

Questions to the Modus: Semantics, Structure and Pragmatics (Based on the Russian National Corpora)*

Victoria V. Kazakovskaya¹, Maria V. Khokhlova²

¹Russian Academy of Sciences, Institute for Linguistic Studies, Tuchkov Per., 9, St. Petersburg, Russia

²Department of Mathematical Linguistics, Saint-Petersburg State University, Universitetskaya Nab., 11, St. Petersburg, Russia

Abstract- The paper deals with the analysis of so-called modus questions, which are interpreted as interrogative constructions appealing to the intentional or mental state of the addressee — her / his opinion, knowledge, evaluation or explanation. The main semantic types, formal characteristics and structural patterns of modus questions including their quantitative properties and characteristics in a natural conversation will be discussed.

Index Terms- Interrogatives, modus semantics, question models, pragmatics, Russian text corpora.

1. Introduction

The paper deals with the analysis of so-called modus questions (henceforth *MQs*), which are interpreted as interrogative constructions appealing to the intentional or mental state of the addressee — her / his opinion, knowledge, evaluation or explanation (cf. dictum questions, or questions to the dictum [1]).

Prototypical *MQs* are represented by constructions containing explicit modus frame with verbs of propositional attitude, e.g.

What / How do you: SG : PL *think / suppose / believe / reckon* etc. (...)?,

Why do you: SG : PL *think / suppose* etc. *that* (...)?

in the form of *WH*-questions:

(1) — *Where do you think he can be hanging about? — How can I know? — Koršun responded petulantly.*

‘Kak dumaeš, gde on možet boltat’sja? — A ja znaju? — nedovol’no otkliknulsja Koršun.’ [G. Praškevič. Aleksandr Bogdan (2001)].

Cf. *MQs* in the form of *YES / NO*-questions:

Do you: SG : PL *think / suppose / believe / reckon* etc. *that* (...)?,

including their following variants

(2) — *Do you really consider him to be the greatest person? — asked me a man whom we call by convention in our story Petrov. — Of course, the greatest, and what else?*

‘ — Neuželi vy ego, pravda, sčitajete veličajšim? — sprosil menja tot, kogo my v našem rasskaze uslovno nazyvajem Petrovym. — Konečno, veličajšim, a kakim že ješče? [V. Vojnovič. Delo № 34840 (1999)].

Within discourse sequences *MQs* can occupy not only an initiative position (see examples 1, 2 above), but also a reactive one, in which their pragmatic potential is essentially extended, e.g.

(3) — *But you have the State Exam in foreign language! — (a) How do you know? — she wondered.*

‘ — U vas že gos po inostrannomu! — A ty otkuda znaeš? — udivilas’ ona.’ [N. Gorlanova. Filologičeskij amur (1980)].

The main semantic types, formal features and structural patterns of *MQs* including their quantitative properties and characteristics in a conversation will be discussed as well as reactions given by the participants of the dialogues.

The analysis shows that *MQs*, although not being very frequent in dialogues of native Russian speakers nevertheless represent a certain trait of modern spontaneous discourse. The *WHAT / HOW do you think*-model (henceforth *WHAT / HOW-model*) proves to be the most prototypical one amongst the mentioned models, whereas interrogative constructions appealing to the reasoning of the addressee’s point of view (*WHY-models*) as well as stylistically marked ones are less common.

The obtained results can be applied in various fields of natural language processing, e.g. dialogue systems.

2. Materials and Methods

The research focuses on the development of the models describing these questions on the basis of the Russian National Corpus (a subset of 123 million tokens) [2]. The search in the system is based on the morphological annotation combined with lemmata and word forms.

Lexico-syntactic models [3] can be used for describing patterns involving lemmas or word forms, part-of-speech tags, characters, and other attributes in an annotated corpus. While writing lexico-syntactic models we used regular expressions and the IMS Corpus Workbench query language.

The search of the system is based on morphological annotation combined with lemmata and word forms. For example, the pattern $[lemma="kak"] \ []\{0,5\} \ [word="sčitaješ" \ | \ word="sčitajete"] \ []\{0,15\} \ [lemma=" \?"]$ describes constructions with the interrogative-relative pronoun *kak* (‘how’ / ‘what’) and verb *sčitat* (‘consider’ / ‘think’) in both 2SG or 2PL forms, with the distance between them being up to 5 words: cf. *kak ty sčitaješ, (...)* ‘how / what do you consider (...)’ — *kak ty vsjo-taki sčitaješ, (...)* ‘how do you ever / still consider...’ — *kak že ty vsjo-taki sčitaješ, (...)* ‘how do you ever still consider (...)’.

* The research was carried out with the financial support of the Russian National Foundation in 2015 (grant 14-18-03668).

The restriction $[/0,15\}$ means that there are up to 15 words between *sčitaješ*: 2SG or *sčitajete*: 2PL ‘you consider’ and the end of the sentence (a question mark).

3. Analysis Results and Their Discussion

A. The Main Semantic Types of Modus Questions and Their Quantitative Properties in the Corpora

Questions appealing to the propositional viewpoints of the addressee are inherent to the process of communication involving finding out opinion, supposition and evaluation of the speakers [4]. Semantic differentiation of MQs based on the typology of propositional or so-called intensional predicates allows distinguishing questions in different modus. We will focus on the most typical questions belonging to the various spheres of the mental (i.e. so-called epistemic) modus: 1) MQs of supposition, 2) MQs of reasoning, and 3) MQs of authorship.

MQs of supposition can be described as interrogative sentences appealing to the intentions of the addressee — their opinion, hypothesis, supposition about an object:

What / How do you: SG : PL think / suppose / believe / reckon etc. (...)?;

Up to you, according to you...?

Such MQs can be interpreted as a model with the *HOW*-element (*HOW-model*) (see an example 1 above).

MQs of reasoning ask the addressee to explain the viewpoint that are not evident or clear for the speaker: *Why do you: SG : PL think / suppose etc. that (...)?;*

(4) — *And why do you still think that they shot to Marusja?*

‘— A počemu vy vsjo-taki sčitajete, čto streljali imenno v Marusju?’ [T. Ustinova. *Bol’šoje zlo i melkije pakosti* (2003)].

By MQs of authorship the speaker tries to discover the source of information or the way the information was found:

Who told you, that (...)?

From what source do you SG : PL know etc. (see an example 3a).

This type of questions is usually used if the information is known to be reliable but the speaker is sure the addressee cannot be the author of this information.

The whole range of MQs was restricted to the most characteristic models (or constructions) that include modus frames of mental semantics with the prototypical intentional predicates (particularly, *suppose* ~ ‘predpolagat’, *consider* ~ ‘sčitat’, *think* ~ ‘dumat’) in the second person singular and plural forms inherent to the Russian conversation.

Table 1 represents the models of MQs found in the Russian National Corpus (1950—2014).

TABLE I Models of MQs in the Russian National Corpus

Semantic type of MQ	Number	%
MQ of supposition	2251	54
MQ of reasoning	331	8
MQ of authorship	1584	38
Total	4166	100

B. Modus Questions of Supposition: Some Formal Characteristics, Structural Features and Quantitative Properties

MQs of supposition include frames that are introduced by the main parts of the compound sentences, clauses or simple sentences being in postposition to the main utterance of the speaker.

MQs of supposition represented by *HOW-model* prove to be the most numerous type among the questions to the modus (54%, see table 1). Table 2 shows quantitative characteristics of the variants of *HOW-model*.

TABLE II MQs of supposition in the Russian National Corpus

Model	Number (%)
What / How do you: SG : PL think	76
What / How do you: SG : PL suppose	4
What / How do you: SG : PL reckon	20
Total	100

Plural form prevails over the singular one (53% vs 47%) amongst all verb forms with one exception (*HOW do you think-model*).

MQs with frames occupying the main part of the compound sentence and simultaneously the absolute beginning of the initiative reply (52%) proved to be the most frequent, e.g.

How do you think (...)?

How do you suppose (...)?

How do you reckon (...)?

The analysis of the replies given to the *HOW-model* in an initiative position shows that the addressees tend to react to these questions (85%) giving a verbal reply (79%) [4]. Among them informative replies dominate. This fact makes frames with *HOW-model* a rather successful tool that stimulates the addressee to reveal her / his point of view. 10% of the replies repeat the modal part of the question.

Thus MQs of supposition that highlight the importance of the partner’s opinion and value of her / his position for the speaker can be ascribed to the MQs in their primary (or main) function. The lack of knowledge (and consequently answer) by the speaker himself represents the distinctive trace of such questions. In this case the structural redundancy of the modus can be explained on the pragmatic level.

C. *Modus Questions of Reasoning: Some Formal Characteristics, Structural Features and Quantitative Properties*

In text corpora the *WHY-model* can be both compound and simple sentences with plural forms dominating (54%).

TABLE III MQs of reasoning in the Russian National Corpus

Model	Number (%)
Why / What do you: SG : PL think	73
Why / What do you: SG : PL suppose	2
Why / What do you: SG : PL reckon	25
Total	100

D. *Modus Questions of Authorship: Some Formal Characteristics, Structural Features and Quantitative Properties*

This kind of interrogative constructions tend to occupy the postposition in replies consisting of multiple sentences (52,2%).

TABLE IV MQs of authorship in the Russian National Corpus

Model	Number (%)
Who told you	48
From what source do you SG : PL know	52
Total	100

Formal patterns of this MQ are represented by partial questions with compound and simple sentence structures occupying a reactive and rarely initiative position within dialogue units. MQs of authorship lack the redundancy (in opposite to *HOW-model*), and close relation to the category of evidentiality not grammaticalised in Russian [4, 5].

Besides stating the authorship these types of MQs often express surprise (even unpleasant) by the fact that the speaker is well informed and also disagreement with an utterance that can be disputed, e.g.

(5) *Losja's face began to break out with blemishes, his eyes flashed angrily: — Who told you such nonsense? — Then why can't they release him?*

‘Litso Losja pošlo pjatnami, glaza serdito vspykhnu: — Kto tebe skazal takuju erundu? — Togda počemu ego ne vypuskajut?’ [M. Petrosjan. *Dom, v kotorom...* (2009)].

This MQ belongs to the intermediate sphere (between the centre — questions in the dictum — and the periphery — questions in the modus) of functional and semantic category of questioning [6]. The fact that the semantic component in the focus of question can be interpreted as a part of the physical world (i.e. as a part of the dictum) is an argument for including MQ of authorship to the given sphere.

E. *Reactions to the Modus Questions: Comparative Analysis*

Reactions to the modus questions found in corpora can be grouped by their features (Fig. 1).

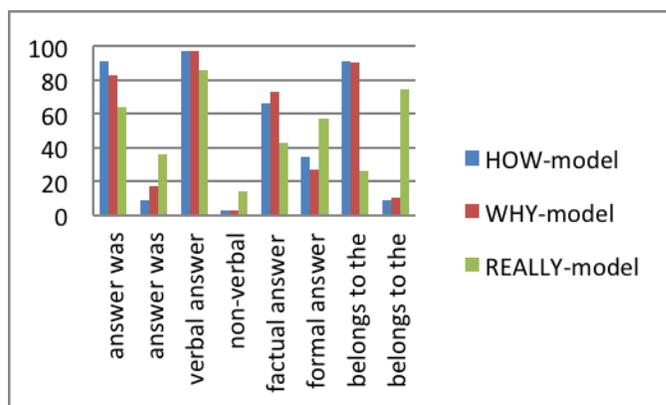


Fig. 1. Answers given to modus questions

The addressees tend to give explicit reactions to the questions rather than to ignore them and prefer a verbal way of answering, e.g.

(6) — *Granny, tell me, is it true that those who swear then have their teeth fall? — And who told you that? — Uncle Pasha.*

‘— Babulja, skaži, a eto pravda, što te, kto rugajutsja, u togo potom zuby vypadajut? — Eto kto tebe skazal? — Djadja Paša.’ [V. Durnenkov. *Mir molitsja za menja* (2005) // «Majskije čtenija», 2002, №7].

The most successful for getting an answer proved to be the *HOW-model*, and the least successful was the *REALLY-model*, which is some kind of *YES/NO*-questions.

Non-verbal or latent reactions implying hesitation or uncertainty such as shrugging shoulders or other gestures are quite infrequent (about 3% in common).

When answering the questions based on the *HOW-model* the speakers either agree or disagree with the addressers not repeating modus or dictum of the interrogative utterances.

The dominant answers to the *WHY-* and *REALLY-models* contain a reasoning of the opinion, clarification or evaluation.

4. **Conclusions and further work**

The analysis shows that MQs, although not being very frequent in dialogues of native Russian speakers nevertheless represent a certain trait of modern spontaneous conversation.

The most widespread types of MQs in Russian dialogue are questions to the mental modus introduced by partial (pronominal) questioning and appealing to the epistemic status of the addressee (propositional aims of knowledge, opinion and his motivation or source of information).

In such situations MQs, occupying the initiative position, perform their primary or main functions. Secondary functions are often used in the reactive position and aid in keeping up the dialogue or have so-called phatic meaning.

MQs of supposition are structurally redundant but necessary from the viewpoint of successful communication with the model *How do you think?* being the most frequent among them (76%). The *HOW-model* proves to be the most prototypical one amongst the above-described models, however, being sometimes redundant. Constructions appealing

to the reasoning of the addressee's point of view as well as stylistically marked ones are less common. Along with MQs of reasoning MQs of authorship are necessary elements for an informative dialogue. And thus their redundancy is minimal.

MQs are found to be essential not only for successful communication of adults but also for the 'adult – child' dialogues [6]. Being a special semantic type of caregiver questions, which appeal to children's mental advancement, they become some kind of a trigger for the development of their 'subjective sphere' (children's theory of mind) [7, 8]. Thus the results obtained can be used in dialogue systems to create more natural questions and can also serve as a theoretical base for both cognitive and proper linguistic investigation of children's mental development via the special language tools of their caregivers

References

- [1] Ch. Bally, *Linguistique générale et linguistique française*. Leroux, Paris: PRESS, 1932.
- [2] Russian National Corpus, <http://ruscorpora.ru>.
- [3] M. V. Khokhlova, "Razrabotka grammatičeskogo modulja ruskogo jazyka dlja spetsial'noj sistemy obrabotki korpusnyh dannyx," *Vestnik Sankt-Peterburgskogo gosudarstvennogo universiteta. Ser. 9. Filologija, vostokovedenije, žurnalistika*, vol. 2, pp. 162–169, 2010 [in Russian].
- [4] V. V. Kazakovskaya, "Modusnyje voprosy (k probleme izbytočnosti)," *Acta Linguistica Petropolitana: Trudy Instituta lingvističeskikh issledovanij RAN*, vol. 6, pp. 225–242, 2010 [in Russian].
- [5] A. Aikhenvald, *Evidentiality*. Oxford: OUP, 2005.
- [6] V. V. Kazakovskaya, *Questions and Answers in "Adult – Child" Dialogue*, Moscow: URSS, 2011 [Vopros i otvet v dialoge "vzroslyj – rebenok"].
- [7] V. V. Kazakovskaya, "What helps children to acquire "the language of ToM"?", *Budapest Central European University Conference on Cognitive Development*, pp. 123–124, January 2013.
- [8] V. V. Kazakovskaya, "Functions of Questions in Adult–Child Dialogues (based on Russian)," *X IASCL*, p. 58–59, July 2005.