



NP2 copular agreement in Czech presentational contexts

Kučerová, Ivona
Onea, Edgar

(Citation)

Papers from the Second International Workshop on the Syntax of Predication and Modification:43-58

(Issue Date)

2026-02-25

(Resource Type)

conference proceedings

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/0100501355>

(URL)

<https://hdl.handle.net/20.500.14094/0100501355>



NP₂ copular agreement in Czech presentational contexts*

Ivona Kučerová and Edgar Onea
McMaster University | University of Graz

Abstract

We discuss a class of classificatory NP-NP copular clauses, such as *This tree is a linden tree*. They display NP₂ φ -feature agreement (copular and inter-sentential anaphoric agreement), even though NP₂ is a predicate. We argue that this agreement is mediated by a complex covert pronoun in the Middlefield topic position. The proposal thus predicts that only languages in which the covert pronoun can be merged in the Middlefield position may have NP₂ agreement. Second, even if a language allows NP₂ agreement, NP₂ agreement arises only when no other element occupies the Middlefield topic position. Moreover, our proposal has semantic consequences. We argue that semantic double-classification of an individual (as a tree and as a linden tree, in our example) must involve two derivational steps (two applications of extended conceptual covers, Aloni 2001).

Keywords: NP-NP copular clauses, NP₂-agreement, topic, Middlefield, classification.

1. The empirical challenge

1.1. Copular clause primer

Since the seminal work of Higgins (1979), formalized in contemporary terms in Mikkelsen (2005), scholars distinguish at least three types of NP-NP copular clauses: **predicational** clauses, as in *Mary is a doctor*, **specificational** clauses, as in *The problem is the children*, and **identificational** clauses, as in *This doctor is Max*. The differences can be nicely demonstrated in Czech, a Western Slavic language in the centre of our investigation, as Czech displays morphosyntactic properties that highlight the structural differences among these three types of copular clauses (Bartošová 2017, a.o.). In this

* Ivona Kučerová's research for this paper was supported by the Social Sciences and Humanities Research Council of Canada (Insight Grant #435-2025-1183 "Nominal integration in associative, comitative and accompaniment structures: A comparative syntax study"). Edgar Onea's research for this paper was supported by the Austrian Science Fund (FWF) (Project: "SFB Language between Redundancy and Deficiency", number F1003, Subproject 2, "Grammatical Gender in Discourse and Grammar."). We thank the audiences at the Bielefeld-Leiden Comparative Syntax Workshop and the 2nd International Workshop on the Syntax of Predication and Modification for their questions and helpful suggestions. Special thanks to Jutta Hartmann, Sjeef Barbiers and Marcel den Dikken.

language, only nominals in Nominative case may trigger φ -feature agreement, and, conveniently, subject NPs of copular clauses must be in Nominative. Moreover, the copular verb (a form of *be*) is always overt and agrees in number and person in all tenses, and crucially, it also agrees in (grammatical) gender in the past. The latter property makes it possible to track which nominal within a copular clause is the source of φ -feature agreement. As for predicate NPs, they may alternate between Nominative and Instrumental case, that is, between two structural cases that reflect aspect and event properties of the copular clause. Very roughly, only stage-level predicates display this case alternation. Crucially, individual-denoting NPs never display this case alternation.

With this background in mind, let us consider an example of a predication clause, (1). Here, NP₁ (*Marie*) is **individual-denoting**, and NP₂ (*professor*) is a predicate. As the example demonstrates, the copular agreement in predication clauses is always with NP₁. This NP is in Nominative. NP₂ cannot trigger agreement and may alternate between Nominative and Instrumental.

- (1) *Marie byla/ *byl profesor/ profesorem*
 Marie.F.SG was.F.SG/was.M.SG professor.M.SG.NOM/INST
na univerzitě.
 at university
 ‘Marie was a university professor.’

Specificational and identificational clauses are crucially different in that their NP₂ is **individual-denoting**, and in some languages, it may trigger agreement (Moro, 1991, a.o.). Czech is such a language. An example of a specificational clause is given in (2). Here, NP₂ specifies or provides an answer to a question raised by NP₁, which gives rise to implementations where NP₁ is construed as a concealed question or a proposition (Romero 2005, Heycock 2012, a.o.). NP₂ is also the source of agreement and must be in Nominative. NP₂ alternates in case.

- (2) *Ten problém/ problémem jsou děti.*
 the.m.sg problem.M.SG.NOM/INSTR are children
 ‘The problem is the children.’

In identificational clauses, NP₂ assigns some form of a referential identity to NP₁. This can be modeled, for example, via a special semantic operator, such as the one proposed in Percus and Sharvit (2024). As we can see in (3), in these clauses the copular agreement is triggered by NP₂. In these clauses, NP₁ is mostly in Nominative. Instrumental requires a special scenario, such as a role-playing scenario; otherwise, the temporal restriction that Instrumental enforces on the referential identity matching is judged as infelicitous (e.g., Uličný 2000), (4).

- (3) *Ten doktor byla Marie.*
 this.M.SG doctor.M.SG.NOM was.F.SG.NOM Marie.F.SG.NOM
 ‘This doctor was Marie.’

- (4) *Tim doktorem byla Marie.*
 this.M.SG doctor.M.SG.INSTR was.f.SG.NOM Marie.F.SG.NOM
 ‘Marie was (=played/ personified) the doctor.’

Existing accounts of NP₂ agreement implement the overarching intuition that in the relevant languages NP₁ of identificational and specificational clauses is not structurally accessible to agreement. This may be because the nominal is not the underlying subject and it only moved to the subject position (Mikkelsen 2005, Dikken 2006, a.o.), or because it corresponds to a more complex structure that lacks φ -features (e.g., a concealed proposition or a concealed question; Romero 2005, Heycock 2012, a.o.). Under these accounts, the crosslinguistic variation between NP₁ and NP₂ agreement is then reduced to differences in the feature composition of the probe, for example, as in Béjar and Kahnemuyipour (2017, 2018).

1.2. Enter Czech and German classificatory clauses

In addition to these three types, Czech (and also German) exhibits a distinct class of copular clauses which has previously been characterized as a class of identificational clauses (Higgins 1973, Mikkelsen 2005). They are typically used in a presentational context but are not necessarily constrained to it. We call them *classificatory clauses* to distinguish them from typical identificational clauses with an individual-like NP₂. In this paper, we only discuss Czech data because they display gender and number agreement in all relevant environments. Consider the data in (5).

- (5) a. *Tenhle strom byl/ byla lípa.*
 this tree.M.SG was.M.SG/was.F.SG linden_tree.F.SG
 ‘This tree was a linden tree.’ (e.g., while pointing at a tree stamp)
- b. *Tenhle film byl/ byla komedie.*
 this movie.M.SG was.M.SG/was.F.SG comedy.F.SG
 ‘This movie was a comedy.’ (e.g., while pointing at an old movie poster)
- c. *Tahle sazenice byla/ byl dub.*
 This seedling.F.SG was.F.SG/was.M.SG oak.M.SG
 ‘This seedling was an oak.’ (e.g., while pointing at a pile of seedlings roughed up by a puppy)

The first property that distinguishes these clauses from their better-established counterparts is that both NPs must be in Nominative. There is no case alternation we have observed with the other types of copular clauses. Second, these clauses exhibit **variable agreement**. That is, the copula can agree either with NP₁ or with NP₂. There is no evidence that the difference in agreement would correlate with movement of one of the NPs. Third, the φ -features (person, number, gender) of an inter-sentential anaphor can match the φ -features of either NP₁ or NP₂, as demonstrated in (6). Czech is a *pro-drop* language. We include a pronominal subject in these examples, but the more natural version is with the subject being dropped and its features only being reflected on the verb.

- (6) *Tenhle strom byl/ byla lípa.*
 thistree.M.SG was.M.SG/F.SG linden_tree.F.SG
 ‘This tree was a linden tree.’
- a. *(On) byl vysoký.*
 (he) was.M.SG tall.M.SG
 ‘It[=the tree] was tall.’
- b. *(Ona) byla vysoká.*
 (she) was.F.SG tall.F.SG
 ‘It[the linden tree] was tall.’

These three properties are entirely unexpected once we consider an additional critical fact, namely, that in classificatory clauses, NP₂ is always a predicate. Unlike in other copular clauses where NP₂-agreement has been associated with individual-denoting properties of NP₂, in classificatory clauses NP₂ is not a referential expression or an existential quantifier of a sort. This can be simply demonstrated by the observation that NP₂ in these clauses does not display any scope interactions with any known operator. The pseudo-Czech examples in (7) demonstrate the readings we do not get.

- (7) a. This tree is not a linden tree.
 ≠ There is a linden tree such that this tree is not it.
- b. Every tree is a linden tree.
 ≠ There is a linden tree such that that every tree is it.

Note that even if we were to assume that the copula construction (PredP') was a scope island, this would still not help, because it is well known that indefinites should be able to ignore scope islands. Hence, any analysis that uses the assumption that NP₂ is referential to derive agreement with NP₂ seems empirically problematic.

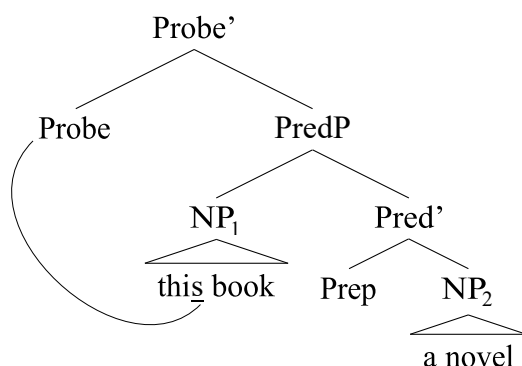
We are now in a position to formulate the puzzle and our research questions. First, in other structural environments in Czech, predicates do not trigger agreement in gender and number. Instead, default agreement, i.e., agreement in neuter singular, is attested with them. However, in classificatory clauses, the copula seems to agree with a predicate. This is our first puzzle. The second puzzle concerns anaphors. In other environments, anaphors referring to non-individuals display default agreement (neuter singular). However, in classificatory clauses, an anaphor seems to agree with a predicate.

These properties also demonstrate how this class of copular clauses differs from its better-established counterparts. That NP₁ is individual-denoting and that NP₂ is a predicate makes these clauses similar to predicational copular clauses. In contrast, the fact that NP₂-agreement is possible and the fact that NP₂ cannot be in Instrumental case makes them look more like identificational clauses.

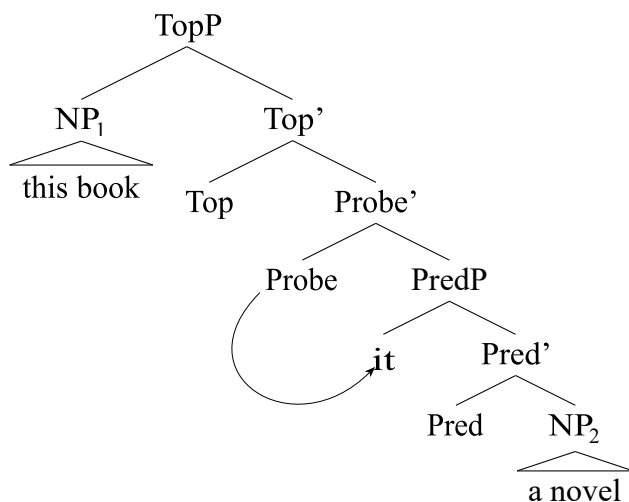
In order to account for these properties, we will argue that we are dealing with a case of a structural ambiguity, and that the agreement variability—both for copula agreement and inter-sentential anaphors—results from two underlying structures. The first structure corresponds to a regular predicational clause. This structure will give rise to an obligatory

NP₁ agreement. The second structure is more complex. We argue that true classificatory clauses (we will define below what exactly we mean by that) are derived within presentational clauses in which NP₂ agreement is mediated by a complex covert pronoun which we will call *it*. This pronoun will inherit features of NP₂ (via concord), and it will inherit referential properties of NP₁ in a binding-like configuration. We will motivate the presence of the covert pronoun by the semantic interpretation of classificatory clauses, and by structural restrictions on when this special type of semantic interpretation arises. A simplified version of the predication structure we will assume is given in (8), and a simplified version of the more complex structure of classificatory clauses is given in (9).

(8) *A simplified predication structure:*



(9) *A simplified presentational structure:*



While in the predication structure, (8), NP₂ predicates directly over NP₁, and NP₁ is the closest accessible goal for the active probe, the presentational structure, (9), is more complex. Here, NP₂ does not predicate over NP₁. Instead, it predicates over the covert pronominal. NP₁ is merged higher in the structure, namely, in the specifier of TopicP, which makes it inaccessible to a low agreement probe. Agreement is then with the covert pronoun, and

indirectly with the φ -features of NP₂.

The consequence of this proposal is that when NP₁ cannot be merged in TopicP either because the language lacks a Middlefield TopicP (as in English), or because the position is occupied by another structural element (we will see examples below), only NP₁ agreement is possible.

The story for the anaphoric agreement will be slightly less straightforward because we need to account for the fact that even predicates can indirectly function as anaphoric antecedents if the pronoun is interpreted as an elliptical definite description that has a presupposition entailed by the initial sentence (Elbourne, 2005). Since such an interpretation requires an accommodation, we will rely on subtle judgement differences between anaphoric expressions with a structurally present antecedent (high ratings) and those where the antecedent needs to be accommodated (lower ratings, slower processing). With this clarification in place, anaphoric agreement displaying φ -features of NP₂ will either result from a direct agreement with *it*, or via the Elbourne-style of accommodation procedure with NP₂.

In addition to discussing a previously unreported type of copular clauses, our proposal reduces crosslinguistic differences in agreement patterns within copular clauses to independently motivated restrictions on the availability of argument positions in the structure, namely, the availability of the Middlefield topic position.

2. The case for a covert pronoun

Our analysis critically relies on the assumption that there is a covert pronominal element in the structure. Since diagnosing covert elements is notoriously difficult, we will first discuss the evidence that points in the direction of such an element. Let us first recall the basic properties of Czech copular clauses. The essential contrast we are trying to explain is between predicational clauses, such as that in (10), in which the copula must agree with NP₁, and classificatory clauses, such as those in (11), where NP₂ classifies NP₁ and the copula may agree either with NP₁ or with NP₂. Crucially, NP₂ retains its predicative properties in both agreement patterns. Note also that the pattern cannot be explained by some form of markedness, as any combination of gender features is possible in these sentences, and the agreement is still variable.

- (10) *Marie* *byla/* **byl* *profesor* *na univerzitě.*
 Marie.F.SG was.F.SG/was.M.SG professor.M.SG at university
 ‘Marie was a university professor.’

- (11) a. *Tenhle strom* *byl/* *byla* *lípa.*
 this tree.M.SG was.M.SG/was.F.SG linden_tree.F.SG
 ‘This tree was a linden tree.’ (e.g., while pointing at a tree stamp)
 b. *Tenhle film* *byl/* *byla* *komedie.*
 this movie.M.SG was.M.SG/was.F.SG comedy.F.S
 ‘This movie was a comedy.’ (e.g., while pointing at an old movie poster)

- c. *Tahle sazenice byla/ byl dub.*
 This seedling.F.SG was.F.SG/was.M.SG oak.M.SG
 ‘This seedling was an oak.’ (e.g., while pointing at a pile of seedlings
 roughed up by a puppy)

The fact that the copula appears to agree with a predicative NP, i.e., NP₂, is unexpected. For predicates do not trigger gender/number agreement in Czech. Instead, they appear in default agreement configurations (neuter singular), as exemplified by (12).

- (12) *Být pracovitý bylo užitečné.*
 to_be hard_ working.M.SG was.N.SG useful.N.SG
 ‘To be hard working was useful.’

Critical insights come from anaphoric agreement properties. Even though the φ -features of an inter-sentential anaphor can match either NP₁ or NP₂, speakers prefer the copular and anaphoric agreement to match (confirmed with several speakers of Czech in interviews), but a mixed pattern is accepted as well (13).

- (13) *Tenhle strom byl/ byla lípa.*
 this tree.M.SG was.M.SG/F.SG linden_tree.F.SG
 ‘This tree was a linden tree.’
 a. *(On) byl vysoký.*
 (he) was.M.SG tall.M.SG
 ‘It[=the tree] was tall.’
 b. *(Ona) byla vysoká.*
 (she) was.F.SG tall.F.SG
 ‘It[=the linden tree] was tall.’

Similar to predicates, Czech anaphors display default agreement (neuter singular) when referring to non-individual denoting phrases. The distinct continuations in (14a) and (14b) demonstrate this. When the antecedent of the anaphor is construed as an individual, then the anaphor agrees in φ -features with the individual-denoting NP (14a). When the antecedent is propositional or situation-denoting, then the anaphor is in neuter singular (14b).

- (14) *Petr chodil s dívkou z Prahy...*
 Petr walked with girl from Prague
 ‘Petr dated a girl from Prague.’
 a. *Ta byla nudná.*
 that.F.SG was.F.SG boring.F.SG
 ‘She was boring.’
 b. *To bylo nudné.*
 that.N.SG was.N.SG boring.N.SG
 ‘It [=dating the girl from Prague] was boring.’

Moreover, as (15) shows, predicative NPs in predication copular clauses cannot function as antecedents of anaphors. This fact also demonstrates that being an NP in Nominative is not a sufficient condition to function as an antecedent of an individual-denoting anaphor.

- (15) *Marie byla profesor na univerzitě.*
 Marie.F.SG was.F.SG professor.M.SG.NOM at university
 ‘Marie was a university professor.’
 a. *Byla úspěšná.*
 was.F.SG successful.F.SG
 ‘She was successful.’
 b. *#Byl úspěšný.*
 was.M.SG successful.M.SG
 Intended meaning: ‘She was successful.’ (can only mean ‘some previously mentioned male was successful’)

Yet, in classificatory copular clauses inter-sentential anaphors can share φ -features with a predicative NP – NP₂. This is entirely unexpected.

Moreover, the observed variability in anaphoric agreement correlates with distinct semantic interpretations. While the anaphor agreeing with NP₁ is compatible with a range of predicates, the anaphor agreeing with the φ -features of NP₂ restricts the interpretation of the individual denoted by NP₁ only to the semantic content classified by NP₂. We demonstrate what we mean first with the English example in (16). As the continuations demonstrate, the meaning of the anaphoric *it* is ambiguous between the book as the physical object and the content of the book (i.e., as being a novel).

- (16) This book is a historical novel.
 a. It is heavy [=the book as a physical object, not the novel as its content]
 b. It is engaging [= the content of the book being a novel]

As (17) demonstrates, when the copula of a classificatory clause agrees with NP₁, we retain the antecedent ambiguity attested in English. However, when the copula and the anaphor agree with NP₂, only the continuation referring to the classified content of the book, that is, the book being a novel, is felicitous (18).

- (17) *Tahle knížka byla historický román.*
 this.F.SG book.F.SG was.F.SG historical novel.M.SG
 ‘This book was a historical novel.’
 a. *Byla těžká.*
 was.F.SG heavy.F.SG
 ‘It was heavy.’ [=the book as a physical object, not the novel as its content]
 b. *Byla nudná.*
 was.F.SG boring.F.SG
 ‘It was boring.’ [= the content of the book being a novel]

- (18) *Tahle knížka byl historický román.*
 this.F.SG book.F.SG was.M.SG historical novel.M.SG
 ‘This book was a historical novel.’
- a. *#Byl těžký.*
 was.M.SG heavy.M.SG
 Intended: ‘It was heavy.’ [=the book as a physical object, not the novel as its content]
- b. *Byl nudný.*
 was.M.SG boring.M.SG
 ‘It was boring.’ [= the content of the book being a novel]

The profile of the anaphoric data strongly suggests that even though it looks like the anaphoric agreement is directly with NP₂, it is in fact with a *new object* created by NP₂ classifying over NP₁.

There is one more piece of data we have not discussed yet. The variable agreement only arises when the copula linearly intervenes between NP₁ and NP₂. When the copula precedes NP₁, the copula must agree with NP₁, as in (19). In this case, anaphoric agreement with NP₂ requires an accommodation of the corresponding antecedent, which is readily available in configurations that yield NP₂ copular agreement.

- (19) *Podle mého názoru, byl/*byla tenhle strom lípa.*
 according_to my opinion was.M.SG/*was.F.SG this tree.M.SG linden_tree.F.SG
 ‘In my opinion, this tree was a linden tree.’
- a. *(On) byl vysoký.*
 He was.M.SG tall.M.SG
 ‘It was tall.’
- b. *??(Ona) byla vysoká.*
 she was.F.SG tall.F.SG
 ‘It was tall.’

A similar word-order pattern has been reported for other NP₂-agreement copular clauses. Specifically, Hartmann and Heycock (2018, 2020) report that Germanic speakers (German, Icelandic, Faroese) require NP₂ agreement in specificational and identificational clauses when the copula intervenes between NP₁ and NP₂, while they allow or even strongly prefer NP₁ agreement in configurations in which the copula linearly precedes NP₁. The Germanic pattern goes against the spirit of existing analyses of NP₂ agreement because they crucially assume that the internal structure of NP₁ does not lend itself to being an accessible Agree goal. Making the copula potentially closer to NP₁ should not make a difference. We will adopt this overall conclusion for our classificatory clauses as well, and instead of assuming that NP₁ simply became a closer goal, we will consider a different structural solution.

We will put forward a proposal that strictly relies on the presence of a covert complex pronoun, and we will argue that what is special about the copula–NP₁–NP₂ orders is that their structure does not allow for this covert pronoun to be merged. Instead, the linear order is only compatible with a regular predication clause, and in turn only displays NP₁ agreement. When, however, the pronoun can be merged, we obtain a presentational structure in which

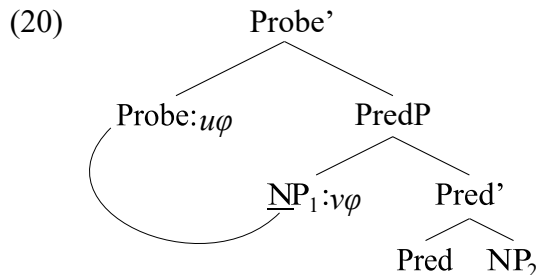
the pronoun inherits φ -features of NP₂, and the copula and the inter-sentential anaphor may indirectly agree with them, giving rise to what looks like NP₂ agreement.

3. The analysis

We argue that the variable agreement attested with classificatory clauses reflects a structural ambiguity. The linear string of these clauses is compatible with an analysis that corresponds to a predication clause, and an analysis that requires a new type of structure – a presentational clause where the agreement is mediated by a covert complex pronoun.

3.1. The predication structure

Let us first consider the predication structure option. With much of the literature on predication clauses, we assume that NP₁ and NP₂ form a small clause. When the φ -Probe merges in the structure, NP₁ is the closest accessible goal. In turn, NP₁ values the unvalued φ -features of the Probe (we will discuss the structural identity of the Probe later in this section). Finally, NP₁ raises to Spec,TP to satisfy the EPP. A simplified structure of the proposed derivation is given in (20). The tree only includes the relevant part of the derivation where we see why NP₁ is the source of agreement.



In terms of semantics, we assume a standard analysis according to which NP₁ is applied as a semantic argument of NP₂. However, our analysis is also easy to restate using silent semantic operators that combine the elements of the small clause, e.g., along the lines of Percus and Sharvit (2024).

The structure we assume has one individual-denoting NP – NP₁ – that can serve as an anaphoric antecedent. However, since an anaphoric pronoun construed as an elliptical definite description would have its presupposition entailed by the first sentence, anaphoric reference using the φ -features of NP₂ is not completely excluded (Elbourne, 2005).

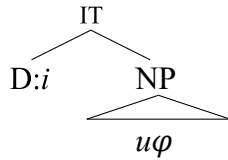
3.2. The presentational structure

Our analysis crucially relies on the intuition that when an anaphor displays φ -features of NP₂, it does not refer to the predicate denoted by NP₂. Instead, it refers to a new object formed by NP₂ classifying over NP₁. This intuition is supported by the data we repeat in (21). When the copula agrees with NP₂ and the anaphor shares features of NP₂, the predication quantifies over the classified content of NP₁, instead of NP₁ itself.

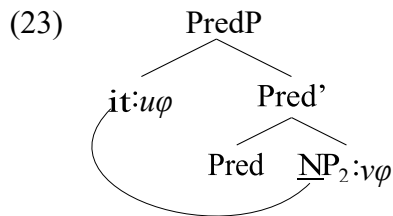
- (21) *Tahle knížka byl historický román.*
 this.F.SG book.F.SG was.M.SG historical novel.M.SG
 ‘This book was a historical novel.’
- a. *#Byl těžký.*
 was.M.SG heavy.M.SG
 Intended: ‘It was heavy.’ [=the book as a physical object, not the novel as its content]
- b. *Byl nudný.*
 was.M.SG boring.M.SG
 ‘It was boring.’ [= the content of the book being a novel]

The intuitive meaning we get is IT (=what this book is) is a historical novel. We model this intuition as a relationship between NP₁ and NP₂ mediated by an additional structural element, which we formalize as a complex covert pronoun – IT. For concreteness, we assume that the covert pronoun is composed of a nominal part (a set of unvalued φ -features – $u\varphi$) and a variable (i ; akin to Kratzer 2009’s minimal pronoun index).

- (22) *The structure of IT:*



This complex pronoun (IT) forms a small clause with NP₂. Once the small clause is formed, the unvalued φ -features of IT get valued by the features of NP₂ by concord (or possibly by Agree, if the pronoun is directly merged with NP₂, prior to the predicative functional structure being merged). The first part of the derivation is schematized in (23).



This initial derivational step captures two of our core empirical generalizations: First, in this structure, NP₂ fully retains its predicative properties, as it predicates over the covert pronoun IT. Second, the structure now contains an individual-denoting element that can become an antecedent of an anaphor, and the anaphor then may indirectly match the φ -features of NP₂.

At this derivational stage, however, the covert IT still contains an unbound variable. We argue that in the presentational structure of classificatory clauses the variable gets bound by NP₁, merged higher in the structure. The question is where.

3.2.1. NP₁ is merged in Spec,TopicP

Our proposal needs to capture two generalizations discussed in section 2. First, since the copula agrees with the features of NP₂, mediated by the covert pronoun IT, the copula – or, more precisely, the Probe that determines the features of the copula – needs to be structurally closer to IT than (the base-generated position of) NP₁. Second, since XP-copula-NP₁-NP₂ orders only allow NP₁ copular agreement (and disprefer NP₁ anaphoric agreement), this structure cannot contain the covert pronoun IT. The core data **are** repeated in (24).

- (24) *Podle mého názoru, byl/ *byla tenhle strom lípa.*
 according_to my opinion was.M.SG/ *was.F.SG this tree.M.SG linden_tree.F.SG
 ‘In my opinion, this tree was a linden tree.’
- a. *(On) byl vysoký.*
 he was.M.SG tall.M.SG
 ‘It was tall.’
- b. ?? *(Ona) byla vysoká.*
 she was.F.SG tall.F.SG
 ‘It was tall.’

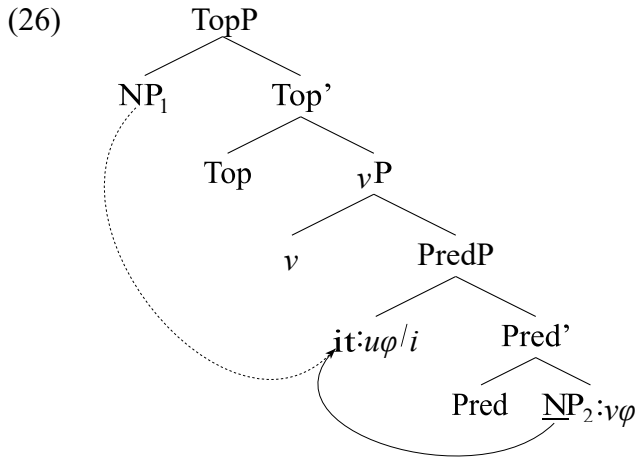
Critically, the restriction on NP₂ agreement arises only when NP₁ remains lower in the structure. As we can see in (25), when NP₁ follows the XP and precedes the copula, NP₂ agreement is possible.

- (25) *Podle mého názoru, tenhle strom byl/ byla lípa.*
 according_to my opinion this tree.M.SG was.M.SG/was.F.SG linden_tree.F.SG
 ‘In my opinion, this tree was a linden tree.’
- a. *(On) byl vysoký.*
 he was.M.SG tall.M.SG
 ‘It was tall.’
- b. *(Ona) byla vysoká.*
 she was.F.SG tall.F.SG
 ‘It was tall.’

The data strongly suggest that when NP₁ stays low, the covert pronoun IT cannot be in the structure. In other words, the pattern suggests that NP₁ is in a complementary distribution with the covert pronoun IT. In contrast, when NP₁ raises, it may raise from a position that is distinct from the base-generated position of IT.

To account for the contrast, we argue that in the NP₁ agreement configurations, NP₁ gets merged in the specifier of TopicP, an A-position located in the Middlefield (e.g., Frey 2000 for German). The covert pronoun IT itself merges in the specifier of the small clause.

The result is that the φ -features of IT get valued by NP₂, and its open variable gets bound by NP₁. The corresponding syntactic derivation is schematically given in (26).



3.2.2. v as the primary φ -Agree probe

So far, we have not addressed the question of the identity of the φ -Agree probe. If our reasoning about the derivational steps is correct, then T cannot be the relevant φ -Agree probe. Instead, we argue that the primary φ -Agree probe in Czech (but also in German we do not discuss in detail here) is v .

If T were the probe, then NP₁ would always be closer, and there would be no NP₂ agreement. This reasoning holds irrespective of whether we only assume downward Agree or whether we allow agreement to take place in specifier-head configurations (or some version of upward Agree). For T to be the probe and for the covert pronoun IT to be structurally closer to T, NP₁ would have to merge higher than T, for example, in the specifier of TP (assuming Agree is strictly downward). However, such a structure would incorrectly predict that NP₁ should never be in a complementary distribution with the covert pronoun IT. Therefore, v must be the primary probe (see e.g. Sigurðsson 2012, Schäfer 2012, Kučerová 2016 for other data patterns and analyses arguing for v being the primary probe in languages that, like Czech, display NP₂ agreement).

3.3. But why the covert pronoun when the predicative structure is possible?

At this point, we still face another question: Why is the covert pronoun IT needed and why can't we always merge NP₁ in the small clause (as it seems to be the case in languages like English)? We propose a semantic answer.

We assume that classificatory clauses in their presentational guise generally answer a question under discussion (in the sense of Roberts 2012 and subsequent literature) such as those in (27).

- (27) This is NP₂.
 a. What is this?
 b. Who is this?

The availability of these two types of questions under discussion (with 27b being restricted to animate referents) precisely mirrors the intuition that the intuitive reading of (27) is (28):

(28) What this is is NP₂.

Simplifying somewhat, we model a question such as (27a) as in (29).

(29) $\{\lambda w.P(w)(this)|P \in \langle s,et \rangle \wedge P \in CC\}$

Thereby, CC is understood as an extended conceptual cover (Aloni, 2001) defined as follows:

(30) CC is a conceptual cover exactly iff:
 $\forall w \in W : \forall d \in D : \exists ! P \in CC : P(w)(d)$

What (30) expresses is that a conceptual cover is essentially a partition of concepts such that for every individual in any world there is exactly one concept that applies to it. In practical terms, we assume that a conceptual cover always only applies to a subset of individuals and is contextually determined. This makes essentially all full taxonomies of the conceptually relevant domain of individuals a conceptual cover. By contrast, the original theory of Aloni (2001) had a stronger requirement that restricted members of conceptual covers to individual concepts.

We further assume a general semantic principle according to which there is a minimal syntactic or pragmatic domain in which the *wh*-word and the descriptions of other elements of that domain cannot belong to different CCs. Consider the example in (31).

(31) This tree is a linden tree.

Suppose we had no IT pronoun and NP₁ and NP₂ were interpreted in the same minimal domain as a presentational construction. The question under discussion would then be: *What is this tree?* And since this tree is already a concept denoting the relevant object, the only possible true answer would be (32).

(32) This tree is this tree.

We follow Gajewski (2002) in assuming that such L-analytical sentences are ungrammatical. Thereby, a sentence is L-analytical if its meaning is tautological or contradictory regardless of which non-logical lexical items are used in the sentence. Precisely this would be the case for (32). The IT pronoun, having no features of its own, circumvents this principle, as no features constrain the CC.

One crucial prediction is that whenever NP₁ is topical the issue does not emerge, as the reference of NP₁ is "handled" at a higher discourse level. This semantic assumption makes two immediate predictions. First, in presentational constructions, the NP₁ must always be topical and will have a trace with no interpreted features relevant to CCs. Second, if the NP₁ is base-generated in a topic position, it can bind an element that has no intrinsic φ -features that are relevant to CCs.

Another prediction is related to the availability of Topic positions and anaphoric agreement. If a language does not have a low argumental Topic position (in the Middlefield),

such a language cannot accommodate a copular clause with two NPs and a presentational pronoun such as the covert pronoun IT we propose. Consequently, such a language only allows NP₁ agreement. This prediction is borne out in languages like English (33).

(33) I am / *are you.

Since this type of language cannot accommodate both NP₁ and a covert pronoun like IT, presentational copular clauses in these languages are expected to only have a presentational pronoun as NP₁ (34).

(34) That/ It was Mary.

5. Conclusions

Our case study demonstrates that NP₂ agreement in copular clauses may be mediated by an additional structural element. If our proposal is on the right track, it opens a real possibility that the linguistic variation in the domain of NP₂ copular agreement reduces to the availability of the mediating element – in our case, the covert pronoun IT. The availability of such a pronoun correlates with an independently attested language property: the availability of a topic position in the Middlefield. No other mechanism is needed.

As for the proposed semantic interpretation, conceptual covers (CCs) are our way to spell out the intuition that these constructions classify the presented item.

References

- Aloni, Maria (2001). *Quantification under Conceptual Covers*. Doctoral Dissertation, University of Amsterdam, Amsterdam.
- Bartošová, Jitka (2017). *Topics in Copular Clauses*. Doctoral Dissertation, McMaster University, Hamilton, ON.
- Béjar, Susana, and Arsalan Kahnemuyipour (2017). Non-canonical agreement in copular clauses. *Journal of Linguistics* 53, 463–499.
- Béjar, Susana, and Arsalan Kahnemuyipour. 2018. Not all phi-features are created equal: A reply to Hartmann and Heycock 2018. *Journal of Linguistics* 54, 629–635.
- Coon, Jessica, and Stefan Keine. 2021. Feature gluttony. *Linguistic Inquiry* 52, 655–710.
- den Dikken, Marcel (2006). Specificational copular sentences and pseudoclefts. In Martin Everaert and Henk C. van Riemsdijk (eds.) *The Blackwell Companion to Syntax Volume IV*, 292–409. Malden, MA; Oxford: Blackwell.
- Elbourne, Paul (2005). *Situations and Individuals*. Cambridge, MA: MIT Press.
- Frey, Werner. 2000. Über die syntaktische Position der Satztopiks im Deutschen. In Kerstin Schwabe, Andre Meinunger and Dieter Gasde (eds.) *Issues in Topics* ZAS papers in linguistics* 20, 137–172. Berlin: ZAS.
- Gajewski, Jon (2002). L-analyticity in natural language. Unpublished manuscript: MIT.
- Hartmann, Jutta M. and Caroline Heycock (2017). Variation in copular agreement in Insular Scandinavian. In Höskuldur Thráinsson, Caroline Heycock, Hjalmar P. Petersen and Zakaris Svabo Hansen (eds.) *Syntactic Variation in Insular Scandinavian*, 233–275. Amsterdam: John Benjamins.

- Hartmann, Jutta M. and Caroline Heycock (2019). (Morpho)syntactic variation in agreement: Specificational copular clauses across Germanic. *Frontiers in Psychology* 10(2994).
- Hartmann, Jutta M. and Caroline Heycock (2025). Feature gluttony in copular clauses in German? New data from Mistaken Identity Sentences. A talk presented at Leipzig, FG Zyklische Optimierung.
- Heycock, Caroline (2012). Specification, equation, and agreement in copular sentences. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 57, 209–240.
- Higgins, R.F (1979). *The pseudo-cleft construction in English*. New York: Garland.
- Kučerová, Ivona (2016). Long-distance agreement in Icelandic: Locality restored. *The Journal of Comparative Germanic Linguistics* 19:49–74.
- Mikkelsen, Line (2005). *Copular Clauses: Specification, Predication and Equation*. John Benjamins Publishing.
- Moro, Andrea (1991). The raising of predicates: copula, expletives and existence. *MIT Working Papers in Linguistics* 15, 119–181. Cambridge, Mass.: MITWPL.
- Percus, Orin, and Yael Sharvit (2024). Copular asymmetries in belief reports. *Natural Language Semantics* 32, 403–430.
- Roberts, Craige (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5, 1–69.
- Romero, Maribel. 2005. Concealed questions and specificational subjects. *Linguistics and Philosophy* 28:687–737.
- Schäfer, Florian (2012). The passive of reflexive verbs and its implications for theories of binding and case. *Journal of Comparative Germanic Linguistics* 15, 213–268.
- Sigurðsson, Halldór Ármann (2012). Minimalist C/case. *Linguistic Inquiry* 43, 191–227.
- Uličný, Oldřich. 2000. *Instrumentál v struktuře české věty* [Instrumental in the Structure of a Czech Clause]. Prague: Nakladatelství Karolinum.