

Referential Choice: Distribution of Subject Types in Russian Aphasic Speech

Anastasiia Romanova

University of Groningen, University of Potsdam, St. Petersburg State University

Being an essential component of discourse structure, reference presents a very interesting topic for investigation. Some studies have been conducted on referential subject types in canonical non-pro-drop, sentence-based languages in both typical and atypical populations. This study addresses distribution of the subject types in narratives in Russian, an optional pro-drop, discourse-based language. Participants of the study include two experimental groups (non-fluent and fluent aphasic speakers) and a control group. The investigated distribution of full noun phrases, pronouns and zero forms in the syntactic position of the subject in narratives of non-fluent aphasic individuals appeared to be similar to those found in non-pro-drop Germanic languages in relation to high subject omission. On the other hand, fluent aphasic speakers, in addition, demonstrate the tendency of excessive use of pronominal subjects, which must be caused by general nomination problems present in fluent aphasia. Thus, it can be seen that distribution patterns of subject types in narratives of both Russian non-fluent and fluent aphasic speakers are deviant from those of non-brain-damaged respondents. However, presumably, different economy strategies are used by these two groups of aphasic individuals.

1. INTRODUCTION

Being an essential component of discourse structure, reference presents a very interesting topic for investigation. This research relates to Chomsky's (1981) distinction between innate linguistic universal 'principles' and language specific 'parameters', and focuses on the Pro-Drop parameter. Languages differ in whether they allow null subjects or not. English and Dutch, for example, are considered to be canonical non-pro-drop languages, which means that the position of the subject is necessarily filled in overtly. On the contrary, most Romance languages (e.g. Spanish, Italian) are called canonical pro-drop languages in which overt subjects carry additional emphasis. Moreover, this classification is also completed with optional pro-drop languages (presumably like Russian) that generally permit subject drop but do not necessarily require it and even apply some discourse constraints to its use (Gordishevsky & Avrutin 2004).

Subject omission is possible in Russian, a topic-prominent and discourse-oriented language, only in certain pragmatic conditions, and no additional emphasis or meaning appear when a subject is preserved. Russian, however, does allow the omission of a subject if some contextual requirements are met: 'items recoverable from the context are frequently omitted on the surface' (Franks 1995: 307). That is to say, subject omission in Russian is the omission of the topical constituent in an utterance. Gordishevsky and Schaeffer (2002) argue that for the subjects in optional pro-drop languages to be empty, their referents must be recoverable either from linguistic or situational context. The presence of the linguistic or situational antecedent in the conversation (i.e. the referent having been mentioned, being at the centre of the discourse, or being clear from the context) provides the conditions for subject omission. Consider an example of such discourse subject omission from the National Corpus of the Russian language (Zdorenko 2009: 2):

- (1) My vstretimsja?
 we meet.1PL
 'Shall we meet?'
- Davajte, kak i dogovarivalisj.
 agreed.PL
 'Let's do it, as (we) already agreed.'

Person agreement is not marked in the verb in the second line of the dialogue. Nonetheless, the subject is omitted because it is identified in the previous clause. The same clause with the subject present would be fully possible, however, as it is neither redundant nor emphatic.

Exploring speech production in aphasic speakers, whose language is impaired due to selective cerebral injuries, is insightful as it can contribute to our knowledge about the mental representations and/or processing of language. Only a few studies compare the distribution of referential devices in subject position in groups of non-brain-damaged speakers and aphasic speakers. It is a well known fact that pronouns tend to be omitted by Broca's aphasic speakers in spontaneous discourse and there is a substantial amount of literature on the subject noun phrase omission in Broca's aphasic speech. The evidence for this is a high noun/pronoun ratio (Saffran *et al.* 1989) and the omission of pronouns in obligatory contexts. Up-to-date investigations of the distribution of full noun phrases (NPs), overt pronouns and zero forms in subject position, carried out with non-fluent (agrammatic) aphasic speakers of canonical non-pro-drop languages, have revealed the following tendencies. Group studies in Dutch and German (de Roo 1999, 2003; Kolk & Heeschen 1990; Tesak & Dittman 1991) show that aphasic speakers omit 0-47% subject pronouns in finite clauses, while healthy speakers omit only 1-5% (de Roo 2003). Such omission cannot be accounted for by the Tense underspecification as both the Tense and the Nominative case are available for the speaker. An analogue of zero subject pronouns in finite clauses in agrammatic speech (particularly in Dutch and German patients), however, is found in normal speech production as well and is called *the Topic Drop phenomenon*. De Roo (2002) speaks of the Germanic Topic Drop characteristic for Dutch, German, English and the Scandinavian languages (e.g. Cardinaletti 1990; Huang 1984). Such findings go in line with 'a reduced temporal window' for sentence processing (Kolk 1995) in agrammatic speakers. As a result, they apply elliptical constructions that are also characteristic of non-brain-damaged speech production.

2. RESEARCH QUESTION

As a topic-prominent language in which the information structure of discourse is important for the organization of an utterance, the Russian language is particularly interesting to look at in respect to the distribution and prevalence of referential devices. Little is known about the referential processes in Russian-speaking aphasic individuals. Therefore, the present study is a starting point in remedying the paucity of research in this area. Several research questions have been set within the framework of this study.

First, the issue of the extent to which Russian aphasic speakers lose/preserve their ability to refer to entities in the syntactic position of the subject is to be addressed in this study by comparing the patterns of subject types in the aphasic and non-brain-damaged control groups' speech. The general body of previous research on the distribution of subject types in aphasic individuals was focused on canonical non-pro-drop Germanic languages. Thus, one of the aims of the present study is to verify the distribution of subject types in a structurally different, optional pro-drop language like Russian (Gordishevsky & Avrutin 2004). Russian non-fluent aphasic speakers could be expected to follow the example of Dutch agrammatic patients who tend to omit subjects (de Roo 2003), for example. They could equally prove that

the referential choice does not present particular difficulty for the aphasic population and follow the normal speech pattern.

Second, previous research done on aphasic speakers has taken only non-fluent (agrammatic) aphasic individuals into consideration. This study aims to test both non-fluent (with impaired syntactic abilities) and fluent (with impaired word-finding) groups of Russian aphasic speakers. Thus, corresponding cross-group comparisons are to show whether the patterns of realizations and omissions of grammatical subjects vary among aphasia types.

3. METHOD

Participants

The present study is based on data from the experimental groups of non-fluent and fluent aphasic speakers with no uncorrected hearing or vision problems. The non-fluent participants comprised seven individuals (one woman and six men) and the fluent group consisted of five aphasic speakers (two women and three men). The patients range in age from 23 to 64 years old, with a mean age of 41 in the non-fluent group and 50 in the fluent group. The patients are aphasic due to a left hemisphere stroke, infarction, or craniocerebral injury. All the patients had been aphasic for at least three months when the tests were carried out. In order to obtain sufficient and interpretable data, no patients with severe forms of aphasia and/or apraxia of speech were tested. The data was collected at the Federal Center for Speech Pathology and Neurorehabilitation in Moscow.

Apart from the experimental groups, a group of non-brain-damaged control speakers was included in the study. A total of fifteen controls (eleven women and four men), all native Russian speakers, participated in the experiment. The age of the control participants ranges from 23 to 72. The mean age is 42. Table 1 summarizes the general information about the study participants.

(Table 1) *General information on the study participants.*

	Number		Mean age	Months post-onset	Handedness
	male	female			
Non-fluent aphasic speakers	6	1	41	> 3	right
Fluent aphasic speakers	3	2	50	> 3	right
Non-brain-damaged speakers	4	11	42	-	right

Materials

The distribution of subject types in spontaneous speech of the aphasic speakers was tested using narratives as a context. Narrative is a universal genre produced by different speakers of different cultures. This form of discourse is often preserved in patients with moderate and mild forms of aphasia and provides a good measure of patients' spontaneous speech.

In order to elicit spontaneous speech in a narrative task, subjects were presented with the twenty-four-page wordless picture book, 'Frog, where are you?' (Mayer 1969). The images tell a story of a boy, a dog and their pet frog that runs away one day. Throughout the story the boy and the dog are trying to find the frog. While they are searching for it in the forest they encounter many other animals. Eventually, they find their pet frog with a mate and many baby frogs. They take one of the baby frogs as their new pet and return home. This book has been quite successful in eliciting speech production in many cross-linguistic aphasiological and developmental studies (e.g. Losh *et al.* 2001).

Procedure

The experiment was carried out in the following way. First, the picture book was introduced to a patient and the researcher gave him/her instructions: 'Here is a book about a boy, a dog and a frog. We will first go through the book and look at the pictures. Then we will start again and I will ask you to tell me the story'. If patients experienced difficulties in producing the narrative, the researcher encouraged them without giving any linguistic cues that could hinder collecting reliable results in the experiment. Errors were not corrected and no feedback was given to the participants. The speech of the patients was digitally recorded and orthographically transcribed by the researcher. The same procedure was followed when testing non-brain-damaged subjects as well.

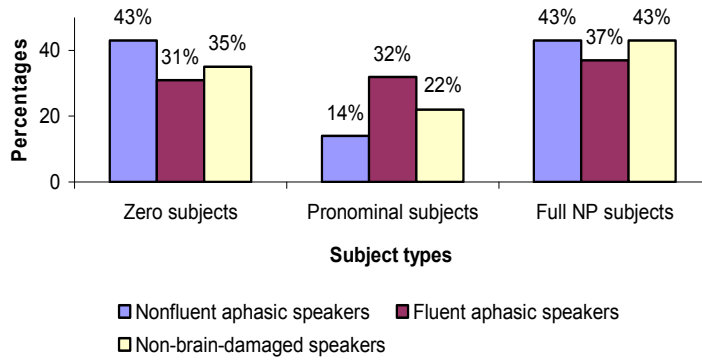
Once transcribed, the narratives underwent the segmentation procedure. Units of segmentation were defined as sequences of speech in which a finite verb is used. Utterance boundaries were determined by intonation contours and pause length. Following the exclusion criteria, all ambiguous forms, false starts, discourse markers (e.g. *skazhem* '(we) will say' = 'Let's say') and all kinds of repetition of a subject within one clause were omitted. Some rare occurrences of both nominal and pronominal subjects within the same finite clause were treated as cases of the nominal subject type. All imperative forms were also omitted from the count as well as subjects in impersonal constructions, such as third person plural forms (e.g. *govorjat* '(they) say'), second person singular, non-referential forms, and others (e.g. *mne kazhetsja* '(it) seems to me'). All the finite clauses that withstood the exclusion procedure were grouped into three categories according to the subject type, namely those containing a full NP, a pronoun, or a zero element as the subject.

Results

Let us consider the referential subject types one by one. Full NP subjects have the highest proportions among all the three subject types in the distribution patterns of all the three groups analyzed. However, the non-fluent aphasic and non-brain-damaged control groups are leading in the use of full NPs in the subject position with the rates of 43% of all the finite clauses, while the fluent aphasic speakers use full NP subjects in 37% of all the finite clauses (see Fig. 1).

Zero subjects are second to full NPs in the frequency of use by the groups tested. The proportions of zero subjects are highest in the non-fluent aphasic group (43% of all the finite clauses), while fluent aphasic and non-brain-damaged speakers have comparable proportions of zero subjects in their narrations (31% and 35% of all the finite clauses respectively).

As for pronominal subjects, they are used to a lesser extent in comparison to the other two subject types overall, and in the non-fluent aphasic and non-brain-damaged groups in particular. However, they are excessively, compared to the other groups, referred to by the fluent aphasic speakers (32% of all the cases, with 14% for the non-fluent and 22% for the non-brain-damaged groups).

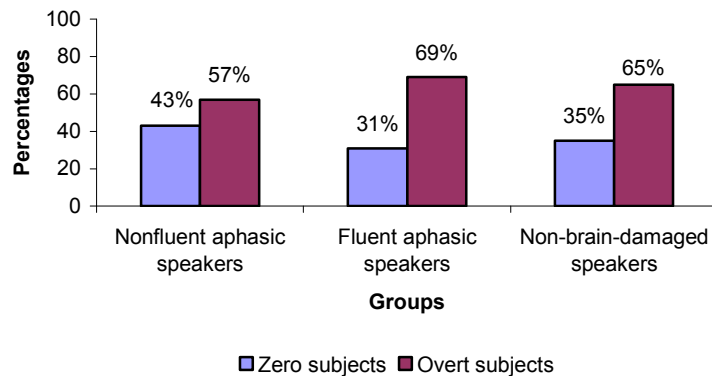


(Fig. 1) *Crossgroup comparisons of the distribution of the subject types (in percentages).*

In order to check for possible associations between the distributions of the subject types across the three groups analyzed, a chi-squared test was used to confirm a significant difference between the distribution patterns in the non-fluent and fluent aphasic groups, and the non-brain-damaged group ($\chi^2 = 62.53$, $df = 4$, $p < .001$). Further paired comparisons of the subject types and the groups of speakers revealed the following tendencies. The distribution patterns are significantly different when we compare the relation of zero and pronominal subjects in the non-fluent aphasic and non-brain-damaged groups ($\chi^2 = 13.33$, $df = 1$, $p < .001$), the fluent and non-brain-damaged groups ($\chi^2 = 9.61$, $df = 1$, $p < .005$), and, finally, the nonfluent and fluent groups ($\chi^2 = 29.99$, $df = 1$, $p < .001$). Compared to the distribution of zero and pronominal subjects in the control group (35% and 22% respectively) that can be taken as a norm, the non-fluent aphasic speakers show a higher proportion of zero subjects (43% of all the finite clauses) and lower proportion of pronominal subjects (14%). The fluent aphasic group, on the contrary, has an increased proportion of pronominal subjects equalling the proportion of zero subjects in their narration (31% and 32% respectively).

Another significant difference in the distribution patterns is observed when comparing pronouns and full NPs in the syntactic position of subject across all three groups consecutively: the non-fluent aphasic and non-brain-damaged groups ($\chi^2 = 6.46$, $df = 1$, $p < .05$), the fluent aphasic and non-brain-damaged groups ($\chi^2 = 10.90$, $df = 1$, $p < .005$), and, finally, the non-fluent and fluent aphasic groups ($\chi^2 = 19.93$, $df = 1$, $p < .001$). Here, the proportions of pronominal subjects are lower than those of full NPs in all the groups. However, the degree of discrepancy between these subject types varies across the participant groups: the largest difference is observed in the non-fluent aphasic group (14% of pronouns among all the finite clauses produced versus 43% of full NPs), the non-brain-damaged group takes the second position in this row (22% versus 43% respectively), whereas the fluent aphasic group demonstrates the comparable percentages for the two subject types (32% versus 37% respectively).

In addition, we compared the distributions of zero subjects and overt (pronouns and full NPs) subjects across the three groups. Chi-squared tests revealed a significant association between group and subject type, specifically when the non-fluent aphasic group was compared to the non-brain-damaged group ($\chi^2 = 7.30$, $df = 1$, $p < .01$), and when the non-fluent aphasic group was compared to the fluent one ($\chi^2 = 10.83$, $df = 1$, $p < .005$). The general tendency of zero and overt subject distribution is consistent for all the groups - overt subjects prevail. However, the degree of such prevalence varies: the largest discrepancy between zero and overt subjects is apparent in the fluent aphasic group (31% of zero subjects in the whole set of finite clauses produced versus 69% of overt subjects) and the least discrepancy is distinctive of the non-fluent aphasic group (43% versus 57% respectively). Fig. 2 illustrates the described variations.



(Fig. 2) Crossgroup comparison of the distribution of zero and overt subject types.

4. DISCUSSION

As shown above, the proportions of full NP subjects are high in each of the groups analyzed. Specifically, the percentages of full NPs in the position of the subject are equal in the groups of non-fluent aphasic and non-brain-damaged speakers (43% of all the finite clauses), which are in turn higher than the proportion of full NP subjects in the speech of the fluent aphasic group (37% of all the finite clauses), as shown in Fig.1. The prevalence of full NP subjects in all the groups tested proves full NPs to be a very stable referential device which is preserved in aphasia.

The proportion of pronominal subjects reaches its highest value in the fluent aphasic speakers (32% of all the finite clauses), which is considerably higher than that of the non-brain-damaged group (22% of all the finite clauses). The percentage of pronominal subjects is the lowest in the non-fluent aphasic group (14% of all the finite clauses produced). The two cross-group comparisons described above bring us to the thought that pronominal subjects may cause the most problems for the non-fluent aphasic Russian speakers. At the same time, they seem to be rather effortless for fluent aphasic individuals, who, according to the mean proportions, tend to use them even more frequently than non-brain-damaged individuals. See the following example from a fluent aphasic individual's speech sample:

- (2) (a) **Oni** prishli v les
they came-PL
'They came to the forest';
- (b) i krichit **etot** auau ... kak vot eto vot, da, praviljno?
scream-3SG this-MASC
'And this one screams auau ... how is this [called], yes, right?'
- (c) **Oni vse** krichat vot eto, v les ili kak eto ...
they all scream-PL
'They all are screaming this, into the forest or how [to say] this';
- (d) **oni** videli
they saw-PL
'They saw';
- (e) eto **ja** zabyla
I forgot-3SG.FEM
'I forgot this';
- (f) kak vot **eto** nazyvaetsja
this-NEUT call-PASS.SG
'How is this called?'
- (g) eto ... vot eti makarony letjat, no vot eti ...

- 'Well ... these spaghetti are flying, but well these...';
 (h) a, znachit, **oni** uvideli pchelok, da?
 they saw-PL
 'And, well, they saw bees, didn't they?'

Pronouns are reduced referential devices and are used only in the case of the referents' activation in the listener's mind. Thus, fluent aphasics, whose speech production is characterized by word-finding problems and distortion in the connections between the meaning and the form of words, might willingly overuse pronominal elements as a strategy.

Paired cross-group comparisons for the proportions of zero subjects reveal another interesting pattern. The percentage of zero subjects used by the non-fluent aphasic individuals (43% of all the finite clauses analyzed) is higher than that of the non-brain-damaged group (35% of all the finite clauses), which is in its turn higher than that of the fluent aphasic group (31% of all the finite clauses analyzed). Overall, the non-fluent aphasic group is characterized by excessive use of zero subjects in the narrations:

- (3) Nu i chego, v grjazj **Ø upali**, nu i sobaka tozhe.
 fell-PL

'And well, they fell into the mud, and the dog as well';

Eto, naverno, zmeja, naverno, ili net, ili derevo, naverno, vot, chto, tishe, **Ø ne znaju**.

not

know-1SG

'This is probably a snake, maybe, or not, or a tree, maybe, well, that, quiet, I don't know';

Ø poshel i **Ø prishel ustaliy**, **Ø ustali**.
 went-3SG.MASC came-3SG.MASC tired got tired-3PL

'He left and came tired, they got tired'.

To check for the difference in the distributions of zero subjects and overt subjects (pronouns and full NPs as a group) we carried out one more cross-group comparison. Differences in distribution patterns are observed between the non-fluent aphasic and non-brain-damaged control groups, and between the non-fluent and fluent aphasic groups. No such difference is found between the fluent aphasic and non-brain-damaged control groups. Looking at the mean proportions of zero and overt subjects in the groups, we see that the percentages of zero subjects are considerably lower than those of overt subjects (see Fig. 2). Thus, all the cross-group differences arise from a discrepancy between the subject types. The fluent aphasic and non-brain-damaged groups have comparable discrepancy between the proportions of zero (31% of the finite clauses in the fluent group, and 35% in the non-brain-damaged group) and overt subjects (69% and 65% respectively). Contrastingly, the non-fluent aphasic group (with 43% of zero subjects and 57% of overt subjects) differs both from the fluent aphasic group (31% and 69% respectively) and the non-brain-damaged control group (35% and 65% respectively). As can be seen, the percentage of zero subjects is higher in the non-fluent aphasic group in comparison to the fluent aphasic and non-brain-damaged groups.

Let us now look more closely at the above-described distribution patterns of zero and overt subjects in terms of the first research question. The non-fluent aphasic group shows clearly different distribution patterns. As non-fluent speakers have considerably more problems with grammatical entities than fluent individuals, a higher percentage of omitted subjects in the non-fluent aphasic speech in the sample of the present study is quite logical. These results are compatible with the findings for canonical non-pro-drop Germanic languages that non-fluent aphasic patients tend to omit subjects, applying (sometimes inadequately) elliptical constructions characteristic of the speech of non-brain-damaged speakers. Thus, the 43% of zero subjects reported in the present study lie within the 0-47%

range of subject omission that de Roo (2003) observed for non-fluent aphasic speakers of Dutch. The data collected in the group of non-fluent aphasic speakers of Russian demonstrate the similarity in the distribution of subject types (patterns of zero and overt subjects) with sentence-oriented non-pro-drop Germanic languages. Due to the choice of a reduced referential device, namely, a zero form, the processing load reduces significantly, which proves to be an effective strategy for non-fluent aphasic speakers (Kolk 1995). Similar to the findings of de Roo (2002), we can call such a subject omission strategy of normal speech production (present in language of non-brain-damaged speakers and overused by non-fluent aphasic individuals) a Topic Drop.

Now let us address the second research question raised in the present study, namely the difference in referential choice between non-fluent and fluent aphasic groups. The high percentage of zero subjects and low proportion of pronominal subjects characteristic of the non-fluent aphasic group is not observed in the fluent aphasic group. Moreover, the mean percentages of the subject types within the latter group do not differ significantly. No difference is found in the distribution patterns of zero and overt subjects in the fluent aphasic (31% of zero subjects and 69% of overt subjects in the sample of finite clauses) and non-brain-damaged groups (35% of zero subjects and 65% of overt subjects). This proves that, unlike non-fluent aphasics, fluent aphasic speakers do not resort to zero subjects in their speech more often than non-brain-damaged individuals. Instead, they tend to overuse pronominal subjects (in comparison to the other two groups given). The results obtained are in line with general observations of fluent aphasia, namely the preservation of grammatical and syntactic structures in speech output and mainly coherent speech production. The excessive usage of pronominal subjects corresponds to the general nomination problems in fluent aphasia.

5. CONCLUSIONS

This study investigated the distribution patterns of full NPs, pronouns and zero forms in the syntactic position of subject in narratives of aphasic speakers and non-brain-damaged individuals. Primarily, the aim was to study for the first time the distribution pattern of subject types in non-brain-damaged and agrammatic speakers of Russian, an optional pro-drop discourse-based language. The findings of the present research revealed similarities between the distributions of referential expression of syntactic subjects in Russian and previously studied non-pro-drop Germanic languages. In particular, the data analyzed showed high proportions of zero subjects in the speech of the non-fluent aphasic speakers. In this way, some positive evidence has been obtained on a Russian dataset for the tendency of subject omission by non-fluent aphasic speakers, noted for Germanic languages.

Besides adding Russian into the cross-linguistic framework of research on referential choice, the present study encompasses groups of both non-fluent and fluent aphasic speakers. The aim is to track any existing similarities and differences in the use of referential devices in the position of the subject between the two experimental groups. The group of fluent aphasic speakers demonstrates even proportions of the subject types within the group. No overuse of subject omission, observed in the non-fluent aphasic group, is found here. In addition to quite a high proportion of zero subjects, fluent aphasic speakers have yet another economy strategy in the speech production: they tend to excessively assign pronominal referential devices to referents in the position of subject.

Further research is clearly needed, however. First, the cross-linguistic approach to the research on referential choice in the subject syntactic position should be enlarged by more data from optional pro-drop languages. Qualitative investigation of types and contexts of zero subjects use in non-brain-damaged speakers in Russian is also required. This is needed in order to gain more extensive and reliable materials for showing the deviance of non-fluent aphasic speakers from the normal pattern.

REFERENCES

- Cardinaletti, A. (1990). Subject/object asymmetries in German null-topic constructions and the status of Spec CP. In: J. Mascaró, & M. Nespor (ed.). *Grammar in progress: Glow essays for Henk van Riemsdijk*. Dordrecht: Foris.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- De Roo, E. (1999). *Agrammatic grammar. Functional categories in agrammatic speech*. PhD dissertation, University of Leiden.
- De Roo, E. (2002). Pronoun omission in Dutch and German agrammatic speech. In: H. Simon, & H. Wiese (ed.). *Pronouns, representation and processing*, (pp. 253-284). Amsterdam: John Benjamins.
- De Roo, E. (2003). Null subject pronouns in Broca's speech production. *Aphasiology*, 17(11): 1057-1072.
- Franks, S. (1995). *Parameters of Slavic morphosyntax*. Oxford University Press.
- Gordishevsky, G. & Avrutin, S. (2004). Optional omissions in an optionally null subject language. In: J. Van Kampen, & S. Baauw (ed.). *Proceedings of GALA 2003*, Vol.1, pp. 187–198. LOT, the Netherlands.
- Gordishevsky, G., & J. Schaeffer (2002). On null subjects in child Russian. *Proceedings of the Third Tokyo Conference on Psycholinguistics*. Tokyo: Hituzi Syobo, 115-137.
- Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15: 321–337.
- Kolk, H. (1995). A time based approach to agrammatic production. *Brain and Language*, 50: 282-303.
- Kolk, H., & Heeschen, C. (1990). Adaptation symptoms and impairment symptoms in Broca's aphasia. *Aphasiology*, 4: 221-231.
- Losh, M., Bellugi, U., Reilly, J., & Anderson, D. (2001). The integrity and independence of evaluation in narratives: Evidence from children with Williams syndrome. *Narrative Inquiry*, 10(2): 1–26.
- Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.
- Saffran, E., Berndt, R., & Schwartz, M. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, 37: 440-479.
- Tesak, J., & Dittmann, J. (1991). Telegraphic style in normals and aphasics. *Linguistics*, 29: 1111-1137.
- Zdorenko, T. (in press). Subject omission in Russian: A study of the Russian National Corpus. In: S. T. Gries, S. Wulff, & M. Davies (ed.). *Language and Computers. Corpus-linguistic Applications: Current Studies, New Directions*. Amsterdam/New York, NY, VI, p. 260.

Anastasiia Romanova

romanova.anastasiya@hotmail.com