

Wh-questions in Balkar

Anastasiia Voznesenskaia*

Abstract. This paper deals with the properties of *wh*-questions in Balkar. It is shown that *wh*-in-situ structures in Balkar are island insensitive (with an exception of coordinate structures). I discuss the complement/adjunct asymmetry regarding intervention effects. I also consider embedded multiple *wh*-structures. In this paper, I discuss a puzzle that the Balkar data presents to the prominent theories of *wh*-questions, which do not explain the properties it shows.

Keywords. Turkic; Balkar; *wh*-in-situ; island effects; intervention effects

1. Introduction. (Karachay-)Balkar is a Turkic language mostly spoken by Karachays and Balkars in the Kabardino-Balkarian and Karachay-Cherkessian Republics of Russia. The data in this paper was collected during fieldwork in Verkhnyaia Balkaria in Kabardino-Balkaria.

This paper discusses the properties of *wh*-questions in Balkar. *Wh*-phrases in Balkar do not obligatorily move to the left periphery and mainly occupy a preverbal position (1).

- (1) a. *üj-nü kim işle-gen-di?*
house-ACC who build-PFCT-3SG
‘Who built the house?’
b. *Alim ne-ni işle-gen-di?*
Alim what-ACC build-PFCT-3SG
‘What did Alim build?’
c. *Alim üj-nü qacan işle-gen-di?*
Alim house-ACC when build-PFCT-3SG
‘When did Alim build the house?’

Since Balkar doesn’t manifest obligatory *wh*-movement this paper considers the data along the lines of the approaches to *wh*-in-situ languages.

These approaches can be classified into two general types: those which analyze *wh*-in-situ structures as involving covert movement (the covert movement approach) and those which don’t postulate any kind of movement (the in-situ approach). Huang (1982) argues that *wh*-phrases move at LF and, hence, syntactically occur higher than they are pronounced. The in-situ approaches use other interpretive mechanisms to derive questions in *wh*-in-situ. An influential approach is based on Rooth-Hamblin alternative computation: *wh*-elements generate a set of focus-semantic values which are then passed up the structure to calculate the focus-semantic value of a clause (Hamblin 1973, Rooth 1992, Beck 2006). Another in-situ approach uses unselective binding: *wh*-phrases are considered on a par with indefinites, they are bound by a question operator in C (Heim 1982, Pesetsky 1987). A different in-situ approach is the choice function approach that Reinhart 1998 argues for. Under this analysis *wh*-expressions are viewed as standard existentials interpreted via choice functions (functions applying to a non-empty set and yielding an individual element). Reinhart argues that the wide scope of *wh*-elements is due to the quantification over choice functions. In this paper I present some Balkar data which cannot be straightforwardly accounted for by any of the approaches above.

* The work was supported by an RFBR grant #19-012-00627a

Author: Voznesenskaia Anastasiia, Lomonosov Moscow State University (anastvozn@gmail.com).

In this paper, I examine the properties of Balkar *wh*-questions and discuss the theoretical puzzles that these data present. The paper is structured as follows: section 2 discusses *wh*-phrases in embedded arguments, section 3 deals with island effects, section 4 discusses intervention effects, section 5 shows data on embedded multiple-*wh* and section 6 concludes the paper.

2. Embedded arguments. *Wh*-phrases in embedded arguments can have matrix or embedded scope. Below are examples of structures with a *wh*-subject, object and adjunct in an embedded clause. The sentences in (2)–(4) can have both i. (matrix scope) and ii. (embedded scope) translations depending on the prosody (all examples in this paper are not echo-questions).

(2) NOMINALIZATION

Alim [fatima-ni zaš-i-ni **ne-ni** žaz-Ɂan-i-n] ešt-gen-di
 Alim Fatima-GEN son-3-GEN what-ACC write-PFCT-3-ACChear-PFCT-3SG
 i. ‘What did Alim hear that Fatima’s son wrote?’
 ii. ‘Alim heard what Fatima’s son wrote.’

(3) FINITE

Alim [**kim** kitab zaz-dı dep] ešt-gen-di
 Alim who book write-PST COMPhear-PFCT-3SG
 i. ‘Who did Alim hear wrote a book?’
 ii. ‘Alim heard who wrote a book.’

(4) Alim [fatima-ni zaš-i-ni **qacan** kitab žaz-Ɂan-i-n] ešt-gen-di
 Alim Fatima-GEN son-3-GEN when book write-PFCT-3-ACC hear-PFCT-3SG
 i. ‘Alim heard that Fatima’s son wrote a book when?’
 ii. ‘Alim heard when Fatima’s son wrote a book.’

Since there’s evidently no obligatory overt *wh*-movement in Balkar I am going to consider it among *wh*-in-situ languages. *Wh*-phrases can also be scrambled to the left periphery of the main clause. However, *wh*-scrambling undergoes the same restrictions as the scrambling of the corresponding XPs, which lets us assume that there is nothing *wh*-specific to it, which is why *wh*-scrambling data is beyond the scope of this paper.

3. Island effects. According to the covert movement approach the *wh*-phrase still undergoes movement which means it should be subject to the island constraints. However, this is not the case in Balkar.

In a complex NP, a *wh*-phrase can have matrix scope (5)–(7). Sentences (b) are given to show that these island effects occur with overt movement.

(5) SUBJECT

a. sen [[**kim** kitab žaz-Ɂan-i] žanliq-ni] ešt-gen-se?
 you who book write-PFCT-3 news-ACC hear-PFCT-2SG
 ‘You heard the news that who wrote a book?’
 b. ^{??}**kim** [[sen kitab žaz-Ɂan-i] žanliq-ni] ešt-gen-se?
 who you book write-PFCT-3 news-ACC hear-PFCT-2SG

(6) OBJECT

- a. sen [[fatima-ni zaš-i **ne-ni** žaz-ɤan-i] žaŋliq-ni] ešt-gen-se?
 you Fatima-GEN son-3 what-ACC write-PFCT-3 news-ACC hear-PFCT-2SG
 ‘You heard the news that Fatima’s son wrote what?’
- b. ***ne-ni** [sen fatima-ni zaš-i žaz-ɤan-i] žaŋliq-ni] ešt-gen-se?
 what-ACC you Fatima-GEN son-3 write-PFCT-3 news-ACC hear-PFCT-2SG

(7) ADJUNCT

- a. sen [[fatima-ni zaš-i **qacan** kitab žaz-ɤan-i] žaŋliq-ni] ešt-gen-se?
 you Fatima-GEN son-3 when book write-PFCT-3 news-ACC hear-PFCT-2SG
 ‘You heard the news that Fatima’s son wrote a book when?’
- b. #**qacan** sen fatima-ni zaš-i kitab žaz-ɤan-i žaŋliq-ni
 when you Fatima-GEN son-3 book write-PFCT-3 news-ACC
 ešt-gen-se
 hear-PFCT-2SG

An anonymous reviewer points out regarding (7) that *when* can behave on a par with *wh*-arguments like *what*, while *how* and *why* may not be able to be extracted. While it should be important to look at what (7)a would look like with *why/how*, this example presents an interesting piece of data in that (7)a is possible while overt extraction of *qacan* is not (7b).

Wh-phrases can also appear inside adjuncts. However, there appears an argument/adjunct asymmetry: (8a) with a *wh*-object is fine while (9a) is ungrammatical. That is, adjuncts seem to present weak islands in this language. Sentences (b) are again given to show that these island effects occur with overt movement.

(8) OBJECT

- a. Asijat [kerim **ne-ni** boja-ɤan-da] kitap oqu e-di?
 Asijat Kerim what-ACC paint-PFCT-LOC book read COP-PST
 ‘Asijat was reading a book when Kerim was painting what?’
- b. ***ne-ni** asijat [kerim boja-ɤan-da] kitap oqu ed-i
 what-ACC Asijat Kerim paint-PFCT-LOC book read COP-PST

(9) ADJUNCT

- a. *Asijat [kerim xuna-ni **qajda** boja-ɤan-da] kitap
 Asijat Kerim fence-ACC where paint-PFCT-LOC book
 oqu ed-i
 read COP-PST
Int. ‘Asijat was reading a book when Kerim was painting a fence where?’
- b. ***qajda** asijat [kerim xuna-ni boja-ɤan-da] kitap
 where Asijat Kerim fence-ACC paint-PFCT-LOC book
 oqu ed-i
 read COP-PST

In a coordinate structure, neither first nor second conjunct can be a *wh*-phrase.¹

¹ During fieldwork on Buryat I have stumbled upon the fact that, although all the other island effects do not present with *wh*-in-situ questions there, the Coordinate Structure Constraint still stands. This makes me wonder if there is

- (10) a. *Asijat [kim bla kištɨʁ]-i-n bek sude
 Asijat who and cat-3-ACC very loves
Int. ‘Asijat loves who and her cat very much?’
 b. *Asijat [it bla kim]-i-n bek sude
 Asijat dog and who-3-ACC very loves
Int. ‘Asijat loves her dog and who very much?’

	wh-phrase	embedded	scrambling
Coordinate Structure	subject object	– –	
Complex NP	subject	+	–
	object	+	–
	adjunct	+	–
Adjunct	object	+	–
	adjunct	–	–

Table 1: Summary of the data on island effects

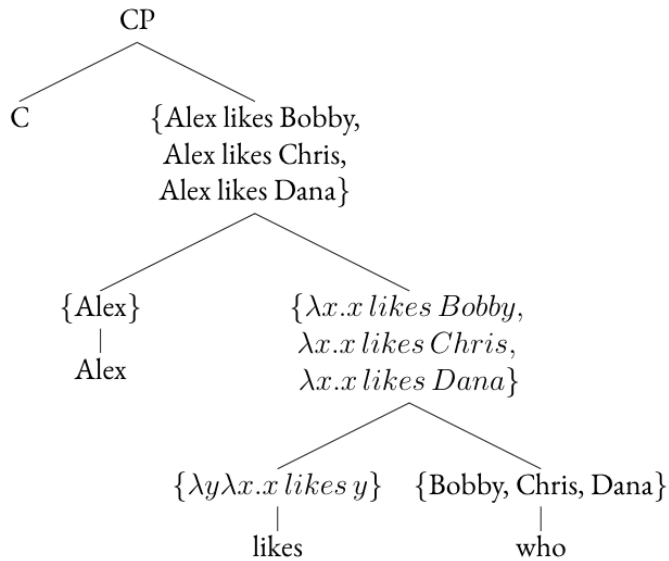
The lack of island effects makes a covert movement analysis very problematic for the Balkar *wh*-questions. If there is covert movement present, then the *wh*-phrases undergoing it would still be subject to any island constraints present in the language. This section had provided evidence that even when a structure shows island effects with overt movement, it might not show the same effect with *wh*-in-situ questions. This is contrary to the predictions of the covert movement approach.

4. Intervention effects. According to the alternative computation approach (Hamblin 1973, Rooth 1992, Beck 2006) *wh*-phrases (along with focused phrases) generate a set of alternatives. A focus-semantic value is computed for each node in the structure and is then evaluated by a question complementizer (or Q) at which point focus semantic values enter ordinary semantics.

something special about this particular constraint that distinguishes it from all the others. This idea is also presented in e.g. Kubota & Lee (2015).

- (11) A toy LF of question interpretation via Rooth-Hamblin alternative computation
(Kotek 2014:23)

Who does Alex like?



However, according to Beck (2006), if the *wh*-phrase gets evaluated within the scope of Q by another focus-sensitive operator, intervention effects occur. The focus-semantic value of the *wh*-phrase gets reset to the ordinary semantic value, which is undefined for *wh*-phrases, and neither the structure's ordinary semantic nor focus-value is defined. This is supposed to lead to ungrammaticality (12). However, given the Balkar data below, this is not true in all cases.

- (12) A *wh*-phrase in situ may not be c-commanded by a focusing or a quantificational element.
(Beck 2006:3)

Significantly, Balkar questions with *wh*-objects seem to lack intervention effects, as I have shown with several of the focus-sensitive operators mentioned in Beck (2006). For all operators, I show a root question and, also, an embedded one to make sure that the operator c-commands the *wh*-phrase.

4.1. ONLY. Both questions where *only* intervenes between the *wh*-phrase and the question complementizer (a), and those where the *wh*-phrase is scrambled over *only* are equally well-formed.

- (13) a. **quru** alim **ne-ni** biledi?
only Alim what-ACC knows
b. **ne-ni** **quru** alim biledi?
what-ACC only Alim knows
'What does only Alim know?'

- (14) a. **quru** alim [fatima-ni qiz-i **ne-ni** et-gen-i-n]
 only Alim Fatima-GEN daughter-3 what-ACC make-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
- b. **ne-ni** **quru** alim [fatimani qiz-i ~~neni~~ et-gen-i-n]
 what-ACC only Alim Fatima-GEN daughter-3 make-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
 ‘What does only Alim know that Fatima’s daughter cooked/made?’

4.2. EVEN. Both questions where *even* intervenes between the *wh*-phrase and the question complementizer (a), and those where the *wh*-phrase is scrambled over *even* receive the same evaluation by the speakers. The question marks in (15) show that the sentence was found odd by the speakers but not unacceptable.

- (15) a. Alim **oquna** **ne-ni** biledi?
 Alim even what-ACC knows
- b. **ne-ni** alim **oquna** biledi?
 what-ACC Alim even knows
 ‘What does even Alim know?’
- (16) a. ?Alim **oquna** [fatima-ni zas-i **ne-ni** žaz-kan-i-n]
 Alim even Fatima-GEN son-3 what-ACC write-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
- b. ?**ne-ni** alim **oquna** [fatimani zaš-i ~~neni~~ žaz-kan-i-n]
 what-ACC Alim even Fatima-gen son-3 write-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
 ‘What does even Alim know that Fatima’s son wrote?’

4.3. ALSO. The results for *also* are parallel to those for *even*.

- (17) a. Kerim **da** **ne-ni** biledi?
 Kerim and what-ACC knows
- b. **ne-ni** kerim **da** biledi?
 what-ACC Kerim and knows
 ‘What does also Kerim know?’
- (18) a. ?Kerim **da** [fatima-ni zas-i **ne-ni** žaz-kan-i-n]
 Kerim and Fatima-GEN son-3 what-ACC write-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
- b. ?**ne-ni** kerim **da** [fatimani zaš-i ~~neni~~ žaz-kan-i-n]
 what-ACC Kerim and Fatima-GEN son-3 write-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
 ‘What does also Kerim know that Fatima’s son wrote?’

4.4. NPIs. For the NPI *bir adam da* both (a) and (b) are again equally acceptable.

- (19) a. **bir adam da ne-ni** bil-me-i-di?
 one man da what-ACC know-NEG-IPFV-3SG
 b. **ne-ni bir adam da** bil-me-i-di?
 what-ACC one man da know-NEG-IPFV-3SG
 ‘What does nobody know?’
- (20) a. **bir adam da** [fatima-ni zaš-i **ne-ni** zaz-kan-i-n]
 one man da Fatima-GEN son-3 what-ACC write-PFCT-3-ACC
 bil-me-i-di?
 know-NEG-IPFV-3SG
 b. **ne-ni bir adam da** [fatima-ni zaš-i ~~ne-ni~~ zaz-kan-i-n]
 what-ACC one man da Fatima-GEN son-3 write-PFCT-3-ACC
 bil-me-i-di?
 know-NEG-IPFV-3SG
 ‘What does nobody know that Fatima’s son wrote?’

As can be seen, an important prediction of the alternative computation approach does not hold for Balkar. Therefore, it is reasonable to look for another approach when trying to make sense of the data above. Similar language profiles with the lack of both island and intervention effects are shown by Hong (2004) for Korean and Soltan (2012) for Egyptian Arabic. Both authors propose an unselective binding analysis: the question operator in C unselectively binds the *wh*-phrases in its scope. Reinhart (1984) argues that unselective binding is inadequate and puts forth a choice function analysis: *wh*-phrases are treated as indefinites. They denote choice functions and are existentially quantified over to achieve wide scope. Reinhart’s approach seems to make the same predictions regarding island and intervention effects: they predict the lack thereof.

However, Balkar doesn’t lack intervention effects completely. They seem to be in place for adjunct *wh*-phrases.

- (21) ***quru** alim [fatima-ni zaš-i **qacan** kitap zaz-kan-i-n]
 only Alim Fatima-GEN son-3 when book write-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
Int. ‘Only Alim knows that Fatima’s son wrote a book when?’
- (22) *Alim **oquna** [fatima-ni zaš-i **qacan** kitap zaz-kan-i-n]
 Alim even Fatima-GEN son-3 when book write-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
Int. ‘Even Alim knows that Fatima’s son wrote a book when?’
- (23) *Alim **da** [fatima-ni zaš-i **qacan** kitap zaz-kan-i-n]
 Alim also Fatima-GEN son-3 when book write-PFCT-3-ACC
 bil-e-di?
 know-IPFV-3SG
Int. ‘Alim also knows that Fatima’s son wrote a book when?’

- (24) ***bir adam da** [fatima-ni zaš-i **qacan** kitap zaz-~~van~~-i-n]
 one man da Fatima-GEN son-3 when book write-PFCT-3-ACC
 bil-me-i-di?
 know-NEG-IPFV-3SG
Int. ‘Nobody knows that Fatima’s son wrote a book when?’

It is important to notice that (21)–(24) become well-formed if the focus-sensitive operator is removed. Unfortunately, *wh*-adjuncts cannot scramble out of complement clauses so it’s not possible to see if scrambling this *wh*-phrase over the intervener would ‘save’ these sentences.

This piece of data brings more confusion to the mix. There’s been established a complement/adjunct asymmetry with intervention which none of the above-mentioned approaches seem to explain.

5. Multiple *wh*-questions. One more distinction in what the covert movement and the in-situ approaches predict has to do with multiple *wh*-questions. Looking at an embedded multiple *wh*-structure like (25) from the point of view of the covert movement approach it seems logically possible that both *wh*-phrases are pronounced in the embedded clause while one of them has moved to the matrix clause at LF. At the same time, according to the in situ theories described above this should not be possible.

- (25) [C_Q [...*wh*.....*wh*]]

Consider (26). Expectedly, it can be a matrix multiple-*wh*-question (i) and an embedded multiple-*wh*-question (ii). The question is whether one of the *wh*-phrases can have matrix scope while the other has embedded scope. The answer is no, interpretations in (iii) and (iv) are unavailable for (26).

- (26) sen [**kim kim-ni** kör-gen-i-n] ešt-gen-se
 you who who-ACC see-PFCT-3-ACC hear-PFCT-2SG
 i. ‘You heard that who saw whom?’
 ii. ‘You heard who saw whom.’
 iii. *‘Who did you hear saw whom?’
 iv. *‘Whom did you hear who saw?’

To get the kind of interpretation in (26)iii.–iv. one *wh*-phrase must be scrambled to the matrix clause, which is in line with the in-situ approaches.

- (27) a. ***kim** (tünene) sen [~~kim~~ kim-ni kör-gen-i-n] ešt-gen-se?
 who yesterday you who-ACC see-PFCT-3-ACC hear-PFCT-2SG
Int. ‘Who did you hear saw whom?’
 b. **kim-ni** sen [kim ~~kimni~~ kör-gen-i-n] ešt-gen-se?
 who-ACC you who see-PFCT-3-ACC hear-PFCT-2SG
 ‘Who_i did you hear who saw t_i?’

In this case, there is an asymmetry reported for other languages (e.g. Özsoy 2009, Turkish): it is possible to scramble the object to the matrix clause leaving the subject in the embedded clause (27b) but not the other way around (27a). This pattern holds for non-*wh* XP-scrambling in Balkar too which indicated that this has to do with how scrambling works in this language. Independently of what leads to the asymmetry (27b) presents a problem for the unselective binding and choice function approaches: moving, *kimni* leaves a trace in the scope of the embedded question operator.

This structure shows another asymmetry: the adjunct *wh*-phrase is also impossible to scramble to the matrix clause.

- (28) *qajda sen [alim kim-ni kör-gen-i-n] ešt-xem-me?
where you Alim who-ACC see-PFCT-3-ACC hear-PFCT-1SG
*‘You heard whom Alim saw WHERE?’

Multiple *wh* data brings another argument against the covert movement analysis. It does not, however, unambiguously point towards a certain analysis among the in-situ group.

6. Conclusion. A few important properties of the Balkar *wh*-questions have been examined. To summarize, Balkar embedded *wh*-structures are island-insensitive (except for the Coordinate Structure Constraint) and one of the islands, the Adjunct Island, shows an object/adjunct asymmetry. Intervention effects are also lacking for object *wh*-phrases but seem to be present for *wh*-adjuncts. Embedded multiple *wh*-phrases can only be interpreted together in one clause. If one of them is scrambled to the main clause it only yields a well-formed question for *wh*-objects.

The traditional covert movement approach does not appear to work because of the island-insensitivity and the multiple-*wh* facts. The in-situ approaches yet considered also only partially explain these data: the lack of intervention effects for objects contradicts the alternative computational analysis, unselective binding and choice functions seem to struggle with the intervention effects that are there. These data are relevant for the study of *wh*-in-situ since it presents a puzzle in this field which has not been widely discussed for other languages. The considered theories do not explain all the facts. This paper does not present suggest a definite analysis for the facts presented, which leaves developing an account for everything for further research. One more factor to consider should be the preverbal position of the *wh*-phrase which distinguishes Balkar and other Turkic languages from other *wh*-in-situ languages.

References

- Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics*. 14(1). 1-56. <https://doi.org/10.1007/s11050-005-4532-y>.
- Hamblin, Charles Leonard. 1973. Questions in Montague grammar. In Barbara H. Partee (ed.), *Montague grammar*. 247-259. Academic Press.
<https://doi.org/10.1016/B978-0-12-545850-4.50014-5>.
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst, MA: UMass dissertation.
- Hong, Sun Ho. 2004. On the lack of syntactic effects in Korean *wh*-Questions. *The Linguistic Association of Korea Journal*. 12(3). 43-57.
- Huang, C.T. James. 1982. Move *wh* in a language without *wh* movement. *The Linguistic Review*. 1(4). 369-416. <https://doi.org/10.1515/tlir.1982.1.4.369>
- Kotek, Hadas. 2014. *Composing questions: Massachusetts Institute of Technology dissertation*. Cambridge, MA: MIT dissertation.
- Kubota, Yusuke & Lee, Jungmee. 2015. The Coordinate Structure Constraint as a discourse-oriented principle: Further evidence from Japanese and Korean. *Language*. 91(3). 642–675.
- Özsoy, Sumru A. 2009. Turkish as a (non)-*wh*-movement language. *Series B-Skrifter*. 132. 221–232.
- Pesetsky, David. 1987. *Wh*-in-situ: Movement and unselective binding. In Eric Reuland & Alice G.B. ter Meulen (eds.). *The representation of (in)definiteness*. 98-129. Cambridge, MA: MIT Press.

- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics*. 1(1). 75-116.
<https://doi.org/10.1007/BF02342617>.
- Soltan, Usama. 2012. On licensing wh-scope: wh-questions in Egyptian Arabic revisited. In Reem Bassiouney & E. Graham Katz (eds.), *Arabic language and Linguistics*. 99-114. Washington D.C.: Georgetown University Press.