

- McCloskey, James and Kenneth Hale. 1984. 'On the syntax of person-number inflection in Modern Irish'. *Natural Language and Linguistic Theory* 1, 487-533.
- Sigurd, Bengt. 1987. 'Referent Grammar (RG) in computer comprehension, generation, and translation of text (SWETRA)'. *Working Papers* 31, 137-67. Lund: Dept of Linguistics.
- Sigurd, Bengt. 1988 (forthcoming). 'A referent grammatical analysis of relative clauses'. *Studia Linguistica Hafniensia*.

Lund University, Dept. of Linguistics  
*Working Papers* 33 (1988), 103-121

## Some Types of Russian and Polish Interrogative Sentences

– A Preliminary Referent Grammatical Analysis

Barbara Gawrońska-Werngren

### INTRODUCTION

In the preliminary analysis presented in this paper, referent grammatical rules of the type developed in Sigurd 1987 will be applied to some examples of Russian and Polish *yes/no* questions and *wh*-questions. Our aim is to investigate the possibility of computer translation between English and Swedish and the Slavic languages by means of RG – a theory that the computer parser used by SWETRA (Swedish Computer Translation Group, Lund) is based on.

As RG is a phrase structure grammar, inspired by GPSG (Gazdar, Klein, Pullum and Sag 1985), such constructions as interrogative sentences and relative clauses are not analyzed as results of transformations, but as structures containing a topicalized constituent and a defective sentence. This means that a question like *who hit him* is analyzed as (using a simplified notation):

$$\text{esent}(q,T,P,F) \rightarrow \text{enpqs}(P), \text{esdsent}(\_,\_,P,s(\text{subj}(P),\text{pred}(B),\text{obj}(X))).$$

who                      who      who      hit      him

q = question

T = tense

P = the focused constituent

F = functional representation

esdsent = English subject defective sentence

enpqs = interrogative noun phrase (subjective form)

The questions introduced by a finite verb are analyzed as verb defective (or aux-defective); if the objective *wh*-word is fronted (*whom did he hit*), the sentence following the interrogative pronoun is classified as object defective, and so on. The computer translation is based on the functional representation of a sentence,

i.e. on the terms 'subject', 'predicate', 'object', 'adverbial' and 'sentence adverbial'.

The other kind of sentence representation, the so-called categorial representation, based on such terms as subjective or objective noun phrase, finite verb, infinitive etc., allows one to denote the constituents in the same order in which they occur in a particular language; as the function of each variable used in the categorial representation is marked in the functional representation F, a quite adequate translation is possible in spite of language-specific word order rules.

### RUSSIAN AND POLISH YES/NO QUESTIONS WITHOUT INTERROGATIVE PARTICLES

The main difference between the English and Swedish *yes/no* questions on the one hand, and the Russian and Polish on the other hand depends on the fact that the Slavic languages concerned here do not use word order as a primary marker of the interrogative mode. The syntactic order in a Russian or Polish question is relatively free (different constituents may be fronted) and the interrogative function is often marked by the sentence intonation only. This kind of interrogative constructions can be illustrated by the following examples:

- (1) rsent: ty pročítal ètu knigu?  
           you read,prf this book
- (2) rsent: ètu knigu ty pročítal?  
           this book you read,prf                   'have you read this book?'
- (3) psent: (ty) przeczytałeś tę książkę?  
           (you) read,prf this book

r= Russian

p= Polish

The English equivalent of the sentences (1), (2) and (3) is (simplified) represented as:

- esent(q,\_,have,s(subj(you),pred(have,inf(read)),obj(np(dem(this),n(book)))))) →  
           → ev(have),evdsent(\_\_\_\_,have,F).
- evdsent(\_\_\_\_,have,s(subj(you),pred(have,inf(read)),obj(np(dem(this),n(book)))))) →  
           → epro(you),einf(read),enp(np(dem(this),n(book))).

The Russian sentence (1) has the same word order as the corresponding declarative sentence *ty pročítal ètu knigu* 'you have read this book', but the value

of the mode variable (q) excludes the possibility of translation to a declarative English or Swedish sentence. The fronted constituent is the personal pronoun *ty* 'you', which functions as the subject; therefore, *ty* is to be placed in the third slot in the 'rsent', and the whole sentence is analyzed as consisting of a focused noun phrase in the subjective form (realized as a pronoun – 'rpros') and a subject defective sentence ('rdsent'). In example (2), the direct object is focused; thus, the sentence is analyzed into an accusative noun phrase ('rnpa') and a sentence lacking an accusative object ('radsent'). In both cases, the functional representations are analogous to the the description of 'esdsent' presented above, which allows one to identify the functions of all constituents in a correct way:

- (1)  
 rsent(q,\_,ty,s(subj(ty),pred(pročítal),obj(np(dem(ètu),n(knigu)))))) →  
           → rpros(ty),rdsent(\_\_\_\_,ty,F).

- rdsent(\_\_\_\_,ty,s(subj(ty),pred(pročítal),obj(np(dem(ètu),n(knigu)))))) →  
           you       you       read,prf                   this   book  
           → rvt(pročítal),rnpa(np(dem(ètu),n(knigu))).  
           read,prf                                   this   book

- (2)  
 rsent(q,\_,np(dem(ètu),n(knigu)),s(subj(ty),pred(pročítal),obj(np(dem(ètu),n(knigu)))))) → rnpa(np(dem(ètu),n(knigu)),radsent(\_\_\_\_,np(dem(ètu),n(knigu)),F).

- radsent(\_\_\_\_,np(dem(ètu),n(knigu)),s(subj(ty),pred(pročítal),obj(np(dem(ètu),n(knigu)))))) → rpros(ty),rvt(pročítal).  
           book                   you   read,prf                   this

The perfective aspect of the sentence may be expressed by inserting an appropriate value in the tense-slot (after 'q'), but the quite complicated problems of tense and aspect relations will not be discussed in this paper.

The Polish interrogative sentence (example(3)) can be described by formulas analogous to those used in (1), if the personal pronoun *ty* is lexically realized. The use of personal pronouns in their subjective forms is optional in Polish, as the information about the grammatical person is carried by the inflected form of the verb. In general, the subjective personal pronouns are realized only in the cases of emphasis; the constructions where the pronoun is deleted are more common.

In the analysis of Polish 'subjectless' sentences one has to take into consideration the fact that they are functionally equivalent to those containing an appropriate personal pronoun. It would be inadequate to describe the 'subjectless' constructions as consisting of a verb and a verb defective sentence, although the finite verb is the first lexical item occurring in the sentence. The functional representation of sentences lacking lexically expressed personal pronouns must contain this grammatical information about the subject that is necessary for establishing translation; without inserting certain grammatical properties in the subject-slot we would not be able to translate such well-formed Polish questions, as e.g. *czytasz?* '(you)read' – 'are you reading?'.

In the example (3), the verbal form is carrying information not only about the grammatical person, but also about the subject's number and gender (here: singular, male). In the plural, there are different verb forms used for subjects with the properties +male and +human, and for those that do not have this combination of properties, e.g.:

śpiewaliście?  
 (you(plural,male,human)),sing  
 śpiewałyście? 'were you singing?'  
 (you(plural,non-male,human)),sing  
 or:  
 (you(plural,male,non-human)),sing.

As we can see, the referent of the subject can be described with several properties – even in these cases where the subject is not realized as a separate lexical item. Therefore, we suggest that both variants of the sentence (3) can be treated as comprising subject defective sentences. In the case where the pronoun is deleted, the subject-slot will contain the description of the referent, but the lexical entity corresponding to the referent's properties will be marked as an empty set: [ ]. The rule for 'psdsent' (Polish subject defective sentence) will also be provided with an agreement condition saying that the verbal form must have the same number and gender values as the referent of the subject (variables N and G2 represent, respectively, the categories 'number' and 'grammatical gender'; S is used for marking the property 'human' or 'non-human').

The interrogative sentence (3) can be described as:

I  
 $\text{psent}(q, \_, R, F) \rightarrow \text{pnps}(R), \text{psdsent}(\_, \_, R, s(\text{subj}(R), \text{pred}(B), \text{obj}(X)))$ .  
 II  
 $\text{psdsent}(\_, \_, R, s(\text{subj}(R), \text{pred}(B), \text{obj}(X))) \rightarrow \text{pvt}(Y), \text{pnpa}(X),$   
 $\{R = \text{np}(r(R1, A, \_, N, S, G2, \_, \_)),$   
 $\text{plex}(Y, B, v, \text{vt}, \text{fin}, \_, N, S, G2, P),$   
 $\text{plex}(\_, A, \text{pros}, \_, N, S, G2, \_, P, \_)\}$ .  
 III  
 $\text{pnps}(\text{np}(r(R1, A, \_, N, S, \_, G2, \_))) \rightarrow \text{ppros}(X),$   
 $\{\text{plex}(X, A, \text{pros}, \_, N, S, G2, \_, P, \_)\}$ .  
 IV  
 $\text{pnps}(\text{np}(r(R1, A, \_, N, S, \_, G2, \_))) \rightarrow [], \{\text{plex}([], A, \text{pros}, \_, N, S, G2, \_, P, \_)\}$ .

Rule I is the most general one: it says that a single Polish question may contain a subjective noun phrase and a subject defective sentence; the functional subject of the defective sentence refers to an entity marked R. The second rule (II) is to be read as: a Polish subject defective sentence, where the subject refers to the entity R and the predicate has the meaning B, can be realized as a finite transitive verb Y, with the meaning B, and an accusative noun phrase X; the referent of the subject (R) and the verb must have the same N, S and G2 values; furthermore, the value P (grammatical person) of the verb must be identical with the corresponding value of the lexical item (the pronoun) which has the same meaning as the referent R. In our example (3), the referent of the subject can be characterized as:

$R = \text{np}(r(R1, \text{you}, \_, \text{sg}, \_, \_, \text{ma}), \_)$ ,

because the verb *przeczytałeś* has the following properties:

$\text{plex}(\text{przeczytałeś}, m(\text{read}, \text{prf}), v, \text{vt}, \text{fin}, \_, \text{sg}, \_, \text{ma}, p2)$ .

The value p2 is common for the verb and the lexical item (plex) representing the pronoun *ty*:

$\text{plex}(\text{ty}, \text{you}, \text{pros}, \_, \text{sg}, \_, \_, \_, p2, \_)$ .

Rules III and IV express the fact that the subjective noun phrase having a certain referent R can be realized either as a subjective pronoun or as the empty set [ ]. For the purpose of translation, the lexicon must contain such items as this

described in rule IV, i.e. items which are not realized lexically, but which have a meaning and certain grammatical properties. The need for such 'plex's' can be explained by means of the examples given below:

- (4a) śpiewałeś?  
 (4b) śpiewałaś?  
 (4c) śpiewaliście?  
 (4d) śpiewałyście?

The four Polish sentences in (4) have to be translated as 'were you singing?' (or 'did you sing?'; the tense and aspect relations will not be discussed, as mentioned). None of the sentences contains a lexically realized pronoun, but the different inflected forms allow us to identify the N,S and G2 values of the referent of the subject. The verbal items are characterized as, respectively:

- (4a) plex(śpiewałeś,m(sing,past),v,vi,fin,\_,sg,\_,ma,p2).  
 (4b) plex(śpiewałaś,m(sing,past),v,vi,fin,\_,sg,\_,fe,p2).  
 (4c) plex(śpiewaliście,m(sing,past),v,vi,fin,\_,pl,hum,ma,p2).  
 (4d) plex(śpiewałyście,m(sing,past),v,vi,fin,\_,pl,hum,fe,p2)

or:

plex(śpiewałyście,m(sing,past),v,vi,fin,\_,pl,hum,ne,p2)

or:

plex(śpiewałyście,m(sing,past),v,vi,fin,\_,pl,inh,\_,p2).

The common property of the verbal forms – the value p2 (2nd grammatical person) indicates that – according to the agreement condition in II – only the lexical items with the values 'pros' and 'p2' can form correct interrogative sentences in connection with the verbs concerned above. The lexicon contains three such items:

- plex(ty,you,pros,\_,sg,\_,\_,p2,\_)  
 plex(wy,you,pros,\_,pl,\_,\_,p2,\_)  
 plex([],you,pros,\_,\_,\_,p2,\_)

As the subjects of the sentences (4a)-(4d) are realized as [], the plex([],you,pros,\_,\_,\_,p2,\_) will be identified as referring to the subject. The agreement condition says that the meaning of the 'plex' representing the subject and the meaning of the referent (variable A in the referent description) have to

be identical. Hence, the referent of the subject will be correctly identified as having the meaning 'you' and the sentence (4a) will be represented as:

$$\text{psent}(q,_,\text{np}(\text{r}(\text{R1,you,_,sg,_,_,ma}),_),\text{F}) \rightarrow \text{psdsent}(\_,\_,\text{np}(\text{r}(\text{R1,you,_,sg,_,_,ma}),_),\text{F}),$$

$$\text{s}(\text{subj}(\text{np}(\text{r}(\text{R1,you,_,sg,_,_,ma}),_)),\text{pred}(\text{sing,past})).$$

In sentences (4b) – (4d) the referents of the subject will have different N, S and G2 values (according to the description of the verbal items presented above), but the meaning of the referent will still be identified as 'you', and, subsequently, sentences (4a)-(4d) will be translated as 'were you singing'.

In the Russian subject defective sentence (example (1)), the agreement between the subject and the predicate may be handled by a rule analogous to II, i.e. for sentences with transitive verbs, without adverbials – like (1):

$$\text{rdsent}(\_,\_,\text{R},\text{s}(\text{subj}(\text{R}),\text{pred}(\text{B}),\text{obj}(\text{X}))) \rightarrow \text{rvt}(\text{Y}),\text{rnpa}(\text{X}),$$

$$\{\text{R}=\text{np}(\text{r}(\text{R1,A,_,N,_,_,G2}),_),\text{rlex}(\text{Y,B,v,vt,fin,_,N,_,G2,P}),$$

$$\text{rlex}(\_,\text{A,pros,_,N,_,G2,_,P,_,_})\}.$$

The agreement condition in the rule for object defective sentences (example (2)) has the following shape:

$$\text{rdsent}(\_,\_,\text{R},\text{s}(\text{subj}(\text{A}),\text{pred}(\text{B}),\text{obj}(\text{R}))) \rightarrow \text{rpros}(\text{X}),\text{rvt}(\text{Y}),$$

$$\{\text{rlex}(\text{X,A,pros,_,N,_,G2,_,P,_,_}),\text{rlex}(\text{Y,B,v,vt,fin,_,N,_,G2,P})\}.$$

The referent of the lacking constituent is not involved in the agreement rule above, as there is no agreement between the object and the predicate.

Furthermore, the variable S is not used in the agreement conditions for 'rdsent' and 'radsent', because the Russian finite verbs do not express the difference between 'personal gender' and 'non-personal' gender in the plural (the Polish inflection forms *przeczytaliście* – personal gender and *przeczytaliście* – non-personal gender – are both translated as *pročitali* – a form common to all grammatical genders in Russian).

#### Yes/no questions with predicative constructions

The analysis here will be limited to this type of Polish and Russian predicative constructions, which contain a finite form of the copula verb *być* (Polish) or *byť* (Russian) and an adjective; as the copula *byť* is not lexically realized in most sentences in the present tense, we will also deal with Russian copulaless sentences.

The syntactic and inflectional differences between an English question with a predicative construction and its equivalents in Russian and Polish can be considered by analysing the following examples:

English: are you ill?

Russian:

(5a) ty bolen?

you ill,ma

(5b) ty bol'na?

you ill,fe

(5c) vy bol'ny?

you ill,pl

Polish:

(6a) (ty) jesteś chory?

(you) are ill,ma

(6b) (ty) jesteś chora?

(you) are ill,fe

(6c) (wy) jesteście chorzy?

(you) are ill,pl,

personal gender

6d) (wy) jesteście chore?

(you) are ill,pl

non-personal gender

The sentence (5c) can also be used for addressing a single person (the pronoun *vy* may function either as the plural form of the 2nd grammatical person, or as a polite term of address, comparable with the German *Sie*); however, as *vy* is as a rule combined with predicates in plural forms, in the following analysis it will be treated as a plural pronoun, although the logical referent may be singular. The translation of terms of address requires a more detailed analysis, which cannot be presented in this paper.

In the sentences (5a)-(5c) the function of the predicate is accomplished by the adjective in predicative form (Russian adjectives have two kinds of inflection: predicative and attributive). The adjective agrees with the referent of the subject in number and gender (despite the above mentioned use of *vy* as a term of address). The copula verb is not used, as the tense value of all sentences is T=present.

In all sentences, the subject noun phrase (the pronoun) is fronted, so they can be analyzed as containing subject defective sentences. An 'rdsent' with a predicative construction, as in (5a)-(5d), may be described by the following rule:

$$\text{rdsent}(q, \text{pres}, R, s(\text{subj}(R), \text{pred}(B))) \rightarrow \text{ra}(X), \{R = \text{np}(r(R1, A, \_, \_, N, \_, \_, G2), \_)), \\ \text{rlex}(X, B, a, \text{pred}, N, G2, \_, \_, \_)\}.$$

a=adjective

pred=predicative form

$\text{rlex}(\text{bolen}, \text{ill}, a, \text{pred}, \text{sg}, \text{ma}, \_, \_, \_, \_).$

$\text{rlex}(\text{bolna}, \text{ill}, a, \text{pred}, \text{sg}, \text{fe}, \_, \_, \_, \_).$

$\text{rlex}(\text{bolny}, \text{ill}, a, \text{pred}, \text{pl}, \_, \_, \_, \_).$

In the cases where the copula verb is present (past and future tense in Russian), the agreement condition must also provide the correct choice of the verbal form. The rules used for Russian sentences with copula verb will therefore be similar to the formula describing the Polish sentences (6a)-(6d):

$$\text{psdsent}(q, T, R, s(\text{subj}(R), \text{pred}(\text{cop}(B, T), \text{copo}(C)))) \rightarrow \text{pcop}(X), \text{pa}(Y), \\ \{R = \text{np}(r(R1, A, \_, \_, N, S, \_, \_, G2), \_), \text{plex}(Y, C, a, \_, \_, N, S, G2, \_, \_, \_), \text{plex}(X, m(B, T), v, \\ \text{cop}, \text{fin}, \_, \_, N, S, G2, P), \text{plex}(\_, A, \text{pros}, \_, \_, N, S, G2, \_, \_, P, \_)\}.$$

If we consider e.g. the example (6a), we can see that the lexical items fulfil the agreement condition formulated above:

$\text{plex}(\text{ty}, \text{you}, \text{pros}, \_, \_, \text{sg}, \_, \_, \_, \text{p2}, \_).$

or – in the pronounless variant:

$\text{plex}([\ ], \text{you}, \text{pros}, \_, \_, \text{sg}, \_, \_, \_, \text{p2}, \_).$

$\text{plex}(\text{jestes}, m(\text{be}, \text{pres}), v, \text{cop}, \text{fin}, \_, \_, \text{sg}, \_, \_, \_, \text{p2}).$

$\text{plex}(\text{chory}, \text{ill}, a, \_, \_, \text{sg}, \_, \_, \_, \text{ma}, \_, \_, \_).$

The referent of the lacking constituent will be identified as:

$r(R1, \text{you}, \_, \_, \text{sg}, \_, \_, \_, \text{ma}).$

The variable S is not used in this case, but it is of importance for examples (6c) and (6d).

### YES/NO QUESTIONS WITH INTERROGATIVE PARTICLES

The interrogative mode of Russian and Polish sentences may be marked by question particles: *czy* in Polish, *razve*, *neuželi* or *li* in Russian. In transformational analysis, interrogative particles are treated as complementizers, generated directly under S-bar (Radford 1981:173); in English, the complementizer nodes are assumed to be empty in main clauses.

It would be quite difficult to apply this analysis to the Russian sentences with the particle *li*, as *li* can be placed either after the predicate or after another

constituent focused in the question (the use of *li* after a fronted adverbial is very common); contrary to the particles *czy*, *razve* and *neuželi*, *li* cannot precede a question. Hence, it seems impossible to place this particle under COMP in the structural tree proposed by transformational grammar.

Since RG-rules allow a free constituent order in the categorial representation, the interrogative particles can simply be inserted in the positions where they really occur (*li* after the focused constituent, the others – before the sentence); there is no need of inserting the symbols representing the interrogative particles in the functional representation, as the question mode is marked by the mode value 'q'.

The fact that a Polish or Russian interrogative sentence may optionally be preceded by the particle *czy* (Polish) or *razve/neuželi* (Russian) can be expressed by the following simple rule:

$$\begin{aligned} r/psent(q, \_, P, F) &\rightarrow r/pq(X), r/psent(q, \_, P, F). \\ pq(X) &\rightarrow [czy]. \\ rq(X) &\rightarrow [razve]. \\ rq(X) &\rightarrow [neuželi]. \\ r/pq(\_) &\rightarrow \_. \end{aligned}$$

In these formulas, we ignore the slight semantic difference between *razve* and *neuželi* (using *neuželi*, the sender is marking a high grade of doubt; a sentence like *neuželi on pročítal ètu knigu?* could be translated as 'has he really read this book?'; here, we will simplify the analysis and assume, that both *razve on pročítal ètu knigu?* and *neuželi on pročítal ètu knigu?* can be translated as 'has he read this book?'; there is of course a possibility of expanding the formulas, so that appropriate sentence adverbials would be inserted in the functional representation and, consequently, in the translation into English or Swedish).

The use of *li* can be illustrated by the examples (7a)-(7c):

- (7a) videl li ty Ivana?  
see.past you Ivan  
'did you see Ivan?'
- (7b) davno li ty videl Ivana?  
long ago you see.past Ivan  
'did you see Ivan a long time ago?'
- (7c) zdorov li on?  
well he  
'is he well?'

The particle *li* will be marked by a symbol different from 'rq' (as we have to avoid generating such ungrammatical constructions as e.g. *\*davno razve ty videl Ivana?*). Here, we use the symbol 'rql'. The sentences (7a)-(7c) are analyzed as (simplified – we ignore agreement conditions and the condition excluding possibility of using both 'rq' and 'rql' in the same single sentence):

(7a)  
rsent(q, \\_, R, F)  $\rightarrow$  rv(R), rql(L), rvdsent(\\_, \\_, R, s(subj(X), pred(B), obj(Z))).

rvdsent=Russian verb lacking sentence

(7b)  
rsent(q, \\_, R, F)  $\rightarrow$  radv(R), rql(L), radvsent(\\_, \\_, R, s(subj(X), pred(B), obj(Z), advl(R))).

davno li	ty videl ivana davno
a long time	you see ivan a long time
ago	ago

radvsent=Russian adverb lacking sentence

(7c)  
rsent(q, pres, R, F)  $\rightarrow$  ra(R), rql(L), radjsent(\\_, \\_, R, s(subj(X), pred(R))).

zdorov li	on zdorov
well	he well

radjsent = Russian adjective lacking sentence

The optional use of *li* is expressed by the rules:

rql(L)  $\rightarrow$  [li].

rql(\\_)  $\rightarrow$  \\_.

#### WH-QUESTIONS WITH THE RUSSIAN AND POLISH EQUIVALENTS OF 'WHO'/'WHOM' AND 'WHAT'

The Russian interrogative pronouns *kto* 'who', *čto* 'what' and their Polish equivalents *kto* 'who' and *co* 'what' are case-inflected. Their declension forms are shown below:

	Russian		Polish	
nominative	kto	čto	kto	co
genitive	kogo	čego	kogo	czego
dative	komu	čemu	komu	czemu
accusative	kogo	čto	kogo	co
instrumentalis	kem	čem	kim	czym
locative	kom	čem	kim	czym

The genitive forms are not equivalent to 'whose' – they are combined with genitive lacking sentences, in which the genitive noun phrase functions as direct object.

According to the RG-rules, a *wh*-question is analyzed as containing a noun phrase marked 'npq', in a certain case form *x*, and a sentence lacking a constituent in the same case form, i.e.:

$$\text{sent}(q, \_, R, F) \rightarrow \text{npqx}(R), \text{xdsent}(\_, \_, R, F).$$

We will illustrate the use of the rule by applying it on some Russian and Polish examples (without agreement conditions):

#### Nominative:

(8) r: kto pročitál ètu knigu?  
 p: kto przeczytał tę książkę?  
 who read,prf this book

$$\begin{aligned} r/\text{psent}(q, \_, R, F) &\rightarrow r/\text{pnpqs}(R), r/\text{psdsent}(\_, \_, R, s(\text{subj}(R), \text{pred}(B), \text{obj}(Z))). \\ r/\text{pnpqs}(\text{np}(r(R1, A, \_, \_, \_, \_))) &\rightarrow r/\text{pwh}(X), \{r/\text{plex}(X, A, \text{wh}, \_, \_, \_, \text{nom}, \_)\}. \\ r/\text{plex}(\text{kto}, \text{who}, \text{wh}, \_, \_, \_, \text{nom}, \_) & \end{aligned}$$

Functional representation (rdsent):

$$s(\text{subj}(\text{kto}), \text{pred}(\text{pročitál}), \text{obj}(\text{np}(\text{dem}(\text{ètu}), \text{n}(\text{knigu}))))$$

who read,prf this book

#### Genitive:

The genitive form of the object is used after certain verbs, classified as 'vg' – genitive demanding – here *chotel'* (Russian) and *chciąc* (Polish), or when a transitive predicate is denied (sentence adverbial=nix). In both cases, we use the same rule for 'sent', but the rules describing 'gdsent' (genitive defective sentence) must be different for (9) and (10):

(9) r:čego ty chotel?  
 p:czego (ty) chciałeś?  
 what you want,past  
 'what did you want?'

(10) r:čego ty ne pomniš'?  
 p:czego(ty) nie pamiętasz?  
 what you not remember  
 'what do you not remember?'

The rules for 'r/psent' are common for (9) and (10):

$$\begin{aligned} r/\text{psent}(q, \_, R, F) &\rightarrow r/\text{pnpqg}(R), r/\text{pgdsent}(\_, \_, R, F). \\ r/\text{pnpqg}(\text{np}(r(R1, A, \_, \_, \_, \_))) &\rightarrow r/\text{pwh}(X), \{r/\text{plex}(X, A, \text{wh}, \_, \_, \_, \text{g}, \_)\}. \\ r/\text{plex}(\text{čego}, \text{what}, \text{wh}, \_, \_, \_, \text{g}, \_) & \\ r/\text{plex}(\text{czego}, \text{what}, \text{wh}, \_, \_, \_, \text{g}, \_) & \end{aligned}$$

The genitive defective sentences in the examples (9) and (10) have different categorial representations:

(9)  
 $r/\text{pgdsent}(\_, \_, R, s(\text{subj}(X), \text{pred}(B), \text{obj}(R))) \rightarrow r/\text{ppros}(X), r/\text{pvgt}(B).$

Functional representation (rgdsent):

$$s(\text{subj}(\text{ty}), \text{pred}(\text{chotel}), \text{obj}(\text{čego})).$$

you want,past what

(10)  
 $r/\text{pgdsent}(\_, \_, R, s(\text{subj}(X), \text{pred}(B), \text{obj}(R), \text{sadvl}(\text{nix}))) \rightarrow r/\text{ppros}(X), [\text{neg}],$   
 $r/\text{pvt}(B).$

Functional representation (rgdsent):

$$s(\text{subj}(\text{ty}), \text{pred}(\text{pomniš}), \text{obj}(\text{čego}), \text{sadvl}(\text{nix})).$$

you remember what

#### Dative:

Examples:

(11)  
 r: komu ty dal knigu?  
 p: komu (ty) dałeś książkę?  
 to whom you give,past book  
 'to whom did you give the book?'





(15a)

čem on gorditsja?  
 what he be proud,pres  
 'what is he proud of?'

(16a)

kogda èto slučilos'?  
 when it happen,past  
 'when did it happen?'

The computer program must provide the possibility of distinguishing constructions like (13a) (which have to be translated as an English *wh*-pronoun and an object defective sentence) from those like (14a)-(15a) (translated as a *wh*-pronoun and a sentence containing a defective prepositional phrase) and from sentences like (16a), where an appropriate interrogative adverb must be correctly inserted.

We suggest that the term 'instrumentalis defective sentence' (*r/pidsent*') should be used only for such constructions that can be preceded by an interrogative (or relative) pronoun in the instrumentalis form. As question (16a) cannot be formulated as \**čem èto slučilos'*, it must be analyzed in a different way than (13a)-(15a).

On the sent-level, sentences (13a)-(15a) can be described as:

$$rsent(q, \_, R, F) \rightarrow rnpqi(R), ridsent(\_, \_, R, F).$$

The differences can be found in the categorial and functional representations of the defective sentences 'ridsent'.

(13a):

The verb *pol' zovat' sja* is classified as 'rvin' – an instrumentalis-demanding verb, because a sentence without a direct object, like \**on pol' zovalsja* is unacceptable, and the only possible case form of the direct object after *pol' zovat' sja* is the instrumentalis. Thus, the defective sentence can be represented as:

$$ridsent(\_, \_, R, s(\text{subj}(X), \text{pred}(B), \text{obj}(R))) \rightarrow rpros(X), rvin(B).$$

on pol'zovalsja cem	on pol'zovalsja
he use,past what	he use,past

(14a):

The lacking instrumentalis noun phrase is equivalent to an English prepositional phrase and functions as an adverbial. The functional representation used in SWETRA's parser contains three adverbial slots, marked A1, A2 and A3. If one of the adverbials is partially lacking – as in the English equivalent of sentence (14a) – the sentence is treated as containing a defective prepositional clause, marked by the symbol of the lacking constituent and by the number of the adverbial. In the English equivalent of (14a), the defective prepositional clause, lexically realized as the preposition *with*, would be marked as e.g. 'epodpp(R,A3)', where epodpp=English defective pp; A3 indicates that the defective pp and the lacking constituent R function as the adverbial marked A3. Using a simplified notation, we can describe the English sentence as:

$$esent(q, \_, \text{what}, F) \rightarrow enpq(\text{what}), epodsent(\_, \_, \text{what}, F).$$

$$epodsent(\_, \_, \text{what}, s(\text{subj}(\text{he}), \text{pred}(\text{kill}, \text{past}), \text{obj}(\text{him}), \text{advl}(A3))) \rightarrow$$

$$\rightarrow [\text{did}], epros(\text{he}), einf(\text{kill}), eproo(\text{him}), epodpp(\text{what}, A3).$$

Subsequently, one has to apply some rules which make it possible to realize the defective prepositional phrase as a single preposition; the choice of the preposition requires quite complicated formulas.

As the Russian sentence (14a) does not contain any prepositional clause, the information about the lacking constituent R (i.e. the same information that is inserted in 'epodpp' in the rule above) has to be placed in the appropriate slot in the functional representation:

$$ridsent(\_, \_, R, s(\text{subj}(X), \text{pred}(B), \text{obj}(Z), \text{advl}(R, A3))) \rightarrow rpros(X), rproa(Z), rvt(B).$$

on ubil ego čem, A3	on ego ubil
he kill,past him what, A3	he him kill,past

(15a)

The sentence is analyzed analogously to (14a) (the only difference being the fact that the finite verb *gorditsja* belongs to the category 'instrumentalis-demanding verbs'):

$$ridsent(\_, \_, R, s(\text{subj}(X), \text{pred}(B), \text{advl}(R, A3))) \rightarrow rpros(X), rvin(B).$$

on gorditsja čem	on gorditsja
he be proud,pres what, A3	he be proud,pres

(16a)

