

Exceptional Case Marking in Uzbek Complementizer Clauses

Christopher A. Straughn
University of Chicago

1 Introduction

Uzbek is a Turkic language spoken in Uzbekistan, Afghanistan, and surrounding regions. The phenomena discussed in this paper are based on the standard dialect of Uzbek based on the dialect of Tashkent, the capital of Uzbekistan. For the sake of readability, standard Uzbek Latin orthography has been employed.

The title of this paper is perhaps misleading, as the phenomena to be discussed are not so easily defined as being ‘complementizers’ or ‘clauses’. A more appropriate pair of terms might be ‘small clause’ and ‘complementizer-like head of a small clause’. Nevertheless, it is the difficulty in distinguishing these various categories that makes this topic worth studying, and we will keep this in mind as we examine Uzbek examples of the following type:

- (1) *Ular she’r-ni ilohiy deb bil-a-di.*
they poem-ACC sacred COMP know-sc pres-3
‘They know the poem to be sacred.’

The goal of this paper, then, is to provide a description of exceptional case marking (ECM) in Uzbek, which has been poorly documented, and to provide an analysis of these constructions within a generative framework. To accomplish these goals, we will start with some background information on Uzbek and its system of complementation. Following that, we will provide a description of the ECM construction. Previous analyses of ECM in other SOV languages (i.e. Turkish and Japanese) may provide some assistance in analyzing the Uzbek construction, so they will be briefly discussed as well. Finally, the ECM construction itself will be analyzed within a generative framework.

2 A Description of ECM in Uzbek

Uzbek appears to be alone within the Turkic language family in possessing the exact ECM phenomena that will be described below. Similar phenomena have not been mentioned in the grammars of any other Turkic languages, and, when tested, speakers of three Turkic languages: Turkish, Kazakh, and Sakha (Yakut); outright reject these constructions. Uzbek appears to stand alone in having this specific type of ECM. ECM in Turkish, as will be further elaborated, may be found, but it is very different from the type found in Uzbek. In order to understand what makes the Uzbek phenomena so special, we must start with a brief discussion of its complementation system.

2.1 Complementation in Uzbek

Uzbek possesses a number of complementation strategies, of which the ECM construction is just one. The strategy that must first be discussed here involves full complement clauses headed by the complementizer *deb*, which evolved from a converbial form of the verb meaning ‘say’. Unlike certain types of complementizers borrowed from Persian, complement clauses headed by *deb* obey the SOV, head-final word ordering of Uzbek. Structurally, this type of complementizer takes full clauses, with no restrictions on what verbal categories may be present.

- (2) (Man) *John Mary-ni sev-a-di deb eshit-di-m.*
 (I) John Mary-ACC love-PRES-3 COMP hear-PST-1SG
 ‘I heard that John loves Mary.’

In the above example, the order of major constituents is (S)_{lo} S O V] V. Although the subject is optional in this example (Uzbek is pro-drop), the crucial characteristic of verb-finality is clearly shown.

A second typological characteristic of Uzbek is preserved in this example as well: the alignment system. Like other Turkic languages, Uzbek is a nominative-accusative language, with a rather interesting system of differential accusative-case marking. In the above example, all subjects are marked with the (null) nominative case and the object with accusative case. The accusative case is not marked on clauses headed by complementizers. This pattern is the norm in Uzbek, and any exceptions to this norm are usually idiosyncratic. Keeping in mind what complementation usually looks like, we may move on to the ECM construction.

2.2 ECM in Uzbek and its Peculiarities

ECM is distinguished from normal complementation by a number of morphological, semantic, and syntactic features. By far the most obvious is the presence of the accusative case on the subject of the embedded clause, as shown below:

- (3) *John Mary-ni chiroyli deb top-di.*
 John Mary-ACC beautiful COMP find-PST.3
 ‘John found Mary to be beautiful.’

As previously noted, this is unusual, as the accusative case is usually reserved for the objects of verbs, and never for subjects, not even the subjects of complement clauses:

- (4) *John(*-ni) Mary-ni sevadi.*
 John(*-ACC) Mary-ACC loves
 ‘John loves Mary.’
- (5) *John(*-ni) Mary-ni sevadi deb o’yla-y-man.*
 John(*-ACC) Mary-ACC loves COMP think-PRES-1SG
 ‘I think that John loves Mary.’

Although some verbs in Uzbek allow for the assignment of cases other than the accusative to their objects, subjects (aside from the ECM construction) receive only the nominative case.

A second peculiarity of the ECM construction in Uzbek is that it co-occurs with a limited set of verbs, all of which have different meanings from their regular usages. In comparison, standard complementation takes from a much wider range of verbs and the meanings of those verbs is entirely predictable. In the very first example, the verb *bil-*, ‘to know’, might be more accurately translated as ‘to consider’. A brief list of verbs commonly found in this construction are shown below:

Verb	Usual Meaning	Meaning w/ ECM
<i>bil-</i>	know	consider
<i>o’yla-</i>	think	consider
<i>top-</i>	find	find/come to consider

Table 1: ECM-taking verbs

As for the predicate in the embedded construction, it exhibits a number of peculiarities as well. The first of these is the restriction of the predicate to nominal or adjectival forms. Verbal predicates are not allowed in this construction:

- (6) *Man u-ni qiziq/ yaxshi odam/ *tush-gan deb*
 I him-ACC interesting good man *fall-PTCP COMP
bil-a-man.
 know-PRES-1SG
 ‘I consider him interesting/a good man/*to have fallen.’

Another unusual feature of the embedded predicate is that it may not express tense, aspect, mood or any other verbal category, nor may person marking occur. As shown in the following examples, while all of these categories may normally be expressed on nominal and adjectival predicates, they may not be expressed in ECM constructions. (Note that when most verbal categories are added to nominal or adjectival predicates, they must be supported by a copula of the form *e-*.)

- (7) a. *Man kesal-man.*
 I tired-1SG
 ‘I am sick.’
 b. *O doktor men-i kesal deb top-di.*
 That doctor me-ACC sick COMP find-PST.3
 ‘That doctor found me to be sick.’
- (8) a. *Siz shoir e-di-ngiz.*
 You(PL) poet COP-PST-2PL
 ‘You were poets.’
 b. *Men siz-ni shoir deb bil-a-man.*
 I you(PL-ACC poet COMP consider-PRES-1SG
 ‘I consider you to be poets.’

As for the status of the accusative-marked nominal, it behaves simultaneously like a subject and an object. As is expected of subjects, pro-drop may occur.

- (9) (*U-ni*) *Xudo-dan qo'rq-adi-gan odam deb bil-ar*
 (Him-ACC) God-ABL fear-PROG-PART man COMP consider-AOR
e-dim.
 COP-PST-1SG
 'I used to consider him a man who feared God.'

Pro-drop may *not* occur with objects, however, which suggests that the accusative-marked nominal is not a true object.

- (10) (**U-ni*) *ye-gan-man.*
 (**It-ACC*) eat-PERF-1SG
 'I have eaten it.'

The verb *ye-* here is obligatorily transitive; even if its object is pronominal, it may not be dropped.

In addition to the presence of accusative case, however, two other tests suggest that these nominals should be considered objects: passivization and reflexive binding. Under passivization, the accusative nominal may be raised to subject position. This is never possible under true complementation.

- (11) a. *Ular she'r-ni ilohiy deb bil-a-di.*
 they poem-ACC sacred COMP know-PRES-3
 'They know the poem to be sacred.'
 b. *She'r ilohiy deb bil-in-a-di.*
 Poem sacred COMP know-PASS-PRES-3
 'The poem was known to be sacred.'
- (12) a. *Demokaratiya kurash-da chiniq-a-di deb ishon-a-man.*
 democracy war-LOC arise-PRES-3 COMP believe-PRES-1SG
 'I believe that democracy arises in war.'
 b. **Demokaratiya kurash-da chiniq-a-di deb ishon-il-a-di.*
 *democracy war-LOC arise-PRES-3 COMP believe-PASS-PRES-3
 'Democracy is believed that arises in war.'

Reflexive pronouns in Uzbek may not occur in subject position, as they would be improperly bound. This restriction extends to the subjects of embedded clauses where, as in English, a non-reflexive pronoun must be employed if one wishes to refer back to the subject. In the case of the accusative-marked nominal in ECM constructions, reflexive pronouns are allowed.

- (13) *Oyjon u/*o'z-i Ali ko'r-gan deb ishon-di.*
 Oyjon she/*self-3 Ali see-PERF.3 COMP believe-PST.3
 'Oyjon believed that she had seen Ali.'

Because the reflexive pronoun *o'z* is not allowed here, the pronoun *u* “he, she, it, that” may refer either to Oyjon, or to another person. This ambiguity is not present in ECM constructions, as the reflexive pronoun is allowed.

- (14) *Ozarbayjon o'z-i-ni musulmon mamlakat deb bil-a-di.*
 Azerbaijan self-3-ACC Muslim country COMP know-PRES-3
 ‘Azerbaijan considers itself a Muslim country.’

3 ECM in Related SOV Languages

Some scholars have speculated that the Turkic languages, the Mongolic languages, the Manchu-Tungus languages and Japanese and Korean are all united under the Altaic language family. While this relationship is far from undisputed, all five language groups share certain typological features, most notably an unmarked SOV word order. In searching for ECM in other verb-final languages, then, it is logical to begin searching within this hypothetical family. Two of these languages, Turkish and Japanese, have been shown to exhibit exceptional case marking. Before we further analyze Uzbek, it is useful to establish these two languages as points of comparison.

3.1 ECM in Turkish

Turkish is one of the only other Turkic languages for which ECM has been documented. Due to poor documentation of the syntax of these languages, it is unclear whether other Turkic languages exhibit this phenomenon; speakers of Kazakh and Sakha (Yakut) reject both the Uzbek and Turkish types of ECM.

In Turkish, ECM is manifested quite differently from Uzbek, as explained in Zidani-Eroğlu (1997). In Turkish, ECM constructions alternate with regular complementizer clauses headed by null complementizers.

- (15) a. *Siz ben Ali-yi gör-dü-m san-iyor-sunuz.*
 You(PL) I Ali-ACC see-PST-1SG think-sc pres-2pl
 ‘You think that I saw Ali.’
 b. *Siz ben-i Ali-yi gör-dü san-iyor-sunuz.*
 You(PL) me-ACC Ali-ACC see-PST believe-PRES-2PL
 ‘You believe me to have seen Ali.’

In these examples, the differences between Turkish and Uzbek are readily apparent. The first obvious difference is the lack of an overt complementizer in either example. The form of the predicate is also quite different in Turkish, as verbal predicates are allowed and tense may be expressed in these predicates. Note, however, that there is no person marking in the second example where ECM has occurred. Other similarities between the two forms of ECM in these languages are the ability of the accusative nominal to become the matrix subject under passivization and to scramble to certain positions within the matrix clause. Zidani-Eroğlu ascribes the characteristics of the ECM examples to the raising of the subject of the embedded clause to the position of the object in the matrix clause, where it is able to receive accusative case and to behave less like a subject and more like an object.

3.2 ECM in Japanese

The differences between ECM in Turkish and Uzbek make it difficult to make any structural claims about either language in relation to the other. For that reason, we turn now to Japanese, which shares some properties with Uzbek that Turkish does not.

Tanaka (2002) provides a description of Japanese ECM that shows some remarkable similarities to Uzbek. Just as in Uzbek, Japanese ECM constructions are headed by an overt complementizer and the embedded predicates in these constructions may not be verbal.

- (16) *John-ga Bill-o baka-da-to omot-teiru.*
 John-NOM Bill-ACC fool-COP-COMP think-PROG
 ‘John thinks of Bill as a fool.’

Given the similarities between Uzbek and Japanese, it is worthwhile to consider Tanaka’s analysis to gain insight into the best way to analyze Uzbek. According to Tanaka’s account, the major problem with analyzing ECM in Japanese is the presence of the overt complementizer, which appears to be a member of the class C. Per Chomsky (2001), any constituent that is to move out of an embedded clause must pass through the “escape hatch” of Spec,CP. Accordingly, exceptional case marking should be impossible when there exists an overt complementizer, as this should prevent A-movement and case assignment from occurring between clausal layers. Only \bar{A} -movement should be possible when a complementizer is present. As a solution to this problem, Tanaka proposes that in Japanese ECM constructions, the complementizer differs from ordinary complementizers in that its edge is licensed as an A-position, resulting in the following derivation:

- (17) *John-ga* [_{VP} *Bill-o_i* [_{CP} *t’_i* [_{TP} *t_i baka-da*]-to] *omot-teiru*].
 John-NOM [_{VP} *Bill-ACC_i* [_{CP} *t’_i* [_{TP} *t_i fool-COP*]-COMP] think-PROG
 ‘John thinks of Bill as a fool.’

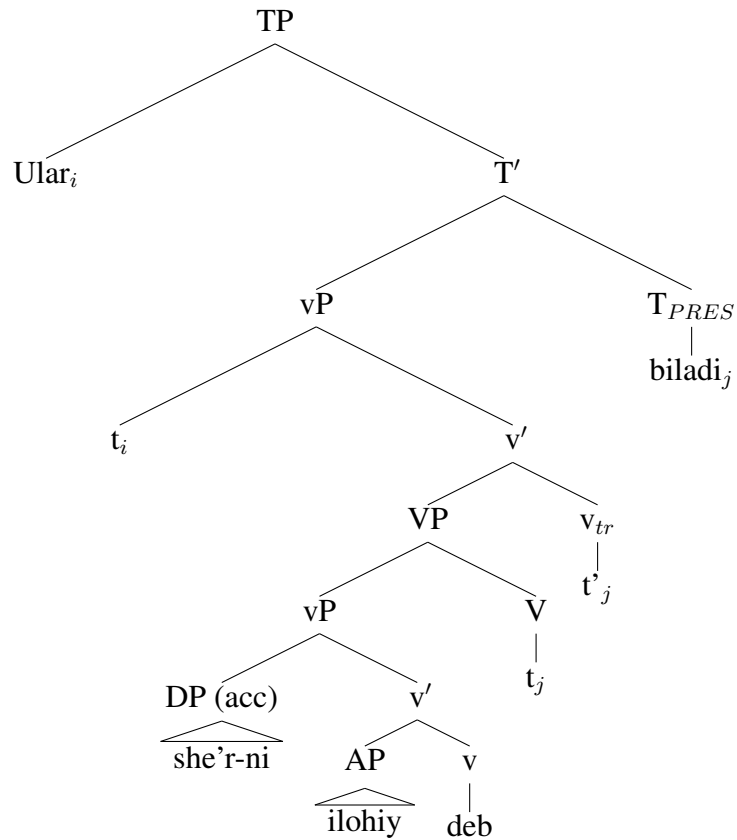
Tanaka’s analysis is problematic for a number of reasons, however. The core argument that certain complementizers license an A-position at their edge is undesirable, even problematic. Although other complementizers in Japanese appear to allow for similar movement (see, for example, Uchibori (2001)), for Uzbek, such a solution would be insufficient, as there is no supporting evidence for an A-position at Spec,CP. Moreover, this solution cannot account for the other peculiarities of ECM in Uzbek, such as the lack of verbal predicates in the embedded clause.

4 An Analysis for Uzbek

At the core of the analysis proposed here is that what appears to be a complementizer in Uzbek is not, in fact a true complementizer, that is, a member of the class C. Although the complementizer *deb* is form-identical with the item heading the ECM constituent, the behavior of the material beneath these items is so different that they must be considered two different morphemes. I propose, instead, that the *deb* of the

ECM construction is merely the overt head of the lowest projection above the predicate, likely vP. Under this analysis, the derivation for the first example (repeated here), is as follows:

- (18) *Ular she'r-ni ilohiy deb bil-a-di.*
 they poem-ACC sacred COMP know-sc pres-3
 'They know the poem to be sacred.'



A major benefit of this analysis is the ability to explain why verbal predicates are not allowed. Verbal predicates in Uzbek are incomplete without some sort of marker of tense or aspect, neither of which is available in this construction. So-called non-finite verb forms in Uzbek are not, in fact, non-finite, but consist of the verb root plus forms bearing aspectual or modal information (Bodrogligeti 2003); the location of these non-finite morphemes at higher points on the tree is evidenced by the ability of non-finite forms to be marked for voice, which under this analysis, is located at the same position as the pseudo-complementizer *deb*.

- (19) a. *qilib* 'having done'
 b. *qilinib* 'having been done'
- (20) a. *qilmoq* 'to do'
 b. *qilinmoq* 'to be done'

Form	Usage	Aspectual/Modal Meaning
-(i)b	Gerund/Converbial	perfective
-a/-y	Gerund/Converbial	imperfective
-gan	Participle	perfective
-moq	Infinitive	irrealis

Table 2: Non-finite verb forms in Uzbek

Further evidence for the aspectual/modal nature of “non-finite” verbal markers in Uzbek may be shown by the use of these forms as predicates. In certain cases, tense may be added to these forms, suggesting that they occupy a space between voice and tense.

- (21) *U tush-gan.*
 He fall-PERF
 ‘He has fallen.’
- (22) *Biz kitob-ni yoz-ib e-di-k.*
 We book-ACC write-PERF COP-PST-3PL
 ‘We had written the book.’

Accordingly, non-finite verb forms may participate in related constructions, such as control. In ECM constructions, however, the presence of *deb* as the voice head prevents verbs from becoming morphologically complete. There there is no such restriction on nouns or adjectives acting as predicates¹.

It is not my desire here to fully discuss the assignment of the accusative case to the lower subject; any number of theories of case may now account for its presence, although they may require the presence of a case head or the movement of the nominal in question to an AGR projection. However, the absence of CP in these constructions makes the presence of case here non-problematic.

5 Conclusions

The account proposed here for ECM in Uzbek is perhaps not so surprising; in many ways it is similar to the one proposed for certain instances of ECM in English. What is surprising, however, is the presence of the overt head *deb*. The lower ‘clause’ in these ECM constructions may be considered a small clause, that is, the linkage of a nominal to a predicate without the intervention of tense or, indeed, any verbal category. It appears to be rare crosslinguistically that small clauses possess overt heads.

Moreover, the solution proposed here, in which there is no true complementizer, likely has implications for the analysis of other languages. Remember that in Turkish, ECM constructions are remarkably similar to regular complement clause constructions. The difference between these two constructions is likely that the complement clause construction possesses a covert complementizer, whereas this

¹Thanks to Thomas Grano for pointing out the logical possibility of non-finite verb forms in ECM constructions.

complementizer is wholly absent in ECM constructions. Japanese appears to be rather more complicated, as some complementizers genuinely do appear to allow A-movement across them. For the examples detailed in Tanaka (2002) and reproduced here, however, the solution proposed here may also work.

Further research into the typological properties of both small clauses and exceptional case marking may reveal more phenomena similar to those found in Uzbek, furthering our understanding of these unusual constructions as a whole.

References

- Bodrogligeti, A. J. E. 2003. *An Academic Reference Grammar of Modern Literary Uzbek*. Muenchen: Lincom.
- Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, ed. by M. Kenstowicz, 1–25. Cambridge: MIT Press.
- Tanaka, H. 2002. Raising to object out of CP. *Linguistic Inquiry* 33.637–652.
- Uchibori, A. 2001. Raising out of CP and C-T relations. *Formal Approaches to Japanese Linguistics* 3.
- Zidani-Eroğlu, L. 1997. Exceptionally case-marked NPs as matrix objects. *Linguistic Inquiry* 28.219–230.

