Russian èto, predication, and big DPs

'Irina Buruki (irine.burukina@nytud.mta.hu)
'Lena Borise (borise@fas.harvard.edu)
'Marcel Den Dikken (dmarcel@nytud.hu)

("'MTA Research Institute for Linguistics, "Eötvös Loránd University, "Harvard University)

1 Introduction

The data under consideration:

Èto copular constructions in Russian of the type ‘(NOM1) èto NOM2’.

(1) a. ‘Èto NOM2’

Èto Ø/byl moj brat.
this.N.SG is.M.SG/was.M.SG my brother.M.SG.NOM
‘This is/was my brother.’

b. ‘NOM1 èto NOM2’

Petja èto Ø/byl moj brat.
Petja.NOM this.N.SG is.M.SG/was.M.SG my brother.M.SG.NOM
‘Petja is/was my brother.’

 Èto = an invariant proximal demonstrative, neuter singular form.

NOM1, NOM2 = nominative DPs.

The copula byt’ ‘be’ (zero in the present tense) must agree with NOM2 (see Section 2 below).

Èto always precedes the copula (if overt, as in past or future tenses), which, in turn, is followed by NOM2. Another nominative DP (NOM1) may optionally precede èto.

Based on the classification in Higgins 1973, ‘èto NOM2’ constructions are identificational ones and are used to answer the question ‘Who is/was that?’, while ‘NOM1 èto NOM2’ constructions are those that have an equative/identity reading and are often used to answer such questions as ‘(And) who is/was Petja?’. Other copular contexts in Russian (i.e. specification and predication) prohibit the use of èto.

1 In some varieties of Russian the present tense form of the copula is jest’, homonymous with the present tense form of the existential verb byt’; thus, occasionally, examples similar to those in (i) can be found. In today’s language, they sound archaic.

(i) a. Sxodstvo èto jest’ real’nost’.
similarity.NOM this is reality.NOM
‘Similarity is the reality.’

b. Xristos jest’ bog i spasitel’.
Christ.NOM is god.NOM and savior.NOM
‘Christ is God and Savior.’ [National Corpus of Russian]

2 See the Appendix for a more detailed discussion of the semantics of èto copular constructions.
Our proposal. Part I.

In such copular constructions èto plays the role of a predicate, while NOM2 is the subject of predication.3 Èto undergoes movement to Spec,TP (predicate inversion).

\[ [_{TP \ èto}, [_{T} \ \{be+R^0\} \ [_{SC \ NOM2} \ [_{R} \ R^0 \ t]]]] \]

Our proposal. Part II.

NOM1, when present, forms a constituent with èto – a ‘big DP’: èto in D^0 and NOM1 in SpecDP.4 Information-structurally, èto and (when present) NOM1 in the constructions under discussion are a clause-internal topic (expressing given information).

\[ [_{DP \ NOM1} \ [_{D'} \ [D^0=\èto] \ [_{NP \ pro}] ProgressiveText]] \]

What we are NOT proposing.

We argue against:

- analyzing èto in ‘(NOM1) èto NOM2’ constructions (especially in the absence of NOM1) as the subject of predication.
- analyzing èto in ‘(NOM1) èto NOM2’ constructions as a dedicated (phrasal) element that has an information-structural function; cf. King (1995) on èto as an element in Spec, FP, Junghanns (1997) on èto as a base-generated internal topic pronoun that adjoins to AgrS.
- analyzing èto as a dedicated functional head on the clausal spine; cf. Geist & Błaszczak (2000), Markman (2008), i.a.
- analyzing NOM1 is all ‘(NOM1) èto NOM2’ constructions as a hanging topic/dislocated element;5 cf. Geist & Błaszczak (2000).

The ‘big DP’ approach further allows us to establish a link between èto copular constructions with a nominal subject and clausal prolepsis.

---

3 Cf. Moro (1997) on English it. The idea that èto is predicative in nature has surfaced in the literature before. According to Geist (2008), equative/identity and identificational copular clauses in Russian are underlyingly predicative, with the predicative reading stemming from the semantic interpretation of èto. Syntactically, Geist (2008) adopts the analysis developed in Junghanns (1997) and Geist & Błaszczak (2000), whereby èto is a clause-internal base-generated topic, and NOM1 is an element dislocated into the left periphery.


5 We argue that in ‘NOM1 èto NOM2’ constructions NOM1 and èto can form a constituent, namely, a big DP; however, we do not dismiss the hanging topic analysis completely. In certain cases, such sentences may have a different structure where NOM1 is merged in the topic position independently of èto. ‘Big DP’ clauses and ‘hanging topic’ clauses are easy to distinguish in complex sentences, since only the former can appear embedded under a non-bridge predicate (see the examples in (16)). In addition to this, the two constructions may have different information-structural and prosodic properties: hanging topic constructions appear to favor contrastive contexts.
2 ‘èto NOM2’ constructions

2.1 Outline of the analysis

In ‘èto NOM2’ constructions èto is the underlying predicate, merged in the complement of the RELATOR head (Den Dikken 2006), and NOM2 is the subject.

(2) \[TP \text{be} [\text{sc} \ \text{NOM2} [\text{R}^0 \ \text{èto}]]]\]
Èto itself is a D\(^0\), with an empty specifier and a silent pronoun as the complement NP:\(^6\)

(3) \[\text{DP} [\text{D}^0=\text{èto}] [\text{NP} \text{pro}] \]

In the course of the syntactic derivation, èto is moved to Spec,TP via predicate inversion (Moro 1997, Den Dikken 2006); an overt copula (in tenses that require an overt copula) is in T\(^0\).

(4) \[\text{TP} \ \text{èto} [\text{T} \ {\text{be}+\text{R}^0} [\text{sc} \ \text{NOM2} [\text{R}^0 \ \text{t}]]]\]

Predicate inversion is obligatory (cf. (1a)):

(5) *Moj brat byl / bylo èto.

my brother.NOM was.M was.N this

Motivation for the predicate inversion: since èto always corresponds to given information, the èto phrase must move to Spec,TP, a position that commonly hosts presupposed/referential material in Russian (cf. Bailyn 2004, Titov 2018). This movement allows NOM2, the new information, to occupy the clause-final position, associated, in Russian, with identificational focus (cf. Pereltsvaig 2004).

2.2 Empirical support

Support for èto being a predicate pro-form.

(6) A: I think that Isabella of France was [the famous wife of Henry VI].

B: Net, èto / *ona byla Margarita Anžujskaja.

No, this / she was Margaret of Anjou

‘No. This (= the wife of Henry VI) was Margaret of Anjou.’

In (6B), the personal pronoun ona ‘she’ cannot be used to refer back to the wife of Henry VI, the predicate of the preamble (6A). Only èto can take a predicate as its antecedent – personal pronouns cannot.

The same picture obtains when the èto kasaetsja... ‘as for...’ test is used to refer back to topics (Reinhart 1982:60, Geist 2008:97): referential expressions can be resumed by personal pronouns, while predicatively interpreted nominal phrases are referred back to with èto:

\(^6\) A reviewer points out that DPs, as schematized in (3), cannot be used predicatively without further assumptions, since the presence of D\(^0\) makes them referential; instead, only NPs can be used predicatively. We follow Partee (1986, 1998) and Geist (2008) and take the operator ident to be responsible for turning the èto-containing DP into one interpreted predicatively. See the Appendix for a discussion of the semantic properties of èto constructions.
(7) (‘Benedict Arnold is a traitor.’)
   a. Čto kasaetsja Benedikta Arnol’da,
      what concerns Benedict Arnold
      ja dumaju, čto *èto/ on predatel’.
      I think that this he traitor
      ‘As for Benedict Arnold, I think that he is the traitor.’
   b. Čto kasaetsja predatelja,
      what concerns traitor
      ja dumaju, čto èto/ *on Benedikt Arnol’d.
      I think that this he Benedict Arnold
      ‘As for the traitor, I think that it is Benedict Arnold.’

The analysis accounts for the following properties of èto copular constructions.

► NOM2 must be nominative and control the agreement. ← As the subject of predication, NOM2 is probed by T⁰ (in a similar way T⁰ agrees downwards with the notional subject in OVS clauses; see Pereltsvaig 2019); cf. also Geist (2008:90).

(8) Èto *bylo / byl moj brat.
    this.N.SG was.N.SG was.M.SG my brother.M.SG.NOM
    ‘This was my brother.’

► Èto cannot appear in predicative copular constructions or in specificational constructions of the kind ‘It/this/that èto NOM’. ← There can only be one (main) predicate per clause.

(9) *Ono / èto / to èto moj brat.
    it this that this my brother.NOM
    Intended: ‘This is my brother.’

► NOM2 cannot refer to a property (cf. also Geist 2008) ← Nominals denoting professions or types of temper are disallowed as NOM2’s, i.e. the subject of predication. Furthermore, they cannot function as the predicate, since that position is already occupied by èto.

(10) a. Petja (po professii) – *(èto) santexnik.
    Petja.NOM by job this plumber.NOM
    ‘Petja is a plumber (by trade).’
   b. Petja (po skladu xaraktera) – *(èto) xolerik.
    Petja.NOM by temper this choleric.NOM
    ‘Petja is a choleric (by nature).’
3 ‘NOM1 èto NOM2’ constructions

Constructions of the type ‘NOM1 èto NOM2’ have the structure parallel to that in (4/5), with one exception:

NOM1 and èto form a ‘big DP’: the demonstrative is in D⁰, NOM1 is merged in Spec,DP, and a silent pro sits in the complement of D⁰.⁷

(11) \[DP \text{NOM1} \left[ D' \left[ D^0=\text{èto} \right] [NP \text{pro}] \right] \]

This analysis accounts for the following properties of the èto constructions under discussion.

► The distribution of the [NOM1 èto] unit is limited: it is unavailable in fragment answers and coordinate structures. ← In Russian ‘big DPs’ the complement of D⁰ must always be silent (pro). The pro must be licensed by being in a local relationship with a structurally present, non-elliptical finite T⁰.⁸

(12) \[DP \left[Tvoj muž\right] \left[D' \text{èto} [NP \text{pro}]\right] \text{(byl)} \text{Petja.}

‘Your husband.\text{NOM} this was Petja’

(13) *\[DP \left[Tvoj muž\right] \left[D' \text{èto} [NP \text{pro}]\right] i \text{moj lučšij drug èto (byl)} \text{Petja.}

‘Your husband.\text{NOM} this and my best friend.\text{NOM} this was P.\text{NOM}

‘Your husband and my best friend is (was) Petja.’

► NOM1 never controls agreement. ← NOM1 is not itself in a relationship with T⁰, NOM2 is the notional subject of predication.


‘My plagues.\text{NOM} Egyptian this was M/were/was.\text{N} Petja.\text{NOM}

‘My plagues of Egypt (i.e. my bane), that was Petja.’

b. \[TP \left[DP \left[Moi kazni egipetskije\right], \left[D' \left[D^0=\text{èto} \right] [NP \text{pro}]\right]\right] \left[T \text{byl} \left[SC \text{[Petja]} \left[R' \text{R}^0 \text{t} \right]\right]\right]\]

► ‘NOM1 èto NOM2’ can readily be embedded under non-bridge predicates (e.g. udivitel’no ‘surprising’). ← NOM1 is not a hanging topic, but a part of the [NOM1 èto] constituent occupying Spec,TP.

7 Compare this behavior, for instance, to that of the ‘big DPs’ in the Romance languages, formed by a referential nominal phrase/pronoun in Spec,DP and a clitic in D⁰ (Torrego 1985, Uriagereka 1995, i.a.). This is as shown below (for ‘weak’ clitics) before the clitic moving out of the ‘big DP’ in (a) and after movement in (b):

(i) a. \[DP\text{-noun phrase/pronoun} \left[D' \text{clitic} [NP \text{pro}] \right] \]

b. \[Lo, \text{vi} \left[DP \text{a} \text{el} \left[D' t, \text{[NP \text{pro}]\right]\right.\right.

he(CL) saw PREP him

‘I saw him.’ (Uriagereka 1995)

8 We assume that, although T⁰ must be structurally present and unaffected by ellipsis, it does not have to be overt. Hence, èto copular constructions are grammatical in the present tense where the copula is silent.
Udivitel’no ‘surprising’ is a non-bridge predicate (Polinsky & Potsdam 2014):

(15) a. Udivitel’no, [čto Petja poceloval Mašu].
    surprising that Petja kissed Maša.ACC
    ‘It is surprising that Petja kissed Maša.’

b. *Mašu udivitel’no, [čto Petja poceloval t].
    Maša.ACC surprising that Petja kissed
    ← A-bar extraction is impossible

c. ?Maša, udivitel’no, [čto Petja poceloval ejo].
    Maša,NOM surprising that Petja kissed her.ACC
    ← a hanging topic with resumption is allowed

‘NOM1 ḥto NOM2’ can be embedded under udivitel’no ‘surprising’:

(16) Udivitel’no, [čto železnyj čelovek – ḥto Toni].
    surprising that iron man.NOM this Tony.NOM
    ‘It is surprising that the Iron Man is Tony.’

► NOM1 ḥto NOM2 constructions are incompatible with the question particle li in embedded interrogative clauses. ← [NOM1 ḥto] is raised to Spec,TP. The question particle li in interrogative clauses obligatorily manifests the C head and must be preceded by the focused constituent (King 1995). Li cannot break into the [NOM1 ḥto] unit; moving NOM1 to Spec,CP via subextraction out of this unit is a violation of criterial freezing. In addition to this, since NOM1 is often not focal, in sentences such as (17) there would be an information-structural incongruity resulting from movement of NOM1 into the pre-li focus position.

(17) Mama sprosila [železnyj čelovek li (*ḥto) Toni].
    Mum asked iron man.NOM Q this Tony.NOM
    ‘Mum asked whether the Iron Man is Toni.’

► ḥto is ungrammatical in wh-questions. ← NOM1 (kto in (18)) can neither be a hanging topic (for information-structural reasons) nor be raised to Spec,CP via subextraction from [NOM1 ḥto] in Spec,TP (because of freezing).9

9 A reviewer points out that ‘Who’s Tony?’ is more naturally translated into Russian as Kto takoj Toni? as opposed to Kto Toni? (the example in (18) with ḥto omitted), and wonders about why takoj is obligatory in such questions. While a detailed discussion of this issue lies beyond the scope of this paper, we would like to suggest the following. First, we partially agree with the reviewer on the judgement. Takoj is preferable in identity questions, however, it cannot be used if an equative clause is expected for an answer; cf. (i) where an appropriate question corresponding to the B2 answer would be Kto Superman?

(i) A: Kto takoj Superman?
    who such Superman.NOM
    ‘Who is Superman?’
    B1: Superman is a DC superhero from the planet Krypton.
    B2: Superman is Clark Kent.

Second, it is our understanding that, in Kto takoj Toni?, takoj modifies the interrogative pronoun forming with it a constituent; note, for instance, that takoj agrees with the question word in gender and number (ii).

(ii) Čto takoj / *takaja Moskva?
4 Prosodic properties of ‘(NOM1) èto NOM2’ constructions

It is commonly assumed that there is an obligatory intonational break or pause between NOM1 and èto, which is indicated by a dash in writing. In speech, however, the prosodic realization of ‘(NOM1) èto NOM2’ constructions is considerably more varied.

In some ‘(NOM1) èto NOM2’ examples, there is evidence for the left edge of a prosodic phrase being aligned with the left edge of èto. It may be realized as a pause or as strong glottal onset on the initial vowel in èto. These are shown in Figures 1 and 2, respectively, with the examples provided in (19).

Note that a prominent pause is only available in strongly emphatic contexts.

(19) a. Mosgorduma — èto tot organ, kotoryj otrazet politicheskoe
    Moscow City Duma this that body which reflects political
    predstavitel'stvo moskvi'cij.
    representation Moscovites.GEN
    ‘The Moscow City Duma is the body that reflects the political representation of
    Moscovites.’ (navalny_2013_045)

b. Minskij protokol ot p'iatogo sentjabrja — èto ramochnaja bumaga.
    Minsk. ADJ protocol from fifth.GEN September.GEN this framing paper
    ‘The Minsk protocol of the fifth of September is a framing document.’
    (lavrov_2015_073)
On the other hand, examples without any pause or evidence for a prosodic boundary are common too:

(20) a. Den’gi – ėto vozmožnost’ ix potratit’.
    money.PL this opportunity they.ACC spend
    ‘Money is the opportunity to spend it.’ (miloserdov_037)

    b. Samaja problema – ėto kadry.
    most problem this human_resources
    ‘The biggest problem is human resources.’ (tv-news_030)
The prosodic break before èto is particularly prominent in contexts that involve **contrast**, some of which also contain **ellipsis**. Here, though, the presence of the pause is likely to be attributable to the presence of contrast and/or ellipsis rather than the nature of the ‘(NOM1) èto NOM2’ construction itself.
(21) (‘Here, two lines are shown. The first one is a predicted curve,’)

\[
\text{drugaja} - \text{èto} \quad \text{srednee} \quad \text{predskazanie} \quad \text{buduščego} \quad \text{razvitija} \quad \text{čislennosti} \\
\text{other} \quad \text{this} \quad \text{mean} \quad \text{estimate} \quad \text{future} \quad \text{development} \quad \text{head-count} \\
\text{čelovečestva.} \\
\text{human_population} \\
\text{‘the other is the estimate of the future growth of human population.’} \\
(\text{academ_rossija_demograf_147})
\]

\[\Rightarrow\] It is not the case that all ‘NOM1 èto NOM2’ contexts present strong evidence for a prosodic break (pause or glottal onset) between NOM1 and èto. The break is often found in emphatic contexts and those involving contrast (which may be accompanied by ellipsis), whereas non-emphatic, non-contrastive, neutral contexts often do not include a prosodic break. The idea that the prosodic break is obligatory in ‘NOM1 èto NOM2’ contexts may be rooted in Russian \textit{orthography}: a dash is mandatory in ‘NOM1 èto NOM2’ contexts between NOM1 and èto, and the common rationale for this rule is that the dash marks a pause.

Orthography may be misleading. In equative constructions without èto, the orthographic rule is to use a dash in positive statements (\textit{Džek – sobaka}, ‘Jack [is a] dog’), but not in the presence of negation (\textit{Džek ne sobaka}, ‘Jack [is] not [a] dog’). Prosodically, these statements are identical, which means that the presence of the dash is not correlated with prosody.

5 \quad \textbf{Implications}

5.1 \quad \textit{Èto} copular constructions and clausal prolepsis

In other contexts the ‘big DP’ may be an argument, like any other DP; this is not precluded by the analysis.

10
(22) a. [Čto my guljaem v parke], èto prijatno. 
    that we walk in park this nice
    ‘It is nice that we are walking in the park.’ (= b)

b. Èto prijatno, [čto my guljaem v parke].
    this nice that we walk in park
    ‘It is nice that we are walking in the park.’ (= a)

In (22a), the embedded CP and èto together form the big DP that serves as the subject of *prijatno* ‘nice’; èto is in D⁰ and the subordinate clause in Spec,DP (23a).

In (22b), the CP is ‘extraposed’; èto is a proleptic pronoun and the CP binds the pro in the complement of D⁰ from its surface clause-final position, establishing the interpretive link between CP and the matrix clause subject thereby (23b).

(23) a. [DP [CP čto my guljaem v parke] [D, D⁰=èto [NP pro]]]

b. [DP [D, D⁰=èto [NP pro]]] … [CP èto my guljaem v parke]

5.2 Èto copular constructions and èto focus constructions

Under the proposed ‘big DP’ analysis, the èto of ‘NOM1 èto NOM2’ constructions is a D⁰.

Cf. èto focus constructions, where èto is the Top head connecting together a question CP (in Spec,TopP) and an answer IP with the focused XP in it (in Comp,TopP); cf. Markman (2008), Burukina & den Dikken (2020).

(24) a. Kogo Petja priglasil (tak) èto Mašu.
    who.ACC Petja invited PTCL this Maša.ACC
    ‘The person whom Petja invited is Maša.’

b. [Top [CP Kogo Petja priglasil] [Top, Top⁰=èto] [IP Mašu, Petja priglasil t.]]

An important parallel between the construal of èto in ‘NOM1 èto NOM2’ constructions and the function of èto as a Top-head: In both, èto is a functional head mediating a relationship between two terms that are in a semantic co-construal relationship: NOM1 and pro in the former, and the topic and the comment in the latter.

→ This establishes a fundamental parallel between these two approaches to èto that have generally been viewed as irreconcilable.

6 Summary

We have shown that:

- In ‘èto NOM2’ constructions èto is the underlying predicate, merged in the complement of the RELATOR head and raised via predicate inversion.

- In ‘NOM1 èto NOM2’ constructions, NOM1 and èto form a ‘big DP’: the demonstrative is in D⁰, NOM1 is merged in Spec,DP, and a silent pro sits in the complement of D⁰.

11
The current approach to the status of èto allows us to make parallels between èto copular constructions and both èto focus constructions and clausal prolepsis.

While a detailed investigation of the prosodic properties of ‘NOM1 èto NOM2’ constructions is required, the examples from the Russian National Corpus show that a pronounced prosodic break at the left edge of èto is a possible but not a necessary feature of the construction.

References:
Appendix. On the semantics of èto copular constructions

Higgins (1973) distinguished between the following four types of copular constructions: identificational, equative/identity, specificational, and predicative. As mentioned in Section 1, ‘èto NOM2’ are identificational and ‘NOM1 èto NOM2’ constructions have an equative/identity reading;¹¹ èto cannot be used in other copular contexts in Russian.

(25) Identificational
Èto moj brat.
this.N.SG my brother.M.SG.NOM

‘This is my brother.’

(26) Equative/identity
Antosha Chekhonte this AntonChekhov

‘Antosha Chekhonte is Anton Chekhov.’

b. Petja – *(èto) moj brat.
Petja.NOM this.N.SG my brother.M.SG.NOM

‘Petja is my brother.’

(27) Specificational
Predatel’ – (*èto) Benedikt Arnol’d.
traitor this Benedict Arnol’d

‘The traitor is Benedict Arnold.’

¹¹ Èto should be omitted on the predicative reading of moj brat ‘my brother’ but it is obligatory on the equative one. This can be shown with the help of the non-restrictive relative clause test (non-restrictive relatives can only modify referential nominals – i.e., not those interpreted predicatively; Heycock & Kroch 1999). A non-restrictive relative sounds more natural in the presence of èto.

(i) a. ??Petja – moj drug, kotorogo ty tože xorošo znaeš.
Petja my friend which you too well know

‘Petja is my friend, whom you also know well.’

b. Petja – èto moj drug, kotorogo ty tože xorošo znaeš.
Petja this my friend which you too well know

‘Petja is my friend, whom you also know well.’
As argued by Geist (2008), equative/identity and identificational copular clauses in Russian are underlingly predicative, due to the semantic interpretation of èto. The proposal is a adaptation of Partee’s (1986, 1998) approach to identification copular clauses, according to which a special operator, ident, shifts the reading of a post-copular nominal in identification copular clauses to a predicative one, in order to ensure that no special semantics of the copula is required for deriving equative/identity clauses:

Geist (2008) suggests that, in Russian ‘(NOM1) èto NOM2’, ident applies to èto instead, which grants èto a predicative reading. Co-reference between NOM1 and èto results from co-indexing.

Information-structurally, both èto and (when present) NOM1 in the constructions under discussion are topics (expressing given information).