

Connecting Information Structure and Discourse Structure through “Kontrast”: The Case of Colloquial Russian Particles –*TO*, *ŽE*, and *VED’*

SVETLANA MCCOY

Rutgers University, Linguistics Department, 18 Seminary Pl., New Brunswick, NJ 08901, USA

Email: smccoy@rci.rutgers.edu

Abstract. The notion of *kontrast*, or the ability of certain linguistic expressions to generate a set of alternatives, originally proposed by Vallduví and Vilkuna (1998) as a clause-level concept, is re-analyzed here as connecting the level of information packaging in the clause and the level of discourse structure in the following way: *kontrast* is encoded at the clausal level but has repercussions for discourse structure. This claim is supported by evidence from the distribution properties of three colloquial Russian particles –*to*, *že*, and *ved’* which are analyzed as unambiguous markers of *kontrast*. Both the placement of these particles at the clausal level and their role in discourse are viewed as consequences of the type of the *kontrast* set and the cognitive status of information marked by each particle.

Key words: discourse structure, information structure, *kontrast*, particles, colloquial Russian.

1 Introduction

The notion of *kontrast*, introduced in Vallduví and Vilkuna (1998, V&V hereafter), establishes a conceptual distinction between two notions that have been conflated in the literature by the term of *focus*: one is *rheme*, a concept which, in opposition with *theme*, belongs to the domain of information packaging (Chafe 1976 *inter alia*) and the other is what the authors label as *kontrast*, the notion covering quantificational phenomena of a more formal semantic nature (Rooth 1985, 1992, Krifka 1991-92, etc.)¹. V&V use the term *kontrast*

as a cover term for several operator-like interpretations of focus that one finds in the literature: identificational foci, exhaustiveness foci, contrastive foci, contrastive topics, and also interrogative wh-words, which have been seen as paradigmatic “foci” by many semanticists and syntacticians...

¹ Cf. more recent applications of the idea of *alternative sets* in Steedman (2001), Kruijff-Korbayová and Webber (2001). See also Molnár (2001) for an attempt to define *kontrast* as a subtype within a broader category of *contrast*; the latter is claimed to represent an autonomous concept of information structure.

The basic idea behind the notion of *kontrast* is the following: if an expression **a** is *kontrastive*, a **membership set** $M=\{\dots,a,\dots\}$ is generated and becomes available to semantic computation as some sort of quantificational domain. (V&V: 83)

V&V provide a convincing set of data from Finnish, Hungarian, and Catalan (with references to English and other languages) to demonstrate that *kontrastiveness* and *rhematicity* must be treated separately. They assign formal features [K:+] for *kontrastive* elements and [K:-] for non-*kontrastive* ones; similarly, [Rh:+] for *rhemes* and [Rh:-] for *themes*. Thus, there are four possible combinations of these features for any given element:

1. [K:+; Rh:+] – this combination has been discussed in literature under the label of *contrastive focus* (also, *identificational focus* and *exhaustiveness focus*);²
2. [K:+; Rh:-] – this combination has been known as *contrastive topics*;³
3. [K:-; Rh:+] – regular, non-*kontrastive* *rhemes*, or *foci*;
4. [K:-; Rh:-] – regular, non-*kontrastive* *themes*, or *topics*.

As pointed out by V&V, in different languages *kontrast* is expressed by different linguistic resources. For example, in English *kontrast* is signaled mainly by prosody while in Finnish it is encoded by syntactic position. Moreover, there is a conflict between a limited set of structural resources in a language and a set of interpretive categories that need to be expressed.

The goal of this paper is to demonstrate that the use of the notion of *kontrast* can be extended from a clause-level notion to one connecting the clause-level structure and the structure of discourse. Even though this set-generating power is encoded on the clausal level, it is one of the linguistic means that holds the discourse together and reflects the speaker's assumptions about the hearer's state of knowledge and attention.

The linguistic evidence for this claim comes from discourse particles of colloquial Russian –*to*, *že*, and *ved'*, which are analyzed as lexemes whose pri-

² As pointed out by V&V, there is an additional variable present here—the scope: [K:+] has a scope and [Rh:+] has a scope, and they need not but may coincide. I will indicate the scope of the *kontrastive* element within the *rheme* by [_{+K/Rheme} ...].

³ V&V (107, fn. 4) make an important point that not all thematic constituents can be *kontrastive*: only *links* (i.e., pointers to a specific file card for the entry of *rheme*) can while *tails* (i.e., the remaining part of *theme*) cannot (see Vallduví 1992 for terminology). The scope of the [K:+] element within the *link* will be marked here by [_{+K/Link} ...].

mary function is to signal, or unambiguously mark, *kontrast* and which are labeled “*kontrastive markers*,” or “*k-markers*” (McCoy 2001).⁴

Until recently, only descriptive analyses were available for this group of particles and particles *–to*, *že*, and *ved’* were labeled “emphatic,” “contrastive,” “intensifying,” “expressive,” “emotional,” “strengthening,” “enunciative,” etc., with further classification of context-dependent multiple meanings or functions for each particle (see Vasilyeva (1972), *Les Particules Enonciatives en Russe Contemporain*, etc.) However, descriptive approaches to these particles have proved inadequate and a search for an underlying meaning of (some of) these particle has been initiated in the following works: Bitextin (1994), Parrott (1997), Bonnot and Kodzasov (1998), Feldman (2001). However, these early studies in the “unifying” direction have their limitations: they are either based on a single framework and/or choose to deal with unifying a single aspect of the particle(s), such as discourse role or cognitive status of the information marked.

In the previous analyses, the placement rules of these particles at the clausal level made reference to either the position of the particle with respect to a prosodically prominent element or to some information structure construct as ‘theme-rheme’, ‘topic-focus’, etc. However, no comprehensive semantic explanation has been offered for explaining what the prosodic marking is used for and what the relationship between prosody and information packaging at the clausal level is. Similarly, the discourse role of these particles has been analyzed as giving rise to multiple implicatures and inference patterns and thus contributing to the cohesion and coherence of discourse.

In this paper it will be shown that such properties of these particles as their position in the clause and their role in discourse are consequences of their two essential properties as *k-markers*:

1. the type of set it marks and the type of membership within this set;
2. the cognitive status of the referents marked by it.

Below, these 3 particles will be given a unifying analysis which integrates the following current frameworks: the theory of “*kontrast*” (V&V); cognitive statuses of referents in discourse (Yokoyama 1986, Gundel, Hedberg, and Zacharski 1993); information packaging on the clausal level (Vallduví 1992); and the Question Under Discussion organization of discourse (Büring 2000)⁵.

⁴ Even though *kontrast* in colloquial Russian can also be expressed prosodically and by word order, this study focuses only on the *lexical* means of expressing *kontrast* (which is primarily determined by the corpora used). A further investigation (and a different type of corpus) is required to analyze the interaction of several linguistic means of expressing *kontrast*.

⁵ For similar proposals see Kanerva and Gabriele (1995), van Kuppevelt (1995, 1996a,b), Ginzburg (1996), and Roberts (1996). Also see Umbach (2001) for the relation-

The data come from a variety of colloquial Russian texts, such as Protassova's corpus of the CHILDES database (MacWhinney 2000) and Zemskaja and Kapanadze (1978). The structure of the paper is the following: in the next three sections, the kontrastive properties and their consequences will be discussed for each particle individually, with particle *-to* given a more detailed analysis than particles *že* and *ved'*. The final section summarizes the findings with respect to what the analysis of these three particles as k-markers reveals about the role of kontrast at the clausal and discourse levels.

2 K-Marker *-TO*

I will first discuss two essential properties of *-to*: the type of set and the cognitive status of information marked by it. Then I will demonstrate how these essential properties determine the position of *-to* in the clause and its role in discourse.

2.1 TYPE OF SET MARKED BY *-TO*

Particle *-to* marks a set of sets of related propositions (equivalently, a set of questions) which is generated by introducing alternatives to a kontrastive element within the *link* and a kontrastive element within the *rheme*. Consider an example from CHILDES:

- (1) [Varja and her mother are looking at a picture of a dog who put her paw on a bear. They first discuss the dog, then start talking about the bear. Varja gets distracted by taking a scoop into her hand. Mother says:] (CHILDES, séance 2)

U tebj**a**-**TO** sovok, a chto u medvedja v lape?

At you-**TO** scoop but what at bear in paw

'YOU(-**TO**) have a SCOOP, but what does the BEAR have in his paw?'

The information structure of the (English version of the) utterance containing *-to* is shown in (2). Below it, in (3), a generalized structure of the utterance with *-to* is given:

(2) [+K/Link You] (-**TO**) have [+K/Rheme a scoop]...

(3) [+K/Link A]-**TO** has property [+K/Rheme X]

The default case for (3) is when the kontrastive (element within the) link *A* belongs to a set of entities: $M_1 = \{A, B, C, \dots\}$ and the kontrastive (element

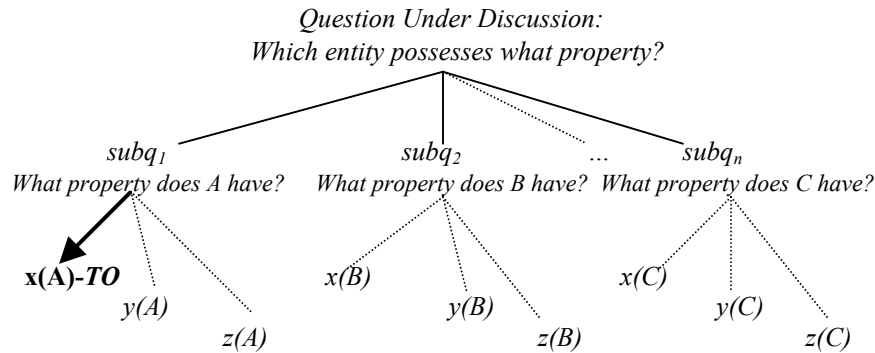
ship between notions of *QUD* and *quaestio* and their application to the analysis of the English *but*.

within the) rheme x belongs to a set of properties: $M_2 = \{x, y, z, \dots\}$.⁶ Thus, the proposition $x(\mathbf{A})$ marked with $-TO$ makes the hearer generate a set M_3 , shown in three alternative ways in (4/5/6). In (4), the set M_3 is represented as a set of sets of propositions, in (5) the same set is shown as a set of questions (following Hamblin 1958/1973 and Karttunen 1977), and in (6) the set M_3 marked by $-to$ is shown graphically:

- (4) $M_3 = \{ \{x(\mathbf{A}), y(A), z(A), \dots\}; \{x(B), y(B), z(B), \dots\}; \{x(C), y(C), z(C), \dots\}; \dots \}$

- (5) $M_3 = \{\text{What is true of } A?; \text{What is true of } B?; \text{What is true of } C?; \dots\}$

- (6) **Set of sets of propositions (= set of questions) M_3 marked by $-TO$:**



So, the primary function of $-to$ is a marker of a set of sets of propositions which is generated by introducing alternatives to the kontrastive link and the kontrastive rheme. While for the proposition containing $-to$ the truth value is asserted, it is not the case with the alternative propositions: they are only made salient with the help of $-to$.

⁶ However, a more marked (i.e., less frequent) scenario is also possible: i.e., when the contrast set for links is a set of properties, while the set of alternatives to the rheme consists of entities, as in the example below:

[Varja can't stop running. Mother is asking: who is the one running?]

(CHILDES, séance 2):

Nu [+K/Rheme kto] [+K/Link begaet] **-TO?**

Well who is-running-**TO**

'As for somebody running(-TO), who is doing this?'

2.2 COGNITIVE STATUS OF REFERENTS MARKED BY *-to*

Particle *-to* marks information (estimated by the speaker to be) known to the hearer but not activated in the hearer's mind at the time of the utterance. How can the speaker assume that the information is also known to the hearer? The speaker can safely assume that the hearer also knows the information if this information is shared through: code (language, culture), encyclopedic knowledge, deixis, or common experience (for more detail, see Yokoyama 1986). The example in (7) illustrates a speech situation where the source of the speaker's assumption about the proposition marked by *-to* being located in the hearer's knowledge set is deixis, while in (8) it comes from the experience shared by the speaker and the hearer:

- (7) [A to B, after a long silence, on the top of a mountain before dawn]
(Vasilyeva (1972:68))

[+K/Link Tišina]-**TO** [+K/Rheme kakaja]!⁷
Quietness-TO what

'How quiet it is!' or

'As for the state of quietness(-**TO**), how quiet it is!'

(Vasilyeva's translation: 'How wonderfully quiet!')

⁷ It may be worthwhile to show how *-to* in this wh-exclamative marks a set of sets of propositions. As pointed out by Grimshaw (1979), exclamative clauses are factive, i.e., they presuppose their propositional content. In the case of (7), this means that some particular state of quietness is presupposed. Moreover, as pointed out by Zanuttini and Portner (1999) and Portner and Zanuttini (2000), "exclamatives convey a conventional scalar implicature indicating that the fact that they express is surprising or noteworthy in some way" (Zanuttini and Portner 1999: 4). In the case of (7), the scalar implicature it conveys is that the particular state of quietness at the moment of the speech event is remarkable in some way (cf. Vasilyeva's translation as 'How wonderfully quiet!'); for example, it can be ranked against some other states of quietness in other deictic dimensions (e.g., in a desert at dawn or in a soundproof chamber with no sound turned on, etc.). Thus, the wh-word *kakaja* 'what' in this exclamative generates a rheme contrast set.

Now, how is the link contrast set created? The presence of particle *-to* in (7) suggests that the set of alternatives brought by the scalar implicature is a part of a set of sets, i.e., that particular scale is compared to other relevant scales. For example, if the proposition in (7) is considered in a context, it is likely that other aspects of the particular situation will be considered next: the hearer might add something like *And how beautiful it is!* (For more detail, see McCoy 2001: 141-144.)

- (8) [Varja practices pulling laces through holes. After several other activities, Mother says:] (CHILDES, séance 5)

Oj, slushaj, a my s tobaj [+K/Link chernen'kij]-**TO** shnurochek
Oh listen but we with you black-TO lace-DIM

[+K/Rheme tak ni razu eshche i ne prodevali], da?
so not once yet even NEG pulled-through, yes

‘Oh, listen, as for the BLACK(-**TO**) lace, you and I haven’t pulled it through yet even once, right?’

To summarize the essential properties of *-to* as a k-marker: with respect to the type of set marked, *-to* marks a set of sets of propositions which differ from each other in the values of the kontrastive link and the kontrastive rheme. With respect to cognitive status of information marked by *-to*, this information is assumed by the speaker to be known to the hearer but not activated in the hearer’s mind at the time of the utterance. In the next subsection I will demonstrate that these two essential properties of *-to* determine its position in the clause and its role in discourse.

2.3 CONSEQUENCES:

POSITION IN THE CLAUSE AND ROLE IN DISCOURSE

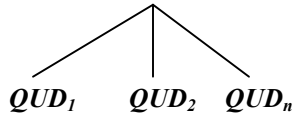
The position of *-to* in the clause is best analyzed with respect to the element marked [+kontrast]. Thus, *-to* is an enclitic to the kontrastive element within the *link*. This is a direct consequence of its kontrastive properties: first, the contrast on the *link* value results in a set of sets of propositions (cf. similar treatment of “contrastive topics” in Büring (2000), etc.); second, the cognitive status of information marked by *-to*—known to the hearer but not activated in the hearer’s mind—is more compatible with the function of the *link* than the *rheme*.

The role of *-to* at the discourse level is also a consequence of being a marker of a set of questions (equivalently, a set of sets of propositions).⁸ In a discourse tree, as in (9), any question under discussion (QUD) can be con-

⁸ Another consequence of the kontrastive nature of *-to* is its multiple implicatures proposed in the literature. Since *-to* is a marker of a set of sets of propositions, it also marks contrast and emphasis, functions as a theme/topic marker or a marker of contrastive topics, implies plurality, etc. The following implicatures are consequences of its second essential property—a marker of information (assumed by speaker to be) known to hearer but not activated in discourse yet: a marker of unexpectedness in addressing a topic; the fact that it adds a tone of familiarity, conversational spontaneity, intimacy, etc.

strained by k-marker *-to* in the way as shown in (6):

(9) **Discourse Situation**



To summarize, the notion of *kontrast* is the core semantic meaning of particle *-to*: by analyzing it as a marker of a set of sets of propositions (equivalently, a set of questions) and as marking referents known to the hearer but not currently activated, its position in the clause and its role in discourse are accounted for.

3 K-Marker *žE*

K-marker *že* differs from *-to* in both the type of set and cognitive status of information marked and, accordingly, in other important properties: position in the clause, role in discourse, etc. However, as with *-to*, *kontrastivness* is the core semantic meaning of this particle and determines its distribution.⁹

3.1 TYPE OF SET MARKED BY *žE*

Particle *že* marks a set of propositions which differ from each other in the value of (at least) one term. The *kontrast* set for *že* contains members which are mutually exclusive: if one proposition is true, the other one(s) is/are false. Thus, the

⁹ Rendering the meaning of utterances containing *že* into English usually involves the use of either some contrastive lexeme, such as *but*, or some prosodic means. Specifically, the English translation of statements containing *že* can employ the so-called “contradiction contour” (Lieberman and Sag 1974; Pierrehumbert and Hirschberg 1990)

Lieberman and Sag (1974: 420) discuss sentences as shown in (i) (their example (8)) which, if pronounced with the special intonation contour, are felicitous only as contradicting some “context proposition.”

(i)

Elephantiasis isn't incurable!

Russian *že* also seems to be comparable to German *doch* and Dutch *toch*; on the German and Dutch particles, see, e.g., Karagjosova (2001a, b) and Zeevat (2000).

relationship among the members of the *že* set often involves (binary) opposition, contradiction, or negation. The kontrast set marked by *že* is generalized in (10) and illustrated with a sample of naturally occurring data in (11):

(10) **The kontrast set marked by *ŽE*:**

$M = \{ X, X' \}$, where $X = \neg X'$
 (X is true if and only if X' is false)

(11) [Varja notices a fly on the windowsill and asks her mother to kill it]
 (CHILDES, séance 4)

*VAR: Ona muxa, muxa.
 *MOT: Muxa, muxa, da.
 *VAR: Ubit', ubit' ee!
 *MOT: Ona **zhe** uzhe ubita.

*VAR: It's a fly, a fly.
 *MOT: A fly, a fly, yes.
 *VAR: Kill, kill it!
 *MOT: It (*že*) is already killed.

Gloss:

Ona **zhe** [_{+K} uzhe ubita].
she že already killed(participle)
 '(But) it (*že*) is already killed.'

In (11), mother's utterance with *že* corrects the presupposition of Varja's previous utterance *Kill it!*, i.e., *The fly is alive*. The members of the kontrast set marked by *že* are mutually exclusive: it is impossible for the fly to be alive (presupposition of Varja's utterance) and to be already killed (mother's utterance containing *že*) at the same time.¹⁰

¹⁰ An interesting case is the use of *že* in wh-questions: how in the world can the kontrast set be mutually exclusive with wh-words? The force of such questions is rhetorical; the question is asked but presupposes that no (reasonable) answer is true, from the speaker's point of view. An example, recorded in a home setting, is provided in (i):

(i) [Father puts away his son's library books and tells him that mother took them back to the library (which she did not do). The boy asks mother if she indeed took them back. She says 'no'. He then says (in the presence of both parents):]

U kogo *že* mne togda sprosit'?
 'Who (*že*) (in the world/the hell...) should I ask then?'

3.2 COGNITIVE STATUS OF REFERENTS MARKED BY ŽE

Particle *že* marks the membership set, one member of which is activated in the hearer's mind at the time of the utterance, while the other (incompatible) member is viewed by the speaker as though it is (or should have been) known to the hearer and should have been activated at this time.

For example, in (11), the activated member of the set is *The fly is alive* (presupposition of *Kill, kill it!*). The other member of the set (is treated by the speaker as though it) should be known to the hearer and, therefore, should be activated—i.e., in the speaker's estimation, there are enough visual cues for the hearer to make the correct conclusion (*The fly is already dead*).

Now let us consider some consequences of the essential properties of *že* as a k-marker.

3.3 CONSEQUENCES:

POSITION IN THE CLAUSE AND ROLE IN DISCOURSE

The placement of particle *že* can be defined with respect to the kontrastive element, which is a propositional term (usually within the *rheme*). The condition of mutual exclusiveness (or binary opposition) on the set marked by *že* makes this particle a good candidate of marking rhematic contrasts, especially cases of kontrastive *verum foci*—i.e., contrast not on the lexical meaning of the verb but on some verbal inflectional category, such as tense, aspect, etc. It is possible to translate utterances containing *že* as cases of rhematic polarity of verum focus;

In (i), the only two reasonable answers ('I should ask mother' and 'I should ask father') have been previously asserted to be false in the context of the utterance. Taking the semantic value of a wh-question to be the set of propositions which could serve as possible (contextually accessible) answers (Hamblin 1973), the set for (i) is given in (ii). The presence of *že* in (i) causes *že* to be applied to the set in (ii), resulting in (iii).

- (ii) { I should ask mother, I should ask father }
- (iii) *že* ({ I should ask mother, I should ask father })
= { I should ask mother, I should ask father } and presupposes that the hearer believes { I should not ask mother, I should not ask father }

Simply put, *že* is defined to take a single proposition and provide a presupposition; however, when *že* is provided with a set of propositions (a question), it is applied to each proposition in the set, creating a presupposition for each. The end result is that for each possible answer, there is a presupposition that the hearer has already rejected it as false; the speaker has provided no felicitous options for answers within this contextually determined restricted set. For more detail see Hagstrom and McCoy (2002).

for example, alternative translations for the utterance with *že* in (11) are *It [=the fly] IS already killed* or *It is ALREADY killed*.

The role of *že* in discourse is also a consequence of its contrastive properties: *že* refers back to a salient element or some unresolved (from the speaker's viewpoint) question in the discourse or discourse situation.¹¹

To summarize, by analyzing particle *že* as a marker of a set of mutually exclusive propositions and as a marker of information which is related through the set relationship to information that is activated in discourse, its distributional properties at the clausal and discourse levels are largely accounted for.

4 K-Marker *VED'*

Etymologically, particle *ved'* is a form of the verb *vedat'* 'to know', from which it has inherited factivity.

4.1 TYPE OF SET MARKED BY *VED'*

Similarly to *že*, particle *ved'* marks a set of propositions which differ from each other in the value of at least one term. However, there are important differences between them: for *ved'* the set membership is restricted to propositions which have illocutionary force of assertions (while *že* does not have this restriction). Also, the members of the set are not mutually exclusive (as in the case of *že*) but rather supplementary; when *ved'* does express contradiction, it is more of the 'yes, but...' type.

The examples in (12-13) illustrate that the difference between *že* and *ved'* comes (partially) from the difference between a mutually exclusive set (*že*) and a supplementary set (*ved'*):

- (12) [Varja and Grandmother are looking at a picture of birds standing on the ground. Varja believes that the birds have fallen down. Grandmother corrects Varja: since they are standing and not lying down, they have not fallen down:] (CHILDES, séance 3)

¹¹ Multiple implicatures associated with *že* at the discourse level can also be viewed as following from its two essential properties. Since *že* is a marker of a restricted set, containing mutually exclusive propositions, it is subsequently a marker of contrast, a marker of emphasis, a marker of contrastive focus, a marker of contrastive topic, it adds a tone of an indisputable argument and is perceived as a verbal attack on hearer, etc. From being a marker of activated information which is to a large extent known to the hearer, it follows that *že* can also be analyzed as a marker of a reference point in the activated domain of reference, as being perceived to be a verbal attack on hearer, as adding a tone of an indisputable and irrefutable argument, and conveying emotions of irritation, annoyance, and impatience, etc.

*VAR: Èta pit'ki.
 %eng: These are birds.
 *VAR: Upai.
 %eng: Fell down.
 *GPP: Oni ne upali.
 %eng: They did not fall down.
 *GPP: Pochemu **zhe** ty думаеш', chto oni upali.
 %eng: Why do you think, that they fell down.
 *GPP: Odna stoit na penechke, drugaja stoit na zemle, a tret'ja stoit na zemle szadi penechka.
 %eng: One is standing on a little stump, the second is standing on the ground, and the third is standing on the ground behind the stump.
 *GPP: Vot tak, szadi penechka.
 %eng: That way, behind the stump.
 *GPP: A pochemu ty думаеш', chto oni upali, oni **zh** ne lezhat, oni stojat.
 %eng: And why do you think, that they fell down, they are not lying, but standing.
 *GPP: Kogda kto-nibud' upadet, on lezhit.
 %eng: When somebody falls, he lies down.

*GPP: [_{+K} Pochemu] **zhe** ty думаеш', chto oni upali.
 'Why (*že*) in the world do you think that they fell down.'

*GPP: A pochemu ty думаеш', chto oni upali, oni **zh** [_{+K} ne lezhat], oni [_{+K} stojat].
 'And why do you think that they fell down, they (*ž*) are not lying down, they are standing.'

In (12), *že* marks a set containing mutually exclusive members *The birds fell down* and *The birds did not fall down* (thus, it conveys emotions of contradiction, correction, and is perceived as a 'verbal attack on the hearer'): Varja believes that *The birds fell down* (since they are standing on the ground and not flying), while the Grandmother refutes that by saying that *The birds did not fall down* (since they are standing and not lying down). Later, this argument is brought to a mediator, Varja's mother (Katja). In presenting the gist of the argument to Katja, the Grandmother uses *ved'* and not *že* (since no opposition is expected from an adult who knows the lexical meaning of *upali* 'fell down'):

(13) [Grandmother re-tells the argument about the "fallen" birds to the mother] (CHILDES, séance 3)

Vot ona govorit, chto èti ptichki upali, a ja schitaju,
PART she says, that these birds-DIM fell-down, and I think

chto oni [_{+K} ne upali], oni [_{+K} stojat **ved'** na nozhkax].
that they NEG fell-down, they are-standing ved' on feet-DIM

‘So she says that these little birds fell down and I think that they did not fall down, they are standing, (*ved'*) you know, on their feet.’

So, in (13) *ved'* marks a supplementary set, the members of which are synonyms rather than antonyms (thus, it constitutes an ‘extension in logic’ and not ‘a verbal attack’): {*The birds did not fall down; The birds are standing on their feet*}.

4.2 COGNITIVE STATUS OF REFERENTS MARKED BY *VED'*

Similarly to *-to*, the particle *ved'* marks information known to the hearer but not activated in the hearer’s mind at the time of the utterance. The difference between *že* and *ved'* comes (partially) from the difference in the cognitive status of information marked. Consider example (14) from Zemskaja and Kapanadze (1978: 52-53; glosses and translation added):

- (14) [Two retired ballet dancers had previously talked about what great actors they used to work with in their youth. Later, one of them realizes that they forgot to mention two of the most famous actors (Čirkov and Čerkasov) and reminds her friend about that]

A: Kogda my s tobj govorili o mjuzik-holle, my govorili: ax kakie byli aktery!
‘When you and I talked about the music hall, we were saying “Oh what actors were there!”’

My s tobj **VED'** ne skazali, čto tam byl Čirkov, tam byl Čerkasov!...
We with you VED' NEG said that there was Čirkov, there was Čerkasov
‘You know, we did not say that there was Čirkov there, there was Čerkasov there!..’

B: [overlapping with the end of the preceding utterance]
Da gospodi! Čerka-a-sov, Čirko-o-v...
‘Oh, my god! Čerka-a-sov, Čirko-o-v...’

In (14), *ved'* marks information known to both the speaker and the hearer which is not activated in the hearer’s mind at the moment: talking about what great actors were in the theater at the time when both of the interlocutors worked there, they forgot to mention two most famous actors.

4.3 CONSEQUENCES:

POSITION IN THE CLAUSE AND ROLE IN DISCOURSE

Similarly to the other two particles, the position of *ved'* in the clause is best defined with respect to the kontrastive element: it is either a proclitic or an enclitic to such element.

The discourse role of *ved'* is determined by the type of set it marks: being a marker of a supplementary set, it relates members of the set to each other by building a super-question above them.¹²

To summarize: *ved'* has been observed to share certain properties with *že* and other properties with *-to*. By addressing the questions of what type of set and what cognitive status of information each of these particles marks, their differences and similarities are explained.

5 Conclusion

The notion of kontrast has been found to be a core semantic meaning of such lexical items as particles *-to*, *že*, and *ved'*: these lexemes are analyzed as unambiguous markers of kontrast. The core semantic meaning of each of these particles has been found to encompass two issues: the type of set and the cognitive status of referents marked by each of these particles. Among important properties of these particles, which are nevertheless treated as consequences of their essential kontrastive properties, are the placement of the particles at the clausal level and their role in the organization of discourse. Their distribution properties and the choice of particles in a particular context is also motivated by the differences in their kontrastive nature. The findings are summarized in Table 1 below.

¹² Also, multiple implicatures of *ved'* can be treated as consequences of its kontrastiveness: since *ved'* is a marker of a (supplementary) set of propositions (assertions), it is a marker of contrast, a marker of emphasis, an assertion marker, etc. Since *ved'* (like *-to*) is a marker of information that is assumed by speaker to be known to hearer but not activated yet, it is a marker of unexpectedness in addressing a topic, a marker of encyclopedic knowledge, and it is perceived as a (polite) reminder, etc.

Table 1 Comparing Kontrastive Markers *-TO*, *ŽE*, and *VED'*

	<i>-TO</i>	<i>ŽE</i>	<i>VED'</i>
ESSENTIAL PROPERTIES:			
Type of set marked	A set of sets of propositions (= a set of questions); requires two [+K] elements: in the <i>link</i> and the <i>rheme</i>	A set of mutually exclusive propositions which differ in the value of at least one [+K] element	A supplementary set of propositions which differ in the value of at least one [+K] element
Cognitive status of referents	<i>Familiar</i> to the hearer but <i>not activated</i> in his/her mind	One member: <i>in focus</i> or <i>activated</i> ; the other member: (treated as) <i>familiar</i> and <i>activated</i>	Same as for <i>-TO</i>
CONSEQUENCES:			
Position in a clause	Enclitic to [+K] element, usually <i>link</i>	Enclitic to [+K] element	Proclitic or enclitic to [+K] element
Role in Discourse	Generates a set of sub-questions (sisters dominated by the same QUD); partially answers QUD and makes salient the other sub-questions.	Refers back to salient element or some unresolved question in the previous discourse	Generates a set by building a super-question above the members of the set

The example of these three colloquial Russian particles suggests that the notion of kontrast has applications beyond the level of the clause to serve as a construct which connects the levels of information structure and discourse structure. Specifically, the type of kontrast set as encoded in the meaning of these particles can be viewed as playing the following roles: at the level of the discourse structure, it signals the relationship of the proposition to the preceding, following or implied propositions in discourse; at the level of information structure, the type of set determines how it is encoded within the clause. For example, if the kontrast set is a set of sets of propositions, there are two kontrastive elements in the clause, one of which is within the link and the other one within the rheme (as in the case of particle *-to*). However, if the kontrast set is comprised of a set of propositions, that requires a single kontrastive element within the clause, which is encoded within the rheme (as in the case of particles *že* and *ved'*). These findings are summarized in Table 2 below:

Table 2 **Kontrast Set in Relation to Discourse Structure and Information Structure**

	-TO a set of sets of propositions	ŽE a set of propositions (mutually exclusive)	VED' a set of propositions (supplementary)
DS	Signals to hearer to treat <i>p</i> as a member of a set of sets of <i>ps</i> (preceding or following <i>p</i> or implied)	Signals to hearer to treat <i>p</i> ₁ and <i>p</i> ₂ as members of set; signals to replace <i>p</i> ₁ with <i>p</i> ₂	Signals to hearer to treat <i>p</i> ₁ and <i>p</i> ₂ as members of set; signals to supplement <i>p</i> ₁ with <i>p</i> ₂
IS	Signals how the set of sets of <i>ps</i> is organized; contains two [+K] elements: [+K/Link] set and [+K/Rheme] set	Contains one [+K] element: [+K/Rheme] alternative set	

A further refinement of this classification is possible if another dimension—the cognitive status of referents in discourse—is taken into account; however, this dimension is not the focus of the current paper (but see Gundel et al. (2001), also current volume, for a notable exception).

In colloquial Russian, lexical means of expressing contrast are used in combination with prosodic prominence and intonation contours. Future research is needed to analyze the interaction of intonation structure with kontrastive particles and relate the two to information structure and discourse structure.

Acknowledgement

The author thanks Catherine Chvany, Paul Hagstrom, Cathy O'Connor, Enric Vallduví, the anonymous reviewers, and the audience of the ESSLLI 2001 workshop for helpful comments.

Data Sources

CHILDES: Other Languages—Russian—Protassova, in MacWhinney, B., 2000, *The CHILDES Database: Tools for Analyzing Talk*, 3d Edition, Vol. 2: *The Database*, Mahwah, NJ: Lawrence Erlbaum Associates.
 Zemskaja, E. A. and L. A. Kapanadze, eds., 1978, *Russkaja Razgovornaja Reč': Teksty*, Moskva: Nauka.

References

Bitextin, A. B., 1994, "Časticy *-to, že, ved'* i vvodnye konstrukcii tipa *kak izvestno* kak sredstva ukazanija na izvestnost' propozicional'nogo sodržanija predloženijslušajuščemu," AKD, Moskovskij Gosudarstvennyj Universitet.

- Bonnot, C. and Kodzasov, C. B., 1998, "Semantičeskoe var'irovanie diskursivnyx slov i ego vlianie na linearizaciju i intonirovanie (Na primere častic ŽE i VED)," pp. 382-443 in *Diskursivnye Slova Russkogo Jazyka: Opyt Kontekstno-Semantičeskogo Opisanija*, K. Kiseleva and D. Paillard, eds., Moskva: Metatekst.
- Büring, D., 2000, "On D-trees, beans, and B-accents," Ms., UCLA.
- Chafe, W. L., 1976, "Givenness, contrastiveness, definiteness, subjects, topics, and point of view," pp. 25-55 in *Subject and Topic*, C. N. Li, ed., New York: Academic Press.
- Feldman, A., 2001, "Discourse markers—Accessing 'hearer-old' information: The case of the Russian že" in *Proceedings of the 27th LACUS Forum*, A. Melby et al., eds.
- Ginzburg, J., 1996, "Interrogatives: Questions, facts and dialogue" pp. 385-422 in *The Handbook of Contemporary Semantic Theory*, S. Lappin, ed., Oxford, UK: Blackwell Publishers.
- Grimshaw, J., 1979, "Complement Selection and the Lexicon," *Linguistic Inquiry* **10**(2), 279-326.
- Gundel, J. K., Hedberg, N., and Zacharski, R., 1993, "Cognitive status and the form of referring expressions in discourse," *Language* **69** (2), 274-307.
- Gundel, J. K., Hegarty, M., and Borthen, K., 2001, "Information structure and pronominal reference to clausally introduced referents," pp. 37-52 in *Information Structure, Discourse Structure and Discourse Semantics*, ESSLLI 2001 Workshop Proceedings, I. Kruijff-Korbayová and M. Steedman, eds.
- Hagstrom, P. and McCoy, S., 2002, "Presuppositions, wh-questions, and discourse particles: Russian ŽE," paper presented at *Formal Approaches to Slavic Linguistics (FASL-11)*, University of Massachusetts-Amherst.
- Hamblin, C. L., 1958, "Questions," *Australasian Journal of Philosophy* **36**(3), 159-168.
- Hamblin, C. L., 1973, "Questions in Montague English," *Foundations of Language* **10**, 41-53.
- Kanerva, J. M. and Gabriele, L. A., 1995, "Intonation and focus layers," pp. 335-346 in *Proceedings of the Northeast Linguistic Society 25 (NELS 25)*, Vol. 1, Papers from the Main Session, J. N. Beckman, ed., University of Pennsylvania.
- Karagjosova, E., 2001a, "Modal particles and the common ground: Meaning and functions of German *ja*, *doch*, *eben/halt* and *auch*," pp. 201-209 in *BI-DIALOG 2001*, P. Kühnlein, H. Reiser, and H. Zeevat, eds.
- Karagjosova, E., 2001b, "Towards a comprehensive meaning of the German *doch*," pp. 131-141 in *Proceedings of the Sixth ESSLLI Student Session*, K. Striegnitz, ed., Helsinki.
- Karttunen, L., 1977, "Syntax and semantics of questions," *Linguistics and Philosophy* **1**, 3-44.
- Krifka, M., 1991-2, "A compositional semantics for multiple focus constructions," *Linguistische Berichte*, Suppl. 4, 17-53.
- Kruijff-Korbayová, I. and Webber, B., 2001, "Information structure and the semantics of 'otherwise'," pp. 67-83 in *Information Structure, Discourse Structure and Discourse Semantics*, ESSLLI 2001 Workshop Proceedings, I. Kruijff-Korbayová and M. Steedman, eds.
- Liberman, M. and Sag, I., 1974, "Prosodic form and discourse function," pp. 416-427 in *Proceedings of the Chicago Linguistic Society* 10.

- McCoy, S., 2001, "Colloquial Russian particles –to, že, and ved' as set-generating ('kontrastive') markers: A unifying analysis," Ph.D. Thesis, Boston University.
- Molnár, V., 2001, "Contrast from a contrastive perspective," pp. 99-114 in *Information Structure, Discourse Structure and Discourse Semantics*, ESSLLI 2001 Workshop Proceedings, I. Kruijff-Korbayová and M. Steedman, eds.
- Parrott, L., 1997, "Discourse organization and inference: The usage of the Russian particles že and ved'," Ph. D. Thesis, Harvard University.
- Les Particules Enonciatives en Russe Contemporain* 1 (1986), 2 (1987), 3 (1988), Paris: Institut d'Études Slaves.
- Pierrehumbert, J. and Hirschberg, J., 1990, "The meaning of intonational contours in the interpretation of discourse," pp. 271-312 in *Intentions in Communication*, P. R. Cohen et al., eds. Cambridge, MA: MIT Press.
- Portner, P., and Zanuttini, R., 2000, "The force of negation in wh exclamatives and interrogatives," in *Negation and Polarity: Syntactic and Semantic Perspectives*, L. R. Horn and K. Yasuhiko, eds. New York and Oxford: Oxford University Press.
- Roberts, C., 1996, "Information structure in discourse: Towards an integrated formal theory of pragmatics," pp. 91-136 in *OSU Working Papers in Linguistics 49: Papers in Semantics*, J. H. Yoon and A. Kathol, eds. Ohio State University, Columbus, OH.
- Rooth, M., 1985, "Association with focus," Ph.D. Thesis, University of Massachusetts, Amherst, MA.
- Rooth, M., 1992, "A theory of focus interpretation," *Natural Language Semantics* **1**, 75-116.
- Steedman, M., 2001, "Information structure and the syntax-phonology interface," *Linguistic Inquiry* **34**, 649-689.
- Umbach, C., 2001, "Relating contrast and contrastive topic: A focus-semantic analysis of 'but'," pp. 175-188 in *Information Structure, Discourse Structure and Discourse Semantics*, ESSLLI 2001 Workshop Proceedings, I. Kruijff-Korbayová and M. Steedman, eds.
- Vallduví, E., 1992, *The Informational Component*, New York: Garland.
- Vallduví, E. and Vilkuña, M., 1998, "On rheme and kontrast," *Syntax and Semantics* **29** (The Limits of Syntax), 79-108.
- van Kuppevelt, J., 1995, "Discourse structure, topicality and questioning," *Linguistics* **31**, 109-147.
- van Kuppevelt, J., 1996a, "Directionality in discourse: Prominence differences in subordination relations," *Journal of Semantics* **13**, 363-395.
- van Kuppevelt, J., 1996b, "Inferring from topics," *Linguistics & Philosophy* **19**, 393-443.
- Vasilyeva, A. N., 1972, *Particles in Colloquial Russian: Manual for English-Speaking Students of Russian*, Moscow: Progress Publishers.
- Yokoyama, O. T., 1986, *Discourse and Word Order*, Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Zanuttini, R., and Portner, P., 1999, Types of clauses: A Case study in exclamatives. Ms., Georgetown University.
- Zeevat, H., 2000, "Discourse particles as speech act markers," *LDV Forum*, 74-91.