The acquisition of relative clauses in Russian and Polish in monolingual and bilingual children

Alla V. Peeters-Podgaevskaja,1 Bibi E. Janssen1 and Anne E. Baker1,2
1University of Amsterdam / 2University of Stellenbosch

In this paper, we present the results on (re)production of relative clauses in a sentence repetition task by Russian and Polish monolingual and bilingual children (3;6–6;6 years). Russian and Polish both involve overt case marking in the relative clause formation, but the case marking in Polish is more regular and transparent. Our results suggest that, when case marking is the only disambiguating cue, its transparency and regularity contributes to a significantly better performance. This explains why the Polish children, both monolingual and bilingual, were more successful in reproduction of subject and object relatives compared to their Russian peers.

Keywords: acquisition of relative clauses, transparency in case marking, Slavic languages

1. Introduction

Relative clauses (RCs) form part of the complex structures in any language and as such have given rise to a considerable amount of research in the field of linguistic typology and language acquisition. The factors that influence the acquisition process have attracted the most attention (Kidd et al., 2007). Cross-linguistic work has indicated that there seems to be an interplay between different factors, such as semantics and pragmatics of head nouns and the discourse frequency of different types of RCs. Structural aspects are also of influence, such as distance dependency, word order, the grammatical role of the relativized elements within the subordinate clause, and case marking.

With regard to the last two structural elements there are contradictory results in the literature. From previous studies we know that, in languages with
post-nominal RCs, subject relatives (SRs) are generally easier to process than object relatives (ORs) (Diesel & Tomasello, 2005). Research on languages with prenominal RCs has, however, yielded mixed results. Some studies have shown a SR preference (Korean: Kwon et al., 2013; Basque: Gutierrez-Mangado & Ezeizabarrena, 2014), while others have reported an OR advantage (Cantonese and Mandarin: Chan, Matthews, & Yip, 2001; Basque: Gutierrez-Mangado, 2011).

The results on the influence of case marking are also mixed. For languages with overt case marking it is not clear whether the inflectionally rich nominal morphology is helpful in the acquisition of RCs. Some studies found a positive effect of overt case marking in comprehension (Polish: Smoczyńska, 1985; Russian: Polinsky, 2008). Others questioned such beneficial effects in comprehension and imitation/reproduction tasks (Polish: Mróz & Flynn, 2010; Romanian: Bențea, 2012; Sevcenco et al., 2012a, 2012b).

It therefore continues to be a topic of research to explore the different factors involved in the acquisition of this complex area of syntax. By investigating comprehension and (re)production of RCs in Russian and Polish by means of a sentence repetition task this study aims to contribute to the evidence on the role of the relativized elements and overt case marking. The two languages to be studied are, on the one hand, similar with respect to richness of their nominal morphology and the structure of RCs. On the other hand, there are some differences in the transparency of the realization of case marking and the structure of RCs. Any differences found between two languages may again indicate the role of these factors.

The studies mentioned above have examined the acquisition and performance of RCs in monolingual speakers. In second language learning, on the other hand, research on RCs acquisition is strongly focused on the performance of adult learners and then within a small set of languages with predominantly poor nominal morphology (e.g., English, Japanese, Spanish: Flynn, 1983; Swedish: Hyltenstam, 1984; Italian: Croteau, 1995; English, Kazach: Vinnitskaya et al., 2003; Japanese, Chinese, Vietnamese, Indonesian: Kanno, 2007). Information on how bilingual children comprehend and produce RCs in their L2 with inflectionally rich morphology is extremely limited. By studying this population, who has already been shown to be slower in acquiring the case marking system (Peeters-Podgaevskaja, 2008; Janssen, 2016), we will be able to examine whether or not and to what extent higher transparency and regularity in case (and gender) marking contributes to faster and more successful acquisition of RCs in general, not only in monolingual children. By comparing bilingual children with monolinguals we expect to find commonalities as well as differences in their developmental paths, preference patterns and error types that will deepen our understanding of bilingual development of syntax.
2. Previous research on child acquisition of relative clauses

The ability to produce and to comprehend RCs is often considered a milestone in language acquisition, whether as a first or second language. Many studies on different languages have shown that the first RCs are already produced by children in their first language between 2;6 and 3;0 years, but up to the age of six years their use in spontaneous speech is not frequent (De Villiers & De Villiers, 1985: 112). The RCs that are produced also often contain errors, such as inappropriate word order, omission of relative pronouns and prepositions in oblique RCs (in which an oblique element is relativized), and substitution of relative pronouns with interrogative pronouns.

Syntactically, child learners face the complex task of learning how to assign the correct grammatical function to the modified noun phrase. The Noun Phrase Accessibility Hierarchy (NPAH), which was introduced by Keenan and Comrie in 1977, predicts the order of relativization: subject > direct object > indirect object > oblique > genitive > object of comparison. According to the authors, NPAH implies that:

if a language can relativize a particular position on the NPAH, then it must also be able to relativize all higher positions. Thus, if a language can relativize direct objects, it must also be able to relativize subjects. (Comrie, 2007: 303)

Three structural features are important in understanding RCs and their acquisition: the head, the element in the main clause modified by the RC; the gap, the element relativized inside the RC; and the filler or complementizer, the head of the RC marked by a relative pronoun. Although the head and the gap have a variety of different syntactic functions, SRs (1) and ORs (2) are considered the most common (the examples with the RC italicized are taken from Diessel and Tomasello (2005: 882)):

(1) The horse that pushed the goat stands on the lion.
(2) The cow that the sheep pushed stands on the kangaroo.

In early production a clear preference for SR above OR has been reported for several languages (e.g., English: De Villiers & De Villiers, 1985; German: Mills, 1985; Dutch: van der Meer et al., 2001; Hebrew: Arnon, 2011; Romanian: Sevcenco et al., 2012a), but not for others (Cantonese and Mandarin: Chan, Matthews, & Yip, 2011). In bilingual children there may be an influence of the other language in this regard. Yip and Matthews (2007) reported on three young Cantonese-English bilinguals (analyzed by means of longitudinal diary reports between ages 1 to 6), who started in their speech production of English with ORs, which is in contrast with the monolingual acquisition pattern in English, but in accordance with the Cantonese
Acquisition of relative clauses in Russian and Polish

pre-nominal relativization strategy. This fact has been interpreted as stemming from the influence of their more dominant language, Cantonese, in which ORs follow canonical SVO word order, whereas SRs have VOS (Yip & Matthews, 2007: 289–291).

It has been argued that the first RCs produced are similar to the structure of a simple (in)transitive sentence. In English, for example, the RC structures are attached to the predicate nominal of a copular clause (3) or an isolated noun phrase (4) (Diessel & Tomasello, 2005: 883).

(3) Here's the tiger that's gonna scare him.

(4) The girl that came with us.

According to Diessel (2007: 312–313), this developmental pattern in English is related to two factors. On the one hand, relatives that are attached to the predicate nominal of a copular clause (3) or an isolated noun phrase (4) are quite frequent in child directed speech because they match the communicative needs of young children. On the other hand, they are much less complex than other types of RCs resembling the structure of simple (in)transitive sentences that denote a single state of affairs and characterize child speech in the early stages.

Interestingly, the comprehension of RCs, and especially ORs, appears in many languages to be more problematic than their production, even at an older age. For example, the study of Duinmeijer (2016) reported that Dutch monolingual children (mean age 8;4) and adolescents (mean age 14;5) experienced difficulty with the correct interpretation of ORs showing results below chance, while their comprehension of SRs was at ceiling (Duinmeijer, 2016: 182). Some scholars ascribe these problems with comprehension to the degree of the dependency distance between filler and gap: the greater the distance between filler and the relativized element in the subordinated clause, the greater the processing load of RCs (the Filler-Gap hypothesis) (Hawkins, 1987). It has been suggested that this is why SRs are generally much easier to acquire, while ORs cause difficulty in production as well as in comprehension (see Kidd et al., 2007, for a discussion of several languages).

Some other researchers argue that problems with ORs can be explained by the fact that young children are not capable of understanding syntactic structures involving wh-movement, that is movement of the relative pronoun or relative phrase from its canonical position to the front of the RC (Friedmann & Novogrodsky, 2004). The result of the movement is a relativization gap at the original position in the sentence structure. When wh-movement takes place without crossing an element having another syntactic role, language learners do not have processing difficulties. The latter occurs when wh-movement takes place across the subject as it is the case in ORs. Thus in sentence (2) the cow in the RC has to be moved across sheep in order to land at the beginning of the clause.
Alla V. Peeters-Podgaevskaja, Bibi E. Janssen and Anne E. Baker

Diessel and Tomasello (2005), however, argue that the preference for SRs can be explained by their automatic activation due to their high frequency in input and by children’s preference for the actor (the highest ranked thematic role) to be expressed by the sentence-initial noun phrase. This preferred subject interpretation results, for example, in the interpretation of ORs as SRs in English and German (Diessel & Tomasello, 2005: 890).

Processing of RCs can be significantly favored and enhanced by different kinds of semantic, pragmatic and morphological cues that can be found in the clause. For example, animacy, distinctive verb agreement, or gender have been shown to assist in disambiguating SRs and ORs in several languages (e.g., Dutch: Mak, Vonk, & Schriefers, 2002; Hebrew: Arnon, 2011; Romanian: Sevcenco et al., 2012b). The same disambiguating role has been observed for proper nouns and pronouns (Warren & Gibson, 2002: 85), and semantic and pragmatic plausibility (Mecklinger et al., 1995).

Frequency is also considered to play a role: higher frequency of SRs and lower frequency of ORs with animate head nouns in discourse may explain the problems in processing of ORs (Mak, Vonk, & Schriefers, 2002; Diessel & Tomasello, 2005). If there is ambiguity between SRs and ORs, that is when neither grammatical cues (word order or agreement) nor semantic and pragmatic cues can facilitate the correct interpretation (see, for example, Dutch material in Gibson, 1998: 56–57; Duinmeijer, 2016: 180–183), a subject interpretation is preferred (German: Mills, 1985; Dutch: Mak, Vonk, & Schriefers, 2002).

For those languages that have case marking, it could be the case that this marking disambiguates syntactic roles at the onset of the RC and thus aids children’s comprehension and reproduction of RCs. On the other hand, overt case marking could be seen as a complicating factor in production, due to its grammatical encoding load. Learning the case marking system does not seem to be problematic for young children acquiring Russian or Polish, but it does appear to have an impact on bilingual children (Russian: Polinsky, 2007; Peeters-Podgaevskaja, 2008; Polish and Russian: Janssen, 2016).

Findings from a few comprehension studies support the idea that case marking at the beginning of the clause aids comprehension. Polinsky (2008) in her study on comprehension of Russian RCs with different word orders (SVO and SOV for SRs, and OVS and OSV for ORs) elicited in a picture matching task, reported that

---

1. Dutch RCs are structurally ambiguous. The word order is the same in SRs and ORs with the verb taking the final position in the subordinate clause. Dutch has no case marking system to indicate which element takes which syntactic role. RCs can be also semantically ambiguous, if the subject and object of a RC have the same features (in terms of animacy, number and gender), and verb semantics does not reveal which noun phrase is likely to have which role (Duinmeijer, 2016: 159).
monolingual Russian children aged 6;6 had no problem in assigning the correct syntactic functions to the nouns that were both animate. They performed at ceiling with more than 90% of accuracy. The same accuracy was observed in Polish children aged six who correctly performed in a picture selection task on different RCs structures (object-subject, prepositional object-subject, object-object, and object-genitive object) (Smoczyńska, 1985: 662–663). However, this study used items involving animate agents and inanimate objects: the semantics therefore possibly made the task easier masking the role of case marking.

In contrast to these studies, Bențea (2012) found no facilitative effect of case marking in comprehension of RCs in Romanian. Overt case marking, expressed by the accusative particle pe in combination with the relativizer care ‘that, which’ and occurring at the immediate onset of the RC, did not help children (divided into three age groups, from 4;0 to 6;11) to disambiguate between ORs and SRs in a picture selection task, in which all nouns were animate. Moreover, the resumptive clitics at the relativization gap that should make this structure even more transparent, did not seem to facilitate the interpretation either. All age groups scored around 25% on ORs. Two major error types were encountered: agent errors and reversed roles (Bențea, 2012: 8).

The findings on comprehension of Romanian RCs from Bențea (2012) are supported by the results from the research on (re)production of Romanian RCs. Sevcenco et al. (2012a) showed that monolingual children do not benefit from overt case marking in their native tongue. In a elicited preference task Romanian children aged 5;0 to 6;11 produced SRs with 92% of accuracy, while ORs reached only 50%. Interestingly, there was also evidence for an avoidance strategy in the adult control group: they produced direct ORs but only in 23% of the contexts requiring them (Sevcenco et al., 2012a: 62).

Similar difficulties with the (re)production of ORs were observed in German where case marking is explicitly encoded on the articles and the ending of a nominal constituent. Since the relative pronoun is marked for case and occurs at the beginning of a RC, it is expected to provide all of the information necessary to determine the relativized syntactic role. However, in a sentence repetition task German-speaking children (mean age 4;5) performed significantly better on SRs than ORs (Diessel & Tomasello, 2005: 892). The preferred subject interpretation resulted, for example, in structure conversion of ORs to SRs. The occurrence of conversion, however, did not directly reflect the lack of grammatical knowledge. The German children sometimes started with the relative pronoun in the nominative case, but then corrected themselves and changed the relative pronoun to the case required in the context (Diessel & Tomasello, 2005: 898).

The few results that exist from bilingual research have shown that bilingual children also experience difficulty with the reproduction of RCs. In an elicited
imitation task from the study of Mróz and Flynn (2010), the English-Polish bilinguals, who were dominant in English, scored very poorly in English with 31% overall accuracy at age 5, and 28% at age 6. Their overall accuracy in Polish, which has case marking, was at floor (8% at age 5 and 7% at age 6).

As the studies discussed above show, the findings on acquisition of RCs in languages with a rich case morphology are contradictory. It would seem logical that overt case marking could be a facilitating factor in comprehension, since it results potentially in disambiguation at the immediate onset of the RC, but there is no sufficient evidence to prove this assumption. There is considerable variability in the analysed data due to differences in methodology (with respect to tasks and procedures, differences in measures (e.g., overall accuracy vs. type accuracy) and the size of subject samples) and language-specific features (e.g., case particles, as in Romanian, vs. explicit case marking on the word ending, as in Slavic languages, etc.). In order to increase the validity and reliability of conclusions, languages with a highly comparable nominal morphology and RC structures need to be studied in a highly comparable setting.

Taken overall, previous results on monolingual and bilingual acquisition of RCs indicate that the predictions of the NPAH hold for languages with post-nominal RCs, but not for those with prenominal RCs. In bilingual children there is some evidence of interference from one language to the other. Case marking does not clearly help monolingual children in comprehension (picture matching/selection tasks) and reproduction of RCs (elicited preference/imitation tasks, sentence repetition tasks) and appears to be of no help to bilingual children.

This study aims to contribute to the knowledge on the role of overt and transparent case marking in comprehension and production of RCs in both monolingual and bilingual children.

3. Goals of this study

As the discussion of the literature in the previous section has shown, there is relatively little information on the acquisition of RCs in languages with rich nominal morphology. That research has not provided a clear answer to the question whether the observed difficulties in child production and comprehension of ORs in languages with post-nominal RCs also hold for languages with rich inflection. The main goal of this study is to investigate the influence of overt and transparent case (and gender) marking on the acquisition of RCs. In this respect Russian and Polish provide an ideal testing ground due to their highly comparable nominal morphology and structure of RCs, on the one hand, but some difference in the transparency and phonological regularity of case suffixes, on the other. Previous
research (Janssen, 2016) has shown that this difference in transparency leads to slower acquisition of case and gender of Russian nouns, compared to Polish. This might also lead to differences in performance in the acquisition of RCs between Russian children and their Polish peers.

However, as remarked earlier, monolingual children are relatively fast in learning the case and gender marking system in both Russian and Polish (at age 3;6–4;0) (see Gvozdev, 2007; Eliseeva, 2015; Smoczyńska, 1985). This makes it difficult to find a range of ability in case marking using only a monolingual group at an age that is possible to test. By studying bilingual children who are slower in acquiring the case marking system (Peeters-Podgaevskaja, 2008; Janssen, 2016) alongside monolingual children the relationship between case marking and acquisition of RCs should become clearer. We will be able to examine whether or not and to what extent higher transparency and regularity in case (and gender) marking contributes to faster and more successful acquisition of RCs.

This study will consider the relative difficulty of SRs compared to ORs whereby it is predicted that in Russian and Polish SRs will be easier to comprehend and to produce. This should also hold for the bilingual children since their second language is Dutch, which also has postnominal RCs, and since neither SRs nor ORs in Dutch need any grammatical adjustment or transformation. The study will also provide further information on types of errors and pattern preferences in the both languages.

4. The structure of relative clauses in Russian and Polish

Russian and Polish show many parallels in relative clause formation. Any position on the NPAH can be relativized. RCs in both languages are based on filler-gap dependencies and involve relative pronouns which cannot be deleted. For example, kotoryj/który ‘that, which, who’ agrees with the head in the main clause in gender and number, while the case marking depends on the syntactic function of gap/filler. In contrast to English, the word order in both types of transitive RCs is identical. However, it is not true for the syntactic functions of the constituents: in SRs the direct object, expressed by the relative pronoun and marked for the accusative, is in the final position of the subordinate clause (SVO) (see 5a); in ORs the subject in the nominative takes the final position (OVS) (see 5b): 2

2. Note, the default word order in the subordinate clause depends on the nature of the nominal constituents: when subjects or objects in the RC are expressed by personal pronouns, the word order is different (SOV and OSV).
Although the canonical word order in SRs is SVO and in ORs OVS, in a simple transitive sentence this latter inversion is not frequent. Despite the fact that Russian and Polish are free word order languages, there is a preference for the SVO order (in the neutral discourse style). In Russian, for example, young children show a very strong bias towards SVO in their speech production (Kallestinova, 2007; Udes, 2013).

In both languages, the most common relative pronoun is kotoryj/który ‘which, that, who’ that belongs to the adjectival declension (it is declined as an adjective) and is marked for gender, case, and number. Despite major similarity in adjectival case endings between Russian and Polish (and hence in the declension of kotoryj/który), two linguistic features make the whole declension of adjectives in Russian more complex and less transparent. In contrast to Polish, the vowels in Russian case suffixes are determined by the final stem consonant and vowel reduction. While Polish unstressed vowels are not phonetically reduced and sound clear, Russian unstressed vowels undergo qualitative change (for example, for the vowels occurring in the case suffixes: /a/ and /o/ are reduced to [ə], and /e/ is reduced to [i]) (Timberlake, 2004). Thus, whereas Polish has only one adjectival declensional paradigm for each gender, Russian has for each gender three paradigms with allomorphic endings (stressed endings with the final hard stem consonant, unstressed endings with the final hard stem consonant, and unstressed endings with the final palatalized stem consonant) (see overview in Janssen, 2016: 40–41).

In relative clause formation, other pronouns can also be used as a relativizer. However, they are less common and differ cross-linguistically in frequency, pragmatics and syntactic usage. For example, RCs with the relative pronoun kakoj/jaki ‘which’, which is also declined as an adjective, refers to objects belonging to a similar category and sharing the same qualities and/or features (Podgaevskaja & Honselaar, 2007: 503). However, the usage of the Polish jaki is fairly rare, while the Russian kakoj is more frequent. Note, that in the standard Russian the demonstrative tot ‘that’ in the main clause is required; in colloquial Russian it can be omitted:

(6) Rus.: Ja kuplju tebe podarok, kakoj zaxočes’.
I buy you gift-m.sg.acc that-m.sg.acc you.want
‘I will buy you a gift whichever you want.’
In informal language, čto/co ‘that’ can be used instead of kotoryj/który. This relative pronoun cannot be modified and refers both to animate and inanimate antecedents. However, there is a difference between the two languages. While the Russian čto is infrequent and occurs only in SRs, the Polish co is very frequent and can be easily used in both types of RCs. In the subject function (7a), co is comparable to the Russian čto. In non-subject RCs (7b) the relative marker co co-occurs with a resumptive pronoun that is then obligatory (Lavine, 2003). Co cannot be modified and only signals the onset of the RC, while the resumptive pronoun, being part of the argument structure of the verb, is marked for case, number, and gender (examples borrowed from Skwarski, 2010):

(7) a. Pol.: Pióro, co leżalo na stole.
pen-n.sg.nom that-n.sg.nom lied on table
‘The pen that was lying on the table.’

b. Pol.: Pióro, co nim pisałam.
pen-n.sg.nom that-it-n.sg.instr I.wrote
‘The pen with which I wrote.’

In colloquial Polish, however, the resumptive pronoun can be omitted in the accusative context, if the antecedent is masculine inanimate and the head of the RC is animate. Compare (8a) (borrowed from Lavine, 2003) with (8b) without resumptive pronoun go:

(8) a. Pol.: Ten samochód, co go Janek widzia wczoraj,
that car that-it-m.sg.acc Janek saw yesterday
zniknąć tajemniczo.
disappeared mysteriously
‘That car that Janek saw yesterday disappeared mysteriously.’

b. Pol.: Ten samochód, co Janek widzia wczoraj, zniknąć
that car that Janek saw yesterday disappeared
tajemniczo.
mysteriously
‘That car that Janek saw yesterday disappeared mysteriously.’

In this respect, the usage of the Polish co resembles the usage of the Romanian care (with and without resumptive pronouns).

In contrast to the pronouns discussed above, the use of the relative kto is restricted in Russian and is extremely rare in Polish. In Russian, a kto-relative clause is mostly centre-embedded and refers only to animate antecedents that are specified by the deictic pronoun tot ‘that’ in the main clause. The relative kto is always associated with masculine and singular.
In sum, the most common and neutral relativizer in both languages is the pronoun *kotoryj/*który that agrees with the head in gender and number, while the case marking depends on the syntactic role of the complementizer in the RC. In Russian, the relative pronoun *kakoj* is a possible, but relatively infrequent alternative to *kotoryj* and shares the same grammatical features. In Polish, an alternative to *który* is the relative pronoun *co* that does not show any grammatical marking and is used with and without resumptive pronouns in both types of RCs. Use of other relative pronouns is restricted by specific contexts and therefore very infrequent.

5. Acquisition of relative clauses in Russian and Polish children

There is not much information available on the acquisition of RCs in Russian and Polish, only a few case studies. In Russian, the first utterances with the relativizer *kotoryj* seem to appear at age 2;6 (comparable with the age in the English data, but lagging behind Hebrew). According to Gvozdev’s data (2007) on one child aged 1 to 9, the first RCs are far from being adult-like: the utterances show many omissions, the order of clauses is deviant with the RC preceding the main clause, and the structures are ill-formed with the head being embedded in the RC itself (as shown in (9)). Target-like sentences appear only at age 5;3 (Gvozdev, 2007: 266).

\[(9)\] Mužički gde, rabotajut kotorye?
\[\text{little.men-pl.nom where work that-pl.nom}\]
\[= \text{Gde mužički, kotorye rabotajut?}\]
‘Where are those little men who are working?’

In contrast to Gvozdev, in the longitudinal data of Eliseeva (2015: 264–265) from one girl aged 1;6 to 3;0, there is more variety in the types of RCs (subject, object, prepositional) and relative pronouns (*kotoryj, kakoj*); no deviation from standard usage was found:

\[(10)\] Vedërko ja našla, vot iz kotorogo ja
\[\text{little.bucket-n.sg.acc I found here from that-n.sg.gen I}\]
\[\text{myla pol.}\]
\[\text{washed floor}\]
\[\text{(2;11)}\]
‘I have found the little bucket I washed the floor with.’

3. For Russian, in free narratives of adult native speakers 40% of all complex sentences feature relativization with *kotoryj* (Rifkin, 2002: 467).
Acquisition of relative clauses in Russian and Polish

(11) Ja tut iskala mašinku, *kakuju* mne nado. (2;10)
I here searched little.car-F.SG.ACC *which-F.SG.ACC* me *needed*
'I have looked here for the little car that I needed.'

Interestingly, although no errors in case marking and gender agreement between the head noun and the relative pronoun were reported by the authors, such errors cannot be excluded when we take into account the fact that the acquisition of adjectives is completed relatively late, after age 3;6 (Gvozdev, 2007: 440).

There is practically nothing known about bilingual acquisition of RCs in Russian and Polish. The spontaneous speech data of Dutch-Russian unbalanced bilinguals aged five to seven, for whom Russian was the weaker language, indicated that the relativizer *kotoryj* does not occur at all (Peeters-Podgaevskaja, 2008: 623). The children systematically and consistently make use of the pronoun *kakoj*, which is quite unusual in the speech of monolingual children:

(12) Ja vstretil mal’čik, *kakoj* tri goda.
I met *boy-M.SG.NOM* *which-M.SG.NOM* three *year*
(=Ja vstretil mal’čika, *kotoromu* tri goda.)
'I have met a boy who is three years old.'

In the production of Polish children (based on a longitudinal study of 10 children aged 1;6 to 6;0) (Smoczyńska, 1985), RCs appear at an earlier age of two years, first in isolated clauses. However, the data are dominated by the indeclinable relative pronoun *co* up to the age of four years (Smoczyńska, 1985: 660):

(13) Ja mam szoferkę *co* ma okno.
I have *cab-F.SG.ACC* that has *window*
'I have a cab that has a window.'

When used, *który* lacks any grammatical marking. Thus, at early stages of the acquisition of Polish RCs *który* seems to be misinterpreted as an equivalent of the indeclinable *co*. This observation is in part supported by Mróz and Flynn (2010: 16). In their study the substitution of *który* with *co* was observed only in the group of the three-years-old monolingual children and only in 1% of all utterances. Contrary to that, the English-Polish bilinguals substituted the *który*-relatives with *co*-relatives in 22% of all cases, regardless of age and without any structural changes in the sentence. Smoczyńska (1985: 663) claimed that this *co*-substitution is because of morphological complexity of *który* which requires case marking and gender/number agreement with its antecedent.

Further, on the basis of the observation of one child aged 3;9, Smoczyńska (1985: 663) argued that even when the child was perfectly able to repeat sentences with *który*, he did not necessarily understand the sentences correctly. In her pilot...
experiment with an act-out task, her subject failed to act-out the sentences he successfully repeated. Smoczyńska (1985: 663) concluded that children at certain age had difficulty not with sentence processing, but rather with information processing. Although this conclusion is drawn on the data of one child, it suggests asymmetries between production and comprehension of RCs in monolinguals.

In sum, on the basis of the very limited amount of previous research Polish children seem to start acquiring RCs earlier than Russian children. However, at earlier stages of acquisition Polish children produce mostly *co*-relatives, which are easier to master due to the fact that the relativizer *co* does not encode any grammatical information. A substantial amount of *co*-substitutions is also found in the bilingual data. This is in line with other studies reporting that very young children (that is with low language proficiency) show a clear preference for RCs which involve the fewest processing constrains in their languages (Dasinger & Toupin, 1994: 479). However, we have no evidence as to (a) whether Polish children are able to produce object *co*-relatives that require resumptive pronouns, coded for gender, case, and number; nor for (b) whether or not and to what extent Polish bilinguals experience difficulties with case and gender encoding on *który*.

Russian children, on the other hand, have to produce *kotoryj*-relatives that are both the most frequent and common in Russian but also full of grammatical information that must be obligatorily expressed. The alternative *kakoj* does not make the task easier as it morphologically behaves like *kotoryj*. Although attested in the monolingual data, *kakoj* is the only relative pronoun that is used by bilinguals. The absence of an indeclinable relative pronoun in Russian might lead to higher rates of grammatical errors and violations of the required constraints, especially in bilingual children. Further, allomorphy in gender and case endings makes the whole nominal system of Russian less transparent and regular. This can be seen as a complicating factor that might cause some delay in acquisition of the structure (Goldschneider & DeKeyzer, 2001).

6. The current study

As was stated in the Introduction, the main goal of this study is to examine the influence of overt and transparent case (and gender) marking on the acquisition of RCs in Russian and Polish. Comprehension and production of these structures will be investigated in both monolingual and bilingual children using a sentence repetition task for Russian and Polish. The study will consider accuracy but also pattern preferences and error types.
7. Method

7.1 Participants

In total, 79 typically developing Russian children, and 78 typically developing Polish children were recruited for this study. There was both a monolingual and bilingual group per language: Russian monolinguals (MoRu, \( n = 41 \), mean age: 4;9) and Russian-Dutch bilinguals (BiRu, \( n = 38 \), mean age: 5;1); and Polish monolinguals (MoPo, \( n = 40 \), mean age: 5;1) and Polish-Dutch bilinguals (BiPo, \( n = 38 \), mean age: 5;4).\(^4\) The monolingual children groups were recruited at preschools in St. Petersburg, Russia, and Warsaw, Poland. All monolingual children were aged between 3;6 and 6;6, typically developing and native speakers of either standard Russian or Polish. The bilingual children were recruited at Saturday/Sunday schools in major cities in the Netherlands. These children also had to be aged between 3;6 and 6;6, typically developing and simultaneous or early sequential bilinguals (with exposure to Dutch being longer than 24 months).\(^5\)

In order to be able to compare the performance of the Russian and Polish monolinguals on RCs specifically, it was important that they were comparable in their general language proficiency. Using the results of the gender and case (genitive and accusative) production tasks (Janssen, 2016), both groups performed at ceiling. However, the Polish children were significantly better than their Russian peers in the gender production task \((U = 499, p = .001)\) and the genitive production task \((U = 475, p = .002)\) as a whole. When considering the genders (masculine, feminine, and neuter) separately, the Polish monolinguals also outperformed the Russian monolinguals on the feminine and the neuter in the accusative production task (feminine: \(U = 607.5, p = .008\); neuter: \(U = 440, p = .001\)) (Janssen, 2016).

\(^4\) In addition, two monolingual adult control groups were tested (with 10 participants in each group), see Janssen (2016).

\(^5\) The relative amount of input, the age of onset and the length of exposure to both languages were determined on the basis of a language background questionnaire that was an adaptation of Questionnaire for Parents of Bilingual Children, LITMUS PaBiQ (Tuller, 2015). For this research, the questionnaire was translated into Russian and Polish. From the Questionnaire, all bilingual participants were exposed to Russian or Polish from birth onwards. Their exposure to Dutch had started at the latest at 3;1 years old. The parents reported to address their children only in their mother tongue. About 25% of the children in both groups preferred to speak Russian or Polish at home, and 38.5% of the Russian bilinguals and 50% of the Polish bilinguals preferred to speak Dutch. Despite the fact that both groups attended Saturday/Sunday schools, the children were not yet literate due to their young age (this was an important selection criterion). Moreover, there was a difference in the amount of formal instruction: the Polish bilinguals attended the Polish schools for three to four hours only two days per month, while the Russian bilinguals attended classes for three hours every weekend (Janssen, 2016).
The general language proficiency of the Russian-Dutch and Polish-Dutch bilinguals was compared in Dutch by using a Dutch sentence repetition task (Van de Scheur, 2012). No significant difference was found between the groups. With respect to the language proficiency in Russian and Polish, the bilingual groups were not homogeneous. However, the distribution of the bilinguals over dominancy groups was highly comparable: each group contained 19 balanced bilinguals; in the Russian-Dutch group 6 children were dominant in Russian and 12 in Dutch, and in the Polish-Dutch group 8 children were dominant in Polish and 11 in Dutch (Janssen, 2016: 137–138). The results of the gender and case (genitive and accusative) production tasks were used to further determine the proficiency level of the bilinguals in their Russian and Polish. The Polish-Dutch bilinguals significantly outperformed their Russian peers on the gender production task ($U = 485.5$, $p = .009$) and the accusative production task ($U = 432$, $p = .010$) (Janssen, 2016).

To summarize, there was no reason to suspect that the groups of children were significantly different in terms of their general language ability. They were further comparable in age and development.

### 7.2 Tasks

Items taken from a Sentence Repetition Task (SRT) were used to examine ability on RCs. The choice of the SRT was determined by a long tradition of elicited imitation tasks that are known for their high reliability as an indicator of language proficiency (see a discussion of advantages and weaknesses of elicited imitation in Lust et al., 1996; Flynn & Foley, 2009: 34–37; Marinis & Armon-Lotem, 2015). An SRT involves analyzing the sentences at different linguistic levels of representation and then using the production system to re-produce these sentences. If the sentences are long enough to make “parroting” impossible, participants must rely on their knowledge of lexicon and grammar to be able to regenerate the sentences. In general, children’s scores on SRTs correlate strongly with their scores on normed language proficiency tasks (see e.g., Meir, Armon-Lotem, & Walter, 2015). In our study, if children have only problems with interpretation of RCs and not with grammatical marking, it would be mostly visible in their structure conversions, simplification of RCs and avoidance strategy, but not in overt case errors. On the other hand, if children make overt grammatical errors in (re)production of RCs, it could point to the fact that either the basic grammar has not yet been internalized, or the target structure has not been fully acquired, or both.

The Russian SRT was based on the subset of the SRep-LITMUS (Marinis & Armon-Lotem, 2015), developed to screen Russian-Hebrew six-to-seven-year-olds (Meir & Armon-Lotem, 2015). Since most of the children tested in our study were younger than 6–7 years, the SRT was shortened, and structures normally
acquired after the age of six years were excluded. The SRT contained diverse structures. Sentences longer than 14 syllables were adjusted to be more suitable for younger subjects (see Marinis & Armon-Lotem, 2015, for a discussion of sentence length). This version was piloted with five-year-old Russian-Israeli children in Israel. It was then translated into Polish in order to make the tasks as parallel as possible (Janssen, 2016).

From these SRTs, only the subsets of items containing SRs and ORs will be further discussed. For each language, eight test sentences were constructed (with the sentence length between 11 and 14 syllables): four for SRs and four for ORs. In all eight sentences the RC followed the main clause and was introduced by a relativizer kotoryj/który ‘which, that, who’. Given the relatively young age of the subjects, the main clause in all these items was a presentational construction with a copular (with a covert copular in Russian) (see sentences 14–17). As discussed in the previous section, such sentences belong to the earliest RCs acquired by children (Diessel & Tomasello, 2005: 884). The OR contained a direct object marked for the accusative case with a salient overt ending in order to avoid any ambiguity in grammatical markers. All stimuli were comparable in terms of number (singular), gender (masculine and feminine), animacy (both the subject and the object were animate), and the number of syllables. Thus, the only overt cue was (gender and) case marking. Because the relative pronoun is at the immediate onset of the RC, it should not increase the working memory load and is expected to facilitate the correct interpretation. The word order and the order of syntactic roles in the RC was canonical: in the SRs the direct object was in the final position of the subordinate clause (SVO) (sentences 14–15); in the ORs the subject took the final position (OVS) (sentences 16–17).

Subject relatives

(14) a. Rus.: Ėto žiraf, kotoryj ljubil verbljuda.
this giraff-em.sg.nom that-em.sg.nom loved camel-em.sg.acc.an
‘This is the giraffe that loved the camel.’

b. Rus.: Ėto krysa, kotoraja tolknula myšku.
this rat-em.sg.nom that-em.sg.nom pushed mouse-em.sg.acc
‘This is the rat that pushed the mouse.’

(15) a. Pol.: To jest baranek, który kochał wielbłąda.
this is lamb-em.sg.nom that-em.sg.nom loved camel-em.sg.acc.an
‘This is the lamb that loved the camel.’

6. As well as SRs and ORs, the following structures were also included in the SRT: SVO with an obligatory preposition, SVO with a free preposition, SVDOIO, SOV word order, OVS word order, Conditionals, Bi-clausal sentences with coordination, Bi-clausal sentences with subordination, WH-questions, and prepositional WH-questions (Janssen, 2016: 81).
b. Pol.: To jest norka, która popchnęła myszkę.
   this is mink-f.sg.nom that-f.sg.nom pushed mouse-f.sg.acc
   ‘This is the mink that pushed the little mouse.’

Object relatives
(16) a. Rus.: Ėto slon, kotorogo ukusil krokodil.
   this elephant-m.sg.nom that-m.sg.acc.an bit crocodile-m.sg.nom
   ‘This is the elephant that the crocodile bit.’
b. Rus.: Ėto devočka, kotoruju kormila mama.
   this girl-f.sg.nom whom-f.sg.acc fed mother-f.sg.nom
   ‘This is the girl whom mother fed.’

(17) a. Pol.: To jest słoń, którego ugryzł krokodyl.
   this is elephant-m.sg.nom that-m.sg.acc.an bit krokodyl.
   crocodile-m.sg.nom
   ‘That is the elephant that the crocodile bit.’
b. Pol.: To jest dziewczynka, którą karmiła mama.
   this is girl-f.sg.nom whom-f.sg.acc fed mother-f.sg.nom
   ‘That is the girl whom the mother fed.’

The eight RCs were randomly distributed amongst the other items of the Russian and Polish SRTs.

In order to establish whether or not the RC subsets in both languages were reliable, reliability analyses were performed on the scale of each relative clause sentence repetition task: the Cronbach’s alpha is reported for Polish and Russian separately. The relative clause sentence repetition tasks for both Russian and Polish were checked with an item-analysis and turned out to have a good internal reliability (for Russian: \( \alpha = .87 \); for Polish: \( \alpha = .84 \)). Deletion of each of the items would reduce the reliability of the test. No outliers in items were observed; all items were therefore included in the analysis.

7.3 Procedure

The whole SRT, including both types of RCs, was administered by an adult female experimenter. All items (total: 48) were pre-recorded and were offered in a PowerPoint presentation with an audio track only on an 11-inch notebook with built-in speakers. All responses were recorded with a voice recorder. Before the test started each child was familiarized with the task by means of an instruction and four practice trials that provided an opportunity for the experimenter to give
explicit feedback to the child, if necessary. After every block of eight sentences the child saw a smiley on the screen and received a sticker in order to encourage his/her participation. Each of the SRTs took about 10–15 minutes to administer, depending on the language level of the child and the child’s age.

7.4 Coding and scoring

The recordings were transcribed by the experimenter. Children’s responses were assigned a score of 1 for a correct target response. Lexical substitutions or alternations (of verbs and nouns) in the target structure that did not violate case and/or gender agreement were considered correct. Substitutions of the target relative pronoun with its alternatives (kakoj for Russian, and co for Polish), when correctly applied, were also assigned a score of 1. All other responses were assigned a score of 0. For the error analysis an additional classification of errors was made. The following error types were distinguished (see Table 1):

Table 1. Overview of error types in Russian and Polish

<table>
<thead>
<tr>
<th>Error types in Russian</th>
<th>Error types in Polish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender/case violation</td>
<td>Gender/case violation</td>
</tr>
<tr>
<td>Kakoj substitutions (incorrect)</td>
<td>Co substitutions (no resumptive pronouns)</td>
</tr>
<tr>
<td>Other substitutions of target pronoun</td>
<td>Other substitutions of target pronoun</td>
</tr>
<tr>
<td>Koto(ra) reduction</td>
<td>Omissions</td>
</tr>
<tr>
<td>Omissions</td>
<td>Simple clause paraphrase</td>
</tr>
<tr>
<td>Simple clause paraphrase</td>
<td>OR-to-SR / SR-to-OR conversion</td>
</tr>
<tr>
<td>OR-to-SR / SR-to-OR conversion</td>
<td></td>
</tr>
</tbody>
</table>

8. Results

All monolingual and bilingual children completed the SRT. From the monolingual Russian RCs subset of 328 items, there was no response in ten cases; from the monolingual Polish RCs subset of 320 items there were 12 such cases. From the bilingual Russian-Dutch subset of 304 items, two sentences were not responded to, and from the bilingual Polish-Dutch subset of 304 items there was no response in 18 cases. The results in terms of accuracy for the children are presented in Table 2.7

We explored main and interaction effects in a repeated measures ANOVA, with RC condition as a within-subjects factor, and with language (Russian or Polish) and bilingual mode (monolingual or bilingual) as a between-subjects

---

7. The adult control groups in both languages performed at ceiling.
We found a main effect of RC condition with a Greenhouse-Geisser correction ($F(1, 152) = 137.225$, $p < .001$, $\eta^2 = .474$) with SRs being (re)produced significantly better than ORs. Furthermore, there was a significant interaction effect for RC condition and bilingual mode with a Greenhouse-Geisser correction ($F(1, 152) = 3.888$, $p = .050$, $\eta^2 = .025$), with the monolinguals performing significantly better than the bilinguals, as well as a significant interaction between RC condition, bilingual mode, and language with a Greenhouse-Geisser correction ($F(1, 152) = 4.073$, $p = .045$, $\eta^2 = .026$).

When exploring the effect of language in the monolinguals and bilinguals separately, we found no significant difference between the two monolingual groups neither in all sentences taken together nor in SRs and ORs considered separately after a Bonferroni correction. This means that the significant interaction effect of RC condition, bilingual mode, and language was due to differences in the bilinguals. The Polish-Dutch children outperformed the Russian-Dutch children on the task as a whole ($F(1, 74) = 8.701$, $p = .004$) and on the SR condition ($F(1, 74) = 9.619$, $p = .003$), but not on the OR condition ($F(1, 73) = 2.841$, $p = .100$). So, as expected, the SR items were easier to perform than the OR items, and the monolinguals were better in the reproduction of RCs than the bilinguals. The effect of Polish vs. Russian was only found in SRs for the bilinguals.

### 8.1 Error analysis

Both the Russian and Polish monolingual and bilingual children made a range of errors. A complete overview of all error types is provided in Tables 3 and 4, but we will only discuss here the most important and frequent types of the errors.
that children made when they attempted the target structures. First, the errors for Russian will be analysed, followed by an error analysis for the Polish groups.

**Russian**

From the total of 160 responses to the SRs, the Russian monolinguals made 22 mistakes. As shown in Table 3, in their responses two types of errors occurred: incorrect case and/or gender co-indexation between the constituents within the subordinate clause and lexical omissions.

**Table 3. Overview of typical errors made by Russian and Russian-Dutch children**

<table>
<thead>
<tr>
<th>Error types</th>
<th>MoRu SRs</th>
<th>BiRu SRs</th>
<th>MoRu ORs</th>
<th>BiRu ORs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender/case violations</td>
<td>7</td>
<td>17</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>Kakoj substitution (total)</td>
<td>7</td>
<td>20</td>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>– correct</td>
<td>4</td>
<td>13</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>– incorrect</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>Other lexical substitutions (total)</td>
<td>1</td>
<td>6</td>
<td>–</td>
<td>11</td>
</tr>
<tr>
<td>– kto/čto</td>
<td>–</td>
<td>3</td>
<td>–</td>
<td>4</td>
</tr>
<tr>
<td>– takoj</td>
<td>–</td>
<td>3</td>
<td>–</td>
<td>5</td>
</tr>
<tr>
<td>Koto(ra) reduction</td>
<td>–</td>
<td>10</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Omissions (total)</td>
<td>7</td>
<td>36</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Simple clause paraphrases (total)</td>
<td>1</td>
<td>14</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>– correct</td>
<td>1</td>
<td>8</td>
<td>–</td>
<td>4</td>
</tr>
<tr>
<td>– incorrect</td>
<td>–</td>
<td>6</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>SR-to-OR conversions (total)</td>
<td>2</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>OR-to-SR conversions (total)</td>
<td></td>
<td></td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>– correct</td>
<td></td>
<td></td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>– incorrect</td>
<td></td>
<td></td>
<td>26</td>
<td>26</td>
</tr>
</tbody>
</table>

The Russian monolinguals obviously had much more difficulty with the analysis and (re)production of ORs than with SRs. From the total of 158 OR responses, 70 sentences were not target-like. As seen from Table 3, there was a strong tendency to convert an OR to a SR, whereby about 70% was grammatically incorrect: both the relative pronoun and the noun constituent in the subordinate clause was produced in the nominative case with verb agreement being preserved. Changes in the opposite direction occurred only twice. Apart from the structure conversion, the Russian monolinguals made relatively many errors in case and gender co-indexation between the relative kotoryj/kakoj, the head noun in the main clause, the head referent in the subordinate clause, and the verb. Especially the masculine genitive kotorogo, pronounced as [kətɔrɔvɔ], was misinterpreted as either feminine
nominative *kotoraja* or neuter nominative *kotoroe* which both sound similar as [katórajə]. This suggests that problems with the analysis of the syntactic functions in the sentence together with low transparency of the genitive masculine ending caused additional processing pressure resulted in surface errors in (re)production. The reduction of *kotoryj* to *kotor(a)*, which occurred eight times, pointed in the same direction: the children avoided any case marking on the relativizer because they had difficulty with dissociation of phonologically comparable forms ([katórajə] vs. [katórajə]) as well as with interpretation of the whole structure. This indicates that overt case marking, if insufficiently transparent, is a complicating factor in (re)production of ORs when the child tries to disambiguate two structures.

The Russian-Dutch children produced only 60 correct SRs from the total of 150 test items: 47 with *kotoryj* and 13 with the alternative *kakoj*, the usage of which was almost three times as higher as in the monolingual data. Interestingly, the items with the relative pronoun in the nominative masculine (*kotoryj*) were produced correctly more than 50% of the time, while the items with the relative pronoun in the nominative feminine (*kotoraja*) were performed only with 28% of accuracy. Among the most frequent error types were lexical omissions, violations in gender and case co-indexation, and reduction of *kotoryj*. In 14 sentences, the omission of the relative pronoun converted the target item to a SVO sentence (with correct and incorrect case and gender marking). The omission of lexical information indicates that the SRT as such was quite demanding for the Russian bilinguals who were often not capable to recall the target structure.

The scores on the ORs were even lower: from 151 ORs only 15 target structures were correctly produced (nine with *kotoryj* and six with *kakoj*). As in the monolingual subset, the most frequent error type was an OR-to-SR conversion. Similar results in both groups pointed to the fact that language proficiency did not influence the ability to correctly analyse the OR target structure. The difference in language proficiency, however, was visible in other types of errors, such as different lexical omissions and simple clause paraphrases, and to a less extent in case and gender violations, and *kotoryj* reduction (see Table 3). In (18–22) some examples of the most frequent error types, extracted from the Russian bilingual data, are presented.

**OR-to-SR conversion**

(18) Ėto petux, **kotorogo pojmala kurica.**
    this rooster-m.sg.nom that-m.sg.acc.an caught-f chicken
    ‘This is the rooster that the chicken caught.’

> Ėto petux, **kotoryj pojmal kuricu.**
    this rooster-m.sg.nom that-m.sg.nom caught-m chicken-f.sg.acc
    ‘This is the rooster that caught the chicken.’
Gender and case violation

(19) Ėto žiraf, kotoryj ljubil verbljudu.
this giraffe that-M.SG.NOM loved-M camel-M.SG.ACC.AN
‘This is the giraffe that loved the camel.’

> Ėto žiraf, kotoryja ljubil verbljudu.
this giraffe-M.SG.NOM that-F.SG.NOM loved camel-M.SG.ACC.AN

Omission of a sentence part

(20) Ėto devočka, kotoruju kormila mama.
this girl that-F.SG.ACC fed mother-F.SG.NOM
‘This is the girl who(m) the mother fed.’

> Devočki i budet mama.
girls and will.be mother

Simple clause paraphrase

(21) Ėto žiraf, kotoryj ljubil verbljudu.
this giraffe that-M.SG.NOM loved camel-M.SG.ACC.AN
‘This is the giraffe that loved the camel.’

> Ėto žiraf ljubil verbljudu.
this giraffe loved camel-M.SG.ACC.AN
‘(This was) the giraffe (that) loved the camel.’

Reduced koto(ra)

(22) Ėto medved’, kotorogo obmanula ptica.
this bear that-M.SG.ACC.AN tricked bird-F.SG.NOM
‘This is the bear that the bird tricked.’

> Ėto medved’, kotora obmanula ptica.
this bear _____ tricked bird-F.SG.NOM

Polish
As shown in Table 4, the total number of mistakes in the Polish data is lower than in Russian.

From 157 responses to the SRs, the Polish children made only 10 mistakes (from which five omissions and only one error on case and gender marking). As expected, the performance on the ORs was significantly worse, but the Polish children achieved 67% of accuracy, which was better than the Russian children managed (56%). One type of the errors was particularly frequent: an OR-to-SR conversion whereby both the relativizer *który* and the noun constituent in the subordinate clause took the nominative case. The same tendency was also observed in the Russian monolinguals.

The Polish bilingual children from the total of 144 SRs responses produced 105 items correctly. This was substantially higher than by the Russian-Dutch
bilinguals (68% vs. 40%). However, this relatively high number of correct responses is partly due to co-substitutions when no case and gender adjustments were needed. Although, as in the Russian-Dutch bilingual data, the most frequent error type here were lexical omissions, the total number of omissions in the Polish task was higher than in the Russian SRT (24 vs. 37). To recall and to reproduce the semantic information seemed to be demanding for the Polish bilinguals even more than for the Russian bilingual children.

As expected, the ORs caused much more problems. Only 22% of all items was correctly repeated, although this was still twice as much as in the Russian data. One third of all errors accounted for the OR-to-SR conversion, followed by different types of omissions and co-substitutions without resumptive pronouns in the subordinate clause (see Table 4). As discussed earlier, in ORs the unmodified co must be used in combination with a resumptive pronoun that, being part of the argument structure of the verb, is marked for case, number, and gender. When co is not accompanied by a resumptive pronoun, the sentence should be interpreted as a SR. However, the syntactic functions of the nouns and case marking must then be adjusted: the subject of the OR must be transformed into the object of the SR by overt case marking. This conversion has not been found in the bilingual data that points to the fact that resumptive pronouns (and their declination) as part of the object relative structure were yet not acquired by the Polish bilinguals. It is, however, difficult to say whether the Polish monolinguals in our study had command of this structure: two cases of co without resumptives are insufficient to draw any conclusion. In contrast to the Russian data, which exhibited more variation in errors, other error types (such as simple clause paraphrases or case and gender violations) were very infrequent in Polish.

<table>
<thead>
<tr>
<th>Error types</th>
<th>MoPo SRCs</th>
<th>BiPo SRCs</th>
<th>MoPo ORCs</th>
<th>BiPo ORCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender/case violations</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Co substitution (total)</td>
<td>2</td>
<td>18</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>– correct</td>
<td>2</td>
<td>18</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>– incorrect (no resumptives)</td>
<td>–</td>
<td>–</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Other lexical substitutions</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>(total)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omissions (total)</td>
<td>5</td>
<td>38</td>
<td>6</td>
<td>37</td>
</tr>
<tr>
<td>Simple clause paraphrases</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3</td>
</tr>
<tr>
<td>(incorrect)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR-to-OR conversions</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>OR-to-SR conversions (total)</td>
<td></td>
<td></td>
<td>30</td>
<td>49</td>
</tr>
<tr>
<td>– correct</td>
<td></td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>– incorrect</td>
<td></td>
<td></td>
<td>29</td>
<td>45</td>
</tr>
</tbody>
</table>
In (23–26) some examples of the most frequent error types, extracted from the Polish bilingual data, are presented.

**OR-to-SR conversion with co (without resumptives)**

(23) To jest słoń, którego ugryzł krokodyl.
this is elephant that-M.SG.ACC.AN bit crocodile-M.SG.NOM
‘This is the elephant that the crocodile bit.’

> To jest słoń, co ugryzł krokodyl.
this is elephant-M.SG.NOM that bit crocodile-M.SG.NOM

**Omission of a sentence part**

(24) To jest słoń, którego ugryzł krokodyl.
this is elephant that-M.SG.ACC.AN bit crocodile-M.SG.NOM
‘This is the elephant that the crocodile bit.’

> Ugryzł krokodyl.
bit crocodile-M.SG.NOM

**Gender and case violation**

(25) To jest dziewczynka, którą karmiła mama.
this is girl that-F.SG.ACC fed mother-F.SG.NOM
‘This is the girl who(m) the mother fed.’

> To jest dziewczynka, który karmiła mama.
this is girl that-M.SG.NOM fed mother-F.SG.NOM

**Simple clause paraphrase**

(26) To jest żyrafa, która widziała zebrę.
this is giraffe that-F.SG.NOM saw zebra-F.SG.ACC
‘This is the giraffe that saw the zebra.’

> Żyrafa widziała zebrę.
giraffe-F.SG.NOM saw zebra-F.SG.ACC
‘The giraffe saw the zebra.’

9. **Discussion**

The main goal of this study was to investigate the role of overt and transparent case (and gender) marking in the acquisition of RCs in Russian and Polish. We hypothesized that case marking could be a facilitating factor in comprehension due to disambiguation of the target structure at the immediate onset of the RC. At the same time, overt case marking could also be seen as a complicating factor in production resulting in surface errors, when the gender and case system has not been yet sufficiently internalised. We also assumed that because of less transparency in adjectival case endings in Russian, Russian children would lag behind their Polish
peers in acquisition of RCs. The results of this study supported our assumption: the Polish children (both the monolinguals and bilinguals) outperformed their Russian peers on the task as a whole, and on the SRs. However, the differences between the monolingual groups were statistically not significant. This can be explained by the fact that the basics of the case and gender system in both languages have been already established in monolingual children. As already discussed, we know from the literature that monolingual children acquire the case and gender marking system in both Russian and Polish at age 3;6–4;0 (Gvozdev, 2007; Eliseeva, 2015; Smoczyńska, 1985). The mean age of our participants was around five years. Moreover, as followed from the gender and case production tasks used as a background measure for this study, language proficiency in the monolingual groups was comparable: only the genitive production task showed a significant difference with the Polish children outperforming the Russian peers, but the genitive was not tested in the RC task (Janssen, 2016). The gender production task and the accusative production task as used in Janssen (2016) showed no significant advantage for the Polish children. However, when considering the genders separately, the Polish children outperformed the Russian peers on the feminine and the neuter. These are the genders where transparency is apparently of great importance, but the neuter gender was also not tested in the RC task (Janssen, 2016).

The difference between the results of two language groups reached statistical significance in the bilingual data. We interpret this finding as suggesting that a more regular and transparent form-to-meaning mapping in the gender and case system in Polish contributes to more successful acquisition of RCs, since other differences between the tasks and groups were levelled out or controlled for: the SRTs were identical for both languages, the mean age of the bilinguals, the distribution of the bilinguals over dominancy groups and other background factors were highly comparable. The Polish bilinguals showed substantially fewer case and gender violations and no reduction of the relative pronoun in both types of the RCs. The significant difference was, however, only seen in SRs, not in ORs. All children in our study had difficulty with (re)production of ORs. Thus, our prediction of the role of transparency of case marking is only in part borne out. This is due to several factors.

Problems with ORs are in line with what we know from other research (Diessel & Tomasello, 2005; Kidd et al., 2007). As we discussed in the section on previous research, the difference in accuracy between SRs and ORs has been attributed by some authors to the Filler-Gap hypothesis: the varying distance between filler and gap increases the processing load because the language learner has to keep unintegrated information in working memory before reaching the verb (Gibson, 1998). However, in Russian or Polish RCs the gap must be identified on the bases of overt case marking on the relative pronoun that determines the syntactic role of the relativized constituent at the beginning of the RC. This should facilitate the initial
correct analysis of these structures and result in comparable performance on both types of RCs. However, our results did not support this. Our results fit better with the explanation proposed by Diessel and Tomasello (2005) who argued that SRs with a copular main clause belong to the earliest acquired syntactic patterns due to the fact that they denote the actor prior to any other thematic role. Despite the fact that Russian or Polish are free word order languages, young children still have a clear preference for the actor taking the sentence-initial position in both comprehension and reproduction, regardless of syntactic features of the sentence and discourse style (as shown in Kallestinova, 2007; Udes, 2013). Our data indeed exhibit a clear misinterpretation of ORs as SRs in approximately one third of all ORs. Moreover, this conversion often meant that the utterances the children produced in the SRT were ungrammatical. Even the monolingual data are characterized by examples of case and gender violations, and avoidance of case marking on kotoryj (for Russian). Apparently, automatic activation of the SVO structure was in competition with processing the correct cue that resulted in avoidance of any case marking. Thus, the advantage in language proficiency in the monolingual children was overridden by the structural complexity.

Nevertheless, we cannot agree with the claim that overt case marking does not facilitate processing and comprehension of ORs at all, as proposed by Bențea (2012) and Sevcenco et al. (2012b) on the basis of data from Romanian. In our study, the degree of transparency does determine the success of acquisition. Although the Russian children in our study performed slightly above chance and were therefore worse than their Polish peers, they were still twice as good as the Romanian children from the study of Bențea (2012). This very poor performance of Romanian children can be explained by language-specific features of case marking in Romanian (i.e., omission of accusative markers in colloquial speech, and low salience and distant placement of resumptive pronouns in ORs).

The re-analysis strategy, mentioned in the study of Diessel and Tomasello (2005), when children corrected themselves and changed the relative pronoun to the case required in the context, was not observed in our study. We found a different type of the re-analysis: in about 10% of all responses on the SRs and in 15% of all responses on the ORs, the Russian bilinguals reanalysed the whole sentence as a simple transitive sentence. This is due to certain structural similarity between SRs and ORs and simple sentences, which children learn before they start producing RCs (Diessel & Tomasello, 2005).

As expected, the monolingual children were significantly better in their command of both types of the RCs than the bilingual children. It is feasible that influence from the other language played an important role in bilingual acquisition of these structures. As mentioned in the section on previous research, Dutch has no overt case marking. This may impede the learning of the case system in Russian
and Polish (Polinsky, 2007; Peeters-Podgaevskaja, 2008). The Dutch relativizer also only marks one gender distinction in the singular (common gender vs. neuter) and is ambiguous in other grammatical contexts. This may lead bilingual children to be less sensitive to the subject-object asymmetry. The Russian bilinguals could not use the Dutch pattern in their Russian, since Russian RCs require a clear gender and case distinction. In Polish, however, with the indeclinable co-relativizer the bilinguals could sometimes transfer the Dutch pattern and thus achieved a higher accuracy in their results on SRs, but not on ORs where the use of resumptive pronouns is required.

Although differing in the total amount of errors, their frequency, and preference patterns, the monolingual and bilingual children did not differ in types of errors (as shown in Tables 3 and 4). This points to commonalities in the acquisition of RCs in both groups, and between the languages involved. Due to the larger number of errors in the bilingual children it was possible to identify and describe these errors. Lexical omissions, which occurred quite infrequently in the monolingual data, seemed to be one of the most frequent error types in the bilingual data. Apparently, the SRT was quite demanding for the Russian-Dutch and Polish-Dutch children, who obviously had difficulty with retaining lexical information in their working memory and reproducing this information immediately afterwards. The type of the RC did not make a difference: in both types of RCs and in both languages a comparable amount of lexical omissions was found. Insufficient vocabulary knowledge must be excluded as an explanation, since the SRT involved only highly frequent words, and the majority of the children were proficient being either balanced bilinguals or even dominant in Russian and Polish.

Simple clause paraphrases were mostly attested in the Russian bilingual data. This indicates that the monolinguals and the Polish-Dutch bilinguals could identify the RC structure quite well, while the Russian-Dutch children apparently had difficulty with RC identification in general. In this respect, the idea of a universal developmental path, observed in other languages, starting with simple transitive sentences before arriving at RCs, is supported by our data. The question is, however, why the Polish-Dutch children did not transform RCs into simple transitive sentences. We are inclined to think that the Polish bilinguals had better knowledge of the structure itself due to the indeclinable co-pronoun, which diminishes the processing and generating load and thus makes the whole structure more attractive to use, facilitating in this way the acquisition of RCs in general.

The use of kakoj in Russian and co in Polish as an alternative relative pronoun was also very suggestive. Although it occurred several times in the monolingual data, it was substantially more frequent in the bilingual data, which is in line with previous observations (Peeters-Podgaevskaja, 2008; Mróz & Flynn, 2010). This means that some bilingual children have already acquired RCs (at least SRs) as a
grammatical structure, but failed to acquire the particular relativizer kotoryj/który. That is why they substituted the common but still less frequent relative pronoun kotoryj/który with the more frequent (also used in interrogative sentences) and thus better known pronoun kakoj/co. The use of co by the Polish bilinguals is in line with findings of Smoczyńska (1985) who observed the dominancy of co-relatives in Polish monolinguals up to the age of four years. In this respect, our Polish bilinguals are in line with monolingual development, but demonstrate a previous stage in acquisition of relative pronouns. In contrast, the Russian bilinguals showed a different preference pattern in acquisition of relative pronouns. The fact that there were numerous examples of the use of the kakoj-pronoun is different from what we know from monolingual child data where this relativizer occurred only occasionally.

However, the use of the relative pronoun co by the Polish bilinguals was overgeneralised. When correct in SRs, the isolated co was not correct in ORs where resumptive pronouns were required. The complete absence of resumptive pronouns that are marked for gender, number, and case indicates that ORs with this particular relativizer have not yet been acquired. Apparently, the children did not perceive these two words as belonging to one structure, probably because of their low salience and regularity: the resumptive pronouns are clitics, they are placed at the relativization gap and thus can be separated from the relative marker co, they vary in form due to different case marking and are frequently omitted in RCs with accusative inanimate objects. The problems with resumptive pronouns in ORs are in line with the data in other languages with the same structure (Hebrew: Arnon, 2011; Romanian: Bențea, 2012; Sevcenco et al., 2012b). It is, however, difficult to say whether or not the Polish monolinguals have acquired this structure as they mostly avoided it in their output.

Finally, our results suggest that the SRT is useful as a general measure indicator of language proficiency in comprehension as well as in production. For monolinguals, we can state that the SRT is not mere imitation, but taps into comprehension ability of the children. OR-to-SR conversions with correct grammatical structures demonstrate the children’s misanalysis of the target structure rather than erroneous knowledge of grammar (the subjects and objects must be overtly marked for case and gender anyway). For bilinguals, it is more difficult to determine whether the errors are due to the lack of grammatical knowledge or misinterpretations of the structures, or both. On the one hand, case and gender violations in SRs and erroneous simple clause paraphrases show that the nominal morphology has not yet been automatized. Omissions of resumptive pronouns in the OR structure in Polish point in the same direction: the morphosyntax has not been fully acquired. On the other hand, a comparable amount of infelicitous OR-to-SR conversions in the monolinguals and bilinguals emphasises difficulty with comprehension of the target structure.
10. Conclusion

To summarize, the present study has shown that case marking, when regular and transparent, facilitates the acquisition of RCs (Polish). Less transparency in case endings slows down the acquisition of case (and gender) morphology and as such negatively affects the acquisition of the tested structures (Russian). This explains the significantly better performance of the Polish children (monolinguals and bilinguals taken together) compared to the Russian children (monolinguals and bilinguals taken together) in whose data much more gender and case violations and avoidance of any case marking on the relative pronoun were observed.

The subject-object conversions and structure simplifications found in the (re-)production of RCs in Russian and Polish support the idea for a general developmental path in acquisition of these structures. The children start with SRs, and at age of five years they are quite successful in correctly producing them, because the SRs with a copular main clause are highly frequent in ambient language and child speech and also because they resemble simple transitive sentences that have already been acquired (although some Russian bilinguals showed that they were not through the stage of simple transitive sentences). Lexical alternatives of the relative pronoun that are easier to acquire such as in Polish, increase accuracy and provide a better access to the acquisition of the whole structure.

The ORs in our study, despite the same sentence structure as the SRs, caused difficulty. This is in line with the general developmental order of acquisition of relatives in languages with postnominal RCs. This is the result of frequency effects, on the one hand, and cognitive and grammatical maturation of children, on the other. However, the more transparent and regular case marking is, the more it facilitates comprehension as well as production (due to a better form-to-meaning mapping that leads to faster acquisition of nominal morphology). This explains why all Polish children taken together were significantly better than their Russian peers.

References


Peeters-Podgaevskaja, A. V. (2008). Problemy osvoenija russkogo jazyka kak vtorogo rodnogo det’mi 5–7 let i sozdanie adekvatnogo učebnogo posobija [Problems of the acquisition of Russian as L2 in children aged 5 to 7, and compilation of an appropriate language course].
In E. de Haard, W. Honselaar, & J. Stelleman (Eds.), Literature and beyond: Festschrift for Willem G. Weststeijn on the occasion of his 65th birthday (pp. 609–627). Amsterdam: Pegasus.


Address for correspondence

Alla V. Peeters-Podgaevskaja
Department of Russian and Slavic Studies
University of Amsterdam
Spuistraat 134
1012 VB Amsterdam
The Netherlands

A.V.Peeters-Podgaevskaja@uva.nl

Publication history

Date received: 18 March 2017
Date accepted: 3 June 2018
Published online: 30 July 2018