whose structure is as follows (adapted from Chomsky 1986)\textsuperscript{16}:

\begin{center}
\begin{tikzpicture}
  \node (AP) {AP};
  \node (him) [below = of AP, xshift = -2cm] {him};
  \node (A') [right = of AP, xshift = 2cm] {A'};
  \node (intelligent) [below = of A', yshift = 2cm] {intelligent};
  \draw (him) -- (AP);
  \draw (A') -- (him);
\end{tikzpicture}
\end{center}

A is a predicate and thus licenses a specifier (Hale and Keyser 1993), the
external argument and so-called subject of the SC. In Minimalist accounts,
external accusative case is assigned to [him] by the node v* to which V
adjoins via head-to-head movement in the course of the derivation. This node
would have a valued (thus, interpretable, following Chomsky 1999)
[Transitivity] feature and unvalued φ-features to check with the external
argument of AP, the closest goal in its c-command domain. While this
acknowledges the fact that [him] is an argument of [intelligent] and not of
[consider] the failure of this proposal is the fact that it does not provide an
explanation for the morphological concord between Prn and A, visible, for
example, in Spanish, Polish, Serbian or in Czech and Slovak (in ECM and AcI
constructions and also in constructions of resultative state):

(41) a. Los considero [los inteligentes] (Sp.)
    \[\text{CL}_{\text{ACCPl}} \text{consider}_1 \text{SgPrslmpt} \ [\text{CL}_{\text{ACCPl}} \text{intelligent}_p]\]
    “I consider them intelligent”

b. Znalazłem go pijanego /* pijanym
    \[\text{pro found}_1 \text{SgPret} \text{him}_{-\text{CLAkkSg}} \text{drunk}_{-\text{CLAkkSg}} /*\text{drunk}_{-\text{InstrSg}}\]

\textsuperscript{16} We will ignore proposals that take SC to lack a head and take the argument and the
predicate as sister nodes in a configuration like (i) since, even within traditional X-bar
theoretical assumptions, it violates endocentricity:

i) [SC [DP] [AP]]
c. Našao sam ga pijanog /* pijanim
   pro found1SgPret him-CLAkkSgm drunk-CLAkkSgm /*drunk-InstrSgm

d. Švejk jí měl pěkně vyčištěnou.
   Švejk her-CLAkkSgf had3SgPret nicely cleanedAkkSgf

e. Mat’ ju našla vyplakanů
   mother her found3SgPret cryingAkkSgf (examples 18 b, c and e from Bailyn & Citko 1999:27, 19 a-c).

This concord was accounted for in the early 90's from the perspective of Agr projections. According to this analysis, there is an AgrA node that would dominate AP, and the external argument of the adjectival SC would rise to take inflectional features in a Spec-Head relation with ConcA₀. This proposal was abandoned by most researchers when Chomsky (1995) argued against Agr projections because they are nothing more than ϕ-features receptacles, without interpretable information for the interfaces, thus implying a violation of Full Interpretation; and Spec-Head relations were replaced by probe-goal relations, which can take place at a distance, Spec-Head configurations being the result of independent EPP-related phenomena.

Krivochen (2010a) intends to recover Abney’s idea of D (D ≠ Determiner) as an Agr node, but with a fundamental difference with respect to traditional agreement projections: the presence of interpretable features. Krivochen (2010a) proposed that the procedural category D comprised four semantically interpretable dimensions:

(42) a. Referentiality
b. Definiteness
c. Person / Number
d. Case

If Referentiality is a dimension, then it can take a negative value (since there is no referent assignment process affecting the SC, as it is not a sortal entity, see Panagiotidis 2010), and the projection would not be a referential DP (i.e., a denotative definite description), but it would keep the agreement ϕ-features. [-Ref] is as interpretable as [+ Ref], so there is convergence at the interface levels without violation of the Principle of Full Interpretation. Every SC, then, would be a DP projection (in traditional X-bar terms, a {D} structure with a procedural D nucleus in our 2011e atomic phrase structure theory), which takes the projection of the A lexical predicate as a complement, with which it forms an extended projection (Grimshaw 1991). Given this scenario, the derivation of the small clause in (40) would be, following Krivochen (2010a):
But this representation is not yet correct in Radically Minimalist terms, since (among other complications) it assumes the fact that A is a primitive category in the grammatical system, in the line of Hale and Keyser (1993 et. seq). Following Mateu Fontanals (2000a, b), Jackendoff (1990, 1997, 2002) and our own work within Radical Minimalism (see specially Krivochen 2011d, e and Krivochen & Kosta 2013:80), we will decompose A into a configuration of the type [P…√], since to ascribe a property to an entity is to conceptually locate that entity within the sphere of the property, from a localist point of view (or asserting its being part of the predicate’s extension, from a logical point of view). Moreover, individual-stage adjectival predication involves an atelic static eventive node, the reason why a [BE] {event} node is necessary for semantic interpretation. We do not need to posit unvalued referential features in D, since we have already abandoned the probe-goal relation and Agree as an operation in the computational system. Such a modification would yield a simpler (and theoretically more correct) representation, namely

\begin{equation}
\{D\}
\end{equation}

In this representation, the root would incorporate à la Baker (1988) onto P and then {event} and D₀, motivated by the P / eventive “defective” character at PF (Hale and Keyser 2002). This process can also be accounted for via our External
Merge Theory of Movement in terms of Merge of a token of the root in the aforementioned nodes, and when Chain Reduction (see Nunes 2004 and Chapter 5 of Krivochen & Kosta 2013) takes place at the Morphological component, our Anti-Spell-Out principle comes into play. For semantic purposes, however, all the properties of the relevant nodes count, and therefore the final interpretation is that of an individual level predicate (by virtue of the nature of P), which applies to a sortal element in an atelic static way (by virtue of the procedural instructions conveyed by the eventive node). \{D\} transforms the whole clause into an argument, nominal in nature and for conceptual purposes.

The representation can also explain the SCs in which the argument is placed after the predicate, as in:

(45) \text{Consideran [SC inteligente a María] (Sp.)}
\text{Consider \footnotesize{\text{3PIIIPresImpf}} [intelligent to Mary]}
\text{“They consider Mary intelligent”}

The SC in this case, has no theme in informational terms, becoming a thetic clause. The ungrammaticality of similar structures in English is derived from interface conditions, namely, from the consideration of the semantic underspecification of Spanish and English roots, which leads us to review the proposed structures. In Spanish, the root $\sqrt{\text{consider-}}$ (and the other roots that can instantiate as verbs which take a SC as a complement) within a “lexical” verbal/eventive level allows the incorporation of yet another root to form a complex predicate: $[\text{V } \sqrt{\text{consider-intelligent}}]$, although this incorporation is not strictly necessary. The root in question is semantically underspecified, but not severely underspecified, so there is no collapse explicature level if the root is left alone in $V_0$. Incorporation enables the construction of an explicature with a more specific eventive reference (in combination with Time and Aspect nodes), while the root in situ leaves a "lighter" main V. The structure (41a), therefore, is (46)

(46) \text{Ellos [\{v } \sqrt{\text{consider-intellig-} } \text{\}  [d [D] [event [BE] [p María \text{\{WITH\} } \sqrt{\text{ intellig-} } ]\text{\}]]]}

In English, on the other hand, it seems to be the case that verbal roots with which SC appear are semantically specified enough to saturate the valence of the terminal node $V_0$, so the addition of another root is not necessary for the construction of an explicature. Therefore, optimally, there is no incorporation, since it would involve extra processing effort with very little benefit, if any at all. The only option left, according to basic economy principles, is to have a “heavy” verb that selects a SC in which the subject has risen for thematic reasons, as mentioned above.
Taking up the idea of how to explain the assignment of Instr to the lower nominal or adjective secondary predicates, we can use the same mechanism of decomposition of lexical categories and incorporation. E.g., The light verb is a kind of semantically defect predicate, which does not have the ability to assign Theta-role to its internal argument in its governing category, and thus, by definition, cannot assign case on its own. It needs to attract the lexical verb. Only then, it can assign theta role to the internal argument and case accusative in SpecvP. There is, however, still a problem how to explain the case assignment of the Instr case to a position, which seems to be either an argument (complement) or adjunct position (on the differentiation between A- and A’-positions from the viewpoint of RM cf. Kosta & Krivochen 2014). The scenario would then be the following:

\[
\text{rabotat'} (V) - \text{rabotat' uchitelem (vP)}
\]

intr. tr.

On the level of system, there seem to be two different verbs: (1) rabotat’ and (2) rabotat' uchitelem, (1) a general meaning ‘to work’, and second a stage level meaning ‘to exercise the profession of teacher’ (2).

Since (2) is a stage level predicate which is true for the time span of the profession only, it must have been derived etymologically and also derivationally from the more general (intr.) lexical verb rabotat’ (not limited to any temporal, modal or aspectual restrictions) via incorporation of the general lexical verb into vP. Semantic decomposition of the verb rabotat' uchitelem is thus an important Zero hypothesis (prerequisite) for my analysis:

Lexical Array: \{√ RABOT-, √ UCHIT'\}

(i) Merge ROOT √RABOT- with lexical VP

(ii) Merge ROOT √UCHIT' with the complement/adjunct of the lexical VP

\--> (iii) Re-Merge (vP) with the head of the light vP
(Merge over Move (Phase model))

(iv) The product of the incorporation assigns theta role to VP (via L-Marking), V assigns per default Instr to the complement/adjunct!
To summarize, we have accounted for the derivation of Small Clauses without the need to resort to pro_{expl} for changes in the argument-predicate phonological order inter-linguistically.

3. Further perspectives of research on predicative Instrumental vs. Nominative

In this contribution I was able to show how the system works in Russian and at least in Spanish and English. This analysis should, however, be extended to other Slavic languages. Since I am working in the domain of comparative Slavic Syntax, I would like to give some further perspectives of research pointing especially at Polish syntax as compared to Russian and Czech later. Since Polish is the second major Slavic language at our department and we are responsible for the teachers programs both in Russian and Polish, I would like to mention the same problem in Polish which has already been addressed in some recent functionally oriented studies on predicative instrumental vs. nominative. Gerd Hentschel is maybe the best-known specialist working on secondary predicative constructions in Polish and other Slavonic languages (cf. a summary in Hentschel 2009). He also mentions studies, where the use of predicative Instrumental vs. Nominative have been assumed to be closely connected to the opposition change of state vs. permanence\textsuperscript{17}. I believe that the long-standing discussions and still ongoing hot debates on this distinction have to do with the unclear status of the definition change of state vs. permanence. Moreover, the facts are more complicated since we often have to do with a dynamic of language evolution and we cannot capture this very fine-grained functional difference synchronically since all is in flow so to say. There is also the fact of language contact which can be seen especially in case of Czech and Sorbian which have a strict loss of instrumental in predicative position (under the influence of German) and on the other hand the tendency to typological language shift in South Balkan languages Macedonian and Bulgarian (which are more or less ‘case-less’), Slovenian (maybe also language contact) and Serbian,

---

Croatian and Bosnian somewhat in between these two influences. For the time being, I thank you very much for your attention.

References


(2010a) Defeating Lexicocentrism. lingBuzz/001130.

(2010b) What Principles and Parameters got Wrong. Ms. ICREA / UAB.


Grimshaw, J. (1991) Extended Projection. Ms, Brandeis University, Waltham, MA.


Online: http://www.uni-potsdam.de/u/slavistik/wsw/habil/habil.htm


(2013a), How can I lie if I am telling the truth? - The unbearable lightness of the being of strong and weak modals, modal adverbs and modal particles in discourse between epistemic modality and evidentiality. In: Thielemann, Nadine & Peter Kosta (eds.). Approaches to Slavic Interaction Amsterdam, Philadelphia: John Benjamins, 167-184.

(in print A) Case and Agree in Slavic Numerals – Valuation of Features at the Interfaces Within a Phase-Based Model (under review, submitted for Zeitschrift für Slawistik).

(in print B) On the Causative/Anti-causative alternation as principle of affix ordering in the Light of the Mirror Principle, the Lexical Integrity Principle and the Distributive Morphology. http://ling.auf.net/lingbuzz/002055


(2010c) Theta Theory Revisited. UNL. Under Review.


(2000b) Why Can’t We Wipe the Slate Clean?. Universitat Autónoma de Barcelona. Departament de Filologia Catalana.


(2011b) Case, Tense and Clausal Arguments in Polish. Ms. Adam Mickiewicz University, Poland.


Stowell, Tim. 1978. What was there before there was there? In D. Farkas et al. (eds.) Papers from the Fourteenth Regional Meeting of the Chicago Linguistics Society, 457-471.


