Vadim Kimmelman, Russian State University for the Humanities[*]

On the Interpretation of èto in so-called èto-clefts

1. Introduction

In Russian, there is a construction with the pronoun *èto* in the sentence-initial position, the so-called *èto*-cleft construction[1]. There are two types of constructions with *èto*: a sentence with contrastive focus (1) (focus *èto*-cleft), and thetic sentence (2) (thetic *èto*-cleft).

(1) Okno   razbil   Vasja?– Net, èto **Petja**[2] (razbil okno).
     window   break   Vasja   no èto Petja broke window
     'Did Vasja break the window? No, it's Pete who broke the window.'

(2) Čto za   šum, čto   slučilos'? – Èto Petja razbil **okno**.
     what prep. noise what happen   èto Petja break   window
     'What's the noise, what happened? Pete broke the window.'

There are two problems for the analysis of these constructions:
Problem 1: What is the syntactic position of *èto* (in generative framework), and is it the same for both (1) and (2)?
Problem 2: What is the semantic status of *èto*, and is it the same for both (1) and (2)?

The first problem has been discussed by King (1993) and Junghanns (1997). King (1993) proposes that èto functions as a specifier of a focus phrase, and that the argument of the focus phrase is the constituent which is focused and therefore moved to the position adjacent to *èto*. Junghanns (1997) provides arguments against this analysis: (i) the focused constituent need not be adjacent to èto (3), (ii) the whole IP can be focused (as in thetic sentences), and (iii) the syntactic status of the moved constituent is not clear.

(3) Petja razbil stakan? – Net, èto Petja **okno**   razbil.
     Petja break   glass   No èto Petja window   break
     'Did Pete break the glass? No, Pete broke the window.'

Junghanns (1997) analyses both thetic and focus èto-clefts by considering *èto* as a base-generated topic, which is merged as a higher adjunct to AgrSP. The argument for this analysis is the fact that *èto* appears in a position lower than that of specifier of CP but higher than anything else in the clause (4). Junghanns proposes that *èto* is not a focus-marker, but it "connect[s] sentences through taking on the function of topic" (Junghanns 1997:182). Focus appears in every sentence, and doesn't depend on the presence of *èto*. Junghanns' analysis also offers a uniformity of syntactic structure for focus èto-clefts and thetic èto-clefts.

(4) Ja znaju, [CP čto [èto [AgrSP Petja   razbil okno]]].
     I know [CP that [èto [AgrSP Petja   break   window]]].
     'I know that it's Pete who broke the window.'

---

[1] It's called "*èto*-cleft" because it resembles the cleft construction found in English and many other languages (but only resembles, as we will see).
[2] **Bold** marks the word with the main phrase (pitch) accent.

This analysis is, however, semantically inadequate for the reasons which will be considered in this paper. A semantic analysis of various constructions with *èto* (not only of the so-called *èto*-clefts) was proposed by Padučeva (1982). Our analysis hinges on Paducheva's proposals but it is cast in a formal framework in the spirit of Rooth (1992).

2. Analysis
2.1 Focus *Èto*-cleft.
2.1.1 Background.

Focus analysis (after Rooth (1985, 1992) and Kadmon (2001)) uses the following definitions and notations:

$[\![S]\!]^{M,g}$ – the ordinary semantic value of a formula S ($=[\![S]\!]^o$).
$[\![S]\!]^{f\text{-}M,g}$, the focus semantic value of a formula S relative to an arbitrary model M and an arbitrary assignment function g, is the set of all ordinary semantic values obtained by replacing the F-marked subformulas in S with variables, and then interpreting the result relative to every assignment g' which is identical to g except perhaps for the values of those variables.

(Kadmon 2001:295-298)

For example, examine the following question-answer pair:

(5)   A: Who broke the window?
       B: [Pete]$_F$ broke the window.

Here the constituent *Pete* is F-marked, so the focus interpretation is as in (6):

(6) $[\![$[Pete]$_F$ broke the window$]\!]^{f\text{-}M,g} = [\![$x broke the window$]\!]^{M,g} = \{[\![$Pete broke the window$]\!]^{M,g}$, $[\![$Mike broke the window$]\!]^{M,g}$, $[\![$John broke the window$]\!]^{M,g}$, …}

This analysis fits the intuition that focus is the answer to a wh-question.

There are so-called focus-sensitive items, for example, the word *only*, which is intuitively connected with focus:

(7)   A: Who broke the window?
       B: Only Pete broke the window.

(8)   A: What did Pete break?
       B: Pete broke only the window.

The first idea is to consider that the set that *only* is restricting is always the set created by replacing the focus-marked constituent with a variable.

(9)   If *only*$[\![S]\!]$ is of type t, then for all models M, w∈W$_M$, w∈$[\![$*only*$[\![S]\!]]\!]^{M,g}$ iff for all p∈$[\![S]\!]^{f\text{-}M,g}$, if w∈p then p=$[\![S]\!]^{M,g}$ ($[\![S]\!]^o$)

Informally, this means that of all alternatives which are created by replacing the F-marked constituent the only alternative which is true is expressed in the sentence with *only*.

But in reality there are some cases when *only* is not connected with focus:

(10)   Father: Is it true that of all things Mother asked you to do at her ball you only managed to introduce Steve to Sue?
       John: Not exactly… [grin] I only introduced [Bill]$_F$ to Sue.

(Kadmon 2001:316)

Here we can see that *only* is not connected with focus. So we have to propose the meaning of *only* as in (11):

(11) If only⟦S⟧ is of type t, then for all w∈W, w∈⟦only⟦S⟧⟧^{M,g} iff for all p∈C, if w∈p then p=⟦S⟧^{M,g}, where C is a contextually determined set of relevant propositions.

To explain this not obligatory ability of focus to be connected with *only* Rooth proposes the *Focus Interpretational Principle* (FIP):

(12) A constituent may be focused
iff
it is contained in a clause S of which the following holds:
somewhere in the environment there is an (independently) salient set of propositions – call it Γ – such that
      (i)   Γ⊆ ⟦S⟧^{f} and
      (ii)   ⟦S⟧^{o}∈Γ and
      (iii)   there is some proposition p s.t. p≠⟦S⟧^{o} and p∈Γ
                                     [from Kadmon (2001: 317-318)]

Consider the following examples:

(13) John only introduced [Bill]$_F$ to Sue.
    = {w: for all p∈C, if w∈p then p=⟦introduce (b)(s)(j)⟧}

The set C (the domain for *only*) is a natural candidate to be Γ. C should be such as:
{'John introduced Bill to Sue', 'John introduced Kim to Sue', 'John introduced Lee to Sue' …}

(14) Father: Is it true that of all things mother asked you to do at her ball you only managed to introduce Steve to Sue?
    John: Not exactly… I only introduced [Bill]$_F$ to Sue.

The domain of quantification for *only* is like {'John only introduced Bill to Sue', 'John politely chatted with aunt Betty', 'John danced with the neighbor's daughters', 'John helped to supervise the evening's program',…}. The FIP can not be satisfied by C (C cannot be Γ), because C⊈ ⟦S⟧^{f}.

    But the ordinary semantic value of "John only introduced Bill to Sue" is obviously contrasted in this context with the proposition "John only introduced Steve to Sue". The explicit binary contrast immediately makes salient a "comparison class" (CC) within which the two propositions are contrasted. It serves as the set Γ required by the FIP.

(15) ⟦I only introduced [Bill]$_F$ to Sue⟧^{f}= {⟦only[introduce(x)(s)(j)]⟧}

(16) CC∈⟦I only introduced [Bill]$_F$ to Sue⟧^{f}; [I only introduced Bill to Sue]∈CC (FIP)

What is important for us in this analysis is the fact that items which are intuitively connected with focus are not obligatorily connected with it.

2.1.2 Focus Èto-cleft – not Focus, but Contrast.

We propose that focus *èto*-cleft expresses not only focus, but also a contrastive meaning. The reason for this is that there are some tests which show the difference between focus and contrast in similar sentences.

1. Focus *èto*-cleft can not serve as an answer to a simple question, while the sentence without *èto* can serve as such an answer:

(17) [Teacher in the class]
    A: Kto segodnja dežurnyj?
        who today   on.duty
    'Who is on duty today?'
    B:Vasja.
    B': #Èto Vasja.

But it is a good utterance when there is another <u>alternative</u> in the pretext:

(18) A: Vasja segodnja dežurnyj.
        Vasja today   on.duty
    'Vasja is on duty today.'
    B':Net, èto Petja segodnja dežurnyj.
        no èto Petja today   on.duty
    'No, Petja is on duty today.'

2. If there is a choice between <u>two alternatives</u> in context, a sentence without *èto* containing one of these alternatives is fully acceptable, but the same sentence with *èto* is hardly acceptable (test from Silva-Corvalán 1982):

(19) A: Kto   segodnja dežurnyj: Petja ili Vasja?
        who today   on.duty   Petja or Vasja
    'Who is on duty today: Petja or Vasja?'
    B: Vasja.
    B': #Èto Vasja.

3. If the relevant alternative is formulated as a <u>question</u> in the pretext, focus *èto*-cleft is not absolutely acceptable (test from Lambrecht (1994)):

(20) A: Vasja segodnja dežurnyj?
        Vasja today   on.duty
    'Is Vasja on duty today?'
    B: ?Net, èto Petja.
          no èto Petja
    'No, Petja is.'

4. Focus on the subject does not imply the presupposition of existence, but in sentences with *èto* there is such a presupposition:

(21) A:Kto   prišel?
        who come
    'Who came?'
    B: Nikto   (ne   prišel).
        nobody not   come
    'Nobody came.'

(22) A: Vasja razbil okno?

Vasja break window
    'Did Vasja break the window?'
    B: *Net, èto nikto   ne   razbival okna.
        no   èto nobody not break   window
    'No, nobody did.'

Consider the fact that it is not only impossible to use negative pronouns in focus *èto*-clefts as in (22), but quantifiers are also unacceptable:

(23) *Èto kto-to   razbil okno.
      èto somebody break window (quantifier of existence)

(24) *Èto každyj   razbil okno.
      èto   everybody break window (universal quantifier)

So we conclude that focus *èto*-clefts imply not only a presupposition of existence, but also some kind of the uniqueness presupposition (this term is used here in an informal fashion).

These tests show that focus *èto*-cleft is used only when there is an alternative in the context, contrasting with the *èto*-sentence. This alternative should be specified in the sense that it cannot be questioned or construed as one of two or more alternatives expressed in the context.

Actually, these restrictions form a scale of acceptability – an answer with *èto* is acceptable when the alternative is affirmative, less acceptable when the alternative is under question, even less acceptable when there are two or more alternatives and nearly unacceptable when there is no alternative in the context.

2.1.3 Formal Analysis.

A first idea could be the following: *èto* in this type of sentences is a contrast marker that makes a restriction on the alternative set Γ from Focus Interpretational Principle (FIP): now Γ should contain only two alternatives: the ordinary value of S and some other alternative which is specified in the context. This restriction accounts for the results of tests 1-3 in the previous section. The second restriction we propose is the requirement for a presupposition of existence and "uniqueness" concerning the ordinary value of S.

But there are some not absolutely perfect examples, where FOCUS and CONTRAST do not coincide. Consider the following dialog:

(25) Vasja usually leaves something at home: a textbook or a copybook. Today he comes to class and the teacher asks Masha (who is on duty today):
    T: Čto Vasja segodnja ne prinjos: učebnik ili tetrad'?
       what Vasja today   not bring   textbook or copybook
    'What hasn't Vasja brought today: a textbook or a copybook?'
    M: Èto [Petja]$_C$ ne prinjos [uchebnik]$_F$, a   Vasja vse   prinjos.
       Èto Petja   not bring   textbook   but Vasja everything bring
    'Petja hasn't brought the textbook, but Vasja brought everything.'

If we try to analyze (25) as an instance of focus, we face a difficulty: what is Γ here? We can see the set which produces contrast: the question was about Vasja, and the answer was about Petja. The focus value of the phrase with *èto* (if we forget about contrast) is the following:

(26) $[\![$ Èto Petja ne prinjos [uchebnik]$_F$ $]\!]^f =$ $[\![$ Èto Petja ne prinjos x $]\!]^o =$ {'Petja ne prinjos tetrad', 'Petja ne prinjos učebnik', 'Petja ne prinjos ručku'…}

Γ should be a subset of ⟦S⟧^{f}, but the natural candidate to be Γ, the semantic value of the question, is not a subset of ⟦S⟧^{f}:

(27) ⟦Čto Vasja segodnja ne prinjos: učebnik ili tetrad'⟧^{o}≈{'Vasja ne prinjos učebnik', 'Vasja ne prinjos tetrad'}⊈ {'Petja ne prinjos tetrad', 'Petja ne prinjos učebnik', 'Petja ne prinjos ručku'…}

Now we see that it is incorrect to say that *èto* makes restrictions on the set of focus alternatives.

To solve this problem we propose to view contrast as an independent mechanism and to treat *èto* as a marker which shows that there is a contrast-marked constituent in the sentence. If the constituent is contrast-marked, then there must be a specified alternative in the context. This alternative should be similar to the *èto*-cleft, except for the value of the contrast-marked constituent, and possibly except for the value of the focus-marked constituent. On the level of focus-value, the contrast-marked constituent must also be replaced with a variable.

In the case when contrast and focus-marked constituents coincide, the contrastive set (specified alternative and the ordinary semantic value of the sentence) is suitable to be Γ, and FIP is satisfied. But when contrast and focus do not coincide FIP should be changed:

Focus Interpretational Principle-2 (taking into account contrast):
(28) A constituent may be focused
   iff
   it is contained in a clause S of which the following holds:
   somewhere in the environment there is an (independently) salient set of propositions – call it Γ – such that
      (i)   Γ⊆ ⟦S⟧^{f} and
      (ii)   <u>There is an x such that if we replace the contrast-marked constituent in S with x to form S' then ⟦S'⟧∈Γ</u> and
      (iii)   there is some proposition p s.t. p≠⟦S⟧^{o} and p∈Γ

Now we can interpret example (25):

0. Contrast is produced by the fact that there is a specified alternative in the context: sentence "Čto Vasja segodnja ne prinjos: učebnik ili tetrad" which differs from the *èto*-cleft only by contrast-marked and focus-marked constituents.

1. ⟦Èto [Petja]C ne prinjos [uchebnik]F ⟧f = ⟦ y ne prinjos x ⟧o=
{'Petja ne prinjos tetrad', 'Petja ne prinjos učebnik', 'Petja ne prinjos ručku'…,
'Vasja ne prinjos tetrad', 'Vasja ne prinjos učebnik', 'Vasja ne prinjos ručku'…,
'Kolja ne prinjos tetrad', 'Kolja ne prinjos učebnik', 'Kolja ne prinjos ručku'…,
…}

2. Γ=⟦Čto Vasja segodnja ne prinjos: učebnik ili tetrad' ⟧o≈{'Vasja ne prinjos učebnik', 'Vasja ne prinjos tetrad'} ⊆ ⟦Èto [Petja]C ne prinjos [uchebnik]F ⟧f

3. There is x such that if we replace contrast-marked constituent with x then ⟦ S' ⟧∈Γ: x=Vasja:
Èto [Petja]C ne prinjos uchebnik⇒ Èto Vasja ne prinjos uchebnik∈Γ=≈{'Vasja ne prinjos učebnik', 'Vasja ne prinjos tetrad'}
<u>FIP is satisfied.</u>

So, what does *èto* do? It shows that the sentence with *èto* contains a contrast-marked consituent. But it is not the only function of *èto* in this case: it also requires the presupposition of existence and uniqueness. We can see that ordinary contrast does not require these presuppositions, but *èto*-cleft does (see examples 21-24). This fact needs further explanation. Here we can say that contrast *èto*-cleft is the mechanism to express simple contrast (mainly between entities: constituents of type e (nouns); or between properties: constituents of type e→t (verbs)).

Considering *èto*-clefts we can say that contrast strictly tends to coincide with focus[3]. But there are also so-called contrastive topics, which can be analyzed in the same way. Kadmon (2001: 388-390), analyzing contrastive topics, proposes a special topic-value, but it seems plausible to analyze all phenomena concerned with contrast in a uniform way (consider also "topic-focus" and "focus-focus" from Roberts 1996 and topic value from Büring 1999). Our procedure of contrast-markedness, as far as we understand, fits also for analysis of contrastive topics without introducing a new level of interpretation.

2.1.4 Other Cases.

It is possible to use contrast *èto*-cleft in a non-contrast context. Consider the following example:

(29) Voobše-to on Tolja Ivanov. Èto on učil menja kurit' v pervom klasse.
     In.fact   he Tolja Ivanov Èto he teach me   smoke in first   grade
     'In fact, he is Tolja Ivanov. He taught me to smoke in the first grade.'

As far as we understand, *èto*-cleft here is used as a construction of characterization: the fact that "he taught me to smoke" is characterizing "him". To say very informally, characterization of a subject and saying something about the subject in contrast to the context are semantically close: the information that, despite our expectation (expressed in the context), this person is doing something is a kind of characterization of this person. It doesn't seem too unnatural that these meanings can be expressed by the same construction. But this case will need more work and we have not really studied it yet.

2.2 Thetic *Èto*-cleft.

We call the second type of *èto*-cleft construction thetic because it can serve as an answer to the question: "What happened" – one usual test for thetic construction. In thetic sentences focus is on the whole proposition and it is impossible to divide the sentence into an overt topic and focus, or topic and comment, or theme and rheme, etc.

It is obvious that thetic *èto*-cleft is not the same as the focus (or contrast) *èto*-cleft, because there is no contrast in this case, and no specified alternative is implied in the context. Consider an example:

(30) Naša Tanja gromko plačet. Èto ona uronila v rečku **mjačik**.
     our Tanja loudly cry   èto she drop   in river ball
     'Our Tanja is crying loudly. She has dropped a ball into the river.'

Here *èto* seems to serve as a marker of the causal relation between the situation of crying and the situation of dropping the ball. If we use *èto*, the relation between the situation described by this sentence and some other situation in the context can only be causal. If we use the same sentence without *èto*, the relation can be different:

(31) Naša Tanja gromko plačet. Èto ona mamu zovjot.
     our Tanja loudly cry   èto she   mother call
     'Our Tanja is crying loudly. She is calling her mother' (She is crying to call her mother).

(32) Naša Tanja gromko plačet. Ona zovjot mamu.
     our   Tanja loudly cry   she call   mother

---

[3] By the way, in the example (14) focus does not coincide with the domain of quantification for *only*, but does coincide with contrast.

'Our Tanja is crying loudly. She is calling her mother' (First she is crying. Then she stops crying and starts calling her mother[4]).

It seems that *èto* here is a particle, not a pronoun as in (33); e.g. it cannot be stressed.

(33) Tanja boljna. Èto slyšal ja včera.
     Tanja is.ill Èto hear I yesterday
     'Tanja is ill. I heard it yesterday.'

But consider the following example:

(34) V sosednej komnate razdalsja šum. Èto Petja razbil okno.
     In next room there.was noise èto Petja break window
     'There was a noise in the next room. (It was because) Petja broke the window.'

Èto might be a kind of particle here, a marker of causal relation between this sentence and the preceding one. Within the framework of Erteschik-Shir (1997) it can be formulated in another way, as: *èto* is used when stage topics of both sentences are equivalent, and the second sentence simply adds information about the state of affairs concerning this stage topic.

(35) Razdalsja šum. – ÈtoPetja razbil okno,- obyasnil Vasja.
     there.was noise èto Petja break window explain Vasja
     'There was a noise in the next room. "Petja broke the window", explained Vasja'

The problem is the following: in (35) the situation of noise is in the context (in the larger text), but not in Vasja's words. This problem does not seem very hard, since we can say that it is not important whether the situation is expressed by someone's words or is present only in the narrative context. But consider the following:

(36) Situation (not text, the real world): Petja and Vasja are walking in the field. Then there is a noise. Vasja is surprised, but Petja says:
     P: Èto sobaki layut.
        Èto dogs bark
     'Dogs are barking.'

This usage seems to be obviously deictic – *èto* refers to (or relates to) the noise, to the situation which is not in the linguistic context, but in the real world. This construction is (again) obviously the same – *èto* can occupy only the sentence-initial position, and it can never be stressed. So, we have to conclude that *èto* is an anaphoric pronoun, and in (30) it is also pronoun, and it refers to the situation of preceding sentence. Moreover, this pronoun seems to be the topic of the sentence.
    This analysis also works for the following case:

(37) Zavtra v školu ne poydu – èto skazal Petja.
     tomorrow in school not go èto say Petja
     '"Tomorrow I won't go to the school" – said Petja.'

Here it seems to be a case of so-called textual deixis – *èto* seems to refer to the utterance/sentence – not to the situation, or, more precisely, to the situation of utterance of this sentence. This case would also constitute a problem for the analysis of *èto* as a particle.
    And if we analyze *èto* as a pronoun, it is possible to say that *èto* is a topic referring to the situation in the preceding context, so this kind of *èto*-cleft is not indeed thetic: there is a topic

---

[4] It is one of the possible interpretations of this sequence.

in it. So, the syntactic analysis proposed by Junghanns (1997) is semantically adequate in this case.

So, we can make the following conclusion: *èto* in thetic *èto*-cleft is a pronoun, and in focus *èto*-cleft it is a particle.

3. Conclusion

We have analyzed a very interesting Russian construction – the so-called *èto*-cleft – and have found out that there are (at least) two different constructions which are very similar on the syntactic level[5] but are semantically absolutely different:

1) Contrast *èto*-cleft, where *èto* is a particle marking the presence of contrast-marked constituent.
2) (Thetic) *èto*-cleft, where *èto* is a pronoun referring to the preceding situation and connected with the situation described in the sentence by causal relation.

In our analysis of contrast *èto*-cleft we have suggested a uniform approach to a formal account of different contrast phenomena, which can be a topic for further research.

References

Büring, Daniel (1999) "Topic." In P. Bosh and R. van der Sandt (eds.), *Focus – Linguistics, Cognitive, and Computational Perspectives,* Cambridge University Press, 142-165.

Erteschik-Shir, Nomi (1997) *The Dynamics of Focus Structure.* Cambridge; New York: Cambridge University Press.

Junghanns, Uwe (1997) "On the So-called èto-cleft Construction" In Martina Lindseth and Steven Franks (eds.) *Annual Workshop on Formal Approaches to Slavic Linguistics. The Indiana Meeting of 1996.* Ann Arbor, MI: Michigan Slavic Publications. Pp. 166-190.

Kadmon, Nirit (2001) *Formal Pragmatics: Semantics, Pragmatics, Presupposition, and Focus.* Blackwell Publishers

King, Tracy H. (1993*) Configuring Topic and Focus in Russian*. Doctoral dissertation. Stanford University.

Lambrecht, Knud (1994*) Informational Structure and Sentence Form*. Cambridge University Press

Padučeva E.V. (1982) "Značenie i sintaksičeskie funkcii slova èto." In *Problemy strukturnoj lingvistiki 1980*, ed. V.P. Grigor'ev, 76-91. Moskva: Nauka.

Roberts, Craige (1996) "Informational Structure in discourse: towards an integrated formal theory of pragmatics." In J. Hak Yoon and A. Kathol (eds.), *Ohio State University working Papers in Linguistics, volume 49.*

Rooth, Mats (1985) *Association with Focus.* PhD thesis, University of Massachusetts, Amherst.

Rooth, Mats (1992) "A theory of focus interpretation" *In Natural Language Semantics 1*, pp. 75-116.

Silva-Corvalán, Carmen (1982) "Subject expression and placement in Mexican-American Spanish." In J. Amastae & L. Elías-Olivares (eds.), *Spanish in the United States: sociolinguistic aspects*. Cambridge University Press.

---

[5] Actually there are also differences in the word order in these constructions, but that goes beyond the scope of this paper.