



11-18-2022

Differential Subject Marking in Kazakh

Eszter Ótott-Kovács
Cornell University

Follow this and additional works at: <https://repository.upenn.edu/pwpl>

Recommended Citation

Ótott-Kovács, Eszter (2022) "Differential Subject Marking in Kazakh," *University of Pennsylvania Working Papers in Linguistics*: Vol. 28: Iss. 1, Article 17.

Available at: <https://repository.upenn.edu/pwpl/vol28/iss1/17>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/pwpl/vol28/iss1/17>
For more information, please contact repository@pobox.upenn.edu.

Differential Subject Marking in Kazakh

Abstract

This paper investigates the distinction between nominative and genitive subject complement clauses in Kazakh (Turkic) to determine whether these clauses have the same syntactic structure or whether they are derived independently using different strategies. Based on novel data, the paper shows that Kazakh morphologically distinguishes anaphoric and unique definiteness (in the sense of Schwarz 2009), and that genitive marking on the subject of complement clauses is determined by anaphoricity of the subject. Therefore, nominative and genitive subjects are in complementary distribution: anaphoric subject DPs are in the genitive, while pseudo-incorporated and unique definite subjects are nominative-marked. This distinction can be accounted for by analyzing the anaphoric subject undergoing a semantically motivated movement to the edge of the complement clause, where it is assigned lexically-governed genitive. The analysis contributes to the cross-linguistic study of morphological case assignment by showing that genitive can be a lexical case, and also by demonstrating that definiteness marking can drive differential case marking.

Differential Subject Marking in Kazakh

Eszter Ótott-Kovács*

1 Introduction

This paper explores the topic of differential subject case marking in Kazakh complement clauses. Variable case assignment has drawn a lot of interest because it provides an excellent testing ground for both linguistic theories of case assignment mechanisms (Baker and Vinokurova 2010, Baker 2015, Levin and Preminger 2015), and for approaches to the architecture of syntactic structure, in particular, phases and syntactic derivation (Gribanova 2018/2019, Bondarenko and Davis 2021). This paper contributes to this growing area of research by presenting novel data relating to subject case alternation in Kazakh.

Kazakh, a Turkic language spoken by about 13 million speakers in Kazakhstan and its neighboring countries, is an agglutinative language with SOV basic word order. Complement clauses serve as arguments for the matrix predicate, and they are case-marked according to their syntactic position in the matrix clause. In (1), the complement clause (in square brackets) is the object of the matrix predicate *esti-* ‘to hear’, and consequently it is in the accusative.¹ The complement clause predicate bears an aspectual suffix, which can be spelt out by *-y/AtIn* (prospective) or *-GAn* (non-prospective), an optional nominalizer in *-LIQ* (its allomorph is *-dig* in (1)), and an obligatory subject agreement marker, which is identical in form with the possessor-agreement suffixes found on possessives. Given the presence of an overt nominalizer and possessive agreement suffixes, complement clauses are widely analyzed as mixed projections where nominal projections embed a verbal core (Borsley and Kornfilt 1999, Kornfilt 2003, Kornfilt and Whitman 2011 for Turkish).

- | | | | | |
|--------|--|-------------|---------------------------|--------------|
| (1) a. | [Agaj-Ø | Tyrkija-dan | kel-gen-(dig)-in] | esti-di-m. |
| | [Professor ² -NOM | Turkey-ABL | come-PRF-(NMLZ)-3POSS]ACC | hear-PST-1SG |
| b. | [Agaj- duŋ | Tyrkija-dan | kel-gen-(dig)-in] | esti-di-m. |
| | [Professor-GEN | Turkey-ABL | come-PRF-(NMLZ)-3POSS]ACC | hear-PST-1SG |
| | ‘I heard that the Professor came from Turkey.’ | | | |
| c. | *[Agaj- du | Tyrkija-dan | kel-gen-(dig)-in] | esti-di-m. |
| | [Professor-ACC | Turkey-ABL | come-PRF-(NMLZ)-3POSS]ACC | hear-PST-1SG |

The paper focuses on the subject case marking in complement clauses, such as in (1), where the embedded clause subject, *agaj* ‘older man’, can either be nominative (phonologically zero) or genitive (*-Nly*). It has been noted in the literature that certain Turkic languages (e.g., Balkar (Bondarenko and Davis 2021)) also allow accusative marking on the complement clause subject. This is not possible in Kazakh, illustrated by the ill-formed (1c).

Variable subject case marking in complement clauses is well-attested across the Turkic language family (Kornfilt 2009, Baker & Vinokurova 2010, Gribanova 2018/2019, Bondarenko and Davis 2021). It is, however, less clear what the relationship is between the subject case marking patterns in these languages, for instance whether genitive (in (1b)) and nominative subject (in (1a)) complement clauses can be derived via one unified analysis, or whether they are the result of independent parallel complement clause forming strategies. To answer this question, one needs to investigate if there is any difference between differentially subject case marked complement

*This project is the result of my fieldwork in Kazakhstan, as well as follow-up elicitations with native speaker consultants. I am deeply indebted to all of my consultants who gave their time and attention to answer my questions. I am extremely grateful to Arailym Raikhankyzy and Ainur Bayekeyeva for their great observations and insights about the data. This paper would have been impossible without them. I am also thankful to Miloje Despić, the members of Cornell SynCirc, and the PLC audience for their invaluable comments. All errors are mine.

¹The possessive suffix has an allomorph ending in *n* (i.e., *-(s)In*) when it precedes the accusative, the accusative is spelt out by a phonologically zero allomorph in this context.

²*Agaj* literally means ‘older male’. *Agaj* (in example (1)) denotes an individual not a property; it stands in for the name of the individual in the context.

clauses. With the exception of Turkish, where the distinction between nominative and genitive subjects appears to be straightforwardly related to specificity (nominative subjects are pseudo-incorporated and interpreted as non-specific, while genitive subjects are specific and move to a higher position) (Kornfilt 2009), the difference between complement clauses with differential subject marking in other Turkic languages remains elusive. Upon casual inquiries, speakers report no major distinction between the meanings or uses of (1a) and (1b). This has led researchers to treat the underlying syntactic structure of nominative and genitive subject complement clauses as *de facto* independently built, as if there were two different strategies available in the language, and the distinction between nominative versus genitive were, effectively, accidental. Based on novel data elicited from native speakers, this paper argues against this treatment of differentially case marked complement clauses, showing that the nominative and genitive subjects are, in fact, in complementary distribution. Consequently, there must be a unified analysis that can account for both nominative and genitive subject complement clauses.

The novel observation this paper presents is that anaphoric definite complement clause subjects are genitive-marked, and in all other instances (with unique definite, and non-specific subjects) nominative is used. Section 2.1 introduces the main theoretical claims behind the two types of definiteness, uniqueness and anaphoricity. Section 2.2 shows, based on novel Kazakh data, that the unique-versus-anaphoric definiteness distinction is morphologically distinguished in the language. Then section 3 proceeds to present data relating to differential subject marking in complement clauses, and demonstrates that differential subject marking is driven by anaphoricity in Kazakh. Section 4 provides an outline of an analysis based on the novel empirical data, and argues that nominative and genitive subject complement clauses have the same syntactic structure, the difference being that the anaphoric subjects move to a higher position where they get genitive case.

2 Uniqueness and Anaphoricity

2.1 Uniqueness and Anaphoricity in a Nutshell

The study of definite descriptions goes back to the very beginnings of modern semantics (Frege 1892, Russell 1905, Strawson 1950, Heim 1982, Roberts 2003, among many others). A recurring theme in the debates surrounding the semantics of definite descriptions has been whether they are better characterized in terms of uniqueness or anaphoricity. This paper picks up the thread with Schwarz's 2009 work showing that languages make use of both unique and anaphoric definiteness, and some languages morphologically distinguish them. One such language is German where the definite article following a preposition can be spelled out differently depending on whether the noun phrase is interpreted as a unique or anaphoric definite (Schwarz 2009). Similar morphological distinctions have since been reported for several languages, such as Thai (Jenks 2015), Mandarin (Jenks 2018), Akan (Arkoh & Matthewson 2013) among others (for a typological overview of definiteness marking see Moroney 2021).

Schwarz 2009 distinguishes different types of definite uses, summarized in Table 1. Unique definites are identifiable for the discourse participants from pre-existing world knowledge (global uniqueness), from the larger situational setting (situational uniqueness), or based on knowledge about the unique part(s) of some entity (part-whole bridging). The referents of anaphoric definites are discourse-old: they are identifiable based on the preceding discourse context. Schwarz recognizes three subtypes of anaphoricity: anaphoric definites, cases of producer-product bridging, and donkey anaphoras.

Uniqueness	Global uniqueness	the president of Russia, the Queen of England
	Situational uniqueness	the desk (when there is only one desk in the room)
	Part-whole bridging	I went to a mosque. The roof was old.
Anaphoricity	Anaphoric	A girl and a boy entered the

		room. The boy was crying.
	Producer-product bridging	Aisha watched an opera yesterday. The librettist was Auezov.
	Donkey anaphora	Every farmer that has a donkey beats the donkey .

Table 1: Types of definite uses (based on Schwarz 2009).

2.2 Uniqueness and Anaphoricity in Kazakh

Definite descriptions in Kazakh have not been the subject of linguistic studies. The most well-studied member of the Turkic language family in this respect is Turkish, which lacks definite articles and expresses definiteness using bare nouns. This has led to a debate in the literature about whether Turkish lacks a DP projection and utilizes a type-shifting operation to express definiteness (most prominently advocated for by Bošković & Şener 2014), or whether the language has a phonologically zero D head (Lyutikova & Pereltsvaig 2015, *inter alia*). This paper is not directly concerned with this debate. Rather, our main goal is to draw attention to the fact that the morphological forms of different definiteness types are understudied both descriptively and theoretically in these languages, and to show that Kazakh morphologically distinguishes two types of definiteness.

Kazakh, similarly to Turkish, does not have definite articles. But in Kazakh, unlike in Turkish, not every type of definiteness can be expressed by bare nouns. As demonstrated below with specific examples, bare nouns can serve as unique definite descriptions, anaphoricity cannot be expressed by a bare noun, and a demonstrative has to be used.

Global and situational uniqueness can be conveyed by a bare noun. In (2), the global unique noun phrase ‘the Russian president’ is a bare noun, while (3) illustrates a situationally unique definite, ‘the table’, spelt out by a bare noun. Global and situational unique definite descriptions are expressed by bare nouns regardless of the grammatical function of the noun phrase, *i.e.*, in subject, object, etc. positions.

- (2) Aisha asks: What’s in the news these days? Aigul says:

Rossija-nuñ **prezident-i** Astana-ga kel-di.
Russia-GEN **president-3POSS** Astana-DAT come-3PST
 ‘The Russian president visited Astana.’

- (3) Aisha and Aigul are in an office, where there is only one desk. They’re waiting in silence, but Aisha suddenly turns to Aigul and whispers:

Stol æte kumbat eken.
desk very expensive COP.3EVID
 ‘I heard that the desk was very expensive.’

Part-whole bridging examples pattern differently from the above-mentioned unique definites: bare nouns are marginal in this context (as in (4a)), and some other strategy is utilized to convey definiteness. My consultants volunteered possessive marking on the noun phrase, as shown in (4b). This is not surprising, given that it is cross-linguistically attested that part-whole bridging constructions are marked by possessive morphology (*e.g.*, in Güilá Zapotec (Arrieta-Zamudia 2021), and in Ch’ol (Carol-Rose Little, *p.c.*)).

- (4) Aisha is telling a story: Yesterday I visited the mosque on Baitursynov St. for the first time...

a. ?**Tæbe** æte kœne eken.
roof very old COP.3EVID
 b. **Tæbe-si** æte kœne eken.
roof-3POSS very old COP.3EVID
 ‘I saw that its roof is very old.’

In contrast, anaphoricity cannot be expressed by bare nouns. Anaphoric definite noun phrases in object position must be marked with a demonstrative. In (5), the anaphoric definite ‘the boy’ cannot be expressed by the bare noun, as in (5a). Instead, the demonstrative needs to be used in this context, illustrated in (5b).

- (5) Aisha is telling a story: A girl and a boy entered the room...
- | | | | |
|----------------------|-----------|-------------|-------------|
| a. # Bala-nu | kefe | teatr-da | kœr-di-m. |
| boy-ACC | yesterday | theatre-LOC | see-PST-1SG |
| b. Ol bala-nu | kefe | teatr-da | kœr-di-m. |
| that boy-ACC | yesterday | theatre-LOC | see-PST-1SG |
- ‘I saw the boy in the theatre yesterday.’

The demonstrative must be used with anaphoric noun phrases in all syntactic positions except in subject position. As demonstrated in (6), when the anaphoric DP is the subject, it can be a bare noun (as in (6a)), or it can be marked by a demonstrative (as in (6b)). The exceptional behavior of anaphoric definite descriptions in subject position is not surprising: the exact same patterns have been described for Mandarin as well, where unique definites are bare nouns while anaphoricity is expressed by a demonstrative, except in subject position where bare nouns are allowed to serve as anaphoric definite descriptions (Jenks 2018). Jenks accounts for the exceptionality of subject position by appealing to the discourse function of DPs in subject position: these are continuing topics whose pragmatic function eliminates the requirement of using the demonstrative (for the formal implementation of this analysis see Jenks 2018).

- (6) Aisha is telling a story: A girl and a boy entered the room...
- | | | | |
|-------------------|-----|----------|-----------|
| a. Bala | on | zas-ta | eken. |
| boy | ten | year-LOC | COP.3EVID |
| b. Ol bala | on | zas-ta | eken. |
| that boy | ten | year-LOC | COP.3EVID |
- ‘The boy is ten years old.’

Producer-product bringing examples, such as (7), cannot be constructed with either bare nouns or demonstratives, so some other strategy needs to be used. My consultants suggested using a relative clause in this context, as in (7b).

- (7) Aisha is telling a story: I went to the theater yesterday. I really liked the opera...
- | | | |
|------------------------|--------|-----------|
| a. #(Ol) ƶazwƶu | Æwezov | eken. |
| (that) writer | Auezov | COP.3EVID |
- ‘The writer is Auezov.’
- | | | |
|----------------------------|--------|-----------|
| b. Opera-nu ƶaz-gan | Æwezov | eken. |
| opera-ACC write-RC | Auezov | COP.3EVID |
- ‘(The one) who wrote the opera is Auezov.’

Finally, donkey anaphoras seem to be exceptional in that they allow bare nouns (8b), demonstratives (8c), or possessive morphology on the DP (8a). While the possessive strategy appears to be the most natural way to express the donkey anaphora, (8b) and (8c) are also acceptable. As donkey anaphora is not central to this paper, I leave it for future research to explain why both bare nouns and demonstratives are acceptable in this context.

- (8) a. Eseg-i bar ærbir fermer **eseg-in** ur-adu.
 donkey-3POSS COP.RC every farmer **donkey-3POSS.ACC** beat-3PRS
 ‘Every farmer who owns a donkey beats his/her donkey.’
- b. Eseg-i bar ærbir fermer **esek-ti** ur-adu.
 donkey-3POSS COP.RC every farmer **donkey-ACC** beat-3PRS
- c. Eseg-i bar ærbir fermer **ol esek-ti** ur-adu.
 donkey-3POSS COP.RC every farmer **that donkey-ACC** beat-3PRS
 ‘Every farmer that owns a donkey beats the donkey.’

Table 2 offers a summary of the Kazakh definiteness marking patterns.

Uniqueness	Global uniqueness	bare noun
	Situational uniqueness	bare noun
	Part-whole bridging	?bare noun, other strategy
Anaphoricity	Anaphoric	subject: bare noun, demonstrative non-subject: *bare noun, demonstrative
	Producer-product bridging	*bare noun, other strategy
	Donkey anaphora	other strategy, bare noun, demonstrative

Table 2: Types of definiteness marking in Kazakh.

3 Subject Case in Complement Clauses

After establishing that Kazakh morphologically distinguishes different types of definiteness, this section turns to the discussion of subject case in complement clauses. The novel observation I present here is that genitive subject case marking is only available on anaphoric definite DPs, while nominative is used elsewhere, such as on unique definite DPs and with pseudo-incorporated subjects.

3.1 Non-Specific and Unique Definite Subjects

The external argument can have a non-definite, number neutral interpretation if the subject is situated in the immediate pre-verbal position, i.e., nothing (e.g., adverbs, direct or indirect objects, etc.) can intervene between them and the verb phrase. Their position suggests that these subjects do not move up to the canonical subject position, to Spec,TP. I take this noun phrase to be pseudo-incorporated (following Öztürk's 2005:32-50 analysis for Turkish). Not surprisingly, pseudo-incorporated NPs cannot be marked with the genitive, as illustrated in (9), where the sentences with nominative (in (9a)) and genitive subjects (in (9b)) are uttered in an out-of-the-blue context. The intended interpretation of the subject is 'some doctor(s)', i.e., non-specific and number neutral. Only nominative can be used felicitously under this interpretation: the genitive-marked subject expresses anaphoric definite meaning, i.e., 'the doctor (who we talked about before)'.

(9) Aisha asks: What's new these days? Aigul says:

- a. [Kala-muuz-ga dæriger-Ø kel-etin-in] esti-di-m.
 [town-1PL.POSS-DAT doctor-NOM come-PRSP-3POSS]ACC hear-PST.1SG
 'I heard that a doctor is going to come to our town.'
- b. #[Kala-muuz-ga dæriger-dij kel-etin-in] esti-di-m.
 [town-1PL.POSS-DAT doctor-GEN come-PRSP-3POSS]ACC hear-PST.1SG

While it does not come as a surprise that pseudo-incorporated subjects are in the nominative, it is striking that certain types of definite DPs can be in the nominative as well. The subjects in (10)-(12) have a unique definite referent, and they require nominative marking (with the caveat that part-whole bridging is compatible with the genitive as well).

The complement clauses in (10) have a global unique definite subject, 'Russia's president', which is discourse-new, as it has not been previously mentioned in the given context. The use of genitive subject case results in infelicity in this context.

(10) Aisha asks: What's in the news these days? Aigul says:

- a. [Rossija-nuñ prezident-i-Ø Astana-ga kel-gen-in] esti-di-m.
 [Russia-GEN president-3POSS-NOM Astana-DAT come-PRF-3POSS]ACC hear-PST-1SG
 'I heard that the Russian president visited Astana.'

- b. #[Rossija-nuŋ prezident-i-niŋ Astana-ga kel-gen-in] esti-di-m.
 [Russia-GEN president-3POSS-GEN Astana-DAT come-PRF-3POSS]ACC hear-PST-1SG

We see the same pattern in the case of situationally unique subjects: in (11), there is a discourse-new subject, ‘the desk’, which is unique in the situation described in the context (there is only one desk in the room). This situational unique definite DP patterns identically as the global unique subject in (10): situational unique definite DPs are also incompatible with genitive subject case marking in complement clauses.

- (11) Aisha and Aigul are in an office, where there is only one desk. They’re waiting in silence, but Aisha suddenly turns to Aigul and whispers:
- a. [Stol-Ø œte kumbat eken-in] esti-di-m.
 [desk-NOM very expensive COP.PRF-3POSS]ACC hear-PST-1SG
 ‘I heard that the desk was very expensive.’
- b. #[Stol-duŋ œte kumbat eken-in] esti-di-m.
 [desk-GEN very expensive COP.PRF-3POSS]ACC hear-PST-1SG

Part-whole bridging subjects are yet again outliers in two ways: they require possessive marking, and they do not pattern with global and situational unique definite complement clause subjects in that they allow nominative *as well as* genitive case marking. In (12), the complement clause subject is the definite description ‘the roof’, which is a unique part of a contextually salient entity. Just as in other part-whole bridging constructions (see (4)), possessive marking is obligatory, illustrated by the infelicitous version without the possessive in (12c). With the possessive-marked bridging subject, both nominative (in (12a)) and genitive (in (12b)) case marking is allowed, although my consultants expressed a preference towards the genitive-marked version.

- (12) Aisha is telling a story: Yesterday I visited the mosque on Baitursynov St. for the first time...
- a. [Tœbe-si-Ø œte kœne eken-in] kœr-di-m.
 [roof-3POSS-NOM very old COP.PRF-3POSS]ACC see-PST-1SG
 ‘I saw that its roof was very old.’
- b. [Tœbe-si-niŋ œte kœne eken-in] kœr-di-m.
 [roof-3POSS-GEN very old COP.PRF-3POSS]ACC see-PST-1SG
- c. #[Tœbe-(niŋ) œte kœne eken-in] kœr-di-m.
 [roof-GEN/NOM very old COP.PRF-3POSS]ACC see-PST-1SG

The subject case marking patterns on pseudo-incorporated and definite subjects are summarized in Table 3. In short, these subjects, with the exception of part-whole bridging, can only be marked with the nominative. Further study is required regarding how possessive marking (on part-whole bridging definites) interacts with subject case marking.

Non-specific subject	NOM	*GEN
Global unique subject	NOM	*GEN
Situational unique subject	NOM	*GEN
Part-whole bridging subject	(POSS+) NOM	(POSS+) GEN

Table 3: Case marking on pseudo-incorporated and unique definite complement clause subjects.

3.2 Anaphoric Definite Subjects

In contrast to unique definite subjects, anaphoric definite subjects only allow genitive marking. In sentence (13) the complement clause subject is anaphoric definite, as ‘the boy’ was mentioned in the preceding context. Nominative subject case marking is disallowed on such anaphoric definite DPs, so genitive is the only available option.

- (13) Aisha is telling a story: A girl and a boy entered the room...

- a. #[Bala-Ø ʒula-gan-uun] kœr-di-m.
 [boy-NOM cry-PRF-3POSS]ACC see-PST-1SG
- b. [Bala-nuŋ ʒula-gan-uun] kœr-di-m.
 [boy-GEN cry-PRF-3POSS]ACC see-PST-1SG
 ‘I saw that the boy was crying.’

Anaphoric definite subjects of the producer-product bridging type pattern the same way: they disallow nominative case marking, as in (14a), and only genitive (in (14b)) can be used felicitously in this context.

- (14) Aisha is telling a story: I went to the theater yesterday. I really liked the opera...
- a. #[ʒazwʃu-Ø Æwezov bol-gan-uun] esti-di-m.
 [writer-NOM Auezov be-PRF-3POSS]ACC hear-PST-1SG
- b. [ʒazwʃu-nuŋ Æwezov bol-gan-uun] esti-di-m.
 [writer-GEN Auezov be-PRF-3POSS]ACC hear-PST-1SG
 ‘I heard that the librettist was Auezov.’

The available case marking patterns with anaphoric subjects are summarized in Table 4. The donkey anaphora is not constructible in the complement clause subject position.

Anaphoric subject	*NOM	GEN
Producer-product subject	*NOM	GEN
Donkey anaphora	(non-applicable)	

Table 4: Case marking on anaphoric definite complement clause subjects.

4 Towards an Analysis

The takeaway from the data presented in the preceding section is that genitive and nominative-subject complement clauses are in complementary distribution. Consequently, it would be undesirable to analyze these as two independent complement clause forming strategies, as it would remain unexplained why two independently formed complement clauses are used in exactly those configurations where the other one is disallowed. First, I take a closer look at the syntactic position of nominative and genitive complement clause subjects. I show that negative polarity item (henceforth, NPI) licensing and scrambling facts indicate that the genitive subject is in a higher syntactic position than the nominative subject.

In Kazakh, NPI items, formed with *eʃ-*, are licensed under the scope of clause-mate negation, as in (15a), or at the clause-edge by clause-external negation in line with the Phase Impenetrability Condition (Chomsky 2001, a.o.). If there is no available licenser in the relevant domain, NPIs are disallowed, illustrated in (15b).

- (15) a. **Eʃkim** ʒumʉs iste-gen **ʒok**.
 NPI.who work do-PRF NEG
 ‘No one has worked.’
- b. ***Eʃkim** ʒumʉs iste-gen.
 NPI.who work do-PRF

Matrix clause negation can only license NPI subjects in the embedded clauses if it is in the genitive, as in (16b). The nominative NPI subject in (16a) cannot be licensed by matrix negation. In keeping with the Phase Impenetrability Condition, it is expected that the edge of the embedded clause is available for matrix clause operators. The availability of NPI licensing with genitive subjects suggests that these subjects are at the edge of the complement clause. Nominative subjects, on the other hand, are in a lower position, and thus they are unavailable for NPI licensing by a matrix clause licenser.

- (16) a. *[E]kim-Ø zımuus iste-gen-in] kœr-me-di-m.
 [NPI.who-NOM work do-PRF-3POSS]ACC see-NEG-PST-1SG
 b. [E]kim-nıı zımuus iste-gen-in] kœr-me-di-m.
 [NPI.who-GEN work do-PRF-3POSS]ACC see-NEG-PST-1SG
 ‘I didn’t see anyone working.’

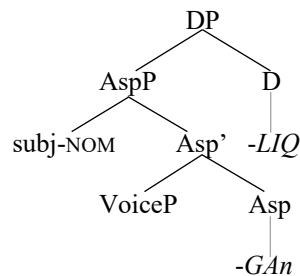
Further support for the positional difference between genitive and nominative subjects comes from scrambling: long-distance scrambling from embedded to matrix clause is only available with genitive but not with nominative subjects. In (17a), the nominative embedded subject cannot scramble to the matrix clause, while the genitive subject in (17b) can.

- (17) a. *Konak-Ø_i Ajfa [t_i Astana-ga bar-gan-un] esti-di.
 guest-NOM_i Aisha [t_i Astana-DAT go-PRF-3POSS]ACC hear-3PST
 b. Konak-tuıı_i Ajfa [t_i Astana-ga bar-gan-un] esti-di.
 guest-GEN_i Aisha [t_i Astana-DAT go-PRF-3POSS]ACC hear-3PST
 ‘Aisha heard that the guest went TO ASTANA.’

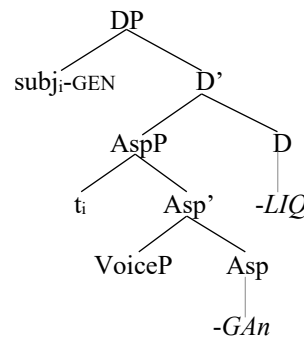
Under the assumption that scrambling takes place in a successive-cyclic fashion, scrambling must proceed through the clause edge. That is, only the phrase positioned at the edge of the embedded clause can move to the matrix clause. As (17b) demonstrates, scrambling is permitted for the genitive subject, indicating that it is at the edge of the complement clause. In contrast, the nominative subject cannot undergo long-distance scrambling, which supports that it is in a lower position.

Following the “mixed extended projection” analysis of nominalizations (Borsley and Kornfilt 1999, Kornfilt and Whitman 2011), I assume that nominal projections embed a verbal core in Kazakh complement clauses. I follow Kornfilt and Whitman 2011 and Gribanova 2018/2019 in analyzing the verbal core of Kazakh complement clauses as AspP, which is selected for by the nominal D head, which can be (optionally) spelt out by *-LIQ*. This is not a trivial choice, but due to space limitations, I refer the reader to the mentioned references for justifying these labels. Based on the NPI licensing and scrambling data, the genitive subject is at the edge of the embedded clause, in Spec,DP, shown in (17b). The nominative subject, conversely, is in the lower Spec,AspP position, given in (17a).

(17) a.



b.



Similar conclusions about the height of the subject have been drawn for other Turkic languages as well: Gribanova 2018/2019 describes the same distribution of nominative and genitive subjects in Uzbek complement clauses, and concludes that genitive subjects are in a higher syntactic position than nominative subjects. Bondarenko and Davis 2021 reach a similar conclusion about accusative and nominative subjects in Balkar. The novel claim this paper puts forth, based on the novel data in Section 3, is that nominative and genitive subjects are in complementary distribution, and thus it is both necessary and justifiable to derive them by one unified analysis. Specifically, it would be undesirable to assume that genitive subject clauses have the structure in (17b), whereas nominative subject complement clauses have a distinct syntactic structure with a CP projection below the DP, as in (18), to block genitive assignment to the subject. While this analysis could account for how nominative can surface in the c-command domain of the D head, it would fail to explain what

drives the complementary distribution between nominative and genitive subject clauses. Under such an analysis, the complementary distribution would be, in effect, accidental.

(18) [DP D [CP C [AspP/TP subject-NOM Asp'/T' [...]]]]

Now, we are left with the question of how to derive the nominative case in the configuration in (17a). Naturally, genitive case originates from the D head, but D cannot assign case via Agree, as nominative subjects would then always be blocked, contrary to fact. A configurational approach to case evaluation would be unfeasible for similar reasons: D is in the same phasal domain as the nominative subject, thus the lower subject should get unmarked (genitive) case. Relatedly, Agree between D and the subject DP cannot drive movement to Spec,DP, because nominative subjects, which are also in agreement relation with D, remain in a lower position.

A potential solution to this dilemma is that subject DPs that have anaphoric definite interpretation must move to Spec,DP as a *semantically* motivated movement. This is not an inconceivable motivation for movement, as Spec,DP is associated with anaphoricity on independent basis. After the subject DP moves to Spec,DP, it receives genitive as a lexically-governed case by the D head. That is, genitive in the language is not an instance of a case assigned via Agree or unmarked case, but a *lexical* case.

5 Conclusions

This paper investigated the distinction between nominative and genitive subject complement clauses in Kazakh to determine if they are generatable by one unified analysis or whether they are the products of independent strategies. Novel data, presented in Section 3, show that Kazakh nominative and genitive complement clauses are in complementary distribution: the genitive-nominative distinction is driven by anaphoricity of the subject DP, since anaphoric subjects are obligatorily genitive-marked and all other subjects (non-specific, unique definite) must be nominative. For simplicity of analysis, it is better to analyze nominative and genitive subject complement clauses as having the same syntactic structure (with anaphoric DPs requiring further movement to the clause edge where they are assigned genitive as a lexically-governed case), and not as two independent syntactic configurations that happen to be in complementary distribution.

By showing that nominative and genitive subject complement clauses have the same syntactic structure, we can gain new insights into how case assignment operates in the language: genitive subject case is not assigned via Agree (or as an unmarked case), rather it is a lexically-governed case assigned under sisterhood. This contributes to the cross-linguistic study of morphological case assignment by showing that genitive can be a lexical case, and also by demonstrating that definiteness marking can drive differential case marking.

References

- Arkoh, Ruby, and Lisa Matthewson. 2013. A familiar definite article in Akan. *Lingua* 123:1-30.
- Arrieta-Zamudio, Ana Laura. 2021. Definiteness in San Pablo Güilá Zapotec. Ms., file:///Users/eo264/Downloads/WSCLA_Definiteness-in-SPGZ_Ana-Arrieta-Zamudio.pdf
- Baker, Mark C., and Nadya Vinokurova. 2010. Two modalities of Case assignment: Case in Sakha. *Natural Language and Linguistic Theory* 28:593-642.
- Baker, Mark C. 2015. *Case: Its Principles and its Parameters*. Cambridge: Cambridge University Press.
- Bondarenko, Tatiana, and Colin Davis. 2021. Insights about cyclic syntax from cross-clausal scrambling and subject case in Balkar. Ms., URL <https://ling.auf.net/lingbuzz/006214>
- Borsley, Robert D., and Jaklin Kornfilt. 1999. Mixed extended projections. In *The Nature and Function of Syntactic Categories*, ed. R. Borsley, 101-131. Leiden-Boston: Brill.
- Bošković, Željko, and Serkan Şener. 2014. The Turkish NP. In *Crosslinguistic Studies on Noun Phrase Structure and Reference*, ed. Hofherr, P. C., and A. Zribi-Hertz, 102-140. Leiden-Boston: Brill.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
- Frege, Gottlob. 1892. Über Sinn und Bedeutung [On sense and reference]. *Zeitschrift für Philosophie und philosophische Kritik* 100:25-50.

- Gribanova, Vera. 2018/2019. Case, agreement and differential subject marking across Uzbek nonfinite clause types. Paper presented at ConCall 3, Indiana University, Bloomington. (Revised in 2019).
- Heim, Irene. 1982. The Semantics of Definite and Indefinite Noun Phrases. Doctoral dissertation, University of Massachusetts, Amherst.
- Jenks, Peter. 2015. Two kinds of definites in numeral classifier languages. *Semantics and Linguistic Theory* 25:103-124.
- Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49(3):501-536.
- Kornfilt, Jaklin. 2003. Subject case in Turkish nominalized clauses. In *Syntactic Structures and Morphological Information*, ed. U. Junghanns and L. Szucsich, 129-215. Berlin and New York: Mouton de Gruyter.
- Kornfilt, Jaklin. 2009. DOM and two types of DSM in Turkish. In *Differential Subject Marking*, ed. H. de Hoop and P. de Swart, 79-111. Dordrecht: Springer.
- Kornfilt, Jaklin, and John Whitman. 2011. Afterword: Nominalizations in syntactic theory. *Lingua* 121(7): 1297-1313.
- Levin, Theodore, and Omer Preminger. 2015. Case in Sakha: Are two modalities really necessary?. *Natural Language & Linguistic Theory* 33(1):231-250.
- Lyutikova, Ekaterina, and Asya Pereltsvaig. 2015. The Tatar DP. *The Canadian Journal of Linguistics / La Revue Canadienne de Linguistique* 60(3):289-325.
- Moroney, Mary. 2021. Updating the typology of definiteness: Evidence from bare nouns in Shan. *Glossa: a Journal of General Linguistics* 6(1).
- Öztürk, Balkız. 2005. *Case, Referentiality and Phase Structure*. Amsterdam-Philadelphia: John Benjamins.
- Roberts, Craige. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26(3):287-350.
- Russell, Bertrand. 1905. On Denoting. *Mind* 14(56):479-493.
- Schwarz, Florian. 2009. Two types of Definites in Natural Language. Doctoral dissertation, University of Massachusetts, Amherst.
- Strawson, Peter F. 1950. On Referring. *Mind* 59(235):320-344

Department of Linguistics
Cornell University
Ithaca, NY 14850
eo264@cornell.edu