Elucidating Nominal Structure in Articleless Languages: A Case Study of Tatar

EKATERINA LYUTIKOVA1 & ASYA PERELTSVAIG2

Moscow State University1; Sholokhov Moscow State University for Humanities1; Stanford University2

1 Introduction

The central question addressed by this paper is whether languages without articles have the same highly articulated functional architecture in noun phrases, including the DP projection. Previous studies focused on Slavic languages (but see Bošković and Şener 2012 on Turkish); some scholars (cf. Progovac 1998, Engelhardt & Trugman 1998, Rutkowski 2002, inter alia) argued in favor of the DP projection, while Bošković (2005, 2008, 2009, 2010; Bošković and Şener 2012) argued against it. Pereltsvaig (2006, 2007, 2013) argued that while some nominals in Russian and other articleless Slavic languages are DPs, others are Small Nominals (SNs) of different sizes. In this paper, we provide novel evidence for the latter position based on another Turkic language, Tatar (spoken by over 5 million in Tatarstan, Russia). Drawing on our fieldwork on one sub-dialect of Tatar (spoken in the village of Kutlushkino), we show that different syntactic constructions call for nominals of different sizes. Moreover, we argue that Differential Object Marking (DOM) in Tatar—unlike in other Turkic languages such as Turkish or Sakha—can only be explained in terms of the amount of

1* Our many thanks to Pavel Graschenkov, Olga Kagan, Ora Matushansky, Sergei Tatevosov, Yakov Testelets for helpful discussions, comments, and suggestions. We are also grateful to our Tatar consultants for their invaluable help. All errors are solely ours. This research has been partially supported by Russian Scientific Foundation (РНФ, grant №14-18-03270 "Word order typology, syntax—prosody interface and information structure in the world's languages" at Sholokhov Moscow State University for Humanities), and Russian Foundation for Basic Research (РФФИ, grant №13-06-00884).
functional architecture in the object: DP objects receive structural (accusative) Case, as in (1a), while SNs (i.e. NPs or NumPs) remain Caseless, as in (1b).


Thus, we rule out alternative analyses based on distinct positions of accusative and unmarked objects or on the semantic interpretation of the object. More generally, we propose that only DPs must receive structural case while SNs are not subject to such Case licensing requirements and may remain morphologically Caseless.

The rest of this paper is organized as follows. In section 2, we show that nominals in Tatar come in different sizes: some are DPs and others are Small Nominals (in the sense of Pereltsvaig 2006). In section 3, we lay out our proposal for DOM in Tatar that accusative-marked (ACC) objects in Tatar are DPs, whereas unmarked objects are Small Nominals. In sections 4 and 5 we argue against the Semantic and Positional Alternatives, respectively. Section 6 concludes the paper.

2 Noun Phrases in Tatar Come in Different Sizes

We assume that the strict ordering of nominal suffixes reflects the order of functional projections in the noun phrase, following Baker’s Mirror Principle (Baker 1985). These suffixes include—in the order away from the noun root—the plural suffix -\textit{lar} in Num°, the ezafe-2 marker -\textit{i/-si} in Poss°, the ezafe-3 marker such as -\textit{im} in D° (note that ezafe-2 and ezafe-3 markers are incompatible with each other, for reasons that are outside the scope of this paper), and case suffix such as -\textit{a} for dative.\footnote{The various suffixes in Tatar are subject to vowel harmony, as well as occasional nasal and/or voicing assimilation.} The word bala-\textit{lar}im-a ‘to my children’ is analyzed as [KP -a [DP -im [NumP lar [NP [N bala-]]]]]. Although Tatar noun phrases are maximally KPs, not all noun phrases are fully projected. Various constructions involve nominals of different sizes: N°, NP, NumP, DP, or KP can be embedded in different constructions.

For instance, the nominal element in complex predicate constructions (CPCs), as in (2a), cannot be ACC-marked, as shown in (2b). Pronouns cannot serve as the nominal element in a CPC (2c); nor does it have room for the plural marker -\textit{lar} (2d) or modifiers of any kind (2e).
Let us now consider nominal (ezafe) constructions. There are three such constructions in Tatar, known as ezafe-1, ezafe-2, and ezafe-3. As we show below, ezafe-1 embeds a bare N°, ezafe-2 embeds a NumP, and ezafe-3 embeds a KP (or a genitive-marked DP).

The embedded element of ezafe-1, which is used typically for materials, is a bare noun, as in (3a). Like the nominal component in CPCs, it cannot be a pronoun (3b), nor can it contain the plural marker -lar (3c), or any modifiers (3d).

In contrast, the embedded element of ezafe-2 is a NumP: like the nominals considered above, it cannot be a pronoun (4a), nor a proper noun (4b). However, it can include the plural marker -lar (4c) or certain modifiers (4d). (The ezafe-2 marker is the suffix glossed here as ‘3’ for 3rd person; it does not show...
agreement.)

(4) a. *min däftär-em 
   I notebook-1SG 
   intended: ‘my notebook’

b. * zefär ata-si 
   (Grashchenkov 2007)
   Zufar dad-3 
   intended: ‘Zufar’s dad’

c. kırsak-li xatîn-nar kijem-e 
   belly-ATR woman-PL clothing-3 
   ‘clothing for pregnant women’

d. bala-lar xastaxanä-se tabib-ı 
   child-PL hospital-3 doctor-3 
   ‘a doctor in a children’s hospital’
   # ‘a doctor of children’s hospitals’

Unlike the ezafe-2 marker, the marker of ezafe-3 shows person and number agreement with the possessor, which can be a pronoun or a proper name, unlike with ezafe-2. We propose that the embedded nominal in ezafe-3 is a genitive-marked DP, or a KP.

(5) a. minem däftär-em 
   I GEN notebook-1SG 
   ‘my notebook’

b. Marat-nųj däftär-e 
   Marat-GEN notebook-3 
   ‘Marat’s notebook’

As it turns out, ezafe-3 itself is a DP; hence, can be embedded only in ezafe-3 but not in ezafe-2.

(6) a. ezafe-2: 
   *[ukuči-nųj däftär-lär-e] papka-si] 
   student-GEN notebook-PL-3 folder-3 
   ‘folder for student’s notebooks’

b. ezafe-3: 
   [[ukuči-nųj däftär-lär-e-neŋ] papka-si] 
   student-GEN notebook-PL-3-GEN folder-3 
   ‘folder for student’s notebooks’

In addition to the constructions described above, Tatar also has a number of attributivizer constructions. Here, only two such attributivizers, which turn a nominal into an attributive modifier, will be considered. Attributivizer -lı selects a bare NP, while attributivizer -ğı selects a KP, specifically a locative-marked one.
Elucidating Nominal Structure in Tatar

(7)  a.  [NP …]-lı  N  
      b.  [KP [DP …]-LOC]-gı  N

The nominal selected by attributivizer -lı cannot be a pronoun (8a), cannot contain the plural marker -lar (8b), but unlike with the nominal element in CPCs or the embedded nominal in ezafe-1, it can contain certain modifiers (8c).

(8)  a.  *ul-lı  čaška  
      it-ATR  cup  
      intended: ‘a cup with it’ (e.g. a blue flower)  
      b.  *kük  čačak-lär-le  čaška  
      blue  flower-PL-ATR  cup  
      intended: ‘a cup with blue flowers’  
      c.  kük  čačak-le  čaška  
      blue  flower-ATR  cup  
      ‘a cup with a blue flower’ OR ‘a cup with blue flowers’

In contrast, the nominal selected by attributivizer -gı is a full-fledged KP: it can be a locative-marked pronoun (9a) or a proper name (9b), even ezafe-3 (9c), which as we have shown above is a DP. It can also contain the plural marker -lar (9d) or modifiers (9e).

(9)  a.  a-n-da-gi  uram-nar  
      it-OBL-LOC-ATR  street-PL  
      ‘its (e.g. the city’s) streets’  
      b.  Kazan-da-gi  uram-nar  
      Kazan-LOC-ATR  street-PL  
      ‘streets of Kazan’  
      c.  Marat-нять  şahär-e-nää-ge  uram-nar  
      Marat-GEN  city-3-LOC-ATR  street-PL  
      ‘streets of Marat’s city’  
      d.  şahär-lär-dää-ge  uram-nar  
      city-PL-LOC-ATR  street-PL  
      ‘streets of cities’  
      e.  zur  şahär-dää-ge  uram-nar  
      big  city-LOC-ATR  street-PL  
      ‘streets of a big city’
3 Our Proposal: Structural Analysis of DOM in Tatar

In the preceding section, we have shown that nominals in Tatar can be projected fully or can be Small Nominals. Here, we argue that the same contrast accounts for the Differential Object Marking (DOM) phenomenon in the language. Our proposal is that the accusative case marker -mı, as in (10a), attaches to a DP. Unmarked objects are thus Small Nominals. More generally, we hypothesize that only DPs must receive structural case, while SNs are not subject to such Case licensing requirements and may remain morphologically Caseless. Three arguments are discussed below in support of this analysis.

Marat car-ACC sell-CONV take-PST
‘Marat bought a/the car.’

Marat car sell-CONV take-PST
‘Marat bought a car/cars.’

The first piece of evidence that suggests that ACC-marked and unmarked objects in Tatar are structurally different involves coordination: as shown in (11a), the two types of objects cannot be coordinated. If the unmarked object precedes the ACC-marked one, as in (11b), the sentence is grammatical, but such examples involve phrasal case marking, further illustrated in (11c).

Marat book-ACC and newspaper sell-CONV take-PST
intended: ‘Marat bought a (certain) book and a newspaper.’

b. Marat [kitap hâm gazet]-ne sat-ip al-dı.
Marat book and newspaper-ACC sell-CONV take-PST
‘Marat bought a (certain) book and a (certain) newspaper.’

c. [mädänijat hâm sângat]-kä bitarafaık
culture and art-DAT indifference
‘indifference to culture and art’

The second argument in favor of our proposal comes from the fact that minimal pairs such as (10) are not always possible. Accusative marker is required whenever the object is or contains some element typically associated with the DP-level. Thus, objects which are pronouns (12a), proper names (12b), or ezafe-3 constructions (12c), as well as objects containing a strong quantifier, such as här ‘every’ or ike…dâ ‘both’ (12d), or a demonstrative bu ‘this’ or şul ‘that’ (12e), must be marked accusative.
Elucidating Nominal Structure in Tatar

(12) a. Marat a-lar-* (nı) kür-de.
   Marat he-PL-ACC see-PST
   ‘Marat saw them.’

   b. Alsu Marat-* (nı) čakir-di.
      Alsu Marat-ACC invite-PST
      ‘Alsu invited Marat.’

   c. Marat Ramil-nej mašina-si-* (n) sat-ip al-di.
      Marat Ramil-GEN car-3-ACC sell-CONV take-PST
      ‘Marat bought Ramil’s car.’

   d. Marat här birem-* (ne) čiš-te.
      Marat every problem-ACC solve-PST
      ‘Marat solved every problem.’

   e. Marat bu mašina-* (nı) sat-ip al-di.
      Marat that car-ACC sell-CONV take-PST
      ‘Marat bought that car.’

Our third argument in support of the structural analysis of DOM in Tatar is that unmarked objects fit the profile of a Small Nominal, as described in Pereltsvaig (2006). First, they cannot have an individuated, specific, partitive, or anaphoric interpretation. For example, only (13a) is an appropriate continuation to ‘We have a cat and a dog’:

(13) a. Min kübesenčä et-ne jarat-a-m.
    I more dog-ACC like-PRS-1SG
    ‘I like the dog more.’

   b. Min kübesenčä et jarat-a-m.
      I more dog like-PRS-1SG
      ‘I like {a dog/dogs} more.’

Second, unmarked objects cannot have wide scope in relation to other quantifiers (14a) or negation (14b).

(14) a. Här ukuči ike kitap uki-di.
    every student two book read-PST
    ‘Every student read two books.’ > 2, *2 >

   b. Marat ike kitap uki-ma-di.
      Marat two book read-NEG-PST
      ‘Marat didn’t read two books.’ Neg > 2, *2 > Neg

While unmarked objects may contain the plural marker -lar, if this marker is absent, the object has a number-neutral interpretation (cf. Pereltsvaig 2013, forthcoming), as in (15a). The presence of the plural marker -lar entails a plural
interpretation, as in (15b). In contrast to unmarked objects, with ACC-marked objects the absence of plural marker -lar does not entail number-neutrality, as shown in (15c).

(15)  
(a) Marat kızıl alma aš-đı.  
Marat red apple eat-PST  
‘Marat ate {a red apple / red apples}.’
(b) Marat kızıl alma-lar aš-đı.  
Marat red apple-PL eat -PST  
‘Marat ate {*a red apple / red apples}.’
(c) Marat kızıl alma-nı aš-đı.  
Marat red apple-ACC eat -PST  
‘Marat ate {a red apple / *red apples}.’

Finally, unlike ACC-marked objects, unmarked objects cannot control syntactic anaphora: they cannot serve as controllers of PRO, as shown in (16a), or antecedents of reflexives or reciprocals, as shown in (16b, c). It should be noted that the ungrammaticality of unmarked objects as controllers of PRO or antecedents of reflexives or reciprocals is preserved regardless of word order permutations; space limitations prevent us from presenting all the possible word order variations.

(16)  
(a) Marat bala-lar.*(ni) džibär-de [PRO uk-ırga].  
Marat child-PL-ACC send-PST study-INF  
‘Marat sent children to study.’
(b) Marat bala-lar.*(ni) džibär-de üze-neñ tu-gan  
Marat child-PL-ACC send-PST SELF.3-PL be.born-PART  
ken-e-neñ bäjräm-lär-e-nä .  
day-3-GEN holiday-PL-3-DAT  
‘Marat sent children to their own birthday parties.’
(c) Marat ber-ber-se-neñ tu-gan ken-e-neñ  
Marat REC-REC-3-GEN be.born-PART day-3-GEN  
bäjräm-lär-e-nä bala-lar.*(ni)čak-ti.  
holiday-PL-3-DAT child-PL-ACC call-PST  
‘Marat invited children to each other’s birthday parties.’

To conclude, ACC-marked and unmarked objects are structurally distinct, as suggested by the fact that they cannot be coordinated. DP objects must be accusative marked, while unmarked objects have all the properties of Small Nominals.

In the following sections, we will consider and dismiss two alternative analyses proposed for DOM in other Turkic languages (and some non-Turkic languages as well): the Semantic Alternative and the Positional Alternative.
4 Semantic Alternative

One alternative analysis, which has been proposed for DOM in Turkish by Enç (1991), as well as for DOM in Hebrew by Danon (2006), correlates the accusative case marking or its absence with the semantics of the object rather than its functional structure. According to this group of analyses, ACC-marked objects are referential, specific, or definite, depending on the particular proposal, whereas unmarked objects are not. As mentioned above, unmarked objects in Tatar have a non-individuated, non-specific, non-partitive, non-anaphoric interpretation. Could it be that the same analysis applies in Tatar? We do not think so.

The data that challenges the Semantic analysis for Tatar involves ezafe-3 construction. As mentioned above, an object which is an ezafe-3 construction must receive structural (accusative) case. However, it may simultaneously receive a non-specific interpretation and take narrow scope in relation to other quantifiers or negation, as shown in (17a, b). For example, the sentence in (17a) can be interpreted as ‘There are two specific poems by Tukay that every student read’ or as ‘Every student read some two poems by Tukay’. Similarly, the sentence in (17b) can be interpreted as ‘There is a photo of Alsu that Marat didn’t see’ or as ‘It is not the case that Marat saw a/any photo of Alsu’. In both examples, it is the second interpretation that is unexpected under the Semantic Alternative.

(17) a. Här ukučı Tukaj- nghị ike śigır-e-*(*n) uki-di. every student Tukay-GEN two poem-3-ACC read-PST
   ‘Every student read two poems by Tukay.’
   2 > ∃ or ∃ > 2

b. Marat Alsu-netinet fotografıa-se-*(*n) kür-me-de. Marat Alsu-GEN photo-3-ACC see-NEG-PST
   ‘Marat didn’t see a photo of Alsu.’
   ∃ > Neg or Neg > ∃

Because the presence of accusative case marking does not guarantee a referential, specific, or definite interpretation, we must reject the Semantic Alternative.

5 Positional Alternative

In the preceding section, we have shown that DOM in Tatar cannot be accounted for purely in terms of the semantics of the object. In this section, we consider—and also reject—another alternative theory that places the burden of explanation on the position of the object in the clausal structure. According to this approach, unmarked objects appear in vP, whereas ACC-marked objects appear higher, outside vP. For example, an analysis along those lines has been proposed for
DOM in another Turkic language, Sakha, by Baker & Vinokurova (2010: 599-602). Furthermore, unmarked objects have been analyzed as pseudo-incorporated into the verb (cf. Massam 2001, Baker 2009). In what follows, we show that although this approach may work for other languages, it is not applicable to Tatar.

The first problem for the Positional Alternative is similar to the challenge we described above for the Semantic Alternative: while unmarked objects behave roughly as expected (i.e. they appear relatively low in the clausal structure), ACC-marked objects do not. In particular, they can appear inside the vP boundary marked by (manner) adverbs such as tiz ‘quickly’. In this respect, Tatar is genuinely different from Sakha, where such sentences are ungrammatical. (The Sakha example (18b) below is from Baker & Vinokurova 2010: 602, their (12b); it is said to be grammatical if the object has contrastive focus.)

(18) a. Tatar
Marat tiz botka-nı aša-dı.
Marat quickly porridge-ACC eat-PST
‘Marat ate porridge quickly.’

b. Sakha
*Masha türgennik salamaat-y sie-te.
Masha quickly porridge-ACC eat-PAST.3sS
‘Masha ate porridge quickly.’

Moreover, although unmarked objects in Tatar have some semantic properties associated with pseudo-incorporated nominals (e.g. their obligatory nonreferential interpretation, obligatory narrow scope, and possible number-neutrality), they do not seem to have a particularly tight syntactic connection to the verb. In fact, if anything is pseudo-incorporated in Tatar, it is the the nominal components in CPCs, considered above. First, the nominal component in CPCs and unmarked objects behave differently in causative constructions: in causative constructions based on CPCs, the causee is marked accusative, as shown in (19a). In contrast, in causative constructions with either ACC-marked or unmarked objects, the cause is ablative, rather than accusative, as shown in (19b, c). This shows that nominal components in CPCs and unmarked objects do not appear in the same structural position.

(19) a. Ukıtučı ukučṭ-nı/*-dan xezmäit-ter-de.
teacher student-ACC/*-ABL work do-CAUS-PST
‘The teacher made the student work.’

b. Min Marat-tan/*-nı kitap al-dir-dı.
I Marat-ABL/*-ACC book take-CAUS-PST
‘I made Marat buy a book.’

c. Min Marat-tan/*-nı kitap-nı al-dir-dı.
I Marat-ABL/*-ACC book-ACC take-CAUS-PST
‘I made Marat buy a (certain) book.’
Second, nominal components in CPCs—like other pseudo-incorporated nominals—cannot be focused by particles such as -gına, whereas unmarked objects can be. For example, the sentence in (20a) can have one of two interpretations, one with focus on the whole verb phrase and the other with focus just on the object. As shown in (20b), the nominal component in CPC cannot be focused by itself.

\[(20)\]
\[
a. \text{Marat bala-ga } jaña \text{ } kitap-}kîna \text{ } uki-di.  
\text{Marat child-DAT new book-EMPH read-PST}  
\text{‘The only thing that Marat did is read the child a new book.’}  
\text{‘Marat read the child only a new book.’}  
\]
\[
b. \text{Äti-se Marat-ka } jaña \text{ mašina}büläk-kîna it-te.  
\text{father-3 Marat-DAT new car gift-EMPH make-PST}  
\text{‘His father only gave Marat a new car as a gift.’}  
\text{NOT: #‘His father gave Marat a new car only as a gift.’}  
\]

Third, unmarked objects can serve as antecedents for discourse anaphora (though not, as you would recall from above, syntactic anaphora), as shown in (21), whereas nominal components in CPCs cannot, as shown in (22). Thus, the second sentence in (22) can mean only that making the student work is useful, not that work itself is useful. The inability to serve as antecedent for discourse anaphora is another hallmark of pseudo-incorporating nominals, suggesting that while nominal components in CPCs are pseudo-incorporated, unmarked objects are not.

\[(21)\]
\[
\text{Sin anarga } \text{kitap ala ala-siŋ.}  
\text{you that.DAT book take.IPfv can.PRS-2SG}  
\text{Häm a-mi matur it-ep ter-ep büläk}  
\text{and that-ACC beautifully make-CONV wrap-CONV gift}  
\text{it-ergä bula.}  
\text{make-INF be.PRS}  
\text{‘You can buy him a book. You can wrap it beautifully and give it to him as a gift.’}  
\]

\[(22)\]
\[
\text{Ukituçi ukuçi-mi } \text{xezmät it-ter-de.}  
\text{teacher student-ACC work do-CAUS-PST}  
\text{#Ul bik fajdaši eš.}  
\text{it very useful matter}  
\text{‘The teacher made the student work. It (making the student work) is very useful.’}  
\text{NOT: ‘The teacher made the student work. It (work) is very useful.’}  
\]
To recap, contrary to the predictions of the Positional Alternative, ACC-marked objects may appear low in the clausal structure, where they structurally compete with unmarked objects. Moreover, unmarked objects can be shown to not be pseudo-incorporated into the verb, whereas nominal components in CPCs fit the profile of a pseudo-incorporated nominal.

6 Conclusion: Functional Architecture, Position and Semantics

In this paper, we argued for the structural analysis of DOM in Tatar whereby DP objects receive structural (accusative) Case, while Small Nominal objects remain Caseless. This serves as an additional argument for the projection of DP in some but not all noun phrases in Tatar, an articleless language. Furthermore, we considered and rejected two alternative approaches to DOM in Tatar: one that places the explanatory burden on the semantics of the object and the other that relies on the position of the object in the clausal structure. We have shown that neither of these alternatives can account fully for the facts concerning DOM in Tatar. The partial overlap between the three possible analyses, however, derives from the fact that Small Nominals lack the DP-layer and consequently lack three things: (a) the semantics of full-fledged DPs (i.e. referentiality), (b) the mobility of full-fledged DPs as they remain invisible to higher Probes searching for [+D], and (c) the ability to be assigned Case. However, such Small Nominal objects are not as low as pseudo-incorporated nominals, such as the nominal components in CPCs. Moreover, we have shown that full DP objects, which are marked accusative, are not necessarily high in the clausal structure and need not be referential.

References

Linguistic Inquiry 16:373-416.
Baker, Mark. 2009. Is head movement still needed for noun incorporation? 
Lingua 119:148-165.
Baker, Mark and Nadya Vinokurova. 2010. Two modalities of Case assignment: 
Elucidating Nominal Structure in Tatar


Pereltsvaig, Asya. forthcoming. On Number and Numberlessness in Languages with and without Articles. In P. Cabredo Hofherr and A. Zribi-Hertz, eds., *Languages with and without articles* [working title].


Ekaterina Lyutikova
Dept. of Theoretical and Applied Linguistics
Moscow State University
Leninskiye Gory, MSU
1st Building for Humanities, 953
119991 Moscow
Russia
lyutikova2008@gmail.com