Optional Infinitives in Russian and Their Implications for the Pro-Drop Debate

Eva Bar-Shalom and William Snyder
Department of Linguistics, University of Connecticut, Storrs

Objectives

In this paper we will argue that evidence from Russian child language favors an analysis of (adult) Russian as a non-pro-drop language (Franks 1995), rather than a pro-drop language (Benedicto 1993). Our argument is based on the distribution of Optional Infinitives (OI's), a construction that is specific to child language. By Optional Infinitive we mean an infinitival verb form used in a root clause, where the target language requires a finite form. OI's have been reported in various languages, including:

(1) English Eve sit floor (Brown 1973).
(2) German: Zahne putzen (Poeppel and Wexler 1993).
(3) Dutch: Ik ook lezen (Weverink 1989).
(4) Hebrew: He lehagid shalom (Rhee and Wexler 1995).
(5) Russian: Bantik zavjazat'.

The acquisitional evidence to be reported here is based on the spontaneous productions of a Russian child, Varya, between the ages of 1 year, 6 months and 2 years, 4 months (1;6-2;4). These data were collected by E. Protassova of the Russian Academy of Education and recorded in the CHILDES Database (MacWhinney and Snow 1985).[1] In addition, we shall report preliminary findings from our on-going longitudinal study of Tanya, a two-year-old, monolingual, Russian-speaking child.

Our study was prompted by observed variation, across languages, in the Optional Infinitive phenomenon. One point of variation is the extent to which different languages exhibit an OI stage at all.[2] In some languages children produce OI's in only a small percentage of their matrix sentences, and produce OI's for only a brief period of time (Phillips 1995, 1996). In Italian, for example, Guasti (1992, 1993) has reported that optional infinitives are present but occur quite infrequently, and disappear

early in the course of language development.[3] Similar findings have been reported for Spanish and Catalan (Grinstead 1993, Sano and Hyams 1994, Phillips 1996). Children acquiring German, Dutch and French, however, produce OI's substantially more frequently and for a longer period of time (Sano and Hyams 1994; Phillips 1996).

The crux of our argument will be a generalization due to Wexler (1995):

(6) A language L has an OI stage iff INFL in that language does not license null subjects.

This generalization is consistent with evidence from a sizable sample of languages, but is perhaps most dramatically supported by Rhee and Wexler's (1995) study of acquisition data for Hebrew. Rhee and Wexler found that Hebrew-speaking children use OI's exclusively in those environments where adult Hebrew prohibits null subjects.

In the following section we will briefly summarize two competing analyses of null-subject phenomena in Russian. We will then present our acquisitional findings for Russian. Finally, we will combine our acquisitional evidence with Wexler's generalization and discuss the implications for the analysis of adult Russian.

The null-subject phenomenon in Russian
Franks (1995) argues that [+/-pro-drop] is not an adequate characterization of the differences between null and non-null subject languages. All of the Slavic languages employ null expletive subjects, and several Slavic languages (e.g. Polish, Serbo-Croatian, Bulgarian) also permit null thematic subjects. Yet, certain Slavic languages, including Russian, generally require thematic subjects to be overt.[4] Franks interprets Russian as a genuinely non-pro-drop language with respect to thematic subjects (in contrast to languages such as Polish or Serbo-Croatian), and argues that the use of null expletive subjects in Russian and in Slavic more generally is due to an independent parametric property.

Benedicto (1993), in contrast, analyzes Russian as a (technically) pro-drop language, even with respect to thematic subjects. In Benedicto's analysis, two conditions are needed for the identification of *pro:* adequate phi-content and an adequate structural configuration. According

to Benedicto, Russian has adequate phi-content, and the structural configuration needed for identification of *pro* in Russian could be obtained by overt V-raising. Yet, there is evidence that Russian lacks overt V-raising. For example, an adverb cannot be placed between the V and a direct object in Russian, in an example such as (7).

(7)   *? Anna …itajet bystro knigi   (cf. Benedicto, p.5).
       Anna   reads   quickly books
       'Anna quickly reads the books.'

A further diagnostic for the presence or absence of verb-movement is the distribution of floating quantifiers. Benedicto observes that in this respect Russian again patterns with non-V-raising languages:

(8)   * Studenty …itajut vse knigi   (cf. Benedicto, p.5).
       students read:3pl:pres   all books
       'The students all read the books.'

Benedicto argues that Russian subject pronouns are not full XP's, but are instead a lexicalization of AGR-S^0. In her analysis, when V fails to raise and identify the phi-content of AGR-S, *pro* is not licensed, and the phi-features are instead lexicalized as an independent X^0 pronoun. In this way, Russian is in some sense an "underlyingly" pro-drop language, and does not employ the full XP subject pronouns found in conventional non-pro-drop languages. Yet, Russian fails to exhibit the surface characteristics of a pro-drop language because it lacks overt V-raising.

The purpose of our acquisitional study was to learn whether Optional Infinitives occur in Russian child language and if so, whether they occur more frequently and/or later than the OI's observed in languages of the Italian type (cf. Phillips 1995). If we assume that Wexler's (1995) generalization (6) is correct, the prediction of (Benedicto 1993) is that Russian children should produce few if any OI's, much as in Italian. The prediction of (Franks 1995), in contrast, is that Russian children should pattern with children acquiring French or German, and produce OI's over a considerable period of time, with a frequency substantially higher than that of Italian children.

Methods and Results

The evidence we examined was drawn from Varya's corpus in the CHILDES database (MacWhinney and Snow 1985), and also from our own, on-going study of a Russian-speaking child, Tanya. The data for Varya were collected by Eketerina Protassova, Varya's mother and a psycholinguist.

Varya

For Varya, we analyzed all the utterances in six of Varya's transcripts (ages 1;6, 1;7, 1;9, 1;10, 2;0 and 2;4), and coded all matrix-clause verbs as [+finite] or [-finite]. Utterances were omitted from our final counts if the verb was an imperative, or if the verb was a non-finite form that is grammatically possible in adult Russian.[5] An utterance was also omitted from the final counts if it was an imitation of a previous adult utterance, if it was an exact repetition of the child's own previous utterance, or if it was not fully intelligible.

The resulting counts are given in Tables 1-6. For each transcript, Varya's mean length of utterance in words (MLU_{word}) was calculated and is provided in the table. This measure serves an approximate index of the child's overall linguistic development (Brown 1973).[6]

TABLE 1 (Varya 1, age 1;6. MLU_{word} 2.60)

+FIN	-FIN
67	14

TABLE 2 (Varya 2, age 1;7. MLU_{word} 2.99)

+FIN	-FIN
115	37

TABLE 3 (Varya 3, age 1;9. MLU$_{word}$ 2.4)

+FIN	-FIN
115	26

TABLE 4 (Varya 4, age 1;10. MLU$_{word}$ 3.10)

+FIN	FIN
247	14

TABLE 5 (Varya 5, age 2;0. MLU$_{word}$ 3.3)

+FIN	-FIN
228	9

TABLE 6 (Varya 6, age, 2;4. MLU$_{word}$ 4.26)

+FIN	-FIN
224	5

Thus, in Varya's corpus, the percentage of [-finite] forms reaches its peak at age 1;7 and then declines: 14 OI's at age 1;6 (17.3%), 37 at 1;7 (24.3%), 26 OI's at 1;9 (18.4%), 14 at 1;10 (5.7%), 9 at 2:0 (3.9%), 5 at 2;4 (2.2%).

Tanya

In April of 1996 we began weekly videotaping sessions with Tanya, age 2;4, the daughter of a Russian graduate student at the University of Connecticut.[7] Tanya is being raised in a nearly monolingual Russian environment; her mother, with whom she spends almost all her time, is a monolingual Russian-speaker. The child's father knows English but speaks Russian at home.

To date, seven of Tanya's videotapes, containing some 1,247 child utterances, have been transcribed and then analyzed in the same way as Varya's transcripts.[8] The results are presented in Tables 7 through 13; only the main verbs of matrix clauses are included. The frequency of OI's in Tanya's seven transcripts is thus 28.0%, 10.0%, 10.7%, 8.0%, 2.6%, 4.0%, and 2.7%. Overall, 22 of Tanya's 232 (matrix-clause, main) verbs, or 9.5%, are OI's. Tanya's use of OI's in particular, and of verbs in general, is somewhat less frequent than Varya's, but the proportion of

Tanya's verbs that are OI's is still considerable, and markedly higher than for Italian children, as discussed below.

Table 7. Tanya, Age 2;5.11, MLU$_{word}$ 1.66

+FIN	-FIN
23	9

Table 8. Tanya, Age 2;5.24, MLU$_{word}$ 1.59

+FIN	-FIN
18	2

Table 9. Tanya, Age 2;6.1, MLU$_{word}$ 2.11

+FIN	-FIN
50	6

Table 10. Tanya, Age 2;6.8, MLU$_{word}$ 2.35

+FIN	-FIN
22	2

Table 11. Tanya, Age 2;7.6, MLU$_{word}$ 3.26

+FIN	-FIN
37	1

Table 12. Tanya, Age 2;7.26, MLU$_{word}$ 2.58

+FIN	-FIN
24	1

Table 13. Tanya, Age 2;8.12, MLU$_{word}$ 2.48

+FIN	-FIN
36	1

Discussion
The analyses of Russian null subject phenomena in (Benedicto 1993) and (Franks 1995) make contrasting predictions for the occurrence of OI's in children's Russian, as described above. In brief, (Benedicto 1993) predicts that Russian children will produce few OI's, much as has been reported for Italian and Spanish. The analysis of (Franks 1995), in

contrast, predicts a higher rate of OI's in Russian than in Italian or Spanish.

Our results in fact provide strong support for the existence of an OI stage in Russian that is considerably more robust than anything observed in Italian. The average percentage of OI's in the six transcripts analysed for Varya is 10.9%; in Tanya's seven transcripts the percentage is 9.5%. In an MLU-based comparison with Martina, the only Italian child in the CHILDES database with frequent OI's (Guasti 1993), Varya and Tanya both employ OI's later in the acquisitional timecourse.[9] Martina's peak use of OI's occurs when her MLU is 2.0 words. By the time her MLU has reached 2.5 words, her frequency of OI's is lower than 1%, where it remains.

In contrast, Varya's peak use of OI's occurs when her MLU is 3.0 words (between the ages of 1;6 and 1;9, and the incidence of OI's remains higher than 2% even when her MLU exceeds 4.2 words.[10] In Tanya's corpus the rate of OI's peaks at the lower MLU of 1.7 words, but much as in Varya's corpus, remains higher than 2% through the end of the available corpus, where the MLU ranges between 2.5 and 3.3 words. Both Varya and Tanya produce OI's with greater than 2% frequency over a period of many months, and produce a considerable number of OI's in absolute terms. By way of contrast, of the three Italian children studied by Guasti, two totaled fewer than ten OI's in their entire corpora, and the remaining child, Martina effectively stopped producing OI's before her MLU reached 2.5 words.[11]

The acquisitional evidence from Varya and Tanya thus provides strong support for the existence of a robust OI phenomenon in Russian. The findings of our study therefore favor the analysis of Russian as a true non-pro-drop language (Franks 1995), rather than a technically pro-drop language (Benedicto 1993).

Notes

1. The CHILDES database includes longitudinal and cross-sectional observational data from children acquiring any of a considerable number of languages. For Slavic, the database currently contains longitudinal data from one Russian child and three Polish children.
2. A second point of variation, concerning the occurrence of overt/null subjects in sentences with OI's, is discussed in (Bar-Shalom, Snyder, & Boro 1996), but

will be set aside here.

3. This point will be elaborated below.

4. In informal speech, it is sometimes possible to omit subject pronouns when their referents are contextually determined (Chvany 1973). This type of subject drop is considered a discourse-related phenomenon, and we will follow Franks (1995) in distinguishing it from "true" subject drop. Also, see (Franks 1995) for a discussion of the occasional possibility of an overt expletive in Russian, but not in Serbo-Croatian or Polish.

5. Imperatives in Russian are subjectless sentences that have their own verb morphology.

6. MLUs in our study were calculated in words, rather than morphemes, following Valian (1991). Valian argues that the word-based measure is more informative when the child's target language is richly inflected.

7. The pilot research reported in this section was made possible by a Faculty Major Research Grant to Dr. William Snyder from the University of Connecticut Research Foundation. The Russian data are simultaneously being analyzed with respect to argument structure and morphological processes of word-formation (cf. Snyder 1995), in a larger project on the syntax-morphology interface in child language acquisition.

8. An additional four videotapes, the first four recorded, have not been transcribed because initially Tanya spoke very little; we believe this was due in part to the child's need to become comfortable with the experimenter and with the presence of a videocamera.

9. MLUs are commensurable across Russian and Italian, because the languages are morphologically quite similar.

10. Note than the incidence of OI's is also low in two other pro-drop languages, Spanish and Catalan: .05 and .03, respectively (Sano and Hyams 1994).

11. Given that longitudinal corpora were available for only three Italian children at the time of Guasti's study, the fact that Martina uses OI's more frequently than the other two is difficult to interpret. One possibility is that all children pass through an OI stage unless OI's are directly blocked by a language-particular property of the morphosyntax in their target language. In this case, a child who masters the relevant morphosyntax of Italian early will never produce OI's, but a child who masters the morphosyntax later may exhibit a brief period of OI production, as observed for Martina. Since Guasti's study, several additional Italian corpora have become available, and our hope is that these data may shed some light on the limited and variable OI phenomenon in Italian.

References

Bar-Shalom, E., Snyder, W., and Boro, J. 1996. Evidence for the

Optional Infinitive Stage in Russian. *In Papers on Language and Processing*. Storrs, CT: University of Connecticut Working Papers in Linguistics.

Benedicto, E. 1993. AGR, Phi-features and V-movement: Identifying pro, *UMASS Occasional Papers.*

Brown, R. (1973). *A First Language*. Cambridge, Mass.: Cambridge University Press.

Franks, S. 1995. *Parameters of Slavic Morphosyntax*. Oxford University Press, New York, New York.

MacWhinney. B. and Show, C. 1985. The Child Language Data Exchange System. *Journal of Child Language* 12:271-296.

Phillips, C. 1995. Syntax at age two: Cross-linguistic differences. In T. Schutze, J. Ganger and K. Broiher (eds). *Papers on Language Processing and Acquisition, MITWPL 26.*

Phillips, C. 1996. Root infinitive are finite. Some implications of cross-linguistic differences in two-year olds' syntax. In Stringfellow, A., Cahana-Amitay, D., Hughes, A. and Zukowski, A. (eds.). *Proceedings of BUCLD 20*. Cascadilla Press.

Pierce, A. 1989. On the emergence of syntax: a cross-linguistic study. Ph.D. Thesis, MIT.

Poeppel and Wexler, K. 1993. The full competence hypothesis of clause structure in early German. *Language* 49:1-33;

Rhee, J. and Wexler, K. 1995. Optional Infinitives in Hebrew. In C. Schutze, J. Ganger, and K. Broihier (eds.). *Papers on Language Processing and Acquisition, MITWPL 26.*

Sano, T. and Hyams, N. 1994. Agreement, finiteness, and the develpment of null arguments. In M. Gonzales (ed.). *Proceedings of NELS 24*, Amherst, MA: GLSA.

Snyder, W. 1995. *Language Acquisition and Language Variation: The Role of Morphology*. Doctoral dissertation, MIT. Cambridge, MA: MIT Working Papers in Linguistics.

Valian, V. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40:21-81.

Weverink, M. 1989. *The Subject in Relation to Inflection in Child Language*. M.A. Thesis, University of Utrecht, the Netherlands.

Wexler, K. 1995. Feature-interpretability and optionality in early child grammar. Paper presented at the Workshop on Optionality,

University of Utrecht, the Netherlands.

Addresses for correspondence:

Eva Bar-Shalom, Department of Linguistics, U-145
University of Connecticut, Storrs, CT 06269 USA
E-mail: barshalo@uconnvm.uconn.edu

William Snyder, Department of Linguistics, U-145
University of Connecticut, Storrs, CT 06269 USA
E-mail: wsnyder@uconnvm.uconn.edu