



Burukina, Irina & Borise, Lena & den Dikken, Marcel. 2024. A 'big DP' analysis of Russian copular constructions with *èto*. *Glossa: a journal of general linguistics* 9(1). pp. 1–37. DOI: <https://doi.org/10.16995/glossa.9440>

# Open Library of Humanities

## A 'big DP' analysis of Russian copular constructions with *èto*

**Irina Burukina**, Hungarian Research Centre for Linguistics & ELTE Eötvös Loránd University, Budapest, Hungary, [irina.burukina@btk.elte.hu](mailto:irina.burukina@btk.elte.hu)

**Lena Borise**, Laboratoire de Linguistique Formelle, CNRS & Université Paris Cité, Paris, France, [lena.borise@cnrs.fr](mailto:lena.borise@cnrs.fr)

**Marcel den Dikken**, Hungarian Research Centre for Linguistics & ELTE Eötvös Loránd University, Budapest, Hungary, [marcel.den.dikken@nytud.hu](mailto:marcel.den.dikken@nytud.hu)

This paper develops a novel analysis of *èto* copular constructions in Russian. First, we address 'NOM<sub>pre</sub> *èto* NOM<sub>post</sub>' constructions, where NOM is a nominative noun phrase. We propose that there are two predication relations established in such clauses: one on the clausal level, with NOM<sub>post</sub> being the subject of predication, and another one within the constituent headed by *èto*. The latter, we argue, constitutes a 'big DP': it contains NOM<sub>pre</sub> in the specifier position and *pro* in the complement position of D<sup>0</sup>; a predication relation is established between the two with the help of the predicativizer *èto*. The analysis is extended to account for '*èto* NOM<sub>post</sub>' constructions. The proposal goes against treating *èto* in '(NOM<sub>pre</sub>) *èto* NOM<sub>post</sub>' constructions as the subject of predication or as a dedicated functional head on the clausal spine; it also explicitly rejects treating NOM<sub>pre</sub> in all '(NOM<sub>pre</sub>) *èto* NOM<sub>post</sub>' constructions as a hanging topic. The proposed analysis captures all the relevant properties of *èto* constructions and makes important predictions about their distribution, the agreement pattern, and the properties of NOM<sub>post</sub>. The 'big DP' approach also allows us to establish a link between *èto* copular constructions and specificational pseudo-clefts.

*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by the Open Library of Humanities. © 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

**OPEN ACCESS**



## 1 Introduction

This paper develops a novel analysis of two types of Russian copular constructions featuring the element *èto* (roughly equivalent to English ‘this’ or ‘it’), illustrated in (1) and (2) below. While these constructions have been addressed in the literature before (Junghanns 1997; Geist & Błaszczak 2000; Markman 2008; Geist 2008, i.a.), some of their properties (e.g., information structure, agreement, and details of interpretation) have remained difficult to account for, and especially difficult to reconcile with each other within the bounds of a single analysis. Our approach incorporates some of the insights from the existing literature, especially Geist (2008). At the same time, we demonstrate that extant analyses cannot account for the full range of data and propose an alternative analysis that has broader empirical coverage.

### 1.1 The two *èto* constructions under study

Two main types of *èto* copular constructions exist in Russian. In the first one, the clause consists of a nominal phrase, followed by *èto* (identical in its form to the proximal neuter singular demonstrative ‘this’), a copula (in future and past tenses), and another nominal phrase, as shown in (1). We label these constructions ‘NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>’, where the subscripts ‘pre’ and ‘post’ indicate whether a nominal phrase precedes or follows *èto*.

- (1) [Petja]<sub>NOMpre</sub> èto byl [moj načal’nik]<sub>NOMpost</sub><sup>1</sup>  
 Petja this was.M.SG my boss.NOM  
 ‘Petja was my boss.’

In the second type of *èto* copular constructions, *èto* is in the clause-initial position, followed by (a copula and) a single nominal phrase, as shown in (2). Accordingly, we label these constructions ‘*Èto* (be) NOM<sub>post</sub>’.

- (2) Èto byl [moj načal’nik]<sub>NOMpost</sub>  
 this was.M.SG my boss.NOM  
 ‘It/this was my boss.’

### 1.2 A global glance at previous scholarship

In the existing literature, *èto* copular constructions have received various treatments. A major dividing line can be drawn between approaches that treat *èto* as a functional head on the clausal spine and analyses that consider it to be the predicate itself. The former typically treat *èto* as a topic head; see e.g., Junghanns (1997), Geist & Błaszczak (2000), Markman (2008).<sup>2</sup> These

<sup>1</sup> In the examples, we use the so-called Scholarly transliteration system for Russian (cf. e.g., Timberlake 2004).

<sup>2</sup> Alternative analyses of *èto* as a functional head include *èto* spelling out a Pred head on the clausal spine that connects the predicate and the subject (in the spirit of Bowers 1993) or *èto* being an overt realization of Tense/Agreement (cf. Citko 2006 on Polish). The main drawbacks of these approaches if applied to the Russian data are discussed by Markman (2008); we refer the reader to her work for more details.

accounts struggle to capture all the syntactic and semantic properties of the *èto* clauses, including the fact that *èto* only appears in identity clauses (as discussed in detail in section 2), the agreement pattern in *èto* copular constructions (as shown in (4) below), and the unavailability of multiple instances of *èto* in a single clause (especially problematic under the assumption that *èto* is merely an exponent of  $\text{Top}^0$ , since a single clause may include multiple topics).

Geist (2008) puts forward an alternative approach, whereby *èto* spells out the predicate. Because it better captures the peculiar behavior of *èto* copular clauses, we take it as a starting point of our discussion. While the analysis that we propose in this paper differs from Geist's in some important respects and does not build directly upon it, the two share the core idea that in copular clauses the *èto* constituent is a syntactic predicate at the clausal level. Thus, the present work may be considered a continuation of this line of research.

### 1.3 The proposal in a nutshell

The novelty of our proposal lies primarily in the argument that there are *two* predication relations established in *èto* constructions: one on the clausal level (much as in Geist's 2008 approach), and another one within the constituent headed by *èto* — which, we argue, constitutes a 'big DP' with a complement and a specifier, and a predication relation established between them with the help of the predicativizer *èto*. This approach will be shown to be restrictive enough to rule out the ungrammatical examples (for instance, examples where  $\text{NOM}_{\text{pre}}$  attempts to control agreement), but at the same time it leaves sufficient room to account for the seemingly exceptional examples of identity/identification constructions that have been overlooked by the existing literature.

### 1.4 The structure of the paper

The paper is structured as follows. Section 2 summarizes the typology of copular sentences and the main proposals regarding the syntax of *èto*-constructions, and identifies the contradictory aspects of the existing work that the current account seeks to address. In section 3, we turn to the ' $\text{NOM}_{\text{pre}}$  *èto* (be)  $\text{NOM}_{\text{post}}$ ' constructions and show that our proposal mobilizing a double predication structure (one within the 'big DP' and the other involving the 'big DP' as the predicate at the level of the clause) can account for them (3.1), and also spell out the workings of predication within the 'big DP' (3.2); the analysis is further extended to '*èto* (be)  $\text{NOM}_{\text{post}}$ ' constructions (3.3). Section 4 provides additional syntactic and prosodic evidence that is compatible with our analysis but would be hard to account for if  $\text{NOM}_{\text{pre}}$  was taken to be a (hanging) topic. In section 5, we discuss the wider implications of our approach. Section 6 concludes the paper.

## 2 Èto and the typology of copular clauses

Setting the stage for the upcoming discussion of the copular clauses with *èto* in Russian, in this section we introduce the commonly accepted classification of copular clauses proposed by Higgins (1973): predicational, specificational, identificational, and identity. We describe each type using examples from English and then proceed by considering data from Russian. Building upon Geist (2008), we first highlight the strengths and limitations of her proposal. After that we expand the dataset to include several more kinds of copular clauses in which *èto* can appear. To anticipate the conclusion, we demonstrate that the use of *èto* is restricted to identificational and identity clauses and is obligatory in those contexts.

### 2.1 Higgins' (1973) classification of copular clauses

Higgins (1973; 1979) has famously put forward a classification of copular clauses, distinguishing the following four types of constructions. Predicational copular clauses, such as *John is tall*, feature a predicational expression (syntactically, the predicate) assigning some property to a referential expression (syntactically, the subject). Next, specificational copular clauses ‘merely say what one is talking about: the subject in some way delimits a domain and the specificational predicate identifies a particular member of that domain’ (Higgins 1979: 213). As Mikkelsen (2005) clarified later, the pre-copular element of a specificational clause behaves as (containing) a variable and the post-copular element provides a value for that variable. Canonical examples of specification (not analyzable as any other type) are *The number of planets is nine* and *What I like is apples*. In these, one of the terms is superscriptional (the header of a list) and the other is specificational (specifying a value for the variable in the superscriptional term).<sup>3</sup> The next type, identificational copular clauses, according to Higgins, are typically used to teach or state the names of people or things, as in *That (man) is John*. They usually contain two nominal phrases: one of them is referential, and the other one (most naturally occurring in the post-copular position) is described by Higgins as identificational (but see below on dependents of this sort also being plausibly characterized as referential). Finally, identity copular clauses (called ‘equative’ in later work) constitute the fourth type. These establish identity between two referential expressions, as in *The morning star is the evening star*.

Higgins considers the possibility of grouping identity clauses together with identificational clauses. Specifically, he notes that referentiality may be understood as a property of the expression itself, and if so, it becomes possible to consider an identificational clause such as *That man over*

---

<sup>3</sup> Higgins does not describe the dependents in specificational clauses as either referential or predicational. Consider, for instance, the specificational clause *The number of planets is nine*. The noun phrase [*the number of planets*] here is the head of the list and [*nine*] specifies the contents of the list, but neither has a precise referent that can be determined without knowing whether the whole proposition is true or false. As a consequence, as Higgins shows, it is impossible to substitute either of these expressions with a synonym that would refer to the same object: *The number of planets is nine. Nine is her lucky number.* → #*The number of planets is her lucky number* (Higgins 1973:201).

*there is Joe Bloggs* an example of identity, too. Although Higgins himself does not fully commit to this idea, we will adopt it in this paper. (See also Mikkelsen 2005 on grouping together identity clauses and identificational clauses with two referential entity-type dependents.) As will be shown in section 2.3, identity clauses in Russian form a natural class with identificational clauses in terms of the distribution of *èto*; in what follows, we will use the term ‘identity’ to refer to both subtypes of constructions.

Higgins also emphasizes that there is no one-to-one correspondence between (in)definiteness and referentiality/predicativity (Higgins 1973: 246), in the context of a discussion of the kinds of grammatical expressions that can be used in various types of copular clauses – proper names, (in)definite NPs, APs, etc.<sup>4</sup> While typically referential expressions include deictic elements, proper names, and definite NPs, and predicational expressions include indefinite NPs, APs, or PPs, at least some indefinite NPs can be referential and at least some definite NPs can be predicative.

To appreciate this, consider first the sentence [*One of the boys*] *was tall*. The indefinite NP here (bracketed) has a specific referent and serves as the semantic and syntactic subject. On the other hand, (3a) shows that a definite NP (bracketed) can denote a property and serve as a predicate: here, John is attributed the property of being the president. Note that (3a) is ambiguous and can also receive an identity reading, if *the president* refers to a specific entity/individual; in this it contrasts with the predicative-only example in (3b). (Thus, only (3a) but not (3b) can be used in response to the question *Which one is the president?*, which requires identification; see also Higgins 1973, Doron 1988 for further discussion of minimal pairs of this sort).<sup>5</sup>

- (3) a. John is [the president].  
b. John is [president].

Importantly for our purposes, this means that the semantic type of a predicate, but not necessarily its grammatical definiteness, determines the type of the copular clause. This will be important for the discussion of *èto* as only appearing in identificational contexts in Russian.

## 2.2 Geist (2008) on *èto* copular clauses in Russian

One of the most syntactically well-developed analyses of *èto* copular clauses in Russian is Geist (2008). Because our approach shares some insights with this analysis, we summarize it below.

Geist (2008) considers clauses of the type ‘NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>’ and argues that NOM<sub>pre</sub> is a dislocated/external topic, while *èto* is an internal topic linked to it. The clause [*èto* (be)

---

<sup>4</sup> Higgins treats (in)definiteness as a grammatical phenomenon. In English, definite nominal phrases are those with a determiner, a demonstrative, or a pronominal possessor, and indefinite NPs are best exemplified by those accompanied by *some* or *any*.

<sup>5</sup> Note that in English, only occupational nouns that are inherently unique (e.g., *president*, *chair*, *Pope*) may occur without an article in the predicate position; we thank an anonymous reviewer for bringing up this observation.

NOM<sub>post</sub>] (excluding the external topic) is treated as predicational: *èto* serves as a predicate and takes NOM<sub>post</sub> as its argument.

A beneficial implication of Geist's proposal (though not one discussed by Geist herself: she only considers *èto* constructions in which the two nominal phrases have matching features) is that NOM<sub>pre</sub> never controls agreement with the copula, and the copula obligatorily displays the phi-features of NOM<sub>post</sub>, as shown by our examples in (4). This agreement pattern follows from the assumption that NOM<sub>post</sub> is underlyingly the subject of predication while NOM<sub>pre</sub> is not in a direct relation with the matrix T<sup>0</sup> head (because, e.g., in Geist's analysis, it is an external topic).

- (4) a. Džordž Eliot *èto* byla/\*byl/\*bylo Meri Enn Evans.  
George Eliot this was.F/was.M/was.N Mary Ann Evans  
'George Eliot was Mary Ann Evans.'
- b. Moi kazni egipetskie *èto* byl/\*byli/\*bylo Petja.  
my plagues Egyptian this was.M/were/was.N Petja.M.NOM  
'My plagues of Egypt (i.e. my bane) were Petja.'
- c. Ty i ja *èto* budet/\*budem lučšaja komanda.  
you and I this be.NPST.3SG/be.NPST.1PL best team  
'You and I will be the best team.'

Geist treats *èto* as a demonstrative pronoun. To make it possible for it to be a predicate, she proposes that the *èto*-constituent has a special type of D<sup>0</sup> head that takes *èto* as its complement and turns it into a predicate of the type  $\langle e, t \rangle$ .<sup>6</sup> The *èto*-predicate establishes an identity between two entity-type items: NOM<sub>post</sub> (via predication) and NOM<sub>pre</sub> (via co-indexation). This is schematized in (5).

- (5) [<sub>S</sub> [<sub>DP</sub> NOM<sub>pre</sub>] – [<sub>IP</sub> *èto* COP NOM<sub>post</sub>]]

The key strength of Geist's (2008) proposal is that it allows for bringing together three semantic types of copular constructions that contain two nominal dependents: (i) predicational ones (cf. in English *John is a teacher (by trade)*), (ii) equative/identity ones (*Mark Twain is Samuel Clemens*), and (iii) specificational ones (*The murderer is John*); note that (ii) is the only type that contains and requires *èto* in Russian. The predicational constructions are taken to be the baseline, while in equative and specificational clauses, different types of type-shifting are applied, in order to combine the two non-property items and derive the corresponding semantics of the clause.

Despite its insightful nature, Geist's analysis leaves aside some important data, and some of the assumptions that it rests upon are not tenable if the full range of data is to be accounted for.

---

<sup>6</sup> Specifically, Geist suggests that the predicativizing D<sup>0</sup> head contains a type-shifting *ident*-operator, as proposed by Partee (1987). In the present paper we follow the same predicativization approach in its general form; however, since our primary focus is on the syntax of the *èto* copular constructions, we remain agnostic as to how exactly the predicativizing effect is achieved in semantics. In principle, our analysis can easily accommodate the *ident*-operator.

First, it is unlikely that  $NOM_{pre}$  always serves as a dislocated/hanging topic, in view of the fact (corroborated in section 4) that  $NOM_{pre}$  has a broader distribution than what is predicted for hanging topics.<sup>7</sup>

Second, as far as *èto* constructions are concerned, Geist (2008) only considers equative/identity clauses that contain two proper names, of the type *Mark Twain is Samuel Clemens*. However, *èto* can also appear in copular clauses with a proper name and a common noun, or with two common nouns, as in *That girl on the left is Maša/my sister*, and it is not immediately apparent how to treat such examples in her analysis. Geist discusses some similar examples (e.g., *Ubijtsa staruxi – Raskol’nikov* ‘The murderer of the old lady is Raskol’nikov’) but claims that they cannot contain *èto* and analyzes them unambiguously as specificational clauses. As we demonstrate in section 2.3, the data are more complex, and *èto* is not ruled out completely in such cases. This prompts a careful re-examination of the distribution of *èto*, which gives rise to the following questions: Is *èto* genuinely restricted to equative/identity constructions? Can it appear in specificational clauses too? Can it ever co-occur with a predicative, i.e. property-denoting (as opposed to referential) item?

To answer these questions, we expand the dataset by adopting the traditional classification of copular clauses developed by Higgins (1973, 1979), introduced in section 2.1, which is broader than that adopted in Geist (2008) and includes identity clauses with all kinds of referential dependents (e.g., *That man is John*) in addition to the equative clauses with two proper names. As mentioned in section 2.1, identity constructions include those with two proper names but are not limited to them, and may involve referential expressions of various types: pronouns, definite nominal phrases (those with a demonstrative pronoun or a possessor), and even indefinite nominal phrases (e.g., those modified by *kakoj-nibud*, *kakoj-to* ‘some’, or *odin iz* ‘one of’). This becomes crucial: as we demonstrate in section 2.3, *èto* is only allowed in the case of identity in a broad sense. We further show that specificational clauses with *èto* are indeed prohibited; however, if presented out of the blue, the same expression can be ambiguous between a referential reading and a property reading (see *the president* examples in (3) above), which obscures the overall picture. To distinguish the possible identity, specificational, and predicational readings, we accompany every key example with a specific context. To bolster our argument that *èto* is only allowed in the case of identity, we also include examples (not previously discussed in the literature on copular clauses in Russian, to our knowledge) that are unambiguously specificational, based on Higgins (1973).

Finally, we bring to light another type of a ‘ $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ’ construction: one in which a predicative adjective is used instead of  $NOM_{post}$ , as illustrated in (6).

---

<sup>7</sup> Although we do not wish to categorically exclude the possibility of a hanging topic analysis for some sentences of this shape, our point is that such an approach is not feasible for the full range of cases. See section 4.

- (6) Šakira èto (bylo) modno.  
 Shakira.F this was.N.SG fashionable.N  
 ‘Shakira was fashionable [e.g., to listen to].’

This construction falls outside the scope of Geist’s analysis because, in her analysis,  $NOM_{post}$  must be a referential nominal expression, given that it is an argument of the equating predicate [*èto*]. Because (6) is a copular *èto*-clause that is otherwise very similar to ‘ $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ’ constructions, it would be advantageous to accommodate examples of this kind as well. As we discuss in section 5.1, our analysis is well equipped to do so.

Despite the crucial differences, our analysis is similar to Geist’s in two respects. First, on our approach as on Geist’s, all ‘ $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ’ constructions have the same syntax, with the constituent headed by [*èto*] being a syntactic predicate at the clausal level. Second, our intuition also is that building an *èto* copular clause involves predicativization – a special operation that turns a referential expression of type  $\langle e \rangle$  into a predicate of the type  $\langle e, t \rangle$ . In contrast with Geist, however, we argue that the [*èto*] constituent is a ‘big DP’, in which *èto* is an exponent of the  $D^0$  head. The complement of  $D^0$  is a silent *pro* whose content is identified by  $NOM_{post}$ , and the specifier position is occupied by  $NOM_{pre}$  – crucially, not an external topic. In the *èto*-clause, a syntactic predication relation is established on two levels: within the ‘big DP’ – between  $NOM_{pre}$  and *pro*, and on the clausal level – between the ‘big DP’ and  $NOM_{post}$ . Such a double-predication approach captures all the relevant properties of the expanded set of *èto* constructions under consideration in this paper and makes important predictions about their distribution.

The next subsection is dedicated to clarifying the data and determining the exact range of copular constructions that can include *èto*. Specifically, we show that, in Russian, the distribution of *èto* is restricted to identity constructions and does not cover specificational and truly predicational clauses.

## 2.3 Fine-tuning the empirical coverage of the analysis: *Èto* and identity

Having introduced Higgins’ classification of copular clauses and shown that it has not yet been comprehensively discussed in the context of Russian data, we provide a detailed discussion of the different types of copular clauses in Russian below. One of our main conclusions is that the distribution of *èto* is restricted to identity contexts (including both identificational and strictly equative clauses).

### 2.3.1 *Èto* and identity/identificational clauses

Identity copular clauses typically answer questions like *Which (one) is John? Who is that person over there?* A natural context for them is observing a group of people/objects (for instance, at a



social gathering or in a photo) and recognizing/identifying them. Another scenario that requires identity clauses is one in which the speaker wants to identify a person/object they are not familiar with with someone or something they know, as in *John? Who is that?*

In (7–10), we provide examples of identity clauses in Russian in several typical contexts, for possible combinations of two referential expressions (deictic, proper names, and (in)definite NPs). They all contain *èto*, which links the two nominal phrases.

*Context 1: We are looking at a group photo and pointing at certain people to identify them.*

- (7) NOM<sub>pre</sub> = definite, NOM<sub>post</sub> = deictic/proper name/definite NP/indefinite NP  
 Ètot paren' sleva èto ja / Petja / moj otec / odin znakomyj.  
 this guy on.the.left this I Petja my father one acquaintance  
 'This guy on the left is me/Petja/my father/some acquaintance.'
- (8) NOM<sub>pre</sub> = indefinite, NOM<sub>post</sub> = deictic/proper name/definite NP/indefinite NP  
 Odin iz parnej sleva èto ja / Petja / moj otec / odin znakomyj.  
 one from guys on.the.left this I Petja my father one acquaintance  
 'One of the guys on the left is me/Petja/my father/some acquaintance.'

*Context 2: We are looking at a group photo and trying to identify a person that we have never met.*

- (9) NOM<sub>pre</sub> = definite, NOM<sub>post</sub> = deictic/proper name/definite NP  
*Which one is your mother?*  
 Moja mama èto ona / ta ženščina sleva / Marina Nikolaevna.  
 my mother this she that.F woman on.the.left M. N.  
 'My mother is her/that woman on the left/Marina Nikolaevna.' (in the latter case, the listener is expected to be able to recognize Marina Nikolaevna)

*Context 3: Somebody mentions a person whose name we do not recognize.*

- (10) NOM<sub>pre</sub> = proper name, NOM<sub>post</sub> = deictic/proper name/definite NP/indefinite NP  
*Peter Parker? Who is he?*  
 P.P. èto on / čelovek pauk / moj otec / odin paren' iz moej školy / kakoj-to  
 P.P. this he Spider Man my father one guy from my school some.SPEC  
 paren' iz moej školy.  
 guy from my school  
 'Peter Parker is him (deictic)/Spider Man/my father/a guy from my school/some guy from my school.'

As mentioned in section 2.1, following Higgins (1973) and Mikkelsen (2005), we take identity clauses to include equative clauses that are used to answer such questions as *Who is who?* (for instance, when playing a game in which players match identical characters). Similarly to the other cases of identity, these copular clauses are used to point out that the two expressions share the same referent.

In Russian, all such clauses feature *èto*, similarly to other identity constructions; this is shown in (11). Note that sentences of this kind are often ambiguous between a strict equative reading and an identification reading: for instance, (11a) can also be used as a response to the question *Which one of them is Mark Twain?* if the listener is expected to be able to recognize Samuel Clemens.

- (11) a. Mark Tven èto Samuel Klemens.  
 Mark Twain this Samuel Clemens  
 ‘Mark Twain is Samuel Clemens.’
- b. *Do not pretend to be my boss. My position is (the same as) your position.*  
 Moja dolžnost’ èto tvoja dolžnost’.  
 my work.position this your work.position  
 ‘My position is your position.’
- c. Sobaki èto sobaki.  
 dogs this dogs  
 ‘Dogs are dogs.’

### 2.3.2 Èto and predicational clauses

Predicational clauses are used to answer questions like *What is John (by trade)? What do you know about John?* These are easily distinguished from the other copular clause types as they do not allow inversion and can be embedded under *consider* without a copula, as in (12b) and (12c), respectively – to be compared with the identity examples in (12d,e).

- (12) a. Peter Parker is a talented student/Spider Man.  
 b. \*A talented student is Peter Parker. predicational  
 c. I consider Peter Parker (to be) a talented student. predicational  
 d. Spider Man is Peter Parker. identity  
 e. I consider Peter Parker \*(to be) Spider Man. identity

The Russian examples in (13) instantiate predicational clauses – as is clear from the fact that inversion in them is disallowed, as shown in (14). Adding *èto* to the sentences in (13) renders them ungrammatical. This demonstrates that in Russian, *èto* is generally prohibited in predicational clauses, for all possible types of subjects and predicates.<sup>8</sup>

- (13) *What do you know about him/Petja/your father/one of the guys?*  
 a. {On / Petja / moj otec / odin iz parnej} (\*èto) vysokij.  
 he Petja my father one from guys this tall  
 ‘He/Petja/my father/one of the guys is tall.’

---

<sup>8</sup> In section 5.1, we discuss some apparent counterexamples, where *èto* co-occurs with a predicational NOM<sub>post</sub>. We suggest that, despite their surface appearance, these clauses should be grouped with identity rather than predicational ones.

- b. {On / Petja} byl {tvoj brat / student}.  
 he Petja was your brother student  
 'He/Petja was your brother/student.'

(14) *What do you know about him/Petja/your father/one of the guys?*

- a. \*Vysokij {on / Petja / moj otec / odin iz parnej}.  
 tall he Petja my father one from guys
- b. \*{Tvoj brat / student} byl on.<sup>9</sup>  
 your brother student was he

Furthermore, adding *èto* to a copular clause with two nominal expressions shifts the interpretation from predicational to identity, as illustrated in (15) and (16). Without *èto*, 'lazybones' in (15) is interpreted as a property of being lazy that is assigned to the subject; thus, (15) is felicitous in response to the question *What do you know about him? Tell me more about what he is like*. With *èto*, 'lazybones' in (16) is interpreted as a referential expression, either the lazybones that we have been talking about or a certain lazybones character (e.g., in a play or in the context of a game in which one member of the group needs to be designated as a lazybones). Consequently, (16) receives an identity reading and can be used to answer the questions *Who/which one is he?* This contrast is similar to the one introduced for English in (3), expressed via the optionality or obligatoriness of the definite determiner.

(15) *What do you know about him/Petja/your father/one of the guys?*

- {On / Petja / moj otec / odin iz parnej} lentjaj.  
 he Petja my father one from guys lazybones  
 'He/Petja/my father/one of the guys is a lazybones.'

(16) *Who is he/Petja/your father/one of the guys?*

- On / Petja / moj otec / odin iz parnej èto (nastoljaščij) lentjaj.  
 he Petja my father one from guys this real lazybones  
 'He/Petja/my father/one of the guys is a true lazybones/the lazybones.'

### 2.3.3 Èto and specification

Finally, the question of whether *èto* is allowed in specificational copular clauses needs to be addressed. None of the unambiguously specificational examples pointed out by Higgins (1973) allow *èto*, as shown in (17):

---

<sup>9</sup> Note that (14b) becomes acceptable under an identity reading, as a response to *Which one was my brother/the student?*, if *èto* is added.

- (17) a. Číslo planet (\*èto) devjat'.  
 number planets this nine  
 'The number of planets is nine.'
- b. Ego rost (\*èto) metr devjanosto.  
 his height this meter ninety  
 'His height is 1m 90cm.'

At the same time, our observation is that many specificational clauses, if supplemented by *èto*, rather than becoming ungrammatical, turn into felicitous identity clauses. This is consistent with the idea, put forward by Higgins (1973), that definite and indefinite NPs can be either superscriptional or referential, and, likewise, proper names can be either specificational or referential. Providing an unambiguous context for example sentences, however, allows for identifying the two readings and matching them with the presence or absence of *èto*.

As an illustration, consider the examples in (18a) and (18b). While on the surface they look almost identical (the only difference being the presence/absence of *èto*), the contexts in which they can be used are strikingly different and, crucially, not interchangeable. Specifically, the identity statement in (18a) can be uttered in the following scenario: the speaker introduced their partner to several people at a party; after the party, their partner says: "You had told me that two of your best friends were going to be at this party. Which of the people we met were they?". In contrast, the specificational (18b) is a felicitous response to the question "Who were your best friends when you were a child?".

- (18) a. Moi lučšie družja **èto** byli Maša i Petja.  
 my best friends.NOM this were Maša and Petja  
 'My best friends were Maša and Petja.'
- b. Moi lučšie družja byli Maša i Petja.  
 my best friends.NOM were Maša and Petja  
 'My best friends were Maša and Petja.'

Geist (2008) also observes that specificational clauses of the type 'NOM<sub>pre</sub> (be) NOM<sub>post</sub>' do not allow for *èto*. To illustrate this claim, she provides the following examples (19; the grammaticality judgments are Geist's).

- (19) a. Ubijca staruxi (\*èto) Raskol'nikov.  
 murderer old.lady.GEN this Raskolnikov  
 'The murderer of the old lady is Raskolnikov.'
- b. Edinstvennyj, kto stal na našu storonu, (\*èto) byla Varvara.  
 only.person who came on our side this was.F Barbara  
 'The only person who defended us was Barbara.'

Interestingly, 7 out of 7 native speakers of Russian that we consulted in an informal survey did not agree with these judgements, and instead thought that the sentences in (19) were equally acceptable with and without *èto*. This is in line with our observation that adding *èto* to a specificational clause may turn it into a well-formed identity copular sentence.

To summarize, *èto* is allowed only in identity clauses. It is ruled out in predicational clauses and unambiguously specificational examples. Apparent cases of *èto* in specificational clauses should be re-classified as identity clauses, as they are felicitous only in identity-establishing contexts (e.g., in response to *which one* questions). In identity clauses, *èto* links two referential nominal expressions, either of which can be deictic, a proper name, a definite NP, or an indefinite NP. Our analysis, presented in the next section, provides an explanation for this restriction on the distribution of *èto*: as a predicativizer, *èto* is required in identity clauses (which contain two referential expressions, neither of which can serve as a syntactic predicate) and is redundant in other types of copular constructions – e.g., in predicational clauses.

### 3 The proposal

Our proposal in a nutshell is that there are two predication relations established in ‘ $\text{NOM}_{\text{pre}} \textit{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ constructions. One is on the clausal level, with the [*èto*] constituent being predicated of  $\text{NOM}_{\text{post}}$ . The other one is established within the [*èto*] constituent, which we argue is a ‘big DP’: its complement position is occupied by *pro*, the specifier position is occupied by  $\text{NOM}_{\text{pre}}$ ; a predication relation between the two DP-internal terms is established with the help of a special functional item, the predicativizer, spelled out as *èto*.

In this section, we introduce the structure for ‘ $\text{NOM}_{\text{pre}} \textit{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ clauses that we propose, discussing its components and the properties of the construction that the analysis successfully captures:  $\text{NOM}_{\text{post}}$  controlling agreement with the copula, and the behavior of ‘ $\text{NOM}_{\text{pre}} \textit{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ construction in questions; we also extend the account to ‘*èto* (be)  $\text{NOM}$ ’ clauses. After that, in section 4 we offer additional evidence against treating all instances of  $\text{NOM}_{\text{pre}}$  as hanging topics. Then, in section 5, we discuss seemingly exceptional examples, where the  $\text{NOM}_{\text{post}}$  position is occupied by a property-denoting expression in the presence of *èto*, and show that they are accounted for by our analysis but not by other existing proposals.

#### 3.1 The clausal level: [*èto*] as the predicate

Descriptively, ‘ $\text{NOM}_{\text{pre}} \textit{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ clauses consist of the invariant element *èto*, which is preceded by one nominative nominal phrase and followed by another one.<sup>10</sup> The copula *byt’* is

---

<sup>10</sup> We know of at least two types of sentences that are similar to the *èto* copular construction discussed in this paper but differ from it in their properties, hence they are not subject to this generalization. In the first type, *èto* is interpreted as a demonstrative pointing to a non-human entity, and  $\text{NOM}_{\text{post}}$  denotes a property ascribed to it. This is illustrated

inserted between *èto* and  $\text{NOM}_{\text{post}}$  in the past and future tenses and is absent in the present tense. This is illustrated in (20).

- (20) a. (*After being introduced to a group of people:*) *Who of them was your boss?*  
 Moj načal'nik èto byl Petja.  
 my boss.NOM this was.M.SG Petja  
 'Moj boss was Petja.'
- b. *Petja? Who/which one was that?*  
 Petja èto byl moj načal'nik.  
 Petja this was.M.SG my boss.NOM  
 'Petja was my boss.'

In line with Geist (2008), we argue that in constructions like (20),  $\text{NOM}_{\text{post}}$  is the underlying subject of predication on the clausal level, while the constituent that includes *èto* is the predicate. The two are combined within a small clause.<sup>11</sup> The constituent that contains *èto* then undergoes movement to Spec, TP — an instance of predicate inversion, described in detail by Moro (1997) and Den Dikken (2006). This is illustrated in (21) with a bracketed example; we will address the status of  $\text{NOM}_{\text{pre}}$  below. The copula (overt or covert) is found in  $T^0$ .

- (21) [<sub>TP</sub> [<sub>DP</sub>  $\text{NOM}_{\text{pre}}$  *èto*]<sub>i</sub> [<sub>T</sub> [<sub>T<sup>0</sup></sub> COP/Ø] [<sub>SC</sub> [<sub>DP</sub>  $\text{NOM}_{\text{post}}$ ] *t<sub>i</sub>* ]]]

Predicate inversion is obligatory: *èto* can never follow the copula, which would correspond to the structure in (21) but without predicate movement. This is illustrated in (22), with (22a) presenting an example where both the subject and the predicate remain *in situ*, and (22b) showing one where  $\text{NOM}_{\text{post}}$  is promoted to Spec,TP instead.

---

in (i), where  $\text{NOM}_{\text{post}}$  carries instrumental case, which is typical of predicate nominals, and the copula obligatorily surfaces in the neuter singular form.

- (i) Èto bylo / \*byl moim košmarom.  
 this.N.SG was.N.SG was.M.SG my nightmare.M.SG.INS  
 'This [thing] was my nightmare.'

In the second type, *èto* is anaphoric to a previously mentioned proposition rather than to an entity. This is shown in (ii), where the postnominal constituent also can be marked with instrumental case.

- (ii) Petja skazal, [èto Maša prinesla knigi]. 'Petja said that Maša brought the books.'  
 Èto<sub>i</sub> bylo važno / važnym / pravdoj.  
 this.N.SG was.N.SG important.N.SG important.SG.INStruth.INS  
 'This (that Maša brought books) was important/true.'

<sup>11</sup> Because nothing in this paper depends on this, we remain agnostic about the internal structure of the small clause. The structural representation in (21), utilizing the expository label 'SC', is meant to be a shorthand for whatever the optimal syntax of small clauses should turn out to be (with candidates including structures mobilizing a functional head as well as the XP-YP analysis of Moro 1997). A structural assimilation of the syntax of predication in the *èto*-DP and in the small clause is technically feasible but not pursued here for simplicity's sake.

- (22) a. \*[Byl [<sub>SC</sub> [<sub>DP</sub> moj načal'nik] [<sub>DP</sub> Petja èto]]].  
 was.M.SG my boss.NOM Petja this  
 (Intended: 'Petja was my boss.')
- b. \*[[<sub>DP</sub> Moj načal'nik]<sub>i</sub> byl [<sub>SC</sub> t<sub>i</sub> [<sub>DP</sub> Petja èto]]].  
 my boss.NOM was.M.SG Petja this  
 (Intended: 'Petja was my boss.')

The obligatoriness of predicate inversion is inextricably linked to the information-structural properties of *èto* constructions. The [<sub>DP</sub> *èto*] constituent represents given information; this is in accordance with the obligatory identity nature of 'NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>' clauses, in which NOM<sub>pre</sub> is identified with NOM<sub>post</sub>: to identify/equate NOM<sub>pre</sub> with NOM<sub>post</sub>, one needs to be familiar with NOM<sub>pre</sub>. In Russian, constituents providing given information, as represented by [+presupposed] and [+referential] features, have been shown to precede others, e.g., in scrambling (Titov 2017). Given that, at the same time, Spec,TP is the natural host for presupposed/referential material in Russian (Bailyn 2004; Titov 2018), we suggest that a similar kind of movement, correlated with the given IS-status of the [<sub>DP</sub> *èto*] constituent, obligatorily applies to this constituent in 'NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>' clauses.<sup>12</sup>

Treating NOM<sub>post</sub> as the notional subject of the clause allows us to derive the central property of identificational clauses that we have been discussing so far (instead of just postulating it): the fact that NOM<sub>post</sub> must be referential, as discussed in detail in section 2.3 in the context of identity copular clauses. As Russian is an article-less language, the distinction between the two kinds of nominal phrases – referential vs. property-denoting/non-referential – is not expressed morphosyntactically, but the non-referential reading can be enforced by generalizing expressions corresponding to English 'by trade' and 'by nature'. Examples of this sort are ungrammatical in the presence of *èto*, as shown in (23) (also demonstrating, once again, that non-identity copular clauses do not allow for *èto*).

- (23) a. \*Petja èto (po professii) santexnik.  
 Petja this by job plumber.NOM  
 ('Petja is a plumber by trade.')
- b. \*Petja èto (po skladu xaraktera) xolerik.  
 Petja this by temper choleric.NOM  
 ('Petja is a choleric by nature.')

---

<sup>12</sup> In connection with the obligatoriness of predicate inversion in *èto* constructions, it may also be helpful to draw attention to the fact that English clefts require *it* to be in the precopular structural subject position (even though arguably *it* originates as the predicate of the focused XP; see Den Dikken 2013, inspired by Moro 1997): in English, while *It is Donald who committed the crime* and *I consider it \*(to be) Donald who committed the crime* are grammatical (the obligatoriness of the copula in the latter showing that predicate inversion is involved), *\*Donald is it who committed the crime* and *\*I consider Donald (to be) it who committed the crime* are not.

In (24), another test for the referential nature of  $NOM_{post}$  is provided. The base sentence is the identity example ‘Petja was your boss’, (24a). A corresponding sentence with the topic marker *čto kasaetsja* ‘as for’, as in (24b), can be constructed and is accepted by native speakers. As argued by Reinhart (1981), only referential nominal expressions can be interpreted as topics; therefore, the acceptability of (24b) suggests that ‘your boss’ cannot be a property, must be interpreted referentially, and, accordingly, should not be analyzed as the predicate. Note that *Petja* in (24b) represents given material, since it picks out one referent from a group of known referents (several people that we have just met); this, according to our proposal, is correlated with *Petja* undergoing inversion and acting as  $NOM_{pre}$ .

- (24) a. *Petja èto byl tvoj načal’nik.*  
 Petja this was your boss.NOM  
 ‘Petja was your boss.’  
 (intended reading: *Oh, so Petja, whom we have just met, was your boss!*)
- b. *We have just met several people. Guess who was my boss and who was my secretary.*  
*Čto kasaetsja tvoego načal’nika, ja dumaju, čto Petja èto byl on.*  
 what concerns your boss I think that Petja this was him  
 ‘As for your boss, I think that Petja was him.’

The fact that  $NOM_{post}$  in ‘ $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ’ constructions is interpreted referentially and not predicatively can also be brought to light (following Heycock & Kroch’s (1999) lead) by adding a non-restrictive relative clause to modify  $NOM_{post}$ . In (25a), *moj drug* ‘my friend’ is potentially ambiguous between being a referential/definite expression (‘this friend of mine’) or denoting a property (‘a friend of mine’). The non-restrictive relative is grammatical only in the identity context (i.e., when *èto* is present), which confirms the referential reading of  $NOM_{post}$ .

- (25) a. *Petja (èto) moj drug.*  
 Petja this my friend  
 with *èto*: ‘Petja is the friend of mine.’  
 without *èto*: ‘Petja is a friend of mine.’
- b. *Petja \*(èto) moj drug, ktorogo ty xorošo znaješ.*  
 Petja this my friend which you well know  
 ‘Petja is the friend of mine, who(m) you know well.’

To recapitulate, using the three diagnostics above – property-denoting phrases like *po professii* ‘by trade’ or *po xarakteru* ‘by temper’ in (23), a topic marker *čto kasaetsja* ‘as for’ in (24b), and a non-restrictive relative clause in (25b) – we can establish that  $NOM_{post}$  must have a referential and not a property-denoting/predicative reading in ‘ $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ’ constructions. The syntactic subjecthood of  $NOM_{post}$  fits well with this.



### 3.2 The ‘big DP’: [NOM<sub>pre</sub> èto] as a constituent

So far, we have argued that in copular clauses of the type ‘NOM<sub>pre</sub> èto (be) NOM<sub>post</sub>’, NOM<sub>post</sub> is the subject, while the constituent that includes èto is the syntactic predicate.<sup>13</sup> Let us now turn our attention to NOM<sub>pre</sub>.

In section 2.3, we discussed that, in identity clauses, both nominal expressions are referential. For NOM<sub>pre</sub>, this can be verified with the help of the *čto kasajetsja* ‘as for’ test, already mobilized in section 3.1:

- (26) a. *Guess which one is my boss and which one is my secretary.*  
 Čto kasaetsja tvoego načal'nika, ja dumaju, čto on èto Petja.  
 what concerns your boss I think that he this Petja  
 ‘As for your boss, I think that he is Petja.’
- b. *Guess which one is Petja and which one is Oleg.*  
 Čto kasaetsja Peti, ja dumaju, čto on èto tvoj načal'nik.  
 what concerns Petja I think that he this your boss  
 ‘As for Petja, I think that he is your boss.’

But despite being referential, NOM<sub>pre</sub> is not a viable candidate for subjecthood. NOM<sub>pre</sub> never controls agreement with the copula in T<sup>0</sup>, as we showed in section 2.2. Two examples showing this are reproduced here in (27) and (28). In both examples, NOM<sub>pre</sub> is plural and NOM<sub>post</sub> is singular, and the copula obligatorily displays the phi-features of NOM<sub>post</sub>.

- (27) a. *Moi kazni egipetskie èto byl/\*byli/\*bylo Petja.*  
 my plagues Egyptian this was.M/were/was.N Petja.M.NOM  
 ‘My plagues of Egypt (i.e. my bane), that was Petja.’
- b. [TP [DP [Moi kazni egipetskie]<sub>D</sub> [D<sup>0</sup> èto] [*pro*]<sub>i</sub>]<sub>k</sub> [T<sub>r</sub> byl [SC [Petja]<sub>i</sub> t<sub>k</sub>]]]
- (28) a. *Ty i ja èto budet/\*budem lučšaja komanda.*  
 you and I this be.NPST.3SG/be.NPST.1PL best team  
 ‘You and I will be the best team.’
- b. [TP [DP [Ty i ja]<sub>D</sub> [D<sup>0</sup> èto] [*pro*]]<sub>k</sub> [T<sub>r</sub> budet [SC [lučšaja komanda] t<sub>k</sub>]]]

Though NOM<sub>pre</sub> is not itself the structural subject of the clause, we propose that it combines with èto in a single constituent of the following shape: [DP [DP NOM<sub>pre</sub>] [D<sup>0</sup> èto [*pro*]]]. To refer to these complex DPs, we use the term ‘big DP’, originally used in Romance linguistics to refer to the constituent that encompasses a nominal expression and a coreferential clitic (Uriagereka 1995;

<sup>13</sup> On neuter pronominal elements being able to serve as predicates of copular sentences (incl. English *it* in (truncated) clefts such as *It is {his smoking/that he smokes} (that bothers me)*), see especially Moro (1997).

2005; Kayne 2005, i.a.).<sup>14</sup> Accordingly, what precedes the copula in the examples in (20) has the structure shown in (29).

- (29) a. [<sub>DP</sub> [Moj načal'nik] [èto [*pro*]]] byl Petja.  
 my boss.NOM this was.M.SG Petja  
 'My boss was Petja.'
- b. [<sub>DP</sub> [Petja] [èto [*pro*]]] byl moj načal'nik.  
 Petja this was.M.SG my boss.NOM  
 'Petja was my boss.'

Within the 'big DP', the specifier position is occupied by NOM<sub>pre</sub> and the complement position is occupied by *pro*,<sup>15</sup> with a predication relation established between the two. As both NOM<sub>pre</sub> and

---

<sup>14</sup> As proposed by Uriagereka (1995; 2005), Torrego (1996), i.a., 'big DPs' in Romance languages are formed by a referential nominal phrase or a pronoun in Spec,DP and a clitic in D<sup>0</sup>. This is shown in (i) (for 'weak' clitics): before the clitic moving out of the 'big DP' in (a) and after the movement in (b).

- (i) a. [<sub>DP</sub> [<sub>D'</sub> clitic [<sub>NP</sub> *pro*]]]  
 b. Lo<sub>i</sub> vi [<sub>DP</sub> a el [<sub>D'</sub> t<sub>i</sub> [<sub>NP</sub> *pro*]]]  
 he.CL saw PREP he  
 'I saw him.' (Uriagereka 1995: 80)

<sup>15</sup> A reviewer asked why the pronoun in the 'big DP' in (29,30) must always be silent. We suggest that this follows from the application of the Avoid Pronoun Principle, or more precisely, "avoid **overt** pronouns"; see originally Chomsky (1981) on PRO and later Kornfilt (1984) extending the principle to *pro*. There is indeed a general tendency in Russian to avoid overt pronouns in certain contexts. One example is the inalienable possession constructions where the possessor is structurally present yet preferably remains silent (see Burukina 2014 for discussion). Another example is instances of hyper-raising illustrated in (i); such sentences are regularly accepted by some native speakers and are attested in corpora, e.g., in the General Internet Corpus of Russian (<http://www.webcorpora.ru/en/>). Importantly, a small survey that we conducted (with 30 monolingual speakers of Russian) showed that those speakers who allow hyper-raising unanimously reject copy-raising, i.e., parallel examples with an overt subject in the embedded clause; see also Milenković (2022) for a similar observation on hyper-raising in Serbian.

- (i) Deti kazalis', budto (\*oni) byli čem-to udivleny.  
 children seemed.PL as.if they were something.INS surprised  
 'It seemed that the children were surprised by something.'

Hyper-raising in Slavic deserves a separate detailed discussion. For the purposes of the present paper it suffices to point out certain parallels between such cases and the copular clauses with *èto*. In (i) the silent subject in the embedded clause essentially serves as a variable that allows for the finite CP to be predicated of the matrix subject (Den Dikken 2017). In the 'NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>' clauses, under our analysis, it is the presence of *pro* in 'big DP', whose content needs to be identified by NOM<sub>post</sub>, that allows to use the DP as a predicate at the clausal level.

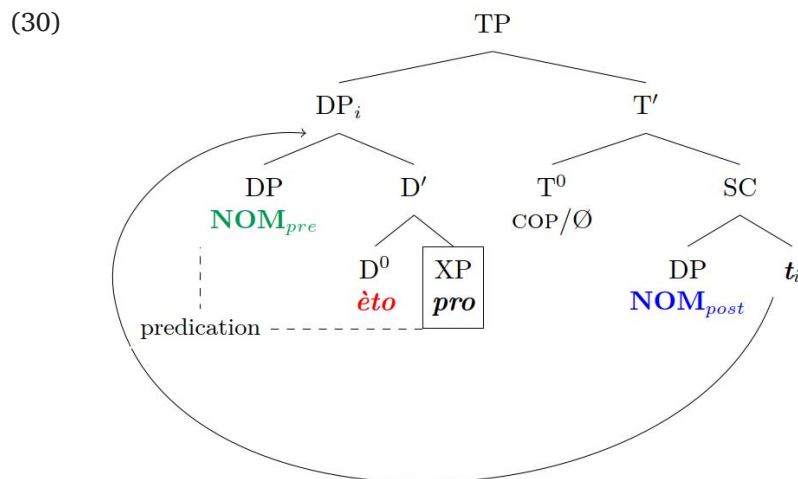
Note that while the examples such as (ii) are not ungrammatical, their information-structural properties point to an underlying structure drastically different from that in (29) and similar to (iii), with the pronoun being focused and *èto* functioning as a focus marker; see Burukina & Den Dikken (2020) for a discussion of the *èto*-focus construction.

- (ii) Petja, èto on byl moj načal'nik.  
 Petja, this he was my boss  
 'Petja, it was him who was my boss (and not Ivan/someone else).'  
 Infelicitous: 'Petja was my boss.' (as a response to *Petja? Which one is he?*)

*pro* are inherently referential, the symmetry must be broken, and we propose that this is done by *èto* in  $D^0$ , which we analyze as a predicate-forming operator, a ‘predicativizer’. Its role is to turn *pro* in its complement position into a predicate and to mediate the structural relation between this predicate and its subject (occupying Spec,DP).<sup>16</sup>

Outside the ‘big DP’, the content of *pro* is contextually determined by  $NOM_{post}$  (‘Petja was it/this, viz., my boss’). We assume that the fact that *pro* is co-indexed with  $NOM_{post}$  allows it to be formally licensed as a silent element and its content identified.<sup>17</sup> The fact that *pro* does not have its content identified within the confines of the ‘big DP’ (causing this ‘big DP’ to contain a free variable) allows, and indeed forces, this ‘big DP’ to be used as a predicate at the clausal level. In this respect, our ‘big DP’ is very much on a par with *tough*-movement constructions: there, too, the occurrence (inside an otherwise saturated constituent) of a variable whose range is left unrestricted inside the infinitival clause is what turns the infinitival clause into a predicate at the level of the copular clause.

Our proposed structure of ‘ $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ’ is shown in (30), with the predicativized *pro* boxed.



- (iii) Èto kašu ja ljublju.  
 this porridge.ACC I like  
 ‘It is porridge that I like.’

<sup>16</sup> An alternative outlook would be to assume that the complement of  $D^0$  is always a predicate and that referential DPs are the result of the establishment of a movement or binding relationship between  $D^0$  and its complement (see, e.g., Longobardi (1994) on proper names and N-to-D movement). On this outlook, the “predicativizing” effect of *èto* on its complement translates into the hypothesis that *èto* can forgo/pre-empt the establishment of such a relationship with its complement. An approach along these lines is empirically equivalent to the one pursued in this paper.

<sup>17</sup> Throughout the paper we use *identify* and *identified* primarily when discussing the process of determining the reference of *pro* within the ‘big DP’. To avoid confusion with the term *identification* as the name of a group of copular constructions à la Higgins (1973), we use *identity* to refer to the latter, as pointed out in 2.3.

The *èto*-headed DP in ‘ $\text{NOM}_{\text{pre}} \text{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ constructions is first merged as the predicate of a small clause, whose subject is  $\text{NOM}_{\text{post}}$ . The *èto*-headed DP then undergoes movement to Spec,TP, which results in the obligatory word order  $\text{NOM}_{\text{pre}}\text{-èto(-copula)-NOM}_{\text{post}}$ . As was shown in (22), the alternatives, without the *èto*-headed DP raising, are ungrammatical.

The hypothesis that  $[_{\text{DP}} \text{NOM}_{\text{pre}} \text{èto}]$  is a constituent provides an explanation for the behavior of ‘ $\text{NOM}_{\text{pre}} \text{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ constructions, e.g., in *wh*-questions, which would otherwise be hard to explain. The relevant empirical observation is that ‘ $\text{NOM}_{\text{pre}} \text{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ constructions are ungrammatical as root *wh*-questions, (31). This is expected on our analysis, since  $\text{NOM}_{\text{pre}}$  (*kto* ‘who’) can neither be merged as a hanging topic – for information-structural reasons: *wh*-words do not make good hanging topics – nor be raised to Spec, CP via sub-extraction from the ‘big DP’  $[_{\text{DP}} \text{NOM}_{\text{pre}} \text{èto}]$  in Spec,TP, because of criterial freezing: the ‘big DP’ has been moved into a criterial position (Spec,TP), and is frozen as a result.<sup>18</sup>

- (31) \**Kto èto Toni?*  
 who.NOM this Tony.NOM  
 ‘Who is Tony?’

We conclude that in the ‘ $\text{NOM}_{\text{pre}} \text{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ constructions under consideration,  $\text{NOM}_{\text{pre}}$  and *èto* form an immediate constituent, a ‘big DP’, as is evident from the behavior of the constructions in *wh*-questions.<sup>19</sup> Furthermore, the agreement facts show that  $\text{NOM}_{\text{post}}$  acts as the subject of the clause probed by T<sup>0</sup>.

---

<sup>18</sup> Note that the intended reading of (31) does not include a pause/comma intonation between *èto* and *Toni*, which would make *Toni* an afterthought and render the structure grammatical.

Also, it should be noted that there are two related types of *wh*-questions that need to be addressed. First, a felicitous rephrasing of (31) is provided in (i), where *takoj* ‘such’ is obligatory – even though its semantic contribution might not be immediately clear:

- (i) *Kto takoj Toni?*  
 who such Tony  
 ‘Who is Tony?’

We propose that in (i), *Toni* acts as the subject, and *kto takoj* is a constituent that corresponds to the predicate. The role of *takoj* ‘such’ is to grant predicative interpretation to the *wh*-word. Cf. also [*Kto zdes?*] *Toni?* (Who here is Tony?), [*Kto iz vas?*] *byl Toni?* (Who of you was Tony?). We suggest that *èto* cannot be used in constructions like (i) due to criterial freezing.

Second, a question like (ii), in contrast with (31), is felicitous. Provisionally, we suggest that *Toni* in (ii) acts as a hanging topic, which is supported by the requirement on discourse-givenness of *Toni* and a prosodic break following it. We leave detailed treatment of examples of this sort for further research.

- (ii) *Toni èto (byl) kto?*  
 Tony this was who  
 ‘Who was Tony?’ (Lit.: ‘Tony – who was that?’)

<sup>19</sup> The proposal that  $[_{\text{DP}} \text{NOM}_{\text{pre}} \text{èto}]$  is a constituent also provides an explanation for the behavior of ‘ $\text{NOM}_{\text{pre}} \text{èto}$  (be)  $\text{NOM}_{\text{post}}$ ’ constructions in embedded yes/no-questions. In yes/no-questions in Russian, the particle *li* is the

### 3.3 Èto clauses without a NOM<sub>pre</sub>

So far in this paper, our attention has mostly been confined to ‘NOM<sub>pre</sub> èto (be) NOM<sub>post</sub>’ constructions, in which èto is preceded by a nominative noun phrase. At this point, we would like to extend the scope of the analysis to ‘èto (be) NOM<sub>post</sub>’ constructions, as illustrated in (32), where èto is again invariant and NOM<sub>post</sub> is a nominative noun phrase, but no additional nominative noun phrase shows up to the left of èto.

- (32) Èto byl rimskij imperator.  
 this was.M.SG Roman emperor.NOM  
 ‘It/this was a/the Roman emperor.’

At the level of the copular clause, the structure of these constructions is identical to that of the other copular clauses with èto, as presented in the previous sections: NOM<sub>post</sub> is the underlying subject of predication, while the constituent that includes èto is the predicate. That the èto-headed DP is once again the predicate of the sentence is supported by the ungrammaticality of predicative copular constructions or specificational constructions of the shape ‘It/this èto NOM<sub>post</sub>’. Èto cannot appear in these constructions, as illustrated in (33). This fits well with our analysis of the èto-headed DP as the predicate, since there can only be one main predicate per clause. For analyses that treat èto as a copula spelling out the T<sup>0</sup> head or a similar functional head

manifestation of the C<sup>0</sup> head (King 1995); it is obligatory in embedded contexts, where it must be preceded by the focused constituent (or, in the absence of a narrowly focused constituent, the verb), as shown in (i).

- (i) Mama sprosila, priexali \*(li) oni včera.  
 Mum asked arrived Q they yesterday  
 ‘Mum asked whether they had arrived yesterday.’

Importantly, in ‘NOM<sub>pre</sub> èto (be) NOM<sub>post</sub>’ constructions, *li* cannot break into the [<sub>DP</sub> NOM<sub>pre</sub> èto] unit. This is illustrated in (iia). On our analysis, this is explained by the fact that [<sub>DP</sub> NOM<sub>pre</sub> èto] is a single constituent, and the C<sup>0</sup> head cannot be wedged between its subconstituents. At the same time, moving NOM<sub>pre</sub> to Spec,CP via sub-extraction out of this ‘big DP’ would incur a violation of criterial freezing: the ‘big DP’ has been moved into a criterial position (Spec, TP) and is frozen as a result. Thus, the only way to express the desired meaning is to use direct speech, as shown in (iib).

- (ii) a. \*Mama sprosila Železnyj čelovek li èto Toni.  
 Mum asked iron man.NOM Q this Tony.NOM  
 ‘Mum asked whether the Iron Man was Toni.’  
 b. Mama sprosila: Železnyj čelovek – \*(èto) Toni?  
 Mum asked iron man.NOM this Tony.NOM  
 ‘Mum asked whether the Iron Man was Toni.’

A reviewer suggested an alternative account for the ungrammaticality of (iia), whereby *li* is considered a focus particle incompatible with the topical NOM<sub>pre</sub>. We agree that this explanation is viable and that the observed pattern may be a consequence of the *independent* fact that there is a tight connection between *li* and focus. Nevertheless, we believe that the data in (i-ii) should be mentioned, as they are potentially relevant for future extended discussions of the information-structural properties of èto copular clauses.

on the clausal spine, examples like (33) posit a problem: under those approaches, the restriction on what can surface as  $\text{NOM}_{\text{pre}}$  would be unmotivated.

- (33) \*Èto / to èto moj načal'nik.  
 this that this my boss.NOM  
 (Intended: 'It/this is my boss.')

The main difference between 'èto (be)  $\text{NOM}_{\text{post}}$ ' and ' $\text{NOM}_{\text{pre}}$  èto (be)  $\text{NOM}_{\text{post}}$ ' is that the specifier position in the 'big DP' is occupied by an overt  $\text{NOM}_{\text{pre}}$  in the latter but not in the former. But èto is still a functional head  $D^0$  within the 'big DP', and its role, in constructions with and without  $\text{NOM}_{\text{pre}}$  alike, is to predicativize the silent pronoun (*pro*) in its complement. The difference between 'èto (be)  $\text{NOM}_{\text{post}}$ ' and ' $\text{NOM}_{\text{pre}}$  èto (be)  $\text{NOM}_{\text{post}}$ ' comes down to the silence/overtness of the subject of predication within the 'big DP': an overt nominative in the latter; a silent element (another *pro*) in the former.

- (34) [[ $\text{DP}$   $pro_m$  èto [ $pro_k$ ]]<sub>j</sub>] byl [ $\text{SC}$  [ $\text{DP}$  rimskij imperator]<sub>k</sub> t<sub>j</sub>]].  
 this was.M.SG Roman emperor.NOM  
 'It/this was a Roman emperor.'

Considering the reference of the *pros* in (34), we propose that  $pro_m$  is essentially  $\text{NOM}_{\text{pre}}$  being pro-dropped, eligible for omission because it is discourse-anaphoric, similarly to other topical/discourse-anaphoric subjects in Russian (McShane 2009; Tsedryk 2015, i.a.) The discourse-anaphoricity of  $pro_m$  becomes evident when (34) is used in a context, for instance, in response to the questions in (35), where another person or a statue acts as an antecedent.<sup>20</sup>

- (35) *Who was the person<sub>m</sub> who just appeared on the screen?*  
 $pro_m$  èto byl (odin) rimskij imperator.  
 this was one Roman emperor  
 'It/this was the/a Roman emperor.'

As for  $pro_k$  in the complement of 'big DP' in (34), we argue that, just as in other ' $\text{NOM}_{\text{pre}}$  èto (be)  $\text{NOM}_{\text{post}}$ ' constructions, its value is identified by  $\text{NOM}_{\text{post}}$ . Èto predicativizes  $pro_k$ , and a predication relation is established within the 'big DP', with  $\text{NOM}_{\text{pre}}$  in its specifier acting as the referential subject of predication.

---

<sup>20</sup> There is another context in which (34) may be used as an answer:

(i) "Which statue<sub>m</sub> was that of the/a Roman emperor?"

The difference between the two, as far as we can tell, is the following. In (35), 'Roman emperor' is used as an 'identifier' for the person on the screen, unfamiliar to the interlocutor. If the example in (35) were used as the response to (i), 'Roman emperor' would be the 'identified', as the interlocutor wants to establish an identity relation between him and one of the statues. Preliminarily, we suggest that this difference is established pragmatically and does not manifest itself in the syntactic structure, which stays as in (34). We thank Ora Matushansky for bringing up this issue and leave it for future research.

To summarize, we have argued that in copular clauses of the type ‘*èto* (be)  $NOM_{post}$ ’ a predication relation is established between the postcopular nominative DP ( $NOM_{post}$ , the subject of predication) and the *èto*-headed DP (the predicate). We have further argued that *èto*, when taking *pro* as its complement, turns this null proform into a predicate. The null proform’s content is identified via co-indexation with  $NOM_{post}$ , as in identity copular clauses with two overt nominal dependents. Thus, our account uniformly accommodates all kinds of copular clauses where *èto* appears, which is a favorable outcome from a theoretical perspective.

## 4 Against a hanging topic analysis

### 4.1 Non-bridge predicates and conditional clauses

In the literature, it has often been suggested that  $NOM_{pre}$  is base-generated outside of the copular clause as a dislocated/hanging topic; see for instance Geist (2008), Markman (2008), and most recently Seres & Espinal (2019). Furthermore, Geist (2008) uses the hypothesis that  $NOM_{pre}$  is a hanging topic to explain the agreement pattern and the fact that the copula never spells out the features of  $NOM_{pre}$ : as a clause-external topic,  $NOM_{pre}$  is not in a local relationship with  $T^0$ .

In this section we provide evidence against a hanging topic analysis being applicable to (all) examples of the ‘ $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ’ type, by looking more closely at their syntactic behavior in two contexts – when these constructions are embedded under non-bridge predicates and in conditional clauses – as well as their prosodic properties.

To start with embedded contexts with non-bridge predicates, consider the examples in (36). They show that *udivitel’no* ‘surprisingly’ in Russian behaves as a non-bridge predicate, in that it does not allow sub-extraction out of an embedded complement clause without resumption, thus forming a weak factive island (Polinsky & Potsdam 2014).

- (36) a. Udivitel’no, [čto Petja poceloval Mašu].  
 surprising that Petja kissed Maša.ACC  
 ‘It is surprising that Petja kissed Maša.’
- b. \*Mašu<sub>i</sub> udivitel’no, [čto Petja poceloval t<sub>i</sub>].  
 Maša.ACC surprising that Petja kissed
- c. ?Maša, udivitel’no, [čto Petja poceloval eë].  
 Maša.NOM surprising that Petja kissed her.ACC  
 ‘As for Maša, it is surprising that Petja kissed her.’

As is typical for non-bridge predicates cross-linguistically, *udivitel’no* ‘surprisingly’ cannot embed a finite clause with a hanging topic, as shown in (37a). However, *udivitel’no* ‘surprisingly’ can take a clausal dependent of the type ‘ $NOM_{pre}$  *èto*  $NOM_{post}$ ’, as demonstrated by (37b). That (37b) is grammatical falls out directly from a treatment of [<sub>DP</sub>  $NOM_{pre}$  *èto*] as a ‘big DP’ constituent

occupying Spec,TP. It would be incompatible with  $NOM_{pre}$  as a hanging topic, accompanied by an independent functional/pronominal element, *èto*.<sup>21</sup>

- (37) a. \*Udivitel'no, čto Maša, Petja poceloval eë.  
 surprising that Maša.NOM Petja kissed her.ACC  
 (Intended: 'As for Maša, it is surprising that Petja kissed her.')
- b. *Context: And who is this Iron Man? Oh, it is Tony?!*  
 Udivitel'no, [čto železnyj čelovek – èto Toni].  
 surprising that iron man.NOM this Tony.NOM  
 'It is surprising that the Iron Man is Tony.'

Further evidence against necessarily treating  $NOM_{pre}$  as a hanging topic comes from conditional dependent clauses. In English, as Haegeman (2007) shows, conditional *if*-clauses resist DP-topicalization within the embedded clause:

- (38) \*If **these exams** you don't pass, you won't get the degree.

Russian does not allow for hanging topics to be embedded in conditional clauses either. This is shown in (39b), in contrast to an acceptable non-embedded clause with a hanging topic in (39a).

- (39) a. Šon Konneri, on rodilsja v nebogatoj sem'e.  
 Sean Connery he was.born in modest family  
 'Sean Connery, he was born in a modest family.'
- b. Esli (\*Šon Konneri,) on rodilsja v nebogatoj sem'e, ego talant vpečatljaet eščë  
 If S. C. he was.born in modest family his talent impresses even  
 bol'se.  
 more  
 'If he was born in a modest family his talent is even more impressive.'

In contrast, a ' $NOM_{pre}$  èto  $NOM_{post}$ ' can felicitously be embedded in a conditional clause, which shows that it does not contain a hanging topic. This is illustrated in (40).

- (40) *Context: The speaker watched a James Bond movie; they knew that one of the actors was Sean Connery but could not identify any of the actors as him.*  
 Esli Šon Konneri èto byl Džejms Bond, to ego talant vpečatljaet eščë bol'se.  
 if S. C. this was James Bond then his talent Impresses even more  
 'If Sean Connery was James Bond, his talent is even more impressive.'

---

<sup>21</sup> While we do not discard entirely the possibility of the pre-*èto* DP being an external topic in **some** sentences, the grammaticality of embedding ' $NOM_{pre}$  èto (be)  $NOM_{post}$ ' under non-bridge predicates, as in (37), is an insurmountable problem for generalized hanging-topic approaches.



The grammaticality of (40) in contrast to (39b) tells us reliably that the noun phrase referring to Sean Connery in (40) is not a hanging topic (as the bulk of the existing proposals for *èto* copular clauses would treat it). As a non-hanging topic, *Šon Konneri* will have to originate somewhere inside the clause. With the *èto*-phrase and *Džejms Bond* already occupying the only two phrasal positions within the copular clause, this forces *Šon Konneri* into originating inside either of these two nominals – and since proper names have no DP-internal dependents (which takes the nominal of *Džejms Bond* off the table), this leaves the DP of *èto* as the only home for *Šon Konneri*. Furthermore, the fact that our analysis maps *Šon Konneri* and *èto* in (40) onto a constituent guarantees that the two of them together precede the copula *byl*; breaking up this constituent by subextracting *Šon Konneri* from it and topicalizing it separately would create syntactic difficulties no matter whether this were done after raising the complex *èto*-DP into Spec,TP (‘criterial freezing’; fn. 19) or prior to this (cyclicality), so the unity of [*Šon Konneri èto*] as a single syntactic constituent must arguably be preserved.

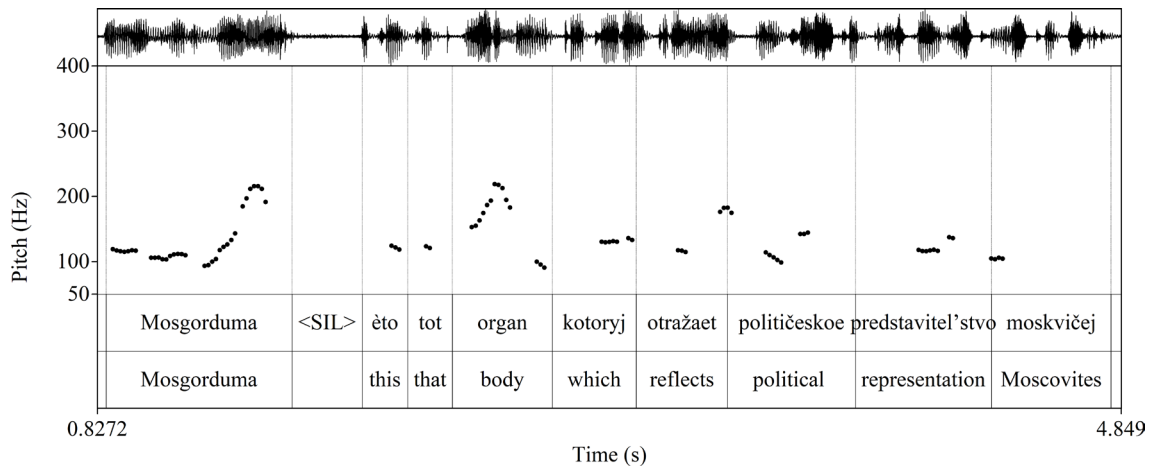
The constituent status of  $\text{NOM}_{\text{pre}}$  and *èto* is also compatible (perhaps surprisingly, at first blush) with the prosody of ‘ $\text{NOM}_{\text{pre}}$  *èto* (be)  $\text{NOM}_{\text{post}}$ ’ constructions. We show this in the next section.

## 4.2 Prosodic properties

It is commonly assumed that there is an obligatory intonational break or pause between  $\text{NOM}_{\text{pre}}$  and *èto*, which might be regarded as undermining the ‘big DP’ analysis of  $\text{NOM}_{\text{pre}}$  and *èto*. We find, however, that there is considerable variability in the prosody of ‘ $\text{NOM}_{\text{pre}}$  *èto* (be)  $\text{NOM}_{\text{post}}$ ’ constructions, and that, while a prosodic break between  $\text{NOM}_{\text{pre}}$  and *èto* is possible, it is certainly not required. Our analysis of a random sample of ca.150 tokens of ‘ $\text{NOM}_{\text{pre}}$  *èto* (be)  $\text{NOM}_{\text{post}}$ ’ constructions from the Multimedia sub-corpus of the Russian National Corpus (<https://ruscorpora.ru/new/search-murco.html>) shows that  $\text{NOM}_{\text{pre}}$  and *èto* are often produced without a prosodic break.

More specifically, we find that in some ‘ $\text{NOM}_{\text{pre}}$  *èto* (be)  $\text{NOM}_{\text{post}}$ ’ examples, the left edge of *èto* is indeed aligned with the left edge of a prosodic constituent, which may be realized as a pause and/or a strong glottal onset on the initial vowel in *èto*. It should be noted, though, that a prosodic break before *èto* is most commonly found in contrastive or strongly emphatic contexts, especially those involving ellipsis, as well as very formal contexts, such as speeches. This is illustrated in (41) and **Figure 1**, an example extracted from an interview with Alexei Navalny, which is pronounced with strong emphasis on *Mosgorduma* ‘Moscow City Duma’. As **Figure 1** shows, there is a perceptible pause between *Mosgorduma* and *èto* (labelled <SIL>).

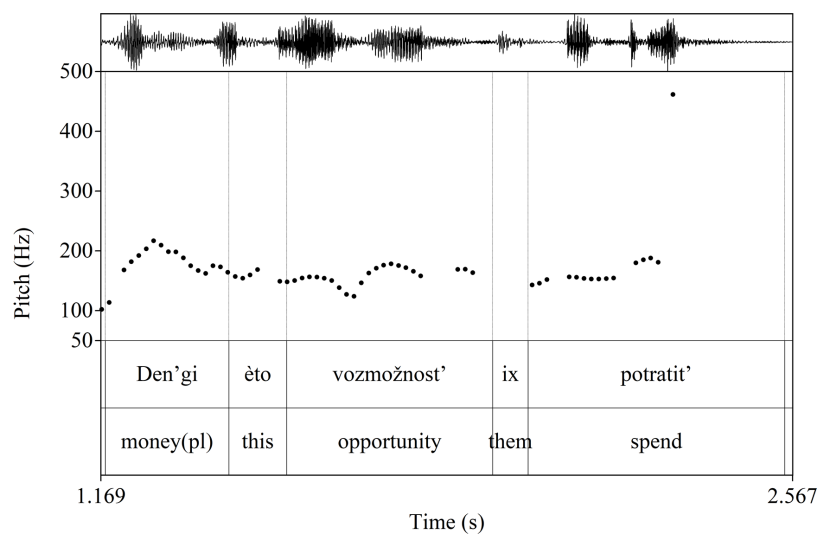
- (41) **Mosgorduma** *èto* **tot organ**, kotoryj otryažet političeskoe predstavitel’stvo  
 Moscow City Duma this that body which reflects political representation  
 moskvičej.  
 Moscovites.GEN  
 ‘The Moscow City Duma is the body that reflects the political representation of  
 Moscovites.’ (*navalny\_2013\_045*)



**Figure 1:** An example of a prosodic break between  $NOM_{pre}$  (*Mosgorduma*) and *èto*.

Nevertheless, examples without a pause or any evidence for a prosodic boundary before *èto* are also commonly attested in the corpus. Two subtypes of this pattern can be distinguished. In the first one, which is especially common in informal and/or rapid speech, no audible pauses or other prosodic boundary cues can be identified. This pattern is illustrated in (42) and **Figure 2**, extracted from an interview with Pyotr Miloserdov. The pronunciation of this example is entirely natural and, in contrast with the example in **Figure 1**, does not convey extra emphasis on any of the constituents.

- (42) Den'gi èto vozmožnost' ix potratit'.  
 money.PL this opportunity they.ACC spend  
 'Money is the opportunity to spend it.' (*miloserdov\_037*)

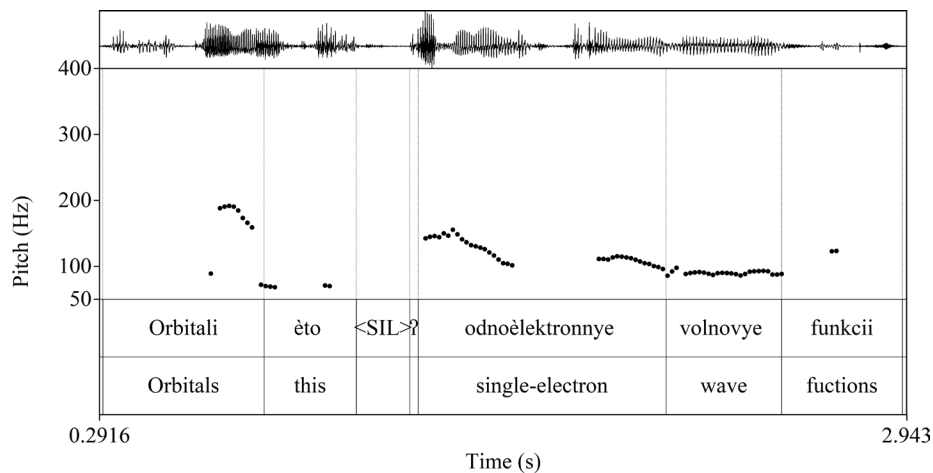


**Figure 2:** An example with no prosodic break between  $NOM_{pre}$  (*den'gi* 'money') and *èto*.

The second subtype of examples that do not contain a prosodic boundary before *èto* in fact provide evidence for a prosodic boundary to the **right** of *èto*, which may be manifested by (i) a pause, (ii) a strong glottal onset on the initial vowel of the word following *èto*, and/or (iii) final lengthening on *èto*. These prosodic strategies were found in a variety of pragmatic contexts. Cues (i) and (ii) are illustrated in (43) and **Figure 3**, where *èto* is followed by a pause, and the first vowel in *odnoèlektronnye* ‘single-electron’ is realized with a strong glottal onset, labelled with ‘?’.

The example is extracted from a mini-lecture on popular science by Professor Artem Oganov.

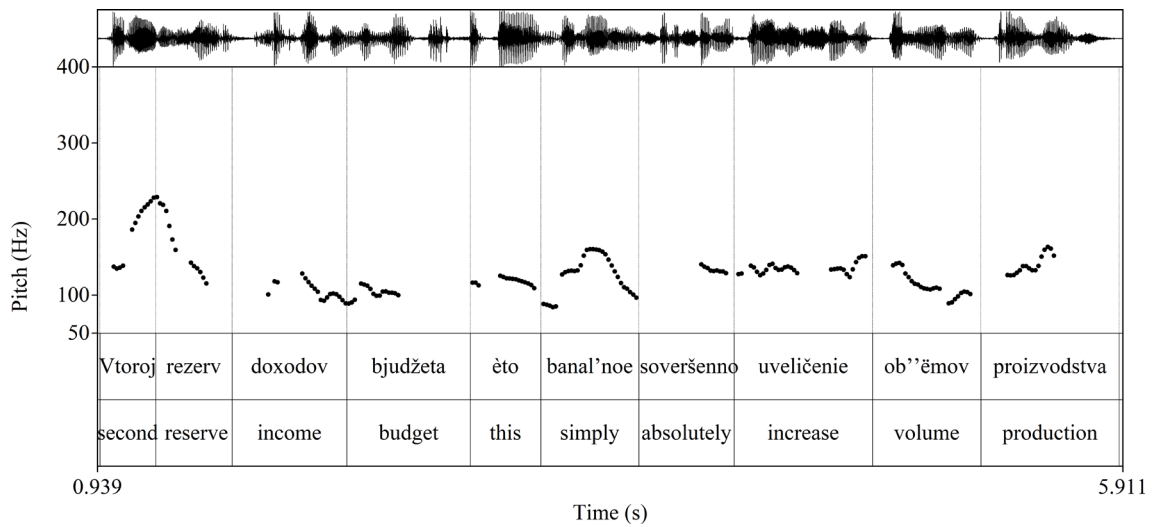
- (43) Orbitali èto odnoèlektronnye volnovye funkcii.  
 orbital.PL this single-electron wave functions  
 ‘Orbitals are single-electron wave functions.’ (oganov\_kvant\_025)



**Figure 3:** An example with a pause after *èto* and a strong glottal onset on the following vowel.

Finally, cue (iii), final lengthening on *èto*, is illustrated with an example from an interview with Sergey Glazyev, shown in (44) and **Figure 4**; the extra duration, received by the final vowel in *èto*, is especially apparent in the waveform.

- (44) Vtoroj rezerv doxodov bjudžeta èto banal'noe soveršenno uveličenie ob"ëmov  
 Second reserve income budget this simply absolutely increase volume  
 proizvodstva.  
 production  
 ‘The second reserve of budget income is simply an increase in production volume.’  
 (glazyev\_008)



**Figure 4:** An example with final lengthening on *èto*.

We hypothesize that the presence of a prosodic break between  $NOM_{pre}$  and *èto* in ' $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ' constructions may be attributable to the expression of contrast or emphasis, or be a pragmatic marker of a formal context. It certainly does not represent an intrinsic prosodic property of the construction itself. ' $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ' constructions have a range of prosodic realizations, including ones in which  $NOM_{pre}$  and *èto* as a unit are prosodically separated from the rest of the clause and likely form a prosodic constituent.

## 5 Implications

The 'big DP' analysis of ' $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ' constructions proposed in this paper also sheds light on the properties of *èto* found in other contexts. Here, we discuss two such constructions: (1) *èto* copular clauses with a predicational dependent (an AP), and (2) specificational pseudoclefts of the question–answer pair type. To the best of our knowledge, (1) has not been previously mentioned in the literature, while (2) remains largely understudied.<sup>22</sup> While we do not aim at providing complete analyses for these phenomena here, we hope to establish a clear parallel between them and the ' $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ' clauses under discussion that may further be explored by future research.

<sup>22</sup> We would like to suggest that the third phenomenon that may bear a structural similarity to ' $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ' clauses is clausal prolepsis, as shown in (i). Here, the position of  $NOM_{pre}$  is occupied by a CP (a fully saturated proposition), and the position of  $NOM_{post}$  is occupied by a property-denoting adjectival/nominal phrase.

(i) [<sub>DP</sub>[<sub>CP</sub> Čto my guljaem v parke], èto] prijatno.  
 that we walk in park this nice  
 'That we are walking in the park is nice.'

We leave a dedicated discussion of these examples in the context of our proposal for further research.

## 5.1 Èto clauses with a predicational dependent

In section 2.3 we argued that the ‘NOM<sub>pre</sub> èto (be) NOM<sub>post</sub>’ clauses can only be of the identity type; in section 3 we have further shown that NOM<sub>post</sub> typically is a referential nominal phrase. However, examples of ‘NOM<sub>pre</sub> èto (be) NOM<sub>post</sub>’ sentences can be found in which the postcopular constituent is represented by unambiguously predicational expressions – an adjectival phrase or a prepositional phrase. Èto is still present and still fails to exhibit number/gender concord with NOM<sub>pre</sub>, always remaining neuter singular. Two such examples, the first featuring a postcopular AP (*modno* ‘fashionable’) and the second a postcopular PP (*na zavtrak/obed* ‘for breakfast/lunch’), are given in (45).<sup>23</sup>

- (45) a. Šakira \*(èto) bylo/\*byla modno/\*modnaja.<sup>24</sup>  
 Shakira.F this was.N.SG/F.SG fashionable.PRED/F  
 ‘Shakira was fashionable [e.g., to listen to].’
- b. V moëm detstve, kaša \*(èto) bylo/\*byla na zavtrak, a  
 in my childhood porridge this was.N.SG/F.SG on breakfast and  
 sup \*(èto) bylo/\*byl na obed.<sup>25</sup>  
 soup this was.N.SG/M.SG on lunch  
 ‘In my childhood porridge was for breakfast and soup was for lunch.’

These examples contrast with ordinary predicational copular clauses: here, èto is prohibited, concord/agreement with NOM<sub>pre</sub> is required, and the interpretation changes – the property is now attributed directly to NOM<sub>pre</sub>. This is shown in (46).

<sup>23</sup> The pattern exemplified in (45) resembles the so-called ‘pancake sentences’ in Scandinavian languages (ia) (see Faarlund 1977; Källström 1993, *i.a.*) and cases of lack of number/gender concord in copular clauses in Brazilian Portuguese (ib) (see Duek 2012 and references therein). While at this point we confine ourselves to an in-depth examination of the Russian data, we believe that the ‘big DP’ analysis that we propose can in principle be applied to other languages to account for similar examples.

- (i) a. Pannekaker er godt. [Norwegian]  
 pancakes.PL are good.NEUT.SG  
 ‘Pancakes are good.’
- b. Banana importada/\*o é caro/cara. [Brazilian Portuguese]  
 banana.F imported.F/\*M is expensive.M/F  
 ‘Imported bananas are expensive.’

<sup>24</sup> The form *modno* here belongs to a special group of so-called predicatives (Zolotova 1982, Bonch-Osmolovskaja 2003, Say 2013, *i.a.*). They usually end with the inflection *-o*, also characteristic of adverbs and neuter singular forms of short-form adjectives. As argued by Letuchiy (2017), predicatives are adjectival in their nature, since adverbs generally cannot serve as primary predicates. However, there is an ongoing debate about whether predicatives and short forms of adjectives are actually the same; see Bonch-Osmolovskaja (2003), Zimmerling (2003), Say (2013), Letuchiy (2017), Burukina (2019) for more data and argumentation. For the purpose of the present discussion, it suffices to say that semantically *modno* denotes a property and morphosyntactically it is invariant.

<sup>25</sup> Some of the native speakers that we consulted prefer examples with a PP, such as (45b), to contain an explicit contrast. The contrastive interpretation may suggest that (45b) contains hanging topics. However, the analysis that we

- (46) a. Šakira (\*èto) byla/\*bylo modnaja/\*modnoe/\*modno.  
 Shakira this was.F.SG/N.SG fashionable.F.SG/N.SG/PRED  
 ‘Shakira was fashionable [herself].’  
 = ‘Shakira was a fashionable person.’
- b. Kaša byla/\*bylo na zavtrak.  
 porridge was.F.SG/N.SG for breakfast  
 ‘The porridge was for breakfast [e.g., yesterday].’  
 = ‘There was porridge for breakfast.’

Existing analyses of *èto* constructions are ill-equipped to capture the data in (45). For instance, under Markman’s (2008) approach,  $NOM_{post}$  is always an ordinary predicate, and the prediction is that an AP in a context like (45a) should exhibit concord with the subject, and the overt past tense copula should spell out the subject’s  $\varphi$ -features. Under Geist’s (2008) approach, whereby  $NOM_{post}$  is always the main subject of predication and  $NOM_{pre}$  is a hanging topic, we expect no  $NOM_{post}$  to ever be a property-denoting expression. Our double-predication analysis, whereby a predication relation is established both within the clause and inside a ‘big DP’, leaves just enough room to accommodate the seemingly exceptional examples.

We have previously proposed that  $NOM_{pre}$  and *èto* form a ‘big DP’, whose complement position is occupied by *pro*. The content of *pro* is determined by  $NOM_{post}$ , and *èto* turns it into a predicate (by turning it into an expression of the type  $\langle e,t \rangle$ ). Both operations – determining the content of *pro* and semantic predicativization – happen at LF, and we would like to suggest that they may in principle occur in two different relative orders, as follows.

In the ‘ $NOM_{pre}$  *èto* (be)  $NOM_{post}$ ’ examples with two referential nominal phrases considered in the previous sections, the interpretation of the content of *pro* within the ‘big DP’ always took place first and predicativization by *èto* followed. As a result, the referential *pro* could only be co-indexed with a  $NOM_{post}$  of the type  $\langle e \rangle$ . But now imagine that predicativization takes place first. In this scenario, *èto* turns *pro* into a predicate of the type  $\langle e,t \rangle$ , and, after that *pro* must get linked to an expression of the same type outside the ‘big DP’; naturally, such an expression is an adjectival/prepositional phrase. Note that, in an example like (45), *èto* cannot be omitted – because it is still required to predicativize *pro*.

Thus, in (45), *modno* ‘fashionable’ and *na zavtrak* ‘for breakfast’ represent abstract properties and determine the content of the predicativized *pro* within the ‘big DP’. The resulting interpretation

---

propose can accommodate such cases by assuming that *kaša* is merged outside the TP and co-indexed with *pro* in the specifier position of the ‘big DP’ (see section 3.3). Crucially, regardless of the topichood status of  $NOM_{pre}$  in (45b), these examples would be hard to account for within alternative analyses, such as Geist’s.

is close to that of sentences of the type *What Shakira was was fashionable*, where *fashionable* is not predicated directly of Shakira but rather specifies the property associated with her as a phenomenon (as opposed to a person); hence, the perceptible difference in the readings of the examples in (45) and (46).

In syntax, the copula receives the default neuter gender specification. This is unsurprising, as  $NOM_{pre}$  does not show concord with the *pro* within the ‘big DP’ and, not itself being in a direct relationship with  $T^0$  (due to being embedded in the constituent in  $Spec,TP$ ), cannot control agreement at the level of the clause. This is schematized in (47). Accordingly, no concord is possible between *Shakira* (F) and *modno* ‘fashionable’ (N) in examples like (45a).

- (47)  $[_{TP} [_{DP} [_{DP} \check{S}akira] [_{D'} [D^0 \acute{e}to \ [pro]_i]]] \text{ bylo } [_{AP} \text{ modno}]_i]$ .  
 Shakira.F      this      was.N.SG/F.SG fashionable.N/F

To recap, we proposed that the two LF operations that involve the ‘big DP’ – namely, determining the content of *pro* and the predicativization of *pro* – can come in two relative orders, which gives *èto* the potential to co-occur with a property-denoting expression (an AP or PP) in post-copular position in identificational copular clauses. We leave a detailed semantic analysis of these examples for future work.

## 5.2 Specificational pseudo-clefts

Finally, we would like to draw a tentative parallel between the ‘ $NOM_{pre} \acute{e}to$  (be)  $NOM_{post}$ ’ constructions and specificational pseudoclefts of the question–answer pair type, illustrated in (48).

- (48) a. Kogo Petja priglasil, (tak) èto Mašu.  
 who.ACC Petja invited PTCL this Maša.ACC  
 ‘As for who Petja invited, it was Masha.’  
 b. Èto Mašu Petja priglasil.  
 this Maša.ACC Petja invited  
 ‘It was Maša that Petja invited.’

Based on previous work (Markman 2008; Burukina & den Dikken 2020), we take *èto* in examples of this kind to spell out the  $Top^0$  head. Its function is to connect a question CP, located in  $Spec,TopP$ , with the answer TP, which contains the focused XP and is located in the complement of  $TopP$ . This is schematized in (49); the parts of the sentence that are structurally present but elided are marked by a strike-through line.

- (49) a.  $[_{TopP} [_{CP} \text{ Kogo Petja priglasil}] \ [_{Top'} [Top^0 = \acute{e}to] \ [_{TP} \text{ Mašu}_i \text{ Petja } \text{priglasil } t_i]]]$   
 b.  $[_{TopP} [_{CP} \text{ Kogo } \text{Petja } \text{priglasil}] \ [_{Top'} [Top^0 = \acute{e}to] \ [_{TP} \text{ Mašu}_i \text{ Petja } \text{priglasil } t_i]]]$

Although *èto* in ‘NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>’ copular clauses is a D<sup>0</sup> head, not a Top<sup>0</sup> head, there is an important parallel between the two kinds of clauses. In both cases, *èto* is a functional X<sup>0</sup> that mediates a relationship between two terms that are in a semantic co-construal relationship. These are NOM<sub>pre</sub> and *pro* in ‘NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>’ constructions and the topic and the comment in *èto*-focus constructions. This brings closer the two sets of data with *èto* that have generally been viewed as irreconcilable.

## 6 Concluding remarks

In this paper, we offered a novel analysis of ‘NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>’ constructions. First, by analyzing its distribution, we demonstrated that *èto* only appears in identity copular clauses. Then, we defended a uniform syntactic analysis of *èto* copular clauses, which successfully accommodates for a range of copular clauses with various types of NOM<sub>pre</sub>/NOM<sub>post</sub> dependents.

Our approach includes the following analytical components: (i) the *èto*-containing DP is a predicate and undergoes predicate inversion from its base position in the small clause to Spec,TP; (ii) NOM<sub>pre</sub> and *èto* form a constituent, a ‘big DP’, with NOM<sub>pre</sub> in Spec,DP, a silent *pro* in Comp,DP, and *èto* spelling out the D<sup>0</sup> head; and (iii) within the ‘big DP’, another predication relation is established, with *èto* predicativizing *pro* in its complement.

Our analysis of NOM<sub>pre</sub> and *èto* as forming a ‘big DP’ accounts for the fact that NOM<sub>pre</sub> does not act as a hanging topic, the behavior of ‘NOM<sub>pre</sub> *èto* (be) NOM<sub>post</sub>’ constructions in questions, and the prosodic properties of these constructions. Our analysis also explains why NOM<sub>pre</sub> never controls agreement with the copula.

In exploiting the ‘big DP’ hypothesis, our analysis of the Russian data is akin to Grewendorf’s (2008) account of left dislocation in German (which, in turn, takes its inspiration from Vat 1997). Grewendorf makes a point of establishing an explicit connection between the German facts and Italian clitic left-dislocation, the latter involving a clitic, for which the ‘big DP’ had originally been proposed (Uriagereka 1995, etc.). Grewendorf emphasizes that German left dislocation is sometimes possible in embedded clauses introduced by a complementizer, though German is by no means as liberal as Russian *èto* constructions in this regard. Whereas (as we saw in section 4.1) Russian *èto* constructions are legitimate under non-bridge predicates and in conditional clauses, German left dislocation is not (as Grewendorf himself points out). For Grewendorf, the German facts follow on the assumption that the specifier of the ‘big DP’ moves to a high topic position in the left periphery. Although connectivity effects are adduced by Grewendorf as indirect support for the proposed ‘big DP’ structure, German provides no direct evidence for the constituency of the hypothesized ‘big DP’. In Russian *èto* constructions, where movement into the left periphery does not (necessarily) take place, the ‘big DP’ syntax that we have proposed is more directly observable. Indeed, it seems to us that the ‘big DP’ approach is the only analysis



that is compatible with the entire spectrum of the empirical facts reviewed in this paper. To the extent that our analysis is successful in capturing the data discussed, it shows that ‘big DP’ structures are employed in environments other than clitic doubling, and can be used to account for otherwise elusive properties of copular clauses.

Finally, by treating *èto* as a predicativizer, our analysis provides an explanation for how two referential expressions,  $NOM_{pre}$  and  $NOM_{post}$ , can be put together to form a predication structure. At the same time, the analysis allows us to capture cases in which the  $NOM_{post}$  position is occupied by a property-type expression in the presence of *èto* and further suggests interesting parallels with specificational pseudoclefts. It has the potential to be extended to apply to copular constructions in other Slavic languages and beyond.

---

## Abbreviations

We used the Leipzig glossing rules (with some additions): 1 – first person, 3 – third person, ACC – accusative, CL – clitic, COP – copula, F – feminine, GEN – genitive, INS – instrumental, M – masculine, N – neuter, NOM – nominative, NPST – non-past, PL – plural, PRED – predicative, PREP – preposition, PTCL – particle, Q – question particle, SG – singular, SPEC – specific.

## Funding information

This research was supported by grants KKP 129921, K 135958, FK 145985 and PD 146415 of NKFIH, the Hungarian National Research, Development and Innovation Office, which we gratefully acknowledge.

## Acknowledgements

We thank our colleagues at the HUN-REN Hungarian Research Centre for Linguistics and the audiences at FDSL14, FASL30, TMP10, and SLS15 for useful feedback at various stages of this project.

We are also grateful to the anonymous reviewers and the editor at *Glossa* who helped us improve and develop the ideas in this paper. Special thanks are due to the Russian native speakers who provided the judgements. All remaining errors are our responsibility.

## Competing interests

The authors have no competing interests to declare.

---

## References

- Bailyn, John Frederick. 2004. Generalized inversion. *Natural Language & Linguistic Theory* 22(1). 1–50. DOI: <https://doi.org/10.1023/B:NALA.0000005556.40898.a5>
- Bonch-Osmolovskaja, Anastasija. 2003. *Konstrukcii s dativnym subjektom v russkom jazyke: opyt korpusnogo issledovanija*. Moscow: Moscow State University dissertation.
- Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24(4). 591–656.
- Burukina, Irina. 2014. *Syntax of the implicit possessives in Russian*. Moscow: Russian State University for the Humanities MA thesis.
- Burukina, Irina. 2019. Raising and control in non-finite clausal complementation. Budapest: Eötvös Loránd University dissertation.
- Burukina, Irina & Den Dikken, Marcel. 2020. Russian *èto*-focus and *to*-topic sentences as elliptical question-answer pairs. Ms. Hungarian Research Centre for Linguistics/Eötvös Loránd University.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Citko, Barbara. 2006. Small clauses reconsidered: not so small and not all alike. Ms. University of Washington.

- Den Dikken, Marcel. 2006. *Relators and linkers: the syntax of predication, predicate inversion, and copulas*. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/5873.001.0001>
- Den Dikken, Marcel. 2013. Predication and specification in the syntax of cleft sentences. In Hartmann, Katharina & Veenstra, Tonjes (eds.), *Cleft structures*, 35–70. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/la.208.02dik>
- Den Dikken, Marcel. 2017. Predication in the syntax of hyperraising and copy raising. *Acta Linguistica Academica* 64(1). 3–43. DOI: <https://doi.org/10.1556/2062.2017.64.1.1>
- Doron, Edit. 1988. The semantics of predicate nominals. *Linguistics* 26(2). 281–302. DOI: <https://doi.org/10.1515/ling.1988.26.2.281>
- Duek, Karen. 2012. Bare nouns and gender agreement in Brazilian Portuguese. *Proceedings of CLS* 48. 19–21.
- Faarlund, Jan Terje. 1977. Embedded clause reduction and Scandinavian gender agreement. *Journal of Linguistics* 13. 239–57. DOI: <https://doi.org/10.1017/S0022226700005417>
- Geist, Ljudmila. 2008. Predication and equation in copular sentences: Russian vs. English. In Comorovski, Ileana & Heusinger, Klaus (eds.), *Existence: Semantics and syntax*, 79–105. Dordrecht: Springer. DOI: [https://doi.org/10.1007/978-1-4020-6197-4\\_3](https://doi.org/10.1007/978-1-4020-6197-4_3)
- Geist, Ljudmila & Błaszczak, Joanna. 2000. Kopulasätze mit dem pronominalen Elementen *to/èto* im Polnischen und Russischen. *ZAS Papers in Linguistics* 16. 115–139. DOI: <https://doi.org/10.21248/zaspil.16.2000.35>
- Grewendorf, Günther. 2008. The left clausal periphery: Clitic left dislocation in Italian and left dislocation in German. In Shaer, Benjamin & Cook, Philippa & Frey, Werner & Maienborn, Claudia (eds.), *Dislocated elements in discourse: Syntactic, semantic, and pragmatic perspectives*, 49–94. London: Routledge.
- Haegeman, Liliane. 2007. Operator movement and topicalization in adverbial clauses. *Folia Linguistica* 18. 485–502. DOI: <https://doi.org/10.1515/flin.41.3-4.279>
- Heycock, Caroline & Kroch, Anthony. 1999. Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry* 30(3). 365–397. DOI: <https://doi.org/10.1162/002438999554110>
- Higgins, Roger. 1973. *The pseudo-cleft construction in English*. Amherst, MA: University of Massachusetts Amherst dissertation.
- Higgins, Roger. 1979. *The pseudo-cleft construction in English*. New York: Garland.
- Junghanns, Uwe. 1997. On the so-called *èto*-cleft construction. In Bošković, Željko & Franks, Steven & Snyder, William (eds.), *Proceedings of the Sixth Annual Workshop on Formal Approaches to Slavic Linguistics*, 166–190. Stanford: CSLI Publications.
- Källström, Roger. 1993. *Kongruens i svenskan [Agreement in Swedish]*. Göteborg: Acta Universitatis Gothoburgensis.
- Kayne, Richard S. 2005. *Movement and silence*. New York: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195179163.001.0001>
- King, Tracy Holloway. 1995. *Configuring topic and focus in Russian*. Stanford, CA: Stanford University dissertation.

- Kornfilt, Jaklin. 1984. *Case marking, agreement, and empty categories in Turkish*. Cambridge, MA: Harvard University dissertation.
- Letuchiy, Alexander. 2017. *Predicative*. Ms. Moscow.
- Longobardi, Giuseppe. 1994. Reference and proper names. *Linguistic Inquiry* 25. 609–665.
- Markman, Vita. 2008. Pronominal copula constructions are what? Reduced specificational pseudo-clefts. In Chang, Charles & Haynie, Hannah (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 366–374. Citeseer.
- McShane, Marjorie. 2009. Subject ellipsis in Russian and Polish. *Studia Linguistica* 63(1). 98–132. DOI: <https://doi.org/10.1111/j.1467-9582.2008.01155.x>
- Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*. Linguistik Aktuell, vol. 85. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/la.85>
- Milenković, Aljoša. 2022. (Hyper)raising in Serbian. Ms. University of Vienna.
- Moro, Andrea. 1997. *The raising of predicates: predicative noun phrases and the theory of clause structure*. Cambridge; New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511519956>
- Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. In Groenendijk, Jeroen (ed.), *Studies in discourse representation theory and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris. DOI: <https://doi.org/10.1515/9783112420027-006>
- Polinsky, Maria & Potsdam, Eric. 2014. Left edge topics in Russian and the processing of anaphoric dependencies. *Journal of Linguistics* 50(3). 627–669. DOI: <https://doi.org/10.1017/S0022226714000188>
- Reinhart, Tanya. 1981. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica* 27(1). 53–94. DOI: <https://doi.org/10.21825/philosophica.82606>
- Say, Sergey. 2013. On the nature of dative arguments in Russian constructions with “predicatives”. In Kor Chanine, Irina (ed.), *Current studies in Slavic linguistics*, 225–245. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/slcs.146.13say>
- Seres, Daria & Espinal, Maria Teresa. 2019. Russian definitional generic sentences. *Glossa: a journal of general linguistics* 4(1). 59. DOI: <https://doi.org/10.5334/gjgl.760>
- Timberlake, Alan. 2004. *A Reference Grammar of Russian*. Cambridge: Cambridge University Press.
- Titov, Elena. 2017. The canonical order of Russian objects. *Linguistic Inquiry* 48(3). 427–457. DOI: [https://doi.org/10.1162/ling\\_a\\_00249](https://doi.org/10.1162/ling_a_00249)
- Titov, Elena. 2018. *The OVS construction*. Ms. University College London.
- Torrego, Esther. 1996. On the nature of clitic doubling. In Campos, Héctor & Kempchinsky, Paula (eds.), *Evolution and Revolution in Linguistic Theory*, 399–418. Washington: Georgetown University Press.
- Tsedryk, Egor. 2015. Deriving null pronouns: A unified analysis of subject drop in Russian. In *Formal Approaches to Slavic Linguistics* 23. 342–361.

Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26(1). 79–123.

Uriagereka, Juan. 2005. On the syntax of doubling. In Heggie, Lorie & Ordóñez, Fernando (eds.), *Clitic and Affix Combinations: Theoretical Perspectives*, 343–374. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/la.74.14uri>

Vat, Jan. 1997. Left dislocation, connectedness and reconstruction. In Anagnostopoulou, Elena & van Riemsdijk, Henk & Zwarts, Frans (eds.), *Materials on Left Dislocation*, 67–92, Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/la.14.07vat>

Zimmerling, Anton. 2003. Predikativy i kačestvennyye narečija: klassy slov i napravlenija derivacii. In *Rusistika na poroge XXI veka: problemy i perspektivy*, 54–59. Moscow: Jazyki slavjanskoj kul'tury.

Zolotova, G. A. 1982. O kategorii ocenki v russkom jazyke. In *Kommunikativnye aspekty russkogo jazyka*, 274–281. Moscow: Nauka.

