Relative clauses in the L1-acquisition of Turkish

Jaklin Kornfilt* – Kutlay Yağmur** – Gabriella Hermon*** – Özge Öztürk**** – Türkay Yalnız

1. Introduction

This study examines the acquisition patterns of head-final relative clauses (RCs) in Turkish in monolingual1 and bilingual children. It has been claimed that both subject- and object-RCs are acquired by children earlier in Indo-European languages, compared to children acquiring Turkish RCs, because of the lack of wh-pronouns and overt complementizers, the differences between nominalized embedded clauses and finite main clauses, and the supposed complexity of nominalization and agreement morphology. It is claimed in Slobin (1986) that these properties make Turkish RCs difficult to acquire. By using experimental elicitation with monolingual Turkish children and longitudinal naturalistic acquisition data from a Turkish child in the Netherlands, we checked the validity of earlier claims regarding the acquisition of relative clauses. Our experiments as well as the longitudinal corpus show that young children (of comparable age as their counterparts acquiring Indo-European languages) do use RC-morphology and syntax correctly for the most part. While they may use RC constructions less frequently than those counterparts, their competence in this area is comparable. Our results show that Slobin (1986 and related work) and Çağrı (2006) were wrong in claiming that children under 5 acquiring Turkish lack the morpho-syntax for head-final RCs without complementizers and relative pronouns in Turkish.

2. Acquisition of relative clauses in Turkish

Turkish RCs have been described in detail in Underhill (1972), Hankamer and Knecht (1976) and Zimmer (1987), among others. Recent work in the Principles and Parameters
and Minimalism frameworks has been conducted by Kornfilt in a number of papers (e.g. Kornfilt 2000, 2005, 2008, and 2009). While there may be some evidence in Turkish that RCs involve head raising rather than an empty OP movement (as discussed in detail in Kornfilt 2005), researchers agree that RCs involve movement which is island sensitive. Hence, one could talk about a chain being established between the head of the RC and an empty position (the relativized NP position):

(1) Subject-gap RC

\[ \text{\_ \_ kedi-yi kovala-yan} \] k\text{\textit{iz} NOM} mor ol -du

\[ \text{cat-ACC chase-(y)An girl.NOM purple become-PST(-3SG)} \]

'The girl who is chasing the cat turned purple.'

(2) Object-gap RC

\[ \text{kiz-in \_ kovala-di\-g-\_} \] kedi sar\text{\textit{i} NOM} ol -du

girl-GEN chase-DIK-3.SG cat.NOM yellow become-PST(-3SG)

'The cat that the girl is chasing turned yellow.'

(3) PP-gap RC (gap as complement of a "secondary" postposition [PP])

\[ \text{kiz-in \_ u\-st\-\_\-n -den zipla-di\-g -i} \] ç\text{\textit{ocuk} NOM} mor ol -du

girl-GEN top-3.SG -ABL jump-DIK-3.SG child.NOM purple become-PST(-3SG)

'The child over which the girl is jumping turned purple.'

Arguments are less clear for the existence of an empty category-head chain in the third type of RC (which we label Postpositional Phrase gapped RC [PP-G RC]). These RCs have been described and/or analyzed by Kornfilt (e.g., 1977, 1984, and 1991) as possibly containing a null pronoun linked via binding to the head. Turkish exhibits some clear-cut and important differences between subject and non-subject RCs which may be relevant for processing and acquisition. In RCs the verb is nominalized; but while SG-RCs use the nominalizer -y(An), other RCs (both OG- and PP-G as in 2-3 above) use the -DIK nominalizer which is obligatorily followed by a possessive agreement marker (agreeing with the subject of the RC). No agreement is possible in SG-RCs, a fact which is, according to Kornfilt (2000, 2008, and 2009), derived from binding theory. Moreover, subjects of OG-

---

2 The 3rd person singular agreement marking for fully tensed predicates is null in Turkish.
3 Traditionally, postpositions in Turkish which bear overt person and number agreement with their complement are referred to as "secondary postpositions", while regular, i.e., "primary postpositions" bear no agreement marker. (See, for example, Lewis 1967, pp. 92-94.)
4 Technically, the gap is presumably inside a complex Noun Phrase in these constructions, and not inside a PP; due to the fact that the PP-head, i.e. the "secondary postposition", bears an agreement marker, and that the "complement" with which the head agrees bears genitive marking, PPs involving "secondary Ps" look identical to a possessive NP, whose Genitive-marked "possessor" is arguably the specifier rather than a complement of the head (cf. Kornfilt 1984). The same analysis would carry over to PPs involving "secondary Ps"; note that in Turkish, complements (otherwise) never agree with their heads.
and PP-G RCs are case marked with Genitive case, the typical case marking for subjects of nominalized embedded clauses in general (this Genitive subject marking in RCs is shown in (2–3) above). These facts are also cited in Slobin (1986) as the reason for considering OG-RCs as syntactically and morphologically more complex than SG-RCs.

The most detailed acquisition study of Turkish RCs to this date is Slobin (1986). Slobin reports that Turkish RCs are acquired late (after age 5), basing his claim on a careful analysis of a naturalistic corpus of 40 hours. Slobin attributes the (supposed) late acquisition to the fact that RCs (and especially OG-RCs) violate some of the 'Operating Principles' he claims hold in child language. In addition, he believes that children have not sorted out the participial and nominalized forms of clauses involved in relativization in Turkish. It is therefore important to supplement the naturalistic observations with experimental data. The results we present in this paper do not support Slobin’s claim for Turkish RCs being learned late in either subject or object position, but our study clearly supports Slobin’s original conclusion that OG-RCs are more difficult than SG-RCs.

There are very few previous studies on Turkish RC acquisition and an obvious lack of experimental studies. As discussed above, Slobin (1986) reports that very few RCs were produced in his survey of 40 hours of naturalistic data: 37 SG- and 5 OG-RCs. His conclusions were that RCs are not productive before age 5 or 6, since children mainly used other strategies (two separate clauses, or simple sentences) to express the functional equivalence of an RC. Ekmekçi (2000) reports on two studies in Turkish which are relevant. Her first study is an imitation study of adjectival phrase RCs (ages 3–6, with age distribution and exact number of subjects not reported). The 3–4 year olds in her study performed at 54% success rate for SG-RCs (what Ekmekçi calls Rel1) and at 63% with OG-RCs (called Rel2 in her study) with children over 4 “showing a significant progress towards subject-participial RCs” (p. 35). She also states that “the progress in object-participial (Rel2) seems to start earlier”, since children even at age 4 achieve a 64% imitation rate. A second production level experiment was administered (the method is not described in detail), and she concludes that children produced Rel1 (SG-RC) at only 56% at age three to 82% at age 5 (with no breakdown given). As far as Rel2 (OG-RC) is concerned the production rate was 50–55% at ages 5–6. Thus, there is a hint here that children had greater difficulty with OG-RCs than with SG-RCs in production, but since the studies are not described in detail we cannot get a clear picture of how difficult SG-RCs vs. OG-RCs really are in Turkish child language.

In a recent comprehension study (Özge, Marinis & Zeyrek, forthcoming) 37 children (aged 5–8) were tested with a picture comprehension test. Both the control adults and the children performed well on SG-RCs but younger children performed much worse on reversible OG-RCs (60.29%) than SG-RCs (98.16%). Even older children performed worse on reversible OG-RCs (72.69%) than SG-RCs (96.03). As pointed out by the authors, since the task involves picture choice it is hard to determine whether the reduced performance

5 "Reversible RCs" are RCs in which the agent and patient are both animate full NPs, so the sentence is potentially semantically reversible.
with OG-RCs is due to the morphology involved in this construction or due to other factors, but the authors conclude that the lower accuracy in OG-RCs may at least in part be attributed to difficulties with genitive-possessive agreement morphology and the use of the -DIK marker in OG-RCs.

Since there are some preliminary results indicating that OG-RCs pose bigger problems than SG-RCs, a subset of the authors of the present study conducted a series of production experiments to not only confirm the object- versus subject-gap difference but to also address the issue of what the cause may be of such a difference, if found (cf. Hermon, Kornfilt & Öztürk 2010). Comprehension experiments cannot address the issue of whether it is the distance between the head and gap which causes the problem, or whether young children simply do not have the required syntactic and morphological knowledge needed to comprehend RCs. In examining production, and crucially by analyzing the production errors made in each RC condition, we can start to understand whether the nature of the problem children have with OG-RCs is due to increased processing demands or due to a lack of morphological and/or syntactic competence involved in these structures.

3. Design of the elicited production experiments

Because comprehension experiments cannot address the question of whether it is the syntactic distance between the head and the gap in RCs which causes the problem, or whether young children don’t have the required syntactic and morphological knowledge needed to comprehend RCs, elicitation experiments were devised. In examining production, and by analysing the production errors made in each RC construction, we can understand whether the additional problems children have with non-subject RCs are due to increased processing demands or due to a lack of morphological and/or syntactic competence involved. In this way we can confirm the subject-RC versus non-subject RC difference and also find out the cause of the difference.

6 The literature on RC-processing is focused around two competing theories which compute such distance differently. The Filler-gap Linear Distance Hypothesis (LDH) quantifies the distance between the filler, i.e. the head noun, and the gap in terms of the number of intervening referential elements, whereby only elements introducing new discourse "referents" (according to Gibson 1998, 2000, the head noun of an NP and the head verb of a VP) are considered as "intervening" between the filler and the gap. The LDH correctly predicts that in head-initial RCs, SG-RCs are easier to process than OG-RCs. When extended to language acquisition, the LDH predicts that in head-final RCs such as in Turkish, OG-RCs should be easier to acquire than SG-RCs. An alternative account, the Structural Distance Hypothesis (SDH), attributes the relative performance difficulty with different types of RCs to the position of the gap in the hierarchical structure (cf. O'Grady 1997, where it is suggested that structures with gaps that are more deeply embedded inside the structure are more complex, and thus are harder for children). For head-final (as well as head-initial) RCs, the SDH predicts that OG-RCs should be harder to process as well as to acquire than SG-RCs, also predicting that Postpositional gap-RCs should be the hardest in those regards. Our experiments show that the LDH makes wrong predictions with respect to the acquisition of head-final RCs, while the SDH's predictions are correct.
3.1. Materials

For the elicited production experiments, we used the experimental technique developed by Zukowski (2001) to reveal the participants’ grammatical knowledge of subject-gap RCs and object-gap RCs. Eight sets of four pictures were shown on a laptop computer screen and the printed versions of the same pictures were shown, as well. Each set included one base picture which introduced 2 identical characters/objects and a following picture ("question picture"), which contained a change from the base picture to be described by the participant. Of the 8 base pictures, 4 were designed to elicit subject-gap RCs, 2 were designed to elicit object-gap RCs, and 2 elicited PP-gap RCs. In the pictures designed to elicit subject-gap RCs, it was the agent of the event depicted in the base picture that underwent a change. In the pictures for eliciting object-gap RCs and PP-gap RCs, it was the patient that underwent a change.

3.2. Informants

For the elicited production experiments, a total of 20 Turkish speaking children in Istanbul (10 male, 10 female) and 10 adult speakers of Turkish (as a control group) took part in the study. Children’s age varied between 4;1 and 6;2, with a mean age of 4;8. Data were collected in Istanbul in August 2005 and January 2006 by the fourth author of the present paper. The naturalistic data were obtained from a Turkish-Dutch bilingual child growing up in the Netherlands. Doğa has been learning Dutch outside of his home, in kindergarten etc. and is now bilingual. His utterances have been recorded and transcribed by his father, Dr. Türkay Yalnız, from an early age on, and the transcriptions also include remarks on the situations in which those utterances were made.

3.3. Procedures

Each of the 8 base pictures was randomly presented 4 times throughout the experimental study (intermixed with different pictures), and each time the base picture was followed by a question picture to elicit an RC response. (1a) and (1b) is a pair of pictures that were used to elicit a SG-RC:
The first picture above is the base picture which depicts two girls: One girl is chasing a rabbit and the other girl is chasing a cat. The second picture is the changed picture (the question picture), in which one of the girls changes her colour to blue. Once the question picture was presented to the participant the experimenter asked "Which girl turned blue?" to elicit an SG-RC with "the girl" as its head. (2a) and (2b) is a pair of pictures that were used to elicit an OG-RC:

Once the question picture was presented to the participant, the experimenter asked "Which cat turned purple?" to prompt the child to produce an OG-RC. The eight base pictures were presented four times each followed by different question pictures, producing a total of 32 trials. At no point in the discourse did the experimenter use RCs, so as not to prime any particular RC structures. The procedure for the adult participants was exactly the same as for child participants, but there was no uninformed listener; the adult participant simply described the change in the picture to the experimenter by answering the target question. (For further details about the experimental procedures, cf. Hermon, Kornfilt & Öztürk 2010.)

4. Results

In order to get a clear picture about the children’s competence, we analysed informants’ utterances. Basically three sets of results can be reported. First of all, out of all the possible target structures, 88.4% of attempted RCs (508 instances out of 575) were grammatical RCs. Even though these RCs were not always pragmatically appropriate due to rever-
sal errors, they were basically well-formed RCs. Secondly, in 90% of the instances the children produced a RC (574 instances out of 640 possibilities), and only in 10% of the instances did they produce other structures. Thirdly, all the children had appropriate morphological competence. There were no errors of deleting the nominalizers, of lacking agreement morphology on -DIK forms, or having improper agreement morphology with -(y)An forms.

However, the success rate is clearly much higher for subject-gap RCs than for non-subject-gap RCs, and among the latter, higher for object-gap RCs than for PP-gap RCs. The informants successfully produced subject-gap RCs: 91.6% (293 out of 320); direct object-gap: 57.5% (92 out of 160); and "PP-gap": 26.2% (42 out of 160). The rate of ill-formed relative clauses was quite low: 3.1% in expected subject-gap RCs (10 out of 320), 6.9% in object-gap RCs (11 out of 160), and 34.4% in "PP-gap" RCs (55 out of 160). In the same vein, the rate of incorrect conversion of thematic roles were also quite low: 2.8% (9 out of 320) for subject-gap RCs, 12.5% (20 out of 160) for object-gap RCs, and 28.8% (46 out of 160) for "PP-gap" RCs. As a result of avoiding RC constructions, a small number of other structures was produced as well: 2.5% (8 out of 320) for subject-gap RCs, 23.1% (37 out of 160) for object-gap RCs, and 10.6% (17 out of 160) for "PP-gap" RCs.

Grammatical errors were only marginally higher in object-gap RCs than in subject-gap RCs. But avoidance of RCs was greater in the non-subject-gap targets. For the elicited production tasks, we conclude that the higher success rate with subject-gap RCs in a head-final RC language like Turkish with SOV-order can be predicted only by the SDH for processing. (The LDH predicts, as mentioned earlier, that the object-gap Cs in head-final RC languages should be easier than subject-gap RCs, and PP-gap RCs are predicted to be as difficult as subject-gap RCs. Our results contradict these predictions.) The structural account correctly predicts both the contrast between subject-gap RCs and object-gap RCs, and the contrast between object-gap RCs and PP-gap RCs.

We found no errors involving the obligatory agreement morphology (in our examples, 3rd person singular agreement) with -DIK forms in the entire data set, and agreement was never used with the -(y)An form. Young children, contrary to Slobin and Aksu-Koç & Slobin's claims, control the participial morphology needed for both subject-gap and object-gap RCs; they know when agreement is required, and they also know about the need for genitive case on subjects of these nominalized clauses.

Upon examination of the structures obtained from the bilingual informant, it is clear that some structures emerge much earlier than we would have anticipated, if our anticipation were based on the relevant claims in the literature. While it may be the case that Doğa is somewhat precocious in his linguistic and intellectual development, some of the examples show that he underwent some of the typical stages that have been described for the acquisition of Turkish and other languages. The data show clearly that Doğa started producing RCs in well-formed as well as pragmatically and semantically appropriate ways at ages completely comparable to his counterparts acquiring head-initial RCs in Indo-European languages. We do observe examples of subject RCs at an earlier stage than non-subject RCs (2;4 for the first clear-cut, successful, genuine subject RC), but even non-subject RCs start emerging at a relatively young age, the earliest being at 2;9.
Subject RCs:

noise make-(y)An car noise make-(y)An car leave-PST
'The car that was making noise. The car that was making noise left.' (Age: 2;3)


(5) Kirmizi... Şapka-si ol -ma -yan hostes -i at-ti-m.
red hat -3.SG be-NEG-(y)An stewardess-ACC throw-PST-1.SG
'I threw out the red stewardess who didn’t have a hat.' (Age: 3;1)

(Doğ'a’s father hears him say that he threw out the stewardess (a toy); however, he points out that there is a toy stewardess still around, upon which Doğ'a tells him that he threw out another toy stewardess: a red one missing a hat.)

Non-subject RCs:

(6) Ozay -in ver -diğ -i geyik o taraf -ta!
Ozay-GEN give -DIK -3.SG deer that side -LOC
'The deer which Ozay gave (me) is on that side!' (Age: 2;8)

(Doğ’a’s uncle, Özay, had given him a toy reindeer as a present, and his father had taken it off his bed (cot); with this utterance, Doğ'a was asking his father to pick it up from the floor and to hand it to him.)

(7a) Doğ'a: [\[Emre-nin i getir-diğ -i\] dondurma-nin], Doğ'a
Emre-GEN bring-DIK-3.SG ice cream-GEN Doğ'a
belki ___i külüğ-in -l ye -r.
perhaps cone-3.SG-ACC eat-AOR (3.SG.)
'Doğ'a may perhaps eat the cone of the ice cream that Emre brought' (Age: 3;1)

cone-3.SG-ACC Doğ-a-dimin.-1.SG
(Correction of pronunciation by the father): 'Its külah (instead of \[küla:\]) (=‘cone’), my dear Doğ'a.'

(7c) Doğ'a: Külah-in -l. Doğ-a belki külah-in-1 ye-r.
cone-3.SG-ACC Doğ-a perhaps cone-3.SG-ACC eat-AOR (3.SG.)
(Repeats the correction in a correct way): 'Its külah. Doğ'a may perhaps eat its cone.'

(7d) Türkay (father): Evet.
yes

Doğ'a’s first utterance with the object-RC was uttered spontaneously, while having dinner, with a thought for the dessert afterwards.

It is thus clear that Doğ'a’s competence as well as performance are solid with respect to both subject-RCs and object-RCs – and this at a fairly young age.

Even if we were to assume that Doğ'a is precocious, he is just somewhat ahead of his
peers, but not unusually so. Note that his linguistic output displays some properties typical of his age cross-linguistically. For example, in (7), he refers to himself by his name, and refers to himself in the third-person singular, instead of the first person singular. Furthermore, he drops the $h$ in the word for *kulah* 'cone' in his spontaneous utterance (but is able to imitate it correctly after his father’s correction). Dropping $h$’s and referring to themselves in the third person is observed in other children of similar age, too.

With respect to RCs, too, some of Doğa’s utterances are different from what an adult would have said; e.g.

(8) klarnet çal -an alet -i ver! üfli -yce -m!
clarinet play -(y)An instrument -ACC give blow -FUT -1.SG

Target: Either: klarnet gibi çal -in -an alet -i ver!
clarinet like play-PASS-(y)An instrument-ACC give
‘Give (me) the instrument that is played like a clarinet!’

Or: klarnet gibi çal-diğ -im/-in alet -i ver!
clarinet like play-DIK -1.SG/2.SG instrument-ACC give
‘Give (me) the instrument which I/you play like a clarinet!’

Here, Doğa’s "mistake" is very similar to some of the errors we saw with the monolingual children during our elicitation experiments: the RC in (8) is formally correct, but does not fit the meaning to be expressed; the utterance can be viewed as an instance of thematic reversal: the clarinet does not play but is played, or someone plays it. This is exactly the kind of “error” we had found in the course of our elicitation experiments: subject-RCs which were formally well-shaped but were reversals of thematic roles, whereby the “correct” RC would have been either an object-RC (as in the second option above), or a subject-RC with a passive (as in the first option above). Given that Doğa started producing “correct” object-RCs soon thereafter, this may well be viewed as a stage where Turkish children do have the competence for such RCs but are still hampered by processing difficulties – a stage to be soon replaced by one where the grammatical competence is completed by the “correct” production and processing, as well.

5. Conclusions

The greater structural embedding of a gap (as in object-gap RCs and PP-gap RCs) seems to cause added difficulties in both production and comprehension regardless of linear distance between the head and the gap, and theories of production will have to come up with a coherent production hypothesis to account for these facts. Our current conclusion is that young children do possess the syntactic and morphological competence for RCs at an age comparable to their counterparts acquiring head-initial languages, and that their performance errors in production and comprehension are due to processing issues whose exact nature still await ultimate answers, but which are similar across languages and are not due to Turkish-specific syntactic (“iconicity”) or morphological (case, agreement, no-
minalization) difficulties. Clearly, further studies are needed to determine the exact reason for why children produce subject-gap RCs with greater accuracy across languages and language types.

As also shown by the naturalistic acquisition data, Turkish speaking children are able to produce both subject and object relative clauses at much younger ages than generally claimed in the literature. In addition to elicitation tasks, we also need more naturalistic data to reflect on the actual real life production forms (and rates). Finally, Slobin’s (1986) and Slobin & Aksu-Koç’s (1985) earlier claims with respect to acquisition of relative clauses require full revision.

References


