A-2   Raising to Quirky Subject in Tatar[*]

Chihiro Taguchi

Nara Institute of Science and Technology

`taguchi.chihiro.td0@is.naist.jp`

Keywords— Tatar, Turkic, syntax, morphology, raising, quirky subject

Abstract

Tatar (< Kipchak < Turkic) has subject-to-subject raising predicates where the raised argument is marked by nominative, dative and genitive case, of which the latter two are "quirky (i.e., non-nominative) subjects" (Sigurðsson 1992). Poole (2016) proposed the Quirky Subject Hierarchy, a three-level hierarchy of the subjecthood of quirky subjects, and classified languages with quirky subject into three categories: Hindi type, Icelandic type, and Laz type. However, the Tatar data in this paper demonstrate that a language may fit more than one categories; a raised subject with dative case-marking in Tatar displays the Icelandic-type characteristics, and a raised subject with genitive displays the Laz-type. I propose that these properties are not language-specific but lexicon-specific.

1 Introduction

The Tatar language is a Kipchak language of the Turkic language family. It is an agglutinative language with several nominal case suffixes. The canonical word order is SOV, and the modification order is AN.

This paper deals with two kinds of subject-to-subject raising (SSR) predicates whose subject is marked by genitive and dative. (1)-(3) are examples of raising constructions in Tatar, (1) with a nominative subject which is a typical way of marking a subject, (2) with a dative, and (3) with a genitive, all having by and large similar deontological meaning. Section 2 outlines what raising is and how it is analyzed in general theoretical linguistics. In section 3 it is argued that the raised arguments with the three kinds of case-marking are all "quirky" subjects accompanied by non-canonical cases. Then, based on the Tatar data, section 4 presents a counterargument against the quirky subject hierarchy (QSH) proposed by Poole (2016).

(1) *Marat   kit-ärgä   tiješ.*
    PN.NOM leave-INF obligatory
    "Marat must leave."

(2) *Marat-qa kit-ärgä   kiräk.*
    PN-DAT   leave-INF necessary
    "Marat must leave."

(3) *Marat-niŋ kit-äse   bar.*
    PN-GEN   leave-FP.3 exist
    "Marat must leave."

---

2 Raising

2.1 Overview

Raising is a syntactic phenomenon commonly found among languages. Compare the two English sentences below:

(4) John seemed to leave.

(5) John tried to leave.

Although they have a similar construction, the predicates show different syntactic operations. The type of the former verb SEEM is called a raising verb, and the latter one TRY is a control verb. The most conspicuous difference of them is that SEEM does not subcategorize for a subject as its core argument, while TRY does. The matrix subject of (4) is in fact the subject of LEAVE, which is "raised" from the subordinate clause to the matrix clause. For clarifying the movement, (6) contains t (trace) in the position where it is supposed to be before the raising. On the other hand, "John" in (5) is the subject of TRY as well as of LEAVE, the latter of which is controlled by the matrix subject. The controlled subject of LEAVE is expressed as PRO in (7).

(6) John$_i$ seemed [t$_i$ to leave].

(7) John$_i$ tried [PRO$_i$ to leave].

Raising to a subject prevents the subject position from being empty and thereby suffices the Extended Projection Principle (Halle & Marantz 1993), which stipulates that the specifier position of a finite TP (i.e. subject) must be filled. The fact that any subject is not subcategorized for by SEEM is also confirmed by a paraphrase of (4), where an expletive IT substitutes in the subject position as in (8).

(8) It seemed that John left.

It is reported that there are two types of raising: subject-to-subject raising (SSR) and subject-to-object raising (SOR). SSR is exemplified in (4), and an example of SOR is shown in (9). The subject of LEAVE is "Mary", but the argument is raised to the object position of the matrix clause.

(9) John believes Mary$_i$ [t$_i$ to have left].

2.2 Raising in Tatar

A non-modal example of (1)-(3) is (10), in which the subject of LEAVE is marked by nominative. When this phrase is embedded within another matrix clause as an infinitival construction (here the matrix predicate can be *tiješ*, *kiräk* and *bar*), the subject is raised to the matrix clause because infinitives do not assign a nominative case in Tatar. As a result of raising, a morphological case of the raised subject is assigned by the matrix predicate.

(10) *Marat   kit-ä.*
     PN.NOM leave-PRS.3
     "Marat leaves."

For instance, the syntactic structure of (1) is (11).

(11) *Marat*ᵢ [IP tᵢ *kit-ärgä*] *tiješ.*

The fact that this construction is not a flat structure but raising can be confirmed by the passivization test. The difference is evident in the semantic representations of sentences (12)-(15). As in (12) and (13), the semantic relationship of the core arguments is the same even after passivization. In contrast, for control predicates in (14) and (15), the semantic role is switched and their meanings are no longer similar at all.

(12) *Marat   xat-nï   jaz-arɣa   tiješ.*
PN.NOM letter-ACC write-INF obligatory
"Marat must write the letter.": $write\langle marat, letter\rangle$

(13) *xat   Marat taraf-ïn-nan jaz-ïl-ïrɣa   tiješ*
letter.NOM PN   side-3-ABL   write-PASS-INF obligatory
"The letter must be written by Marat.": $write\langle marat, letter\rangle$

(14) *Marat   xat-nï   jaz-arɣa   tïrïš-tï.*
PN.NOM letter-ACC write-INF try-PST.3
"Marat tried to write the letter.": $try\langle marat, \lambda x.write\langle x, letter\rangle\rangle$

(15) *xat   Marat taraf-ïn-nan jaz-ïl-ïrɣa   tïrïš-tï.*
letter.NOM PN   side-3-ABL   write-PASS-INF try-PST.3
"The letter tried to be written by Marat.": $try\langle letter, \lambda x.write\langle marat, x\rangle\rangle$

Whether the raised arguments are in fact subjects or not will be discussed in section 3.2.

3   Quirky Subject

3.1   Overview

Quirky subjects refer to subjects with non-canonical case-marking (particularly non-nominative), typically reported in Icelandic (Sigurðsson 1992).

(16) Þeir   hjálpuðu   okkur.
they.NOM help.PST.3.PL we.DAT
"They helped us."

(17) Okkur   var   hjálpað.
we.DAT be.PST.3.SG help.PP
"We were helped."

This construction is contrasted by so-called "quirky-like" construction, epitomized by German in (18) and (19). Although these sentences look similar to the Icelandic counterparts (16) and (17), Zaenen et al. (1985) pointed out that *uns* in (19) is not a subject but an object, i.e., not Spec,IP.

(18) Sie   halfen   uns.
they.NOM help.PST.3.PL we.DAT
"They helped us."

(19) Uns   wurde   geholfen.
    we.DAT be.PST.3.SG help.PP
    "We were helped."

This fact is empirically confirmed by means of subjecthood tests (Sigurðsson 1992). For instance, while the dative subject in Icelandic can bind PRO in the subordinate clause in (20), the pseudo-subject of the German sentences cannot as in (21)[1].

(20) Við   vonuðumst   til [að verða   hjálpað].
    we.NOM hope.MED.PST.1.PL for [to become.INF help.PP]
    "We hoped to be helped."

(21) *Wir   hofften   [geholfen zu werden].
    we.NOM hope.PST.1.PL [help.PP to become.INF]
    "We hoped to be helped."

3.2   Quirky subject in Tatar raising constructions

To determine that the raised arguments observed in section 1 is a subject, we shall now employ the subjecthood tests introduced in Sigurðsson (1992) to Tatar. Below we examine the Anaphoric Binding test and the PRO test applied to Tatar; although Sigurðsson (1992) also mentions the Reduced Conjunction test, it cannot be applied to Tatar because the subject does not have to be morphologically overt, and it is difficult to judge whether the subject was reduced due to the conjunction or just covert.

3.2.1   Anaphoric Binding Test

In (22), the raised subject *Marat* is marked by dative, and it is possible to anaphorically bind the reflexive pronoun *berüze*, which morphologically appears in the default nominative form. This is also the case for the raised argument marked by genitive as in (23).

(22) *Marat-qa$_i$ [berüz-e$_i$ kit-ärgä]   kiräk.*
    PN-DAT   [oneself-3.NOM leave-INF] necessary
    "Marat must leave by himself."

(23) *Marat-niŋ$_i$ [berüz-e$_i$ kit-äse]   bar.*
    PN-GEN   [oneself-3.NOM leave-FP.3] exist
    "Marat must leave by himself."

This shows that the raised genitive and the raised dative have some subjecthood property.

3.2.2   PRO Test

In Sentence (24), the unpronounced PRO is assumed to be case-marked by dative assigned by the predicate *kiräk*, and is bound by *Marat*. Likewise, PRO in (25) is bound by the matrix subject, in which case it is more obvious that PRO is assumed to be marked by genitive by virtue of the third person possessive suffix -*ï(n)* in the verbal noun. It should be noted here that the verb *bul-* is selected instead of *bar* seen in (3), because *bar* cannot be inflected and *bul-* is an alternative to it in such cases.

---

[1]The sentences are cited from Sigurðsson (1992) to which I added glossing.

(24) *Marat*ᵢ [[PROᵢ *kit-ärgä kiräk*] *bul-u-ɣa*] *künder-el-de.*
PN.NOM [[PRO.DAT leave-INF necessary] be-VN-DAT] persuade-PASS-PST.3
"Marat was persuaded that he had to leave."

(25) *Marat*ᵢ [PROᵢ *kit-äse bul-u-ïn-a*] *künder-el-de.*
PN.NOM [PRO.GEN leave-FP be-VN-3-DAT] persuade-PASS-PST.3
"Marat was persuaded that he would have to leave."

3.2.3   Reduced Relative Test

Reduced relative clauses (RRC) are relative clauses that do not use a relative pronoun including optionally dropped relative pronouns. Poole argues that in RRC the relativized argument has to be in the subject position, and makes an assumption that "If XP can be relativized in reduced relatives, XP is a subject" (2016: 7). In English, for example, Sentence (26) is grammatical, while (27) and (28) are not when the same meaning is intended.

(26) Johnᵢ [_ᵢ giving the cattle the hay]

(27) *The cattleᵢ [_ᵢ giving John the hay]

(28) *The hayᵢ [_ᵢ giving John the cattle]

In Tatar, the grammaticality of RRC varies between the relativized dative subject (29) and the relativized genitive (30). The phrase (29) is controversial as to its grammaticality among native speakers, and they prefer avoiding this structure. In contrast, phrase (30) was accepted as a grammatical one. Therefore, while the acceptability of (29) remains unclear, the genitive argument in (30) is tested to be a subject.

In (30), *ixtimal bul-* "be possible" is employed instead of *-äse bar*, because the elicitation was difficult with the latter expression.

(29) ??_ᵢ *kit-ärgä kiräk bul-ɣan keše*ᵢ
leave-INF necessary be-ADJVZ person
"The person who had to leave"

(30) _ᵢ *kit-ü ixtimal-ï bul-ɣan keše*ᵢ
leave-VN possibility-3 be-ADJVZ person
"The person who might leave"

3.3   Raising to Quirky Subject: Uyghur

SSR with a quirky case-marking is also reported in Uyghur, a Karluk language of the Turkic language family (Asarina 2011). In (31), the verb in the subordinate clause *oqu-š* is a verbal noun form, i.e., a non-finite form, therefore it cannot assign a nominative case to its subject. As a result of raising, The morphological genitive case is assigned to the raised argument by the third person possessive suffix *-i*, while the argument syntactically plays a role of the matrix subject. The raised argument is the subject, because *bu ehtimalda* "probably" modifying the entire sentence can be inserted between *Ötkür-niŋ* and *oqu-š-i*, whereas it is impossible for a control predicate *muhim* "important" as in (32). Therefore, the first argument *Ötkür-niŋ* in (31) is located in Spec,IP position, while that in (32) is not[2].

---

[2]It is probably in Spec,CP.

(31) *Ötkür-niŋ*ᵢ *bu+ehtimalda* tᵢ *oqu-š-i   kérek.*   (from Asarina (2011))
     PN-GEN   probably   read-VN-3 necessary
     Ötkür probably must read.

(32) ***Ötkür-niŋ*ᵢ *bu+ehtimalda* tᵢ *oqu-š-i   muhim.*   (from Asarina (2011))
     PN-GEN   probably   read-VN-3 important
     intended: "It is probably important for Ötkür to read.

4 Quirky Subject Hierarchy

4.1 Overview

Through a cross-linguistic study of quirky subject, there seems to be a hierarchy about the properties of quirky subject. Poole (2016) proposes the quirky subject hierarchy (QSH) (33) which categorizes three properties of quirky subjects, and names them Hindi type, Icelandic type, and Laz type respectively. The hierarchy is alternatively expressed in Table 1. For example, a Laz-type language which allows for quirky subject with reduced relativization (lacking CP layer in contrast to a finite relative clause) must also allow for the preceding two properties.

(33) QSH: SOA Binding $\gg$ PRO Binding $\gg$ Reduced Relatives (cited from Poole (2016))
         Hindi type   Icelandic type   Laz type
     For a hierarchy $p_1 \gg p_2 \gg \cdots \gg p_n$, a DP has property p_{i+1} only if it has property p_i

Table 1: QSH

	SOA Binding	PRO Binding	Reduced Relatives
Hindi type	✓	*	*
Icelandic type	✓	✓	*
Laz type	✓	✓	✓

Table 2 is a table based on the partial typological distribution of languages in terms of QSH types given by Poole (2016). Laz-type languages are rarely found and controversial languages are put in parentheses.

Table 2: Typology of QS (Poole 2016)

Hindi type	Hindi, German, Basque, Gujarati, Hungarian, Kannada, Korean, Malayalam, Russian, Spanish
Icelandic type	Icelandic, Faroese, Tamil, Telugu
Laz type	Laz, (Marathi, Tamil, Telugu)

4.2 QSH types are not language-specific

According to this categorization, Tatar is categorized into two types of the QSH. A dative subject in SSR can bind subject-oriented anaphora (SOA) as in (22), and can bind a PRO as in (24). Likewise, the genitive subject passes both the Binding diagnostic (23) and the PRO

diagnostic (25). From this observation, SSR predicates requiring lexical dative case-marking on the subject belong to the Icelandic type, whereas the ones requiring lexical genitive case-marking belong to the Laz type. The distribution is represented in Table 3, and the ambiguous typological position is expressed in Table 4.

Table 3: QSH and Tatar QS

	SOA Binding	PRO Binding	Reduced Relatives
Hindi type	✓	*	*
Icelandic type	✓	✓	*
Laz type	✓	✓	✓
Tatar dative subject	✓	✓	??
Tatar genitive subject	✓	✓	✓

Table 4: Ambiguous classification of Tatar

Hindi type	Hindi, German, Basque, Gujarati, Hungarian, Kannada, Korean, Malayalam, Russian, Spanish
Icelandic type	Icelandic, Faroese, Tamil, Telugu, (??Tatar)
Laz type	Laz, (Marathi, Tamil, Telugu, ??Tatar)

This result empirically demonstrates that the order of the QSH is valid for the Tatar raised quirky subjects. Notably, however, it also tells us that one language may contain several quirky subjects occupying different positions on the QSH, and that the degree of subjecthood of a quirky subject is not language-specific but rather lexically (or morphologically) determined. For this reason, it is wrong to label a language with one of the three types suggested by Poole (2016). In order to evaluate the properties of a language's quirky subjects, it is necessary to look through the comprehensive dataset of the quirky subjects.

Abbreviations: 3 — 3rd person (possessive for nouns); ADJVZ — adjectivizer; D — dative; FP — future participle; G — genitive; INF — infinitive; MED — mediopassive; N — nominative; PASS — passive; PL — plural; PN — personal name; PP — past participle; PST — past tense; VN — verbal noun (nominalizer for verbs)

References

Asarina, A. A. (2011). Case in Uyghur and beyond. PhD thesis, Massachusetts Institute of Technology.

Halle, M. & A. Marantz. (1993). Distributed morphology and the pieces of inflection. In K. L. Hale & S. J. Keyser (eds.), *The view from Building 20: Essays in Honor of Sylvain Bromberger*. MIT Press, pp. 111-76.

Marantz, A. (2000). Case and licensing. In E. Reuland (ed.), *Arguments and case: Explaining Burzio's generalization*. John Benjamins, pp. 11-30.

Poole, E. (2016). Deconstructing subjecthood. Ms., University of Massachusetts Amherst.

Sigurðsson, H. A. (1992). The case of quirky subjects. *Working papers in Scandinavian syntax*, 49. Department of Scandinavian Languages, Lund University.

Zaenen, A., H. Thráinsson & J. Maling. (1985). Case and grammatical functions: the Icelandic passive. In *Natural Language and Linguistic Theory* 3(4). Springer Verlag, pp. 441-483.