ТЕОРИЯ, МЕТОДОЛОГИЯ И МЕТОДЫ

O. A. Oberemko
SPEECH ACTION GROUNDS OF ALGORITHMIC CODING ANSWERS TO AN OPEN-ENDED QUESTION: A CASE OF DEFINING A GOOD INSTITUTION

SPEECH ACTION GROUNDS OF ALGO-RITHMIC CODING ANSWERS TO AN OPEN-ENDED QUESTION: A CASE OF DEFINING A GOOD INSTITUTION

ОБЕРЕМКО Олег Алексеевич — кандидат социологических наук, доцент департамента социологии факультета социальных наук Национального исследовательского университета «Высшая школа экономики», Москва, Россия.
E-MAIL: ooberemko@hse.ru
ORCID: 0000-0002-7967-2533

OBEREMKO Oleg Alekseevich [1] — Ph.D., Associate Professor
E-MAIL: ooberemko@hse.ru
ORCID: 0000-0002-7967-2533

[1] Department of Sociology, Faculty of Social Sciences, National Research University Higher School of Economics, Moscow, Russia

Аннотация. Упростить процедуру и повысить надежность кодирования ответов на открытые вопросы без потери содержания — одна из задач массовых опросов, в которых стремятся уловить динамику мнений. В статье на материале одного из омнибусов ВЦИОМа предложена процедура анализа данных, позволяющая (1) эксплицировать синтактико-семантическую структуру ответа, ожидаемого проектировщи-

Abstract. To reduce complexity and increase reliability of coding answers to open-ended questions are among the main targets of the survey methodology. Using the answers to an open-ended question recorded by the interviewers during a Russian national sample survey, the article demonstrates a procedure, which allows (1) explicating the syntactic-and-semantic structure of the answers, expected by the asked

ком вопроса, (2) реконструировать репертуар вопросов, которым соответствуют зафиксированные ответы респондентов, (3) диагностировать коммуникативную адекватность спроектированного вопроса, (4) уточнять инструкции интервьюерам, что из ответа обязательно записывать, и кодировщикам, как кодировать, и (5) в терминах несложной (школьной) грамматики разрабатывать коды для систематического предварительного кодирования ответов на открытые вопросы. Показано, что качественная (логико-семантическая) перспектива на коммуникацию в массовом опросе может быть полезна для выявления различий между языком (языками) респондентов и языком исследователя.

question designer, (2) reconstructing a repertoire of the questions, which were actually answered, (3) diagnosing communicative adequacy of the question as it was designed for a survey, (4) specifying instructions for interviewers how to record and for codifiers how to code the answers to an open-ended question, and (5) elaborating a comprehensive framework of school grammar categories to formal systematic (pre)coding answers to open-ended questions. The qualitative (logical-and-semantic) perspective on sample survey communication, it is shown, is extremely useful to grasp crucial differences between the respondents' tongue(s) and the pollster' one.

Problem statement

The open-ended question is a unique device to get the public opinion spoken by respondents with their own «tongues» in the sample surveys, to «permit answers not limited to recorded alternatives», and reveal «subtle and often valuable information about reality *from the point of view of the respondent (italics.— O.O.)*» [Montgomery,

Crittenden, 1977: 235]. Crucial for the quality of the whole enterprise of survey, open-ended questions coding is described contradictorily: as a «process often involving some fairly sophisticated judgments» and as «a routine operation that rapidly loses interest for many people» at the same time [Woodward, Franzen, 1948: 253]; a statement about fundamental role of the coding may be followed with a confession that «for most it is the uneventful… delay that comes between the early enthusiasm of data gathering and the excitement of data analysis»; a necessity to keep high quality control over the coding may juxtapose with routine use of «hired hands» for coding [Kammeyer, Roth, 1971: 60], etc.

It seems that the authors about open-ended questions rather prefer unifying and eliminating than revealing specificities in the respondents' perspectives, focusing on fighting against coding errors while the answers ascribed to the predetermined standard categories [Montgomery, Crittenden, 1977; Woodward, Franzen, 1948; Kammeyer, Roth, 1971; Morris, 1990], or even on arguing to replace totally open-ended questions with closed-ended ones to make the data processing easier. Shortly, in the survey practice the advantage of the open-ended questions is out of use and tended to be further reduced, presumably, due to high complexity and low reliability of the data processing.

What is going on with an open-ended question shaped in the terms of the pollster (researcher) tongue during survey interviews — this matter is not often focused by methodological analysts; the pollsters almost never know whether an open-ended question does always function as a standardized stimulus. The very idea to reconstruct recursively the wordings of the *really* answered question according to the recorded answers came to us while it was perceived that some thoughtful answers are hardly to take as a direct reaction to the stimulus-question.

So our aim is to reduce complexity and increase reliability of the coding answers to the open-ended questions process. We suppose that first of all, before starting coding we need to find out whether all the collected answers were actually answered to the same designed question, and, if not, to make a preliminary classification of every collected answer depending on whichever question was actually answered.

Here below are analyzed the answers to the open-ended question recorded by the interviewers during a Russian national sample survey: 'How do you think, what major requirements must meet a school, so it could be told good?'[1] We have in mind reducing complexity and increasing reliability of the process of coding answers to open-ended questions. We suppose that before starting coding we need to find out whether all of the collected answers were actually answered to the same designed question, and, if not, to provide a full repertoire of the questions actually answered. We presuppose that shaping the answer even to an adequately understood open-ended question a respondent often reshapes the latter in a convenient way to express her/his opinion in her/his genuine tongue[2]; and these (re)shaped answers often differently un/fit the open-ended question-stimulus.

---

[1] Originally: Как Вы считаете, каким главным требованиям должна отвечать школа, чтобы её можно было назвать хорошей?

[2] Cognitive interviews show that, actually, the same process starts while thoughtfully answering closed-ended questions but it interrupts intentionally by imposing to choose from the given cues.
As for 'genuine' (tongue), we put such a strong expression neglecting possible over-cooperation by the respondents in

This article develops a logical and semantic (cognitive) procedure to classify the questions actually answered. The procedure is to combine rigorous algorithms and heuristics, and strictly limit use of the latter. Such a combination, we guess, will allow obtaining more reliable data and minimizing loss of semantic information. Developing the procedure helps us:

1) to explicate the syntactic-and-semantic structure of the response, expected by the designed question,

2) to reconstruct the repertoire of the questions actually answered by respondents, and, consequently, to code the collected answers in the terms of the answered questions;

3) to define a communicative in/adequacy of the designed question according to the ratio of the answers matching the expected (designed) response;

4) to offer verisimilar explanation to the deviations from the expected response;

5) to reasonably eliminate from the analysis inadequate responses, which do match to semantically different questions;

6) to offer empirically grounded recommendations on the elements of the answers indispensable to record.

Empirical data

The answers to an open-ended question on criteria of a «good school» from a national survey (pencil and paper interviews) by WCIOM are analyzed [3]; the data matrix contains 759 utterances from 334 respondents selected with a filter question. Some presuppositions were inferred from the shape of the data matrix: (1) interviewers encouraged respondents to voice several criteria (more than one); (2) while inputting, the complex answers were fragmented into the utterances with a single criterion; (3) the shaped SPSS matrix allowed storing not more than 5 utterances.

The analyzed open-ended question was penultimate in a thematic block of questions about school choice; the block was prefaced by a filter question (whether the respondent's family has a school pupil) and continued with a series of closed-ended questions, asking to mark the importance of the listed 7 criteria in choosing a school «for your (grand)child: the proximity of the school to the home, availability of special classes, etc. The next was the open-ended question to be analyzed further:

How do you think, what major requirements must meet a school, so it could be told good? [4]

Obviously this thematic block is shaped so as to facilitate understanding by the respondents: the filter question bolts out inexpert respondents; a series of the closed-ended questions makes a respondent 7 times to recall the different aspects of the choosing a school situation. So the respondent has all the possibilities to get ready to understand adequately and answer a presumably difficult open-ended question; if

accepting the 'pollster's perspective', neither the perspective of the 'generalized other' (nor of the 'majority') because we believe the question is not a sensitive one.

[3] The description of the survey, see: [Shkoly, kotorye…, 2008].

[4] Originally: Как Вы считаете, каким главным требованиям должна отвечать школа, чтобы её можно было назвать хорошей?

so, any deviation from the expected answering mode is caused by any other reason but misunderstanding.

Theoretical and methodological frame of the answered questions reconstruction

Two basic elements constitute the theoretical and methodological frame of the answered questions reconstruction: the first one informs us with the optics how to see a conversation as an interaction in institutional settings; the second one gives us the landmarks where to search deviations from the expected answers.

The first foundation is the theory of conversational implicature by H. P. Grice, which is based on a so called cooperative principle, an indispensable one for any successful conversation:

> *«Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction. This purpose or direction may be fixed from the start (e. g., by an initial proposal of a question for discussion), or it may evolve during the exchange; it may be fairly definite, or it may be so indefinite as to leave very considerable latitude to the participants (as in a casual conversation). But at each stage, SOME possible conversational moves would be excluded as conversationally unsuitable. We might then formulate a rough general principle which participants will be expected (ceteris paribus) to observe, namely: Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged. One might label this the COOPERATIVE PRINCIPLE».* [Grice, 1975: 45].

Here a sociological analyst of survey conversations has some important points: a conversation is taken as (1) a succession of *connected* remarks, (2) produced with *a common set of purposes, or at least a mutually accepted conventions or reciprocally supported expectations*. These two characteristics make reasonable the very idea to reconstruct the actually answered questions from the recorded answers.

Specifically for a survey interview, which is in no way a *casual conversation* (3) its purpose or direction is *strictly defined* from the start by the interviewer (who actually represents the researcher). The asymmetric nature of a survey conversation points the uniquely privileged purposes as common and the uniquely privileged direction to be mutually accepted; they are predetermined by the survey general methodology and by the specific tasks of a specific survey. (4) In the perspective of the survey tasks, all the unsuitable conversational moves should be deprived of support and kept out the records. Eliciting the answered question from recorded answers, we presuppose that (5) the respondent and the interviewer both *make their conversational contribution such as it is required,* according to the *cooperative principle*.

The very *cooperative principle* is explicated with 4 categories: Quantity, Quality, Relation, and Manner. Each category is operationalized with their maxims:

1) Quantity or sufficiency of information: to avoid both deficit and overinformativeness, equally risky to confuse or mislead the partner[5];

---

[5]   H. P. Grice put it in 2 maxims: «1. Make your contribution as informative as is required (for the current purposes of the exchange). 2. Do not make your contribution more informative than is required.» [Grice, 1975: 45].

2) Quality or truthfulness: to say what one believes to be true, or to say on one's own behalf[6];

3) Relation or relevance: to keep the topic (with the only maxim: 'Be relevant.' [Grice, 1975: 46];

4) Manner or being «perspicuous»[7].

Mutual supporting the expectations of the parts is based on willingness of each other to coordinate one's own speech actions[8]. A deviation of a single maxim (presumption) will challenge to the whole cooperative principle, and enhances risks to interrupt communication.

These maxims are consonant with any sociological theory, which takes *Vergesellschaftung* as interactions of agents considering mutual expectations in a certain institutional setting. Basically, H. P. Grice considers any conversation as a socially constructed pattern, i. e. as an invariant being constantly negotiated and kept sustainable through variety of realizations[9].

The theory of conversational implicature by H. P. Grice has been found useful in empirical studies of question-answer paradigm [Beaver, Brady, 2008] and recognized as adequate for researching sample survey conversations [Sudman, Bradburn, Schwarz, 2003]. Such recognition is reasonable, because mutual willingness to coordinate one's own speech acts is basic for a sample survey as a democratic social institution.

In our case the categories and their maxims are supported with the general rules of running sample survey interviews: uniform pattern informed by the instructions and conventionalities of a formal conversation; all the questions are clear and relevant for the respondents; interviewers record the answers accurately, so the analysts receive the collected opinions with no serious distortions; interviewers are to collect socially significant data and their social function is considered by respondents while they dis/ agree to support the conversation. These categories, maxims and conventionalities constitute an analytic frame and presumably prevent our reconstruction from meaningful deviations.

The reconstruction consists of just simple operations permanently used by any participant in everyday questions and answers conversations. The only specificity is that in everyday conversations, while shaping a question to ask, we try anticipating the answer. Our reconstruction is made in the reverse direction: we search the meaningful questions, which fit the recorded answers. As in case of any part of an everyday questions and answers conversation, these maxims constitute an analytic frame and presumably prevent our reconstruction from meaningful deviations, being helpful for triangulating.

As an example we take the abridged answer «добротой учителей», which means «kindness of the teachers» in the instrumental case — «with kindness of the teachers».

---

[6] H. P. Grice composes a supermaxim and two specific maxims — 'Try to make your contribution one that is true': '1. Do not say what you believe to be false. 2. Do not say that for which you lack adequate evidence.' [Grice, 1975: 46].

[7] With 4 maxims: «1. Avoid obscurity of expression. 2. Avoid ambiguity. 3. Be brief (avoid unnecessary prolixity). 4. Be orderly.» [Grice, 1975: 46].

[8] Here is rather referred to 'social action' by M. Weber than to 'speech act' (by numerous authors) — *O.O.*

[9] The notions of invariant and variety of realizations allow using the theoretical frame of the transformational generative grammar by N. Chomsky.

The procedure consists of 3 steps: firstly, checking whether the answer matches to the presumably asked question; secondly, developing the abridged answer into a syntactically complete utterance (with grammatical subject and predicate), thirdly, finding semantically and syntactically adequate question to the developed answer:

> **Initially given (presupposed)**
>
> Question, presumably asked by the interviewer: How do you think, what major requirements must meet a school, so it could be told good?
>
> **Reconstruction steps**
>
> (1) Developing the incomplete answer: Добротой учителей [отличается хорошая школа (от плохой и средней)][10] — With kindness of the teachers [a good school differs (from bad or average ones)].
>
> (2) Inference: the recorded answer does not match to the designed question wording.
>
> (3) Reconstructed question: [Чем отличается хорошая школа (от плохой и средней)?] — [What does a good school differ with (from bad or average ones)]

This example demonstrates how useful recording the nouns in an oblique (here instrumental) case is. Similar cognitive operations are typical for developing a questionnaire when a researcher accepts the role of respondent [Maslova, 1990: 77], being constitutive in the cognitive analysis of questionnaire [Schwarz, 2007].

The other basic element of the theoretical frame of the reconstruction is the Topical Structure Analysis (theory of topic–comment structure of a sentence; Functional Sentence Perspective). Every sentence (explicite aut implicite) has two parts: (1) topic, or theme, which refers to 'what the sentence is about' and to «what is known or at least obvious in a given situation and from which the speaker proceeds in his discourse,» and (2) enunciation, or comment, or rheme, which refers to 'what is said about' the theme and adds to the discourse new (unknown from the perspective of addressee) information for which the sentence is produced [Mathesius, 1967: 239].

Similarly, discussing «the logical structure of question» in a self-administered questionnaire, Olga M. Maslova distinguishes 3 elements: «The first element is an information supposedly equally known to the person who asks and to the person who is asked, and from whom an answer is expected; this is presumably the common point of reference for the asking one and the answering one, a platform for conversation. The second element indicates the deficit in the known information, what is to be known. The third element denotes the field the answer to be looked for, defines the room to search missing information, marks the ground the answer to be built on» [Maslova, 1990: 69]. The third element occurs only in closed-ended questions.

Our open-ended question presupposes as known the existence of schools and the possibility to qualify some of them as good. The informational deficit (rheme) is indicated by the interrogative pronoun «what», which is added with the room to search the answer: a good school is to be described in the terms of *the major requirements it must meet*. This room is defined in limits, and is not (strictly) defined in possible contents; so this room *offers*, or even чч, a respondent to express her/his opinions in her/his own tongue, i. e. *freely*.

---

[10] Here and below, the bracketed text is the result of reconstruction.

Taking the conversation as a unit, the question rheme refers to its interrogative part; the answer rheme refers to the utterance containing the properties of a good school. Excluded from the original wording of the question its phatic component 'How do you think'[11], the topical structure of the designed dialogue is shown in Fig. 1. It contains the expected invariant for the answer rheme; all the other variants should be considered deviant from the cooperative principle.

	Rheme	**Theme**
Question:	**What** major requirements must meet	a school, so it could be told good

	Theme	**Rheme**
Expected answer**:**	[In order a school could be told good, it]	[must meet **such** major requirements]

Figure 1. Topical structure of the designed question-answer conversation

The cooperative principle is often added with a principle of economy, when the words combinations from the asked question are used in the answer as ready building blocks. B. M. Gasparov links such repeats with a «reproductive strategy of verbal behavior» [Gasparov, 1996: 57], and N. D. Arutyunova called them «dialogical citations» [Arutyunova, 1990].

So any question (1) offers the asked person a room to express freely her/his opinions in her/his own tongue, and at the same time (2) imposes an expected frame of answering with ready-to-serve lexical and syntactical units. When accepted the help (echoing the thematic parts of the question), a respondent can focus on the task how to complete the available space, i. e. to express her/his opinions completely in her/his own language but in an externally (with the question) defined — semantic and, broadly, social space. (Happily for the mass survey enterprise, a respondent cooperatively answering an open-ended question has no chance to do without her/his own language.)

Here we can shape a practical rule of thumb how to tell deviations from mere variances of the expected pattern: renouncements from (the «help» in) reproducing *thematic* limits are to consider deviation; any walk 'indoors' the *rhematic* space, while given thematic limits accepted, denotes freedom of expression. The former are to exclude from a further analysis, and the latter are worth to be coded and «noded».

Echoing the thematic parts of the question, a respondent makes the question *closer-ended*, as in the following example:

— **What** *is your name? — My name is* **so-and-so**.

The absence of a list of names does not make the question absolutely open-ended, because a respondent has a strictly narrow room to shape the answer in her/his own tongue, dramatically rejecting the one of the researcher. A free room can be found in no answer or in a metaphorical one, but the both ways are to be deviations off the cooperative principle. This situation is common for a social interaction.

---

[11] It may be omitted from the analysis for its function is reduced to establishing contact with the addressee.

Now we can review our steps of reconstructing the real conversation from the perspective of a 'good' respondent who takes the trouble to look for the genuine expression of her/his opinion in a more or less narrow room of an open-ended question. Reviewing our procedure of reconstructing is aimed to become closer to an algorithmic way of thinking.

Typology

Now we proceed to review the classified typical transformations of the designed question-answer conversation, which were elicited from the procedure exposed above, see the Table 1 below. This table is composed with no references to any Grammar of the Russian Language, but exceptionally with empirical references to the database of 759 utterances from 334 respondents.

Columns

Column 1 contains mere signs of types and subtypes.

Column 2 contains typical (the most complete) examples from the data base with recorded answers, as they are available to data analysts. We may infer that the records typically contains the words (*in italics*), which constitute the very 'core' of the rheme — the most significant words, without which an answer loses its sense for an addressee.

The data of Column 2 allow forming Column 3, which represents the reconstructed models of the recorded answers and their invariant markers, which indispensably present (or logically (syntactically) restored) in the empirical data. We can distinguish *verb phrases* and *noun phrases*, the latter requiring much more (sociological) imagination than the former to shape models of the full answer. But analytically (sociologically) this distinction is rather useless dealing with the Russian language, whose usage, at least compared with the Germanic and Roman languages, tends to neglect (omit) the verbs to focus on the nouns.

We can see that all the verb types are realized using the whole range of the principal verb's meanings, i. e. the whole range of the relations between subject and predicate: identity (*be* or its substitutes), possession (*have* or its substitutes), and action (notional verbs).

Grasping the models of answers with their invariant markers, we got able to reconstruct the answered questions in Column 4. Able to restore empirical answer-and-question exchanges, we can tell 3 ways of thinking: normative, descriptive and actionist thinking. This classification relies on the modality, mood and aspectuality of the verbs-predicates.

Basic (designed by researcher) thinking, type 0, corresponds to the idea that Good School must meet wishful requirements.

Normative thinking is expressed:
— in type 1 with «must + any convenient verb in the infinite form» and with 'school' or any convenient attribute or actor as the grammar subject of the utterance;
— in type 2 with impersonal predicate translated as «it is needed» necessarily continued with a subordinate clause in the subjunctive mood where 'school' or any convenient attribute or actor as the grammar subject of the utterance having its own predicate.

We can see that with the respondents' tongue(s) both types 1 and 2 draw the same pictures where school is thought:
— as an elementary entity or actor, which is imposed with responsibility for wishful things presence;

— as a social space, where wishful things must be or some actors must act in a wishful way.

So the only difference between types 1 and 2 concerns the choice between the modal verbs 'must' and 'need', being the difference of the predicate between infinite form and subjunctive mood just syntactically, i. e. here meaninglessly predetermined. So actually we can merge types 1 and 2 into a single Normative Type; its normativity is probably informed by the thinking of basic question.

General Questions of the normative thinking may sound as follows:

What must

What is needed to   *be (done) in a school, so it could be told good?*

Descriptive thinking of type 3 just ascertains a wishful state of affairs in the same syntactical and grammatical ways, as in the **normative thinking**, except the substitution of the modal verbs formulae with more simplistic using of predicates in the (Russian language analogue of) Present Indefinite Tense. This is the only descriptive type where verbs-predicates are explicitly recorded or/and sounded systematically.

Sociologically, the difference between normativity and descriptivism may be interesting, only if a stratifying capacity is suspected in it; otherwise the Occam's razor is to be used. So ignoring the difference between normativity and descriptivism we can elicit the 7 preliminary variants of the 'actually answered' answers to the basic question about a good social institution.

A good school is represented as a **thing** or **actor** of itself, and:
- a. a good school is of a certain (wishful) kind: *clean; close to home*;
- b. a good school has something: *modern equipment*;
- c. a good school does something (to somebody): *gives good knowledge to its pupils*.

A good school is represented as a **space** where a **thing** or an **actor** *just is* or *is of a certain kind*, and an **actor** does something (to somebody):
- d. inside a good school there is just something (or somebody): *feeding; [guards]*;
- e. inside a good school there is something (or somebody) of a certain kind or with a certain characteristics: *free of charge education*; *every pupil with individuality*;
- f. inside a good school somebody does something to somebody: *teachers give children good education*.

If we accept that 'a good school has…' is semantically synonymous to 'inside a good school there is…', variants d – e could be seen as syntactical elaborations of variant b. This new look allows perceiving that only variant f represents a good school as a scene for actors to perform their actions, being the space interpreted as a simple receptacle of attributes in variants d – e, so they can be cut, having saturated variant b.

All the types 4 and 5 utterances were recorded verbless in the database, and have no other predicates. However the specific verb-predicate of type 4 is easily guessed because of a very specific mark — instrumental case of the noun recorded; the answered question is easily reconstructed:
— What does a good school differ with?
— *[A good school differs] with good reputation.*

Type 5 contains just nouns (or noun phrases with a main noun) in the Nominative (Common) Case. Absence of a full-meaning verb leaves to the answers of the type

the only possible meaning — a mere existence of the referents to the recorded nouns, which makes a school of a good kind:

— What makes a good school?

— *The teachers' moral authority [does].*

In terms of a–f-classification, the verbless nouns types 4 and 5 semantically fit to b-group: 'a good school has smth'. Now we can provide a new and thrifty typology of way of thinking about a good school:

A. A good school is of a certain kind: *clean; close to home*;

B. A good school has something (or somebody) (of a certain kind): *modern equipment*; *[guards]*; *education — free of charge*; *every pupil — individuality*;

C. A good school does something (to somebody): *gives good knowledge to its pupils*.

F. Inside a good school somebody does something to somebody: *teachers give children good education*.

This new typology ignores both the variances in predicates modes and syntactic specificities of utterance composing. So we bravely tell ABCF *invariants* of the *common* thinking about a good school (or even about a good social institution in general).

The last type 6 lexically and grammatically expresses starting or finishing an action, i. e. it expresses *aspectuality*, and at the same time retains a normative shade. This mixture bears an actionist look at a certain single school or generally at the school education and demands of urgent particular changes.

Theoretically, the *aspectual thinking* can totally replicate the ABCF *invariants* of the *common* thinking:

AAsp. A good school is needed urgently to become: *clean; close to home*;

BAsp. A good school is urgently needed: to get (or throw out) *modern equipment*; to hire (to fire) *[guards]*; to start (stop) giving *education free of charge*; to raise (kill) in *every pupil individuality*, etc.

Again, from the sociological perspective, the difference between *common* and *aspectual* thinking may be interesting only if a stratifying capacity is suspected in it; otherwise the Occam's razor is to be used. In this case all the utterances of type 6 will be distributed between ABCF classes.

Prerequisites for algorithmisation

The revealed above speech action grounds compose a comprehensive framework of school grammar categories to formal systematic (pre)coding answers to open-ended questions. Reviewing all the empirically available transformations of the designed invariant of a survey conversation demonstrates the syntactical and semantical limits for deliberate searching respondents' own tongues answering the researcher's question. The Russian language, being synthetic with a differentiated system of flexes, gives advantages to reveal some prerequisites for formalistic algorithmisation in analyzing samples of answers to open-ended questions, where the analytical languages lacks systematic linguistic markers.

Some of such linguistic markers can mark and contribute to explicate a specific manner of talking about the subject-matter in question. Rich data in content may allow saturating the manner of talking to the status of (sub)discourse; in this case we enlarge the room of interpreting the same data.

All the linguistic markers revealed can be easily formalised, and the utterances containing the *direct / explicated* markers of the 1st range can be automatically / algorithmically classified (framed).

The rest utterances with no *explicated* markers (a) can be classified either according to secondary markers of semantic (not syntactical) nature, (b) either can be classified non-alternatively, i. e. to be put into several (more than only one) classes according to the cooperative principle, especially to its Quantity and Relation maxims. Classifying utterances difficulties will show necessary specifications for interviewers instructions how to make short records and for codifiers how to code the answers to an open-ended question; first of all, all the *verbal forms* sounded by the respondents and the *oblique cases of key nouns* should be scrupulously recorded.

Having got a linguistically grounded framework for classifying the utterances, we move closer to *sociologically* (or of other discipline) *relevant* primary categorisations; the ultimate relevancy of categories is defined by the tasks of the study.

Firstly, we can distinguish 3 ways of thinking: normative, descriptive and actionist, — possibly reflecting socially different attitudes to and different characteristics of representation of the social object in question.

For in our case, a bias to normative thinking is communicated by the Question, the descriptive and actionist manners can have a substantial power of differentiating (stratification) the respondents' sample; because the respondents choosing the descriptive or actionist manner should have overcome the normative bias, while looking for one's own tongue to answer. The descriptive manner may mean indifference to the topic, and the actionist manner, on the contrary, may mean a vehement involvement in the problem.

Secondly, we have an analytic framework of the social object in question. In our case we have the lists of the relevant actors and their (un)wishful actions, of the involved things and their qualities; all of them take their parts in making a school — a social object — good or/and bad in the public opinion. These lists express the public expectations expressed in their own categories, which can be easily transformed into the codebook. These codes will better reflect the data, than any *ab ovo* and *ad hoc* formed codebook.

Starting with typology of respondents' transforming the asked question, we are close to achieve an empirically tested framework of all the possible responses to the asked question, which meet the cooperative principle by H. P. Grice. Such a framework is useless when the open-ended questions are asked to test previously formulated hypotheses; and when researcher finds important to check similarity between the respondents' representation of the social object in question and her/his own, a previously elaborated framework of the object will save time and forces.

For sociology is interested in a finite number of social objects, and ask a finite number of typical open-ended questions about the finite number of social objects, accumulation of such analytic frameworks is finite in time. Such accumulation promises to reduce complexity and increase reliability of coding answers to open-ended questions.

Further elaboration will indispensably include at least (a) empirical survey data of answers to open-ended questions with different types of communicative structure, and (b) 'case grammar' and 'frame semantics' by Ch. J. Fillmore [Fillmore, 1968].

*Table 1. **Transformations of the designed invariant of the conversation: thesaurus***

Types	Data Base typical examples	Models of the recorded answers and their invariant MARKERS[12]	Fitted (reconstructed) question	Core idea (focused meaning)
1	**2**	**3**	**4**	**5**
		MUST MEET REQUIREMENTS		**Basic thinking**
0	Only 8 examples just partly correspondent to the expected model	a Good School **MUST MEET** some / somebody's **REQUIREMENTS**	What major requirements must meet[13] a school, so it could be told good?	Good School must meet wishful requirements
1		**MUST + Verb(INF)**[14]		**Normative thinking**
1a	[a school must][15] 1) *be clean*; 2) *have modern equipment* 3) *give good knowledge*	A **school MUST** **1) BE(INF)** somewhat **2) HAVE(INF)** Smth **3) DO(INF)** Smth	What must a school be / have / do, so it could be told good?	School as an elementary entity or actor, placed with responsibility for wishful things presence
1b	[in a school] 1) *feeding must be* [organized] 2) *education must be free of charge* 3) *every pupil must be an individuality* 4) *teachers must give children good education*	In a school 1) an Attribute **MUST BE(INF)** 2) an Attribute **MUST BE(INF)** Smwhat 3) an Actor **MUST BE(INF)** Smwhat 4) an Actor **MUST DO(INF)** Smth (to Smbd)	What or who must 1) be 2—3) and what must they be like 4) do (smth) in a school, so it could be told good?	School as a social space where wishful things must be and/ or wishful actors must do wishful things
2		**IT IS NEEDED + Verb(SUBJ)**[16]		**Normative thinking**
2a	[it is needed that a school] would 1) [be] *close to home* **2) no examples** 3) *give decent education*	**IT IS NEEDED** that a school 1) **BE**(SUBJ) somewhat 2) **HAVE** (**SUBJ**) Smth 3) **DO**(SUBJ) Smth	What is needed that a school would be like / do, so it could be told good?	School as an elementary entity or actor, placed with responsibility for wishful things presence
2b	[it is needed that in a school] 1) *feeding would be* 2) *feeding would be better* **3) no examples** 4) [some they] *would give knowledge* [to the pupils]	It **IS NEEDED** that in a **school** 1) an Attribute **BE(SUBJ)** 2) an Attribute **BE(SUBJ)** Smwhat 3) an Actor **BE(SUBJ)** Smwhat 4) an Actor **DO(SUBJ)** Smth (to Smbd)	What is needed that would be (like) / be done in a school, so it could be told good?	School as a social space where wishful things must be and/ or wishful actors must do wishful things

Types	Data Base typical examples	Models of the recorded answers and their invariant MARKERS[12]	Fitted (reconstructed) question	Core idea (focused meaning)
1	2	3	4	5
3		**Verb(PRES IND)**[17]		**Descriptive thinking**
3a	[A good school] 1) *is safe / repaired* **2) no examples** 3) *prepares a lot of alumni with honor*	A good school 1) **BE(PRES IND) + ADJ / PART II** 2) **HAVE(PRES IND)** Smth 3) **DO(PRES IND)** Smth	What is a good school like?	School as an elementary entity or actor, which already is of wishful kind or have or do wishful things
3b	[A good school is that, where**]** + [they] *feed children properly* 4) *teachers relate well to children*	A good school is that, where 1) an Attribute **BE(IND)** 2) an Attribute **BE(IND)** Smwhat 3) an Actor **BE(IND)** Smwhat 4) an Actor **DO(IND)** Smth (to Smbd)		School as a social space where wishful things already are and wishful actors do (have done) wishful things
4—5		**Noun or noun phrases**		**Descriptive thinking**
4	[A good school differs] + *with good reputation*	A good school differs **N(INSTR)**[18]	What does a good school differ with?	Statement of a sinequanon feature; *an adjective with positively differing meaning usually added*
5	*Teachers' moral authority* [makes a school of a good kind]	**N(Nom)**[19] makes a school of a good kind	What makes a school of a good kind?	Statement of a sinequanon feature; *no adjectives at all, just a noun as an important quiddity as it is*
6		**Lexically or grammatically expressed starting or finishing an act**		**Actionist thinking**
	[in a Good School it is needed] + *to stop testing* + *more funding schools*	In a Good School it is needed to stop / start **Verb(INF)** Smth (for Smbd)	What is needed to stop / start doing in a school, so it could be told good?	Statement of necessary changes in a school or generally in school education

[12] Invariant elements of the recorded answers are written in bold capital letters.
[13] Pink denotes the rheme of questions.
[14] Verb(INF) = any verb in the infinitive form.
[15] The elements taken in [] are implicit and reconstructed.
[16] Verb(SUBJ) = any verb in the Subjunctive Mood.
[17] Verb(PRES IND) = any verb in the Present Indefinite.
[18] N(INSTR) = a noun in the Instrumental Case.
[19] N(Nom) = a noun in the Nominative (Common) Case.

References

Arutyunova N. D. Fenomen vtoroy repliki, ili O pol'ze spora (Phenomenon of Reply, or About the Benefits of Dispute) (in Russian) // Logicheskiy analiz yazyka : Protivorechivost' i anomal'nost' teksta (Logical Language Analysis: Inconsistency and Abnormality of the Text). M. : Nauka, 1990. S. 175—189.

Gasparov B. M. Yazyk, pamyat', obraz. Lingvistika yazykovogo sushchestvovaniya (Language, Memory, Image. Linguistics of the Language Existence) (in Russian). M. : NLO, 1996.

Maslova O. M. Vopros kak instrument sbora empiricheskikh dannykh (Question as an Instrument of Collecting Empirical Data) (in Russian) // Metody sbora informatsii v sotsiologicheskikh issledovaniyakh (Methods of Collecting Information in Sociological Researches). Book 1 / Ed. by V. G. Andreenkov, O. M. Maslova. Moskva : Nauka, 1990.

Mathesius V. O tak nazyvaemom aktual'nom chlenenii predlozheniya (On the so called topic — comment structure of a sentence) (trans.; in Russian) // The Prague Linguistic Circle. Moskva: Progress, 1967. P. 239—245.

Shkoly, kotorye my vybiraem (The Schools We Choose) (in Russian) // WCIOM. Press-vypusk (Press Issue). 2008. No. 1035. Sept. 1. Available at: http://www.socpolitika. ru/rus/social_policy_research/applied_research/document8776.shtml (accessed: 25.10.16).

Beaver D. I., Brady C. Z. Sense and Sensitivity: How Focus Determines Meaning. Malden, MA: Blackwell Publishing, 2008.

Fillmore Ch. J. The Case for Case // Bach and Harms (Ed.): Universals in Linguistic Theory. New York: Holt, Rinehart, and Winston, 1968. P. 1—88.

Grice H. P. Logic and Conversation // Syntax and Semantics. Vol. 3: Speech Acts / Ed. by P. Cole, J. Morgan. N.Y.: Academic Press, 1975. P. 41—58. Available at: http://www. ucl.ac.uk/ls/studypacks/Grice-Logic.pdf (accessed: 25.10.16).

Kammeyer K. C.W., Roth J. A. Coding responses to open-ended questions // Sociological Methodology. 1971. Vol. 3. P. 60—68.

Montgomery A. C., Crittenden K. S. Improving Coding Reliability for Open-Ended Questions // The Public Opinion Quarterly. 1977. Vol. 41. No. 2. P. 235—243.

Morris R. J. Occupational Coding: Principles and Examples // Historical Social Research / Historische Sozialforschung, 1990. Vol. 15. No. 1. P. 3—29.

Schwarz N. Cognitive Aspects of Survey Methodology // Applied Cognitive Psychology. 2007. 21 (2). P. 277—287.

Sudman S., Bradburn N., Schwarz N. Thinking about Answers: The Application of Cognitive Processes to Survey Methodology. San Francisco: Jossey-Bass, 2003.

Woodward J. L., Franzen R. A Study of Coding Reliability // The Public Opinion Quarterly. 1948. Vol. 12. No. 2. P. 253—257.