

Languages of the Caucasus

A free refereed web journal for linguistic work on languages of the Caucasus

The Chechen it-cleft construction

Erwin R. Komen

SIL-International

Meertens Instituut (Royal Netherlands Academy of Arts and Sciences)

Erwin_komen@sil.org

ABSTRACT

This paper presents a biclausal construction in Chechen, arguing that it can be analyzed as an *it*-cleft. The construction consists of a main copula clause with a covert or pronominal subject, and a temporal complement that co-indexes with an adjunct position in a relative clause that does not form a constituent with the subject or the complement. A study of the construction in a corpus of newspaper and journal texts shows characteristics that make it stand out in terms of syntax and function: the cleft clause can appear both clause-finally as well as clause-initially, and its function is limited to text-structuring (it is mostly used to mark the start of a text or the transition to a new paragraph). This latter characteristic is exceptional: *it*-clefts in other languages (such as English and Norwegian) are known to be used for text-structuring to some extent, but Chechen is the first language known to only use it for this purpose. This prompts the question whether there are perhaps other language (e.g. from the Nakh-Daghestan family) with similar characteristics.

Keywords

Chechen, *it*-cleft, focus, text-structure, discourse

This is a contribution from *Languages of the Caucasus*, Vol. 1.1. © 2015 Erwin R. Komen.
ISSN 2375-2068

This Portable Document Format (PDF) file may not be altered in any way.

Tables of contents, abstracts, and submission guidelines are available at
escholarship.org/uc/languagesofcaucasus

*The Chechen it-cleft construction*¹

ERWIN R. KOMEN

SIL-International

Meertens Instituut (Royal Netherlands Academy of Arts and Sciences)

Erwin_komen@sil.org

1. Introduction

A large scale study of Caucasian languages still awaits a break-through, but the languages from the Nakh-Daghestanian (N-D) family (alternatively named “[North-]East Caucasian”) in particular show an interesting blend of characteristics that are sometimes unique in the world (see van den Berg (2005) for an introduction into them). The languages have flexible word order (in combination with a rich case system – especially locatives), are head-last and mostly have postpositions. Almost all are morphologically ergative, though this is often mixed with a nominative-accusative system (Baerman et al. 2002; Holisky 1987). Many have multiple noun classes; less than some African languages, but more than the three-way Indo-European gender systems (Berg 2005; Plaster et al. 2013).

This study examines a construction in Chechen, a language from the Nakh sub group of the Nakh-Daghestanian family, that is exemplified in (1).²

¹ The data for this research goes a long way back to the work done by Zacharski and Cowie in collaboration with an anonymous Chechen speaker who provided the English translations. I am grateful to them, and I would also like to thank my Radboud University colleagues Meta Links and Ans van Kemenade, who provided valuable feedback on drafts of this paper. I have also benefitted from comments made by three anonymous reviewers.

² The data used by this paper come from two different corpora. The first corpus is a selection of books that are available in electronic form (Arsanukaev 2008; Ezhaev & Ezhaeva 2007; Nunaev 1991). The second corpus consists of a set of newspaper and journal articles collected by the New Mexico State University, and will be referred to as the “NMSU corpus” (Zacharski & Cowie 2011). The references to the examples in this paper mark their location within the NMSU corpus: the initial “p” or “m” point to the parallel or monolingual part of the corpus, next follows the number of the text, a colon, and then the line number within the text. The Chechen examples follow the Latinized orthography closely resembling the one developed for Ingush (Nichols 2007). The vowels and consonants roughly appear in their IPA forms, with the following exceptions. The *w* on its own represents an epiglottal stop / ʔ/, while it represents an epiglottal fricative / ʕ/ when it follows a

- (1) Hara **cwa butt xaan** ju Noxchiin Respublikan Q'ooman bibli'otekan
 this one month time is Chechen Republic's national library's
 bielxaxuosha de-byysa ca lyerush q'ahwyegu.
 employers day-night not regarding toil
 'It is one month since the employees of the Chechen Republic's national
 library have been working both day and night.' [m00225.1]

The characteristics of the construction are: (a) the body of the construction is a main copula clause ('this/it is one month'), (b) the copula clause is followed by a relative clause ('that employees of the ... library have been working'), (c) the main clause's subject *hara* 'this' is a pronoun with no antecedent in the preceding text (since this is the opening line of a newspaper article), (d) the predicative argument of the main clause ('one month time') coindexes with an adjunct position in the restrictive relative clause, and (e) the relative clause does not form a constituent together with the subject (*'this that the employees of the ... library have been working') nor with the predicative main clause argument (*'one month time that the employees of the ... library have been working').

In this paper, I show that the construction should be regarded as an *it*-cleft, since it fully satisfies a language-independent definition (section 2). The next section (3) presents a corpus-based study on the variation of the *it*-cleft in its syntax and it discusses the function this construction fulfills. This corpus study yields a set of approximately 100 *it*-cleft samples. The clefts vary in a number of interesting ways, such as the presence or absence of an explicit pronominal subject, the presence or absence of a generic noun (such as *xaan* 'time' in the example, which follows the head noun *butt* 'month') in the clefted constituent and the position of the cleft clause with respect to the main clause. The function of the construction is related to text-structuring. Section 4 compares the *it*-cleft in Chechen with its counterparts other languages, revealing that it takes a unique position in two respects: the variation of the cleft clause position, and the the limitation of its function to text-structuring. Section 5 finishes with conclusions and a discussion that points the way to follow-up work.

2. The construction is an *it*-cleft

A convincing argument that the Chechen construction exemplified in (1) is an *it*-cleft could be found by comparing it with a language-independent definition of the *it*-cleft; especially if such a definition were accompanied by a number of necessary and sufficient diagnostics. This section briefly presents such a definition, showing that its accompanying diagnostics are able to include real *it*-clefts, while they exclude similar but other constructions (such as complement ones). Applying the diagnostics to the Chechen construction under question shows that it should be regarded as an *it*-cleft.

voiceless consonant. The *hw* represents a pharyngeal fricative /ħ/, and the *gh* represents the voiced uvular fricative /ʁ/.

- (4) a. “I would have worked weekends if they’d have let me”, Miss Brechin told me. “I loved my work, we all did”, another survivor said. What was so attractive about it? Partly it was a matter of company – despite what the description of Clark’s says about “chatter”, former composers all remember being allowed to talk – “After all, we were on piece work”, as one pointed out.
- b. It was **the company** [that Joanna Martin missed ___ the most when she left to get married]:⁴
- c. “all the girls helped each other, and if you hadn’t the right kind of type, someone helped you.” [BNC EV] 1224]

The syntactic features of the *it*-cleft in (4b) are visible in the building blocks in Figure 1, while the referential features are shown by the arrows.⁵

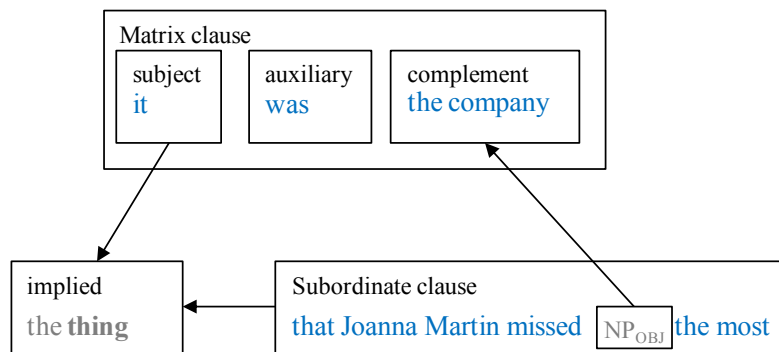


Figure 1 A graphical representation of the *it*-cleft in (4b)

The upper-most building block in Figure 1 is the matrix clause.⁶ Part (3a) of the *it*-cleft definition says that it has to be a copula clause, and Figure 1 identifies the essential elements of such a clause in English: the subject, the auxiliary and the complement. The second building block is a subordinate clause that fulfills the requirements in part (3b) of the *it*-cleft definition: there is co-indexing between the matrix clause’s complement *the company* and the ellipped direct object position in the subordinate clause. The co-indexing relation is indicated by the arrow that starts from the grayed “NP_{OBJ}” and points to the complement. The separation of the matrix clause block from the subordinate (relative) clause is on purpose. The exact nature of the relationship between the two is subject of debate, but there is a consensus that neither the pronominal subject of the matrix clause, nor the complement of that clause serves as the ‘head’ of this relative clause in a simple sense. The subordinate clause is, however, strongly tied to the other elements of the *it*-cleft, as can be seen by the two arrows originating from it: one for the co-indexing with the matrix clause’s complement, and

⁴ The relative-clause examples in this paper have the finite verb in the main clause underlined, the relative clause proper is between square brackets, and the head of the relative clause is in bold face.

⁵ Nested blocks represent syntactic hierarchy, but not all possible hierarchy is shown.

⁶ The copula clause may be as complex as a language allows, and it may contain additional adjuncts or adverbials. A more complex example from English would be: “It must naturally have been the company her co-workers provided.”

one pointing to a *non-overtly expressed* definite NP. It serves as implied head for the subordinate (relative) clause, but it also is the implied antecedent of the matrix clause's subject. The noun heading this NP is one that represents a superset of which the complement is one element, such as "person", "thing" and "activity".

In sum, the arrows in the graphical representation of the *it*-cleft illustrate three essential links characterizing the construction:

- (5) a. a co-indexing from an ellipsed constituent inside the subordinate clause to the matrix clause's complement ("she took up *x*" with *x* = "decorative art-needlework");
- b. an implied NP heading the subordinate (relative) clause ("the thing" heads "that Joanna Martin missed the most");
- c. the NP from (b) serving as antecedent of the matrix clause's subject (the "it" points to "the thing (that Joanna Martin missed the most)").

With the basic ingredients and relations of the *it*-cleft in place, we should see how it distinguishes itself from other, similar, constructions, and what kind of variation exists in cleft types.

2.2. Distinguishing *it*-clefts from other constructions

The characteristics of the *it*-cleft construction that have been discussed in the previous section – both the syntactic as well as the relational ones – are essential in order to distinguish it from other, similar constructions that fall outside the realm of the *it*-clefts. There are two such constructions I would like to compare it with here: (a) the extraposition construction and (b) the copula clause with a complex complement. We will start with the former.

- (6) a. Anyway, it was **good** [that Taff was back to his old self]. [A61:783]
- b. Anyway, it was **his old self** [that Taff was back to __].

The extraposition construction in (6a) looks a lot like the *it*-cleft in (6b), since both consist of a main copula clause, both have a subordinate clause, and in both cases the pronominal subject *it* does not refer back to anything in the preceding context. There is an essential difference, however, and this is illustrated in Figure 2.

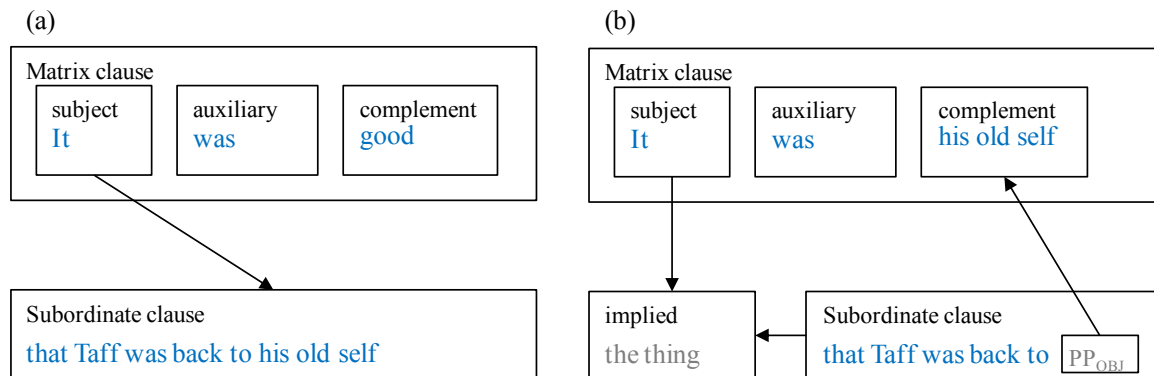


Figure 2 Comparison between an extraposition construction (a) and an *it*-cleft (b)

The graphical displays of the constructions in (6) reveal the difference between them: the subordinate clause of the extraposition construction in (6a) does not have a gap, so there is no arrow from the subordinate clause to the matrix clause's complement, nor is there a non-overtly expressed NP linking the two clauses. The *it*-cleft in (6b) does have a co-indexing link from the ellipted PP's object position in the subordinate clause to the matrix clause's complement, and the two clauses are linked through the non-overtly expressed NP *the thing*.

The difference between *it*-clefts and copula clauses with a complex complement requires a comparison that includes preceding context.

- (7) a. There was someone at the door yesterday.
It was [**my neighbor**, who had a package for me].
- b. That box did not come from the mailman!
It was [**my neighbor**] who had a package for me.

The essential difference between the two constructions in (7) is syntactic in nature, but also involves a difference in linking between the main components, as illustrated in Figure 3.

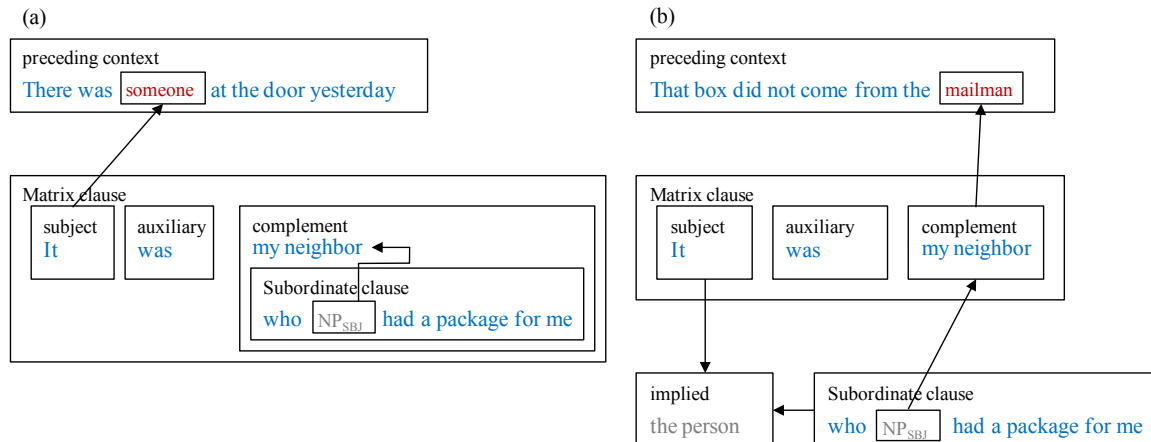


Figure 3 Comparison between a complex complement clause in (a) and an *it*-cleft (b)

The construction in (7a) is nothing more than a copula clause that has a complex NP complement, a noun post-modified by a relative clause: *my neighbour who had a package for me* (see part a of Figure 3). The construction in (7b) looks like (7a), since it has the same subject *it* and the same main verb *was*, but the complement of the main clause is restricted to *my neighbour*; the subordinate clause *who had a package for me* is not a relative clause of the NP *my neighbour* (see part b of Figure 3). The difference between the two constructions only shows up by carefully inspecting the preceding context, which shows whether the pronominal subject *it* has an antecedent or not. The complex complement clause's subject *it* does have an antecedent: the NP *someone* in the immediately preceding sentence. But the *it*-cleft's subject *it* does not have an antecedent

in the preceding context; instead, it points to the non-overtly expressed NP *the person*, which joins it with the subordinate clause.

2.3. Types of *it*-clefts

There are three variations in the *it*-cleft that need to be discussed before we can continue to look at the Chechen *it*-clefts: (a) the adjunct *it*-cleft (b) the predicative *it*-cleft and (c) the informative presupposition *it*-cleft.

The first variant, the adjunct *it*-cleft, is characterized by having a clefted constituent whose co-indexed constituent in the cleft clause is not an argument of the lexical verb in the cleft clause, but an adjunct of, for instance, time (8a,b), location (8c,d) or purpose (8e).

- (8) a. It was **in this year** that Yekuno Amlak, a local chieftain in the Ambasel area, acceded to the so-called Solomonic throne. [Example #45 in Prince (1978)]
 b. It was **the following year** [that the ghost was seen at Althorp for the first time]. [BNC HAF 808]
 c. It was **here** that he died. (Jespersen 1937: 75)
 d. It was **in the woodlanded area** [that a minor road forked left], and it was this road that, as instructed, she steered into. [BNC JYF 258]
 e. It was **because he was ill** that he did not come. (Jespersen 1937: 78)

The clefted constituent in (8d) is the PP *in the woodlanded area*, and this location PP is an adjunct in the cleft clause: *a minor road forked left **in the woodlanded area***. Researchers are divided over the status of adjunct *it*-clefts. Ball (1991) and Patten (2012) do not readily accept them, Jespersen (1937), the first linguist to come up with *it*-clefts, makes mention of them (see 8c,e), and researchers like Hasselgård (2004; 2010) and Gómez González (2000) explicitly accept them. There are good arguments in favour of including constructions like those in (8) as *it*-clefts. First, relative clauses that have an adjunct gap (e.g. [*the time he leaves* _] *is coming near*) have been recognized as genuine relative clauses, and the cleft clause is a kind of relative clause too (Hukari & Levine 1995; Schachter 1973). A second argument relates to the construction's meaning. If the clefted constituent is not interpreted in the cleft clause, then the meaning changes. Saying '*he died*' is a generic statement, but '*he died here*' is much more specific. A clause like '*the ghost was seen at Althorp for the first time*' implies that it was never seen before, while '*the ghost was seen at Althorp for the first time the following year*' does not have that implication. A third argument is that the syntactic building blocks and the connections between them are the same for adjunct *it*-clefts as they are for *it*-clefts that involve an argument of the lexical verb in the cleft clause.

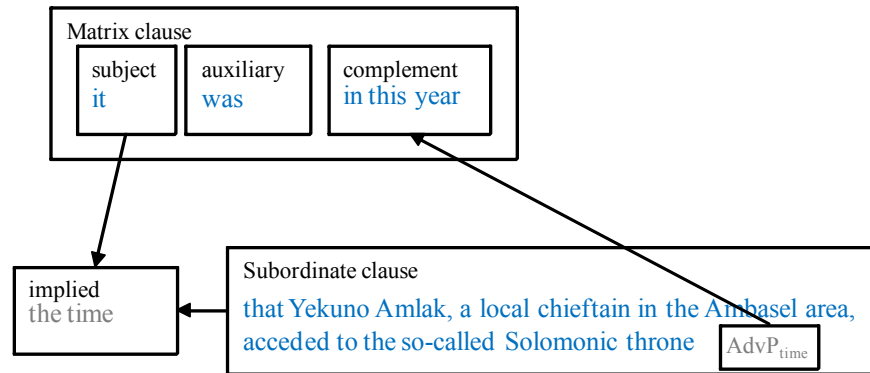


Figure 4 The time adjunct *it*-cleft from (8a)

Figure 4 shows the graphical layout of the construction in (8b), and it in no essential way differs from the graphical representation of the model *it*-cleft in Figure 1. It satisfies the definition of the *it*-cleft provided in (3), since it consists of a matrix copula clause and a subordinate clause, the matrix clause’s subject is pronominal, this subject does not point to something in the preceding context, the complement co-indexes with an ellipsed constituent in the subordinate clause, and both the matrix clause subject as well as the subordinate clause point to an implied NP ‘the time’ that is semantically a superset of the matrix clause’s complement ‘in this year’.

The second *it*-cleft variant that needs to be mentioned here is the ‘predicative *it*-cleft’ (Declerck 1983; Hedberg 1990; Higgins 1979). Since the *it*-cleft builds on the copula clause, and since there are specificational and predicational copula clauses, the *it*-clefts can also be divided into specificational and predicational ones.

- (9) a. The murderer is John.
- b. The murderer is a doctor.
- c. As it is publicly owned, this sector is subject to direct government influence and, frequently, political direction. For practical purposes, it **is the local authorities** [that finance, build and administer these houses], but central government directives, legislation and grants have had direct influence on building rates, rents, finance and sales. [BNC FB2 870]

The specificational copula clause in (9a) specifies the *identity* of the murderer, but the predicational one in (9b) only gives a *characteristic*. The predicational *it*-cleft in (9c) also focuses on a *characteristic* of the authorities that “finance, build and administer” certain houses: they are *local* instead of *central*. Predicational *it*-clefts satisfy the definition in (3). They differ from specificational *it*-clefts in the restriction on the implied noun heading the cleft clause: this noun must be the same as the noun heading the clefted constituent (e.g. ‘authorities’ in (9c)).

The third and final *it*-cleft variant to be discussed in this brief excursion is the ‘informative presupposition’ *it*-cleft, coined as such by Prince (1978). Where an *it*-cleft like (2) is used to express focus (usually contrastive focus) on the clefted constituent against the background of known information in the cleft clause, the informative-

presupposition *it*-cleft has a different make-up in terms of information values: the cleft clause provides the new information.

- (10) It was 10 years ago this month [that young Irwin Vamplew was bopped on the head by a nightstick while smashing windows in Berkeley in order to end the war in Vietnam]. So you can imagine the elation of his parents when he finally emerged this week from his decade-long coma.
 ‘His first words, naturally, where:
 “Down with the Establishment!”’ [Example #46a in Prince (1978)]

The informative-presupposition *it*-cleft in (10) is at the start of a newspaper article, so it is clear there is no textual antecedent for the information provided in the cleft clause. The cleft functions to introduce the information contained in the cleft clause against the time adverbial provided in the clefted constituent. The newspaper story then continues, building on the information provided by the cleft clause. The informative presupposition *it*-cleft satisfies all the requirements of the *it*-cleft’s definition. Its form differs in no essential way from other *it*-clefts; it could readily fit into the graphical display of (8a) in Figure 4.

We may conclude, then, that the definition in (3) satisfies our expectations. It excludes similar but structurally other constructions like the extraposition construction and the copula clause with a complex complement. But in leaving open the syntactic form of the clefted constituent, the status of the co-indexed constituent in the cleft clause and the information statuses of the different parts of the *it*-cleft, it includes adjunct *it*-clefts and informative-presupposition ones. This is crucial for the evaluation of the Chechen construction, as we will see.

2.4. The Chechen construction and the *it*-cleft definition

Given the definition of the *it*-cleft in (3) and its ability to distinguish *it*-cleft from non *it*-cleft, we are now ready to review the Chechen construction and determine if it is an *it*-cleft or not. We will need to show that it satisfies the three necessary and sufficient conditions in (11), which derive straight forwardly from the *it*-cleft’s definition (3).

- (11) *Conditions for an it-cleft*
- a. Structure: the clause containing a cleft construction must be a copula clause supplemented by a subordinate clause that has the form of a relative clause (the “cleft clause”).
 - b. Pronoun: The subject of the copula clause must be an overt pronoun or an ellipited one (in *pro*-drop languages), and it may not have an antecedent in the context preceding the *it*-cleft.
 - c. Co-indexing: The relativized argument or adjunct of the cleft clause must co-index with the complement of the copula clause (the “clefted constituent”).

Since the structural condition (11a) of the *it*-cleft builds on the presence of a ‘copula clause’, we will first need to see how these clauses look like in Chechen.

2.4.1. Copula clauses and subject ellipsis

The main features of Chechen copula clauses are in the agreement pattern and in the ability to have the subject ellipsed (*pro-drop*), as exemplified in (12).

- (12) a. Apti kuotam ju.
 Apti(V) chicken(J) J.PRS⁷
 ‘Apti is a chicken.’
- b. Mylxa du vajn literaturan kyerta gerz? Muott bu-q.
 which is our literature’s main weapon language is-INT
 ‘What is the main weapon of our literature? **It** is language.’ [p34-00002:117,120]

The copula clause in (12a) follows the canonical SOV word order: NP_{subj} – XP_{compl} – Aux, with noun-class agreement between the auxiliary and the complement (which is not unusual, since Chechen is a morphologically ergative language).⁸ The example in (12b) illustrates one type of subject ellipsis: since the subject of the first clause ‘the main weapon’ is identical to that of the second clause, it is not repeated but simply left unexpressed.⁹ This, then, is an example of *pro-drop*, but the ellipsed pronoun in the second clause does have an antecedent in the preceding context: it points to the subject of the first clause.

Example (13) illustrates two situations where a pronominal or ellipsed subject do not have an antecedent in the preceding context.

- (13) a. Vajna massaarna doqqa sovghat du't'amuo hallakbinchu vajn
 to.us all great gift is war destroyed our
 zhimchu maxkahw cwa ghishluo, husam, c'a mettahuottiicha a.
 small in.country one building home house when.restored even
 ‘**It** is a great thing for all of us when even one building, one home or one house
 is restored in our small homeland destroyed by the war.’ [p86-00027:25]

⁷ Glosses: ADJ adjective; ALL allative; B, V, J, D noun class markers; DAT dative; ERG ergative; GEN genitive case; IMPF imperfective past; INF infinitive; INF intensifier; LOC locative; NMLZ nominalizer; OBL oblique case; PL plural; PRS present tense; PST past tense; PSTN past tense with the *-na* suffix; PSTR past tense with the *-ra* suffix; PTC predicational participle; REL relativizer (attributive participle); RFL reflexive.

⁸ Chechen has 6 noun classes, which are signalled on verbs that start with a noun-class prefix. There are 4 possible prefixes (j,v,b,d), and each noun class is defined by the set of noun-class prefixes used in singular and plural. The classes “j-d” and “v-d” are used for feminine and masculine nouns, while the remaining four classes (“j-j”, “d-d”, “b-b”, “b-d”) are used for non-human nouns. The Chechen noun-class agreement is comparable in *function* with gender agreement in other languages. As for the agreement *pattern*: it is only when the complement is *not* an NP that the form of *be* agrees in noun class with the subject.

⁹ The phenomena can roughly be equated to “*pro-drop*”, except that most of the time there is no formal way to recognize the subject. The Chechen complement-verb agreement leaves no agreement suffixes or prefixes that could be linked back to a dropped pronominal subject. Nevertheless, subjects (and sometimes objects) are only ellipsed when they can be inferred pragmatically from the context.

- b. Iza deqq'a vaj noxchii xilarnii, vajna noxchii huma
 It only we Chechens being.and we Chechen thing
 diezarnii aella a daac.
 loving.and said INT not.is
 'It is not purely about us being Chechens and liking our Chechen ways.'
 [p34-00002:35]

The sentence in (13a) is a copula clause with an ellipled pronominal subject that structures roughly like: complement – *be* – when-clause. The ellipled pronominal subject (comparable to English *it*) could be regarded as pointing to the when-clause. The copula clause in (13b) *does* have an overt pronominal subject *iza* 'it', but this subject does not have an antecedent in the text, a situation that is quite similar to the subjects of *it*-cleft constructions.

In sum, we now know how copula clauses in Chechen look like, and we have seen that they may have a pronominal subject that is either overtly expressed or ellipled, and that this subject may link back to something in the preceding context but that it may have no textual antecedent at all too.

2.4.2. Relative clauses

The next element about which the structural conditions of the *it*-cleft in (11a) speak is the presence of a subordinate clause that has the form of a relative clause. Relative clauses in Chechen are always participial ones: the verb that heads the clause appears in a participial form and is inflected in the nominative or oblique case. An illustration of possible head forms is provided in Table 1.

Table 1. *Relative clause heading participial forms*

Verb	Tense	Finite form	Participial form	
			Nominative	Oblique
(copula)	pres, aff	<i>ju, vu, ...</i>	<i>jolu, volu, ...</i>	<i>jolchu, volchu, ...</i>
	pres, neg	<i>jaac, vaac, ...</i>	<i>joocu, voocu, ...</i>	<i>joocuchu, voocuchu, ...</i>
	past	<i>jara, vara, ...</i>	-	-
(auxiliary)	pres	<i>xylu</i>	<i>xylu</i>	<i>xyluchu</i>
	past	<i>xilla</i>	<i>xilla</i>	<i>xillachu</i>
'need'	pres	<i>yeshu</i>	<i>yeshu</i>	<i>yeshuchu</i>
	past	<i>ieshna</i>	<i>ieshna</i>	<i>ieshnachu</i>
'know'	pres	<i>xae'a</i>	<i>xu'u</i>	<i>xu'uchu</i>
	past	<i>xi'na</i>	<i>xi'na</i>	<i>xi'nachu</i>
'go'	pres	<i>jyedu, vyedu, ...</i>	<i>jyedu, vyedu, ...</i>	<i>jyeduchu, vyeduchu, ...</i>
	past	<i>jaxana, vaxana, ...</i>	<i>jaxana, vaxana, ...</i>	<i>jaxanchu, vaxanchu, ...</i>

The copula only has participial forms in the present tense, but there are different variants for the affirmative and the negative. All other verbs (including the auxiliary) have participial forms coinciding with the simple present and past, with the exception

of a few verbs like *xa'a* ‘know’, which have a form for the present participial that differs from the present tense finite form.¹⁰ Class marking (for agreement with the intransitive subject or the transitive object) happens on the first consonant of the verb, but only a limited number of verbs do so.

Important features of relative clauses that are relevant for this paper are (a) agreement and (b) position. There is *case* agreement between the participial heading the relative clause and the head noun: the participial *jolu* in (14a) agrees in nominative case with the head *juq’amettig* ‘relation’, while the participial *jolchu* in (14b) has an oblique case suffix *-chu*, because the head noun *diinatiin* ‘of the animals’ is in a non-nominative case, the genitive.

- (14) a. [*Diesharxuoshca jolu*] **juq’amettig** *ch’oogha* *larjo* *po’eta*.
 Reader.PL.INS J.REL **relation** strongly value.PRS poet.ERG
 ‘The poet values his relationship with the readers very much.’ [p86-00111:9]
- b. *Hinca*, [*Junus vajca voocchu*] **xeenahw**,
 now Yunus(V) we.INS V.NEG.REL.OBL time.LOC(J)
xala du [*yeshush dolu*] **dieshnash** *karuo*.
 difficult D.PRS need.PRS.PTC D.REL word.PL to.find
 ‘Now that Yunus is not with us, it is hard to find the necessary words.’
 [p86-00076:20]
- c. **Avtobus** *gira* *suuna* [*bielxaluoj a iecna, ghaala*
bus see.PSTR 1SG.DAT workmen and take.PST.PTC city
jyedush jolu].
 J.go.PTC J.REL
 ‘I saw a bus taking workmen to the city.’ [m00677:104]

The nominal *class* agreement is completely inside the relative clause: it is between the participial and the nominative-case noun inside the relative clause. This is illustrated by (14b), where the head noun *xeenahw* ‘at the time’ belongs to the *j*-class, but the participial *voocchu*, starting with the prefix *v-*, agrees in class with *Junus* ‘Yunus’, which belongs to the *v*-class. Where the relativized noun is the same as the nominative-case noun inside the relative clause, the class agreement still is with the ‘gap’ left by the relativized noun, as in (14a).¹¹

Relative clauses in Chechen can occur in two positions: canonical (14a,b), where the participial clause immediately precedes the NPs they modify, and extraposed (14c), where the participial clause occurs in absolutely clause-final position. The head noun of the extraposed relative clause can be identified by looking for an NP in the main clause that contains the correct case (nominative if the participial is in the nominative case, and any non-nominative case if the participial is in the oblique), and, where the

¹⁰ Other verbs containing separate participial forms include: *la’a* ‘desire’ (*lae’a* versus *lu’u*), *da’a* ‘eat’ (*do’u* versus *du’u*), *xa’a* ‘sit’ (*xo’u* versus *xu’u*), *gan* ‘see’ (*go* ‘see.PRS’ versus *gu-chu* ‘see.PRS-OBL’).

¹¹ The clause from which the relative is derived would have been *Juq’amettig diesharxuoshca ju* ‘the relationship is with (his) readers’, but when *juq’amettig* is relativized, it is normally no longer visible inside the relative clause.

participial contains a class-prefix and the relativized noun has the nominative case inside the relative clause, by looking for the noun with the corresponding noun-class.

2.4.3. The Chechen construction

The *it*-cleft in (15a) is based on the copula clause *3 butt xaan ju* ‘(it) is three months time’. This kind of copula clause lacks an overt expression of the expletive subject pronoun, just like the one in (12a). There is no overt grammatical subject to the sentence, while the logical subject is the relative clause *that refugees from Samashki have been living without any humanitarian aid*.

- (15) a. **3 butt xaan** ju Semawashkara muhazharsh gumanitarni
 3 month(B) time(I) is from.Samashki refugees(B) humanitarian
 gho doocush wash bolu.
 aid not.having living B.REL
 ‘It is the third month that refugees from Samashki have been living without any humanitarian aid.’ [p86-00085:2]
- b. Semawashkara muhazharsh **qa’a** **battana** gumanitarni
 from.Samashki refugees(B) three.OBL months.DAT humanitarian
 gho doocush wash bu.
 aid(D) D.NEG.PRS.PTC living.PRS.PTC B.REL
 ‘Refugees from Samashki have been living without any humanitarian aid for three months.’

The relative clause is a time-adjunct one, as can be seen by considering its main clause equivalent in (15b). The constituent within this main clause that is relativized is the adjunct *qa’a battana* ‘three months’. The resulting relative clause in (15a) is headed by the temporal NP *3 butt xaan* ‘three months time’.¹²

The construction in (15a) is an *it*-cleft, since it passes the diagnostics defined in section 2.4. According to the *Cleft structure* diagnostic in (11a), an *it*-cleft should consist of a copula clause and a relative clause. The construction in (15a) complies with this. We have seen that the copula construction is *3 butt xaan ju* ‘it is three months time’, and that the relative clause is the remainder of (15a).

¹² The discrepancy between the case of the temporal adjunct in the relative clause (which is dative, according to 15b) and that of the predicative argument in the main clause (nominative, according to 15a) is comparable to preposition deletion in English: “it is **ten o’clock** that the service starts” versus “the service starts **at ten o’clock**”.

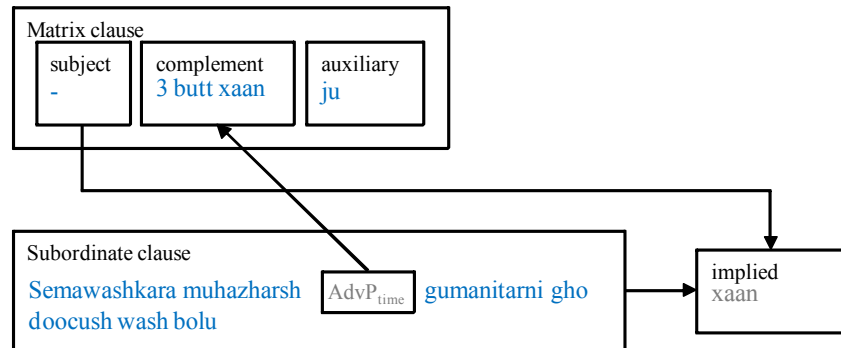


Figure 5 Graphical representation of the Chechen *it*-cleft in (15a)

The *pronoun condition* in (11b) requires the subject of the copula construction to either be an overt pronoun or an ellipsed one, but it should not point back to an entity in the previous context. The construction in (15a) satisfies this condition, since it does not have an overt grammatical subject, and the ellipsed pronominal subject does not link to anything in the preceding discourse. It points, instead, to the implied head of the subordinate clause *xaan* ‘time’, as in Figure 5.

The third condition is the *coindexing* from (11c), which states that the relativized argument or adjunct of the cleft’s relative clause must co-index with the predicative argument of the main clause. This condition holds: the relativized adjunct of the relative clause in (15a) is ‘three months’, and this co-indexes with the main clause complement *3 butt xaan* ‘three months time’.

The combination of *butt* ‘month’ with *xaan* ‘time’ may seem a bit out of place here, but section 3 will show that *xaan* is not always required; it is one of the variations the Chechen *it*-cleft construction allows.¹³ What the presence of *xaan* seems to underscore is the fact that Chechen *it*-clefts is a “predicational *it*-cleft” variant (see 2.3): the clefted constituent is a predication of the cleft clause. Paraphrasing the cleft can take place by a predicational copular clause: “The **time** [that refugees from Samashki have been living without any humanitarian aid] is **3 months**.”

It is safe to conclude that the Chechen construction in (15a) is an *it*-cleft that satisfies all three cleft diagnostics that are defined in section 2.4. The next sections discuss how often this construction occurs in Chechen, in what circumstances, and what purposes they serve.

¹³ The head of the complement *3 butt xaan* ‘three months time’ is *xaan* ‘time’, since there is noun-class agreement between the J-class of *xaan* and the auxiliary *ju* ‘is’; the noun *butt* ‘month’ is of the B-class, and had that been the head of the complement, then the auxiliary would have had the form *bu* ‘is’. The fact that *xaan* ‘time’ is the head of the complement should make it clear that it is not an overt realization of the implied noun *xaan* ‘time’ that heads the subordinate clause and that functions as the anaphor of the ellipsed subject pronoun.

3. Characteristics of the Chechen *it*-cleft

This section discusses a corpus-based study aimed at finding quantitative information about the *it*-clefts in Chechen as well as qualitative information about the variations it comes in.

3.1. Quantifying Chechen *it*-clefts

Quantifying the occurrence of Chechen *it*-clefts can be done using the corpus of texts that has been made available by Zacharski and Cowie (2011), and that have been partly tagged and parsed by Komen (2015). This corpus consists of two parts: a parallel and monolingual one. The parallel part of the corpus contains 324 texts from various sources, where each line of Chechen has been provided with a free translation into English.¹⁴ The monolingual part of the corpus contains 624 texts without translations. The texts have been made available in untokenized plain-text format.

I have transformed these data to *xml* format, and subsequently searched for *it*-clefts using Xquery functions from within the CorpusStudio interface (Komen 2009). I have reviewed and corrected the output of this search manually, and stored the results in a database. An overview of the results is given in Table 2. The search yields a total of 104 *it*-clefts in the “original Chechen” part of the database and 5 in the “translated from English” part, which amounts to 327 and 73 *it*-clefts per 100,000 main clauses respectively.

Table 2. Frequencies of the it-clefts in the two Chechen sub corpora of Komen (2015)

Subject	Clefted constituent	Word order	Illocutionary Force	Original Chechen	Translated from English
overt	time adjunct	canonical	declarative	5 (6%)	0
(none)	time adjunct	canonical	declarative	40 (49%)	0
(none)	time adjunct	reversed	declarative	34 (42%)	3
(none)	time adjunct	canonical	question	2 (2%)	0
(none)	time adjunct	reversed	question	0 (0%)	2
(none)	locative	(any)	(any)	0 (0%)	0
(none)	argument	(any)	(any)	0 (0%)	0
<i>Total</i>				<i>81</i>	<i>5</i>

There are a few observations I would like to make from the results in Table 2. Chechen *it*-clefts can, as would be expected, occur in declarative as well as interrogative forms, but Table 2 makes it clear that the interrogatives are a minority. Another observation is that the number of *it*-clefts occurring in Chechen texts that have been translated from English (these consists of several newspaper articles) appears to be

¹⁴ These free translations have been provided by native Chechen speakers “without an intermediate Russian stage” (Cowie 2011).

considerably lower than the number of *it*-clefts found in the other texts: 73 versus 327 respectively. This shows, perhaps, that the *it*-cleft as it is used in Chechen deviates considerably from its usage in English: Chechen apparently uses *it*-clefts where they do not tend to occur in English, and when a translator (even a native speaker) doesn't see particular constructions in the source text, he tends to not reproduce them in his translation.

Chechen *it*-clefts only appear to have a time phrase, a temporal adjunct, as clefted constituent: no argument or locative adjunct *it*-clefts have been found that conform to the three *it*-cleft diagnostics in section 2.4. There is one interesting observation that needs to be made about the form of the temporal adjuncts serving as clefted constituents in Chechen. We have searched for two possible forms: NPs headed by a time noun in the nominative and in the locative case. This last type of NPs is comparable to English PPs. The noun *sho*, for instance, means 'year', but when it appears in the locative case, such as in *hoqu sharahw*, where *hoqu* is the inflected form of the near demonstrative, it has to be translated with a PP 'in this year'. Interestingly, the temporal adjuncts serving as clefted constituents are all NPs in the nominative case. In this restriction on the grammatical category of the main clause predicative argument, then, Chechen is much more restrictive than a language like English, which allows for a range of predicative argument types.¹⁵

3.2. Variation in the *it*-cleft

The variation in terms of illocutionary force (declarative versus question) as visible in Table 2 coincides with similar variation in English *it*-clefts, but there are a number of other variations in the Chechen *it*-clefts that have no counterpart in Present-day English: (a) the presence or absence of an overt pronominal subject, (b) the order of the relative clause with respect to the main clause, (c) the presence or absence of a generic temporal head noun. We will review these variations in turn.

From the total of 81 original Chechen *it*-clefts, five samples have an overt pronominal subject. One of these is shown in (1), repeated here as (16).¹⁶

- (16) Hara **cwa** **butt** **xaan** ju Noxchiin Respublikan Q'ooman bibli'otekan
 this one monthtime is Chechen Republic's national library's
 bielxaxuosha de-byysa ca lyerush q'ahwyegu.
 employers day-night not regarding toil
 'It is one month since the employees of the Chechen Republic's national
 library have been working both day and night.' [m00225.1]

¹⁵ English allows, for instance, for PP *it*-clefts (e.g. "It was [PP in the summer] that I met her"), and CP *it*-clefts (e.g. "It was [CP because we had left too late] that we needed to hurry").

¹⁶ An anonymous reviewer suggests that the demonstrative pronoun *hara* 'this' in *it*-clefts such as (16) should not be interpreted as an independent demonstrative pronoun (the grammatical subject of the sentence) but as a modifier of the temporal NP. I argue for a subject interpretation of *hara* in these sentences, since such an interpretation seems to do more justice to the referential nature of the demonstrative.

The example in (16) is the opening line of a newspaper article, which makes it clear that the pronominal subject *hara* ‘this’ cannot be anaphoric (its graphical representation is in Figure 6a). Given the marginal number of *it*-clefts with an overt subject pronoun, this structure could be an innovation, in which case we should see it occur relatively more frequently in recent texts, or a remnant of the past, in which case we should see it relatively more often in older texts. The current study only has corpus data from one time period available, so that we cannot look at the dating of the texts. It may be that the use of an overt subject is an innovation, since there is no grammatical need for copula clauses used in *it*-clefts to have an overt subject.

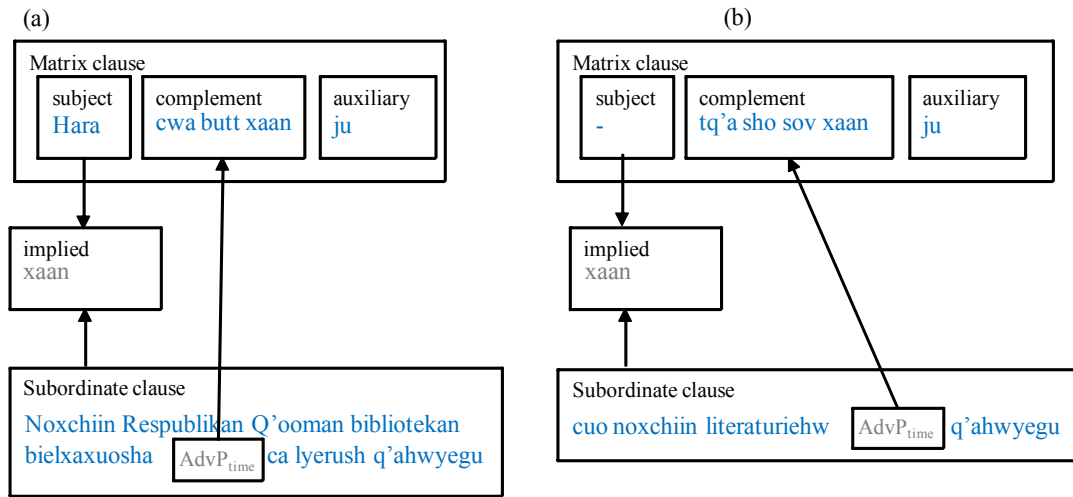


Figure 6 Structure of *it*-clefts with an overt subject (a) and with a reversed order (b)

As for the variation in word order, the relative clause may occur sentence-finally as in (17a), or sentence-initially as in (17b). The sentence-final position of the relative clause is comparable to the English *it*-cleft word order, which also has the relative clause follow the main clause; this is why I will label it “canonical”. The sentence-initial order has no equivalent in English, and I will label it “reversed”, since the order of the main clause and the relative clause is reversed (the structure of this reversed *it*-cleft is shown in Figure 6b).

(17) a. (T'aehwaluonan ojla a jiesh, xaza kyg tuuxush jina ghishluo ju hara.)

Duqa xaan jara ooxa hoqu ghullaqie satesna a.
 much time was we this to.matter hoped &
 ‘(It is a beautifully built building built for the future.)
 We have waited for this for a long time.’ [p86-00185.9]

b. (Sa'iev Wumar literaturiehw kerla stag vaac,)

cuo noxchiin literaturiehw q'ahwyegu **tq'a sho sov xaan** ju.
 he Chechen in.literature toils 20 year more time is
 ‘(Umar Saiev is not a new man in Chechen literature.)
 He has been working hard in Chechen literature for over twenty years.’
 [p86-00184.15]

Even though the word order of “reversed” *it*-clefts like (17b) has the relative clause precede the main-clause’s predicative argument, the two taken together do *not* form a syntactic constituent. An effort to combine the two into one constituent would result in the English equivalent of *‘‘It is the more than twenty years time that he has been working hard in Chechen literature’’. The “reversed” *it*-cleft in (17b), then, conforms to the *it*-cleft diagnostics in section 2.4: all conditions are met, and none of the diagnostics prescribes a particular word order of the main clause. The question obviously arises what the function is of a reversed word order for Chechen *it*-clefts. I will leave this for further research. Suffice it to say that numerically, reversed *it*-clefts occur only slightly less than canonical ones.

The third and final kind of variation on the form of the Chechen *it*-cleft that is to be discussed in this section involves the presence or absence of a generic head noun that is added to the clefted NP. Table 3 gives a more general overview of the variations found in the clefted NPs.

Table 3. Make-up of the clefted NP

#	Attribute	Noun	Modifier	Head	Canonical	Reversed	Total
a	(yes)	<i>sho</i>	<i>sov, gergga</i>	<i>xaan</i> ‘time’	4	3	7
b	(yes)	<i>butt, sho, sahw</i>	(no)	<i>xaan</i> ‘time’	6	2	8
c	(yes)	(no)	(no)	<i>xaan</i> ‘time’	18	24	42
d	(yes)	(no)	(no)	<i>de</i> ‘day’	5	2	7
e	(yes)	(no)	(no)	<i>sho</i> ‘year’	14	7	21
				<i>total</i>	47	38	85

The column “Attribute” contains “yes” in all cases, which indicates that all the clefted NPs need to have some kind of attributive element: a numeral (17b, 18a,b,d,e) or a quantifier (17a, 18c). This underscores the observation made earlier in this paper, that Chechen *it*-cleft are attributive ones. The examples in (18) give an overview of the five NP categories from Table 3.

- (18) a. **36 sho sov xaan ju** [cuo zhurnalistikiehw q'ahwyegu].
 36 year more time J.PRS he journalism.LOC labour.PRS.REL
 ‘He has been working in journalism for over 36 years.’ [m00263.14]
- b. **Di’ sho xaan ju** [shaa Zajna’a quzahw buolx ben].
 four year time J.PRS self Zajna.ERG here work do.PRS.REL
 ‘Zajna has been working here for four years now.’ [p86-00177.15]
- c. **Duqa xaan jaac** [predprijati juxamettahuottuor dwaaduoliina].
 much time J.PRS.NEG enterprise renovating start.PST.REL
 ‘Not long after, he started renovating the enterprise.’ [m00185.4]
- d. **Shi de du** [txo gyyrienan jalta chudierziina devlla].
 two day D.PRS we.PRS.REL autumn.GEN harvest bring.in.PST.PTC finish.PST.REL
 ‘We finished bringing in the autumn’s harvest two days ago.’ [m00470.44]

- e. **5 sho du** [cuo baq'uonashlarjaraŋ organaŋaŋkahw q'ahwyegu].
 5 year D.PRS he law.protection.GEN organisations.LOC labour.PRS.REL
 'He has been working in law-protection organizations for five years.'
 [m00256.7]

Lines “a” and “b” in Table 3, exemplified by (18a,b), cover the 15 clefted NPs that have two nouns in the NP: a noun that specifies the quantity of time (*sho* ‘year’, *butt* ‘month’, *sahwt* ‘hour’), as well as a generic head noun (which can, apparently, only be *xaan* ‘time’). Lines “c”–“e”, exemplified by (18c-e), count the remaining 70 clefted NPs that only have one head noun, and they show that *de* ‘day’ and *sho* ‘year’ can function as head nouns on their own. The columns “Canonical” and “Reversed” divide the occurrences over *it*-clefts in the canonical and reversed order respectively. Table 3 makes clear that the make-up of the clefted NPs does not differ significantly between the canonical and reversed ones.

It appears, then, that only *xaan* ‘time’ can appear as the second noun in an NP and so function as the generic head noun, but the presence of such a generic noun is not required.¹⁷

3.3. The function of Chechen *it*-clefts

A qualitative comparison between Chechen and English *it*-clefts requires us to take a closer look at the function of Chechen *it*-clefts. While *it*-clefts have been recognized as fulfilling a number of different functions in language in general, the two that seem to occur most often are: (a) focusing and (b) text-structuring.¹⁸ This section considers to what extent the Chechen *it*-clefts are used for these two main functions.

Let us first consider the possibility that they function as a focusing device. An argument in favor of this analysis would be that the position of the clefted constituent, the immediately preverbal one, is that of the focused constituent. This is in line with the findings on focus in Chechen as a whole (Komen 2007). However, position as such is not sufficient in this case, since Chechen is an SOV language, and the most natural position for a complement is the preverbal one anyway—focused or not. This is the same problem as that of recognizing object focus from a transitive sentence with SOV word order: the position as such coincides with the unmarked word order, so it does not necessarily point to constituent focus.

If the SOV word order is not a sufficient indication of focus, then the question arises whether there are other indicators of focus-hood. There are a few standard indicators of focus-hood: (a) the presence of focus particles, and (b) the presence of a question word.

¹⁷ The optional presence of the generic head noun *xaan* ‘time’ may play a role in the discussion on the *it*-cleft’s syntax. Patten, for instance, analyzes the English cleft’s pronoun *it* as “part of a definite-like description, equivalent to *the one* or *the thing*” (Patten 2010: 125). Chechen seems to make this connection more explicit.

¹⁸ Functions that have been associated with *it*-clefts are: (a) focus (English) (Jespersen 1937; Quirk et al. 1972), (b) text-structuring (English, Norwegian, Swedish) (Collins 1991; Hasselgård 2004), (c) avoid subjects containing new information (French, Norwegian, English) (Faarlund et al. 1997; Prince 1998), (d) presenting information as a known fact (English) (Prince 1978)

As for focus particles, I would like to briefly discuss three: *bien* ‘only’, *=m* ‘FOC’ and *a* ‘FOC’. The Chechen word *bien* is the equivalent for English ‘only’, but while the English ‘only’ can successfully be used to find out whether a constituent is highlighted or focused, this has so far not been done for Chechen *bien*. The reason may be that *bien* is a negative concord particle, which influences the syntax of the sentence and distorts its use as indicator (Komen 2010).

The clitic *=m* seems to be used for highlighting or focus, but little research has been done on it (Nichols 2011).¹⁹ Interestingly enough, the only place where the *=m* clitic occurs in the database of Chechen *it*-clefts is within the cleft clause, as shown in (19).

(19) (Txojsi’ liicha vyedu... Xi shiila xyly - loomahw lo du deeshash, i bahwaniehw du aexka, daa’ima a booxurg sanna, Terk diestara.)

[Dogha danza =m] **duqa xaan ju.**
rain D.do.PST.NEG FOC much time J.PRS

‘(The two of us go bathing... The river is cold – the snow is melting on the mountain, and this is why the Terek River has swollen up, as it always does in the summer.)

It is a long time that it has been raining.’ (More lit. ‘It’s been a long time since there was (any weather with no) rain.’) [m00129:28]

The preceding context describes how the river Terek has ‘swollen’, and the *it*-cleft sentence in (19) underscores this observation: it is raining all the time; so much that it is a long time ago there was a moment that rain was *not* falling.

The particle *a* can sometimes function as a focus particle too (Good 2003; Kudrinskii 2014; Nichols 1994). Such is the case in examples (13a) and (13b), where it is translated as “even”.²⁰ The corpus of 81 *it*-clefts contains 3 occurrences of *it*-clefts where the intensification particle *a* modifies the clefted constituent. One of these is shown in (20).

(20) Tq’a ysh mella=a sixa xiica jiezash xilla jolu xaan
but they however fast change needing been being time

t'exjaella **shiitta-qojtta sho a** du.
surpassed 12-13 year INT is

‘But the time that they should have been replaced as fast as possible, has now surpassed even twelve-thirteen years.’ [m00249:70]

¹⁹ Nichols analyzes the clitic *=m* in the related Ingush language as marking contrastive focus if the constituent it attaches to occurs clause-initially. Further research is needed to check this out in Chechen, to compare it with contrastive topics, and to find out what the function of the *=m* clitic is when it attaches to a constituent further down in the clause or to a non-nominal clause. A proper investigation into these and other phenomena in Chechen would be greatly facilitated by the presence of a corpus that is syntactically parsed.

²⁰ The particle *a* can have several different functions. It can be used as focus particle (it is referred to as “intensifier” in this usage), but also as negator (for instance turning *huma* ‘thing’ into *humma a* ‘nothing’), as coordinator (for instance *daada a, naana a* ‘father and mother’) and as co-subordinate clause marker (for instance *iza a dina* ‘having done that’).

The combination of a focus particle with the approximate time reference ‘twelve to thirteen years’ sounds a bit awkward in English. There does seem to be some kind of contrast, since the time ‘twelve-to-thirteen years’ is compared with ‘as fast as possible’.

Another indication that a constituent is in focus is, depending on the language, the presence of a question word. Komen’s (2007) study on focus used the presence of a question word as an indicator for focus in Chechen, which is enough reason to try and use it as indicator for focus here too. Only two of the original 81 *it*-clefts have a clefted constituent containing a question word, and these instances are shown in (21a,b).

- (21) a. **Miel** **xaan ju** [vaj karzaxdevlla]?
 how.much time is we put.up.REL
- Ja **miel** **xaan ju** [parghatdovla ghierta, booxush, hwiiza]?!
 or how.much time is get.free to.try saying torment.REL
 ‘How long is it that we have put up (with it)?
 Or how long is it that we torment ourselves, saying we try to get free?’
 [m00300:73-74]
- b. [As horsh dyycu, hwiexado] **miel duqa xa**n ju.
 I these talk.REL teach.REL how much time is
 ‘I have been talking and bringing it up for a long time.’ [p34-00002:21]

The two examples in (21) do contain a question word in the clefted constituent, but it should be noted that both of them are rhetorical questions. The reason they are used is not to elicit an answer, but to convey emotion. As such they do convey a form of intensification. The example in (21b) is from the parallel part of the corpus, and has a “reversed-order” *it*-cleft. The question word again does not serve its role as question-elicitor, but it does convey intensification.

The corpus also provides us with examples like (22), which are from the group of clefts having a clefted NP like line “a” in Table 3, which illustrate that the clefted constituent can have characteristics that are quite unlike those of contrastive focus.

- (22) «Phwarmat» quollajelcha dyyna swa, **30 sho gergga xa**n ju
 Phwarmat created.when since from 30 year almost time is
 so hoqu t’iehw buolxbiesh volu.
 I this on work doing am.REL
 ‘I have been working at this since the creation of “Pharmat” – for about thirty
 years.’ [p86-00064:40]

Whereas constituent focus identifies and enforces one particular variant, explicitly or implicitly contrasted with alternatives (as for example Krifka (2005)), the clefted constituent in (22) contains an approximate time reference, which is an *open* set of alternatives, which makes it much *unlike* focus.

In sum, apart from the focus associated with the preverbal position, which could be labeled as ‘unmarked’, there are a limited number of examples where the clefted constituent of the Chechen *it*-cleft are highlighted, but this never seems to be the main rationale for using a cleft construction.

This brings us to the second possibility for the function of *it*-clefts in Chechen that this research focuses on: that of “discourse segmentation” or “text structuring”. If the *it*-clefts are used to indicate discourse boundaries, then we expect them to occur (a) story-initially, (b) paragraph-initially and (c) story-finally. These possibilities are in line with Johansson’s (2002) interpretation of Swedish data and Hasselgård’s (2004) Norwegian data; they recognize the usage of clefts for “Topic launching”, “Topic linking” and “Summative”.

In order to verify the discourse position of *it*-clefts in Chechen, it is (at least sometimes) better to show a larger stretch of a text, so that we can better judge whether the position of the cleft coincides with a paragraph start, transition or end. This is what has been done in our first example (23), which is an article that contains a forum discussion on the usage of Chechen as the principal language in elementary schools.

- (23)²¹ a. [24]Kati, it is not right for us to come to this magazine’s office and tell them that there is practically nothing being done and that talking is a waste of time.
- b. [25]What is the duty of a magazine? [26]To listen to your, my and their opinions, write them down in some way and deliver them to people. [27]When these guys, another magazine or another newspaper raise an issue, talk about it over and over again, then the government can do nothing else but what it is supposed to do. [28]But nothing happens if we stay away from discussing a problem.
- c. Vaj i dyycush dolu duqa xaan ju.
we.INC that speaking are.REL long time is
We have been talking **for a long time** about it. [p34-00002.29]
- d. [30]What you said in the beginning that two to three grades in school should be in Chechen is the topic we have so far been talking about. [31]It will happen, as long as we keep talking about it. [32]Without giving in. [33]If we speak about it, we should not speak about it superficially. [34]It is not enough to speak about switching elementary classes to Chechen, when it is not understood why it needs to be switched, and it is incumbent on us to provide a foundation for that. [35]It is not purely about us being Chechens and liking our Chechen ways. [36]There is more to it than that.

²¹ [24] Kati, vaj hoqu zhurnalie a daexkina, vaj hoqaerga prakticheski diesh humma a daac, q'amielash dar erna du baexcha, niisa daac. [25] Zhurnalan dieqar hun du? [26] Hwuuna, suuna, hoqaarna, qiechaarna xietash dolchynga la a dyeghna, cwana kiepiehw dwaajaazdina, naaxie dwaqaachuor du-q. [27] Hoqaara, qechu zhurnaluo, gazieta shaa aj'a a aj'ina, jux-juxa diicicha, dan humma a ca xylui, t'aaqqa do wiedaluo shaa diirig. [28] Tq'a diica a ca dyycush wad ditcha, cunax humma a ca xylu. [29] Vaj i dyycush dolu **duqa xaan** ju. [30] Ahw dwaavuolalush aellarg, noxchiin mattahw shi-qo klass xila jiezash xilar, dyycush wash du vaj cq'achunna. [31] Cq'a macca a ghullaq xir du hwuuna cunax, vaj diicichahwana. [32] Q'ar ca lush. [33] Vaj dyycush xilcha, t'exula diica ca dieza. [34] Jyhwanacara shkola noxchiin mattie jaaqqa jieza baxarx tye'ush daac, hunda jaaqqa jieza qietash ca xilcha, cunna bux kechbar - iza vajna t'iehw du. [35] Iza deqq'a vaj noxchii xilarnii, vajna noxchii huma dezarnii aella a daac. [36] I doocurg qin a humnash ma du cigahw.

The paragraphs in (23a-d) are the start of a reaction from one participant in the forum. He addresses the interviewer with “Kati” in paragraph (23a). Paragraph (23b) opens with a typical topic-introducer: a question. Paragraph (23d) likewise identifies a clear change of topic, which is retained as “it” throughout this last paragraph. The line in (23c) contains the Chechen *it*-cleft. It functions as a transition between the previous paragraph (23b) and the next one (23d). The link with the previous paragraph is by the pronoun “it”, which refers to the whole clause “nothing happens if we don’t discuss the problem”. The link with the next paragraph is clear too, because the start of (23d) copies the “we have been talking” element. In sum, the Chechen *it*-cleft here functions as an episode boundary marker.

- (24)²² a. [249]Abdullah: The development of the Chechen language and literature depends mostly, as you said, on a school. [250]The fact that the elementary school should be in the Chechen language is beyond any doubt. [251]Not only elementary school, middle school too should be in the Chechen language. [252]However, as of today, we shall have the financial capabilities to switch only elementary school to the Chechen language.
- b. I would like to say a few words about it, because
 i problemash ooxa tollush dikka xaan ju.
 those problems we.EXC investigating long.INT time is
*we have been studying the problems **for a long time**.* [p34-00002.253]
- c. [254]The elementary school was switched to the Chechen language. [255]It was at the end of the past century. [256]I was the one who paid visits to the Regional Committee at the time of the switch.

There is one more *it*-cleft in this same text which we may consider. Line (24a) starts the contribution of Abdullah, a participant in the forum. This first paragraph gives some background, and (24b) finishes this introduction by announcing that he is going to say “a few words” about this matter. The content of what he then says starts in (24c). Again we see that the *it*-cleft is in a position where it helps finish off one topic, and introduce another one.

A total of 14 *it*-clefts from the corpus (which amounts to 13%) is located at the beginning of a story or report. We can see the English translation of those that occur in the parallel part of the corpus in (25).

²² [249] Abdulla: - Noxchiin muott, literatura qi'ar duqax dolchunna, ahw ma aallara, doozadella du shkolax. [250] Jyhwancara shkola noxchiin mattahw xila jiezar - iza cwaellig cwana aaghuor shiekuonie huma daac. [251] Jyhwancara shkola hwovxa, juqq'iera shkola a xila jieza noxchiin mattahw. [252] Baq'du, vajn taxanleerachu diinahw taruonash xir jaac noxchiin mattie jyhwancara shkola jaaqqa bien. [253] Asa ocunax masiex duosh eer du, hunda aelcha i problemash ooxa tollush **dikka xaan** ju. [254] Jyhwancara shkola noxchiin mattie jaeqqina jara. [255] I dara dwaadaxanchu bweesheran chaqqiengahw. [256] I joqqachu xeenahw obkomie liellarg so vara.

- (25) a. It was **not long ago** that a medical insurance ZAO (closed shareholders company) called "Maks-M" opened another branch in Grozny, at Pervomaiskaya street #85. [p86-00063.2]²³
- b. A team of the Achkhoy-Martan financial department has hoped **for a long time** that a new building would be built. [p86-00027.2]²⁴
- c. (It is **the third month** that refugees from Samashki have been living without any humanitarian aid). [p86-00085.2]²⁵
- d. It is **the 5th year** since the branch of the PTU #113 was opened in the village of Samashki of the Achkhoy-Martan district. [p86-00110.2]²⁶
- e. It has been **at least 25 years** since a literature group called "Shovda" has been working at the newspaper's group "Gums" in the city of Gudermes. [p86-00130.2]²⁷

All of the examples in (25) provide clear opening sentences for a text: they anchor a theme in a timeframe. It is interesting to see that four of the five were translated with English *it-clefts* by the native speakers of Chechen who cooperated in establishing the corpus. What the time adjuncts in the clefted constituents do is establish a link between the whole of the article and the real world. Such a link is a kind of scene-setting, and is usually not something that is developed as topic later on.

The monolingual part of the corpus contains the remaining 9 instances of *it-clefts* that start off a story or report. We have already seen one of them in (16), where it was brought up as illustration of Chechen *it-clefts* having a non-anaphoric pronominal subject. In fact, it should be noted here that four of the five *it-clefts* that use the demonstrative pronoun *hara* 'this' as subject are story-initial ones. The reason for this is probably the avoidance of ambiguity: the near demonstrative *hara* can quite easily link up with something in the previous sentence, or with the previous sentence as a whole, but this is impossible if there is no previous sentence.

What about the "Summative" function Johansson (2002) found for Swedish? Is the Chechen *it-cleft* used for that discourse segmentation function too? The number of times an *it-cleft* is used to *finish* a story is very limited. I have only found one example of this in the whole corpus, and this example is shown in (26).

- (26) Taamasha a baac, **tq'e qojtta sho xaan** ju cuo quzahw q'ahwyegu.
surprise & not.is 20 13 year time is he here toils
'It is no surprise that he has been working here for thirty three years.'
[m00233.11]

²³ **Duqa xaan** jaac ZAO <Maks-M> c'e jolchu medicinie straxovani jaran kompanis Syelzha-Ghaalin Pervomajski uuraman No. 85 jolchu c'iiniewh shien roghiera fili'al dwaajillina.

²⁴ **Duqa xaan** jara T'iehwa-Martan rajonan fin'otdelan kollektivuo kerla ghishluo jarie satyysu.

²⁵ **3 Butt xaan** ju Semawashkara muhazharsh gumanitarni gho doocush wash bolu.

²⁶ Hara 5-**gha sho** du T'iehwa-Martan rajonan Semawashka jyrtahw Syelzha-Ghalin No. 113 jolchu PTU-n fili'al swajillina.

²⁷ **Laxxara 25 sho xaan** ju Gysie ghaalin «Gums» c'e jolchu gazietan redakciehw «Shovda» c'e jolu literaturan kruzhoz bolxbiesh jolu.

The newspaper story that finishes with (26) is a small biography in praise of a doctor called Umar Astamirov, and it speaks of how good he is at his job and how well he relates to patients and people. The concluding remark about the number of years he has been working at this particular hospital is a worthy end of the biography, underlining his dedication to the work, and the hospital commitment to keep him on.

The scope of this article is too limited to discuss all the remaining examples of *it*-clefts in Chechen, but what we have seen so far is that the construction is used as a story-opener (to set the scene for the rest of the story), and that it can function as a paragraph transitioning device in other situations. This “paragraph transitioning” function compares with Johansson’s “Topic linking” one, where one discourse topic (in the clefted constituent) is linked to a subsequent discourse topic (which is in the cleft clause).

4. Comparison with other languages

The characteristics and the function of the *it*-cleft found for Chechen compare well with some of the characteristics of *it*-clefts in Scandinavian languages. But, surprisingly perhaps, English itself has been noted as using *it*-clefts for text-structuring to some degree.

4.1. Norwegian

Gundel (2002) used the book “Sophie’s world” and its Norwegian translation to compare the use of clefts in English and Norwegian, and found that while English predominantly uses the *it*-cleft to express focus on the clefted constituent against the background of a presupposition in the cleft clause, the Norwegian *it*-clefts prefer not to have a cleft clause that is “strongly presupposed”. Gundel notes that 20 of the 23 Norwegian clefts that were not translated as clefts into English were of the “informative presupposition” type: their cleft clauses contain the relatively new information that is being conveyed, as for instance (27a,b).

- (27) a. (Helt innerst gjorde den en brå sving som gjerne ble kalt "Kapteinsvingen".)
 Det var **nesten bare på lørdager og søndager**
 it was almost only on Saturdays and Sundays
 [at det gikk mennesker her].
 that it went people here
 ‘(Furthest in made it an abrupt turn that usually was called "Captain’s bend".)
 It was almost only on Saturdays and Sundays that it went people here.’
- b. (At the end of the road there was a sharp bend known as Captain’s Bend.)
 People seldom went that way except on the weekend. (Gundel 2002 Ex. 26)

Gundel notes that the information in the cleft clause in (27a) and its English equivalent in (27b), the fact that people frequent the place “Captain’s bend”, is new, and that the focal accent would be within the cleft clause. Some of the Norwegian *it*-clefts start a new paragraph:

- (28) a. **Nå** er det [at guden Heimdal får en lys ide].
 Now is it that the.god Heimdal has a bright idea
 b. Then the god Heimdal has an idea. (Gundel 2002 Ex. 28)

The use of an *it*-cleft at this point is not related to focus on the clefted constituent ‘now’, but rather serves to set out the paragraph break in a more overt manner.

Gundel also notes that there are many instances where English *it*-clefts are translated as *it*-clefts into Norwegian too; most of these are instances of stressed-focus *it*-clefts. She further notes, that it would be perfectly grammatical to use *it*-clefts in English or Norwegian in the places where the version of Sophie’s world does not use them. This prompts her to conclude that Norwegian clefts perform the function of “mapping information structure [more] directly onto syntactic structure”: (a) they allow a stronger division into relationally new versus given, which is especially apparent in clefting questions, and (b) they allow relationally new information to remain out of the subject position, something also noted by Faarlund, Lie and Vannebo (1997).

4.2. Swedish

Johansson (2001; 2002) reports an extensive comparative study of clefts used in Swedish versus English. He notes that large part of Swedish *it*-clefts contains proniminials or adverbials as clefted constituent, who are more of the informative presupposition type, as in (29).

- (29) a. Och det var **så** [Jack London och alla de andra kom med].
 and it was thus Jack London and all the others came with
 ‘And that’s how Jack London and all the rest got into the picture.’
 b. Det var **då** [våra bjornbårsdagar kom].
 it was then our blackberry.days came
 ‘That’s when our blackberry-picking days began.’ (Johansson 2001 Ex. 44, 45)

The example in (29a) has the “summative” function: it is used at the end of a story (or the end of a part of a story) to convey a conclusion. The cleft in (29b) functions more as the beginning of a paragraph, introducing the thematic activity of “blackberry picking” against the background of a clearly set-out temporal point of departure. What both have in common is their text-structuring function, and the fact that they translate well as reversed *wh*-clefts into English, a construction of which Collins (1991: 145) already noted that it appears to have a text-structuring function.

4.3. English

Hasselgård (2004) takes up and extends Johansson’s research with data from Present-day English. Regarding the correlation between the type of clefted constituent and the type of cleft construction Hasselgård re-emphasizes the findings by Collins (1991: 112) and Prince (1978: 899) that informative presupposition *it*-clefts tend to have a clefted constituent that is an adverbial. This is completely in line with what has been found for Chechen: this language only has time-adverbials as clefted constituents, and all of the resulting clefts are of the informative presupposition type. Hasselgård

notes several functions of *it*-clefts in English, such as the one of “transitioning”, illustrated by (30b).

- (30) a. The Villa Somalia which was Siad Barre’s official residence in Mogadishu still lies abandoned, guarded by a handful of young men from the United Somali Congress the rebel force which took control of Mogadishu at the end of January,
- b. But it was **in one of the office buildings** [that I discovered the letters], thousands of them, addressed to His Excellency President Mohammed Siad Barre but all unopened.
- c. I picked up one from Britain. It had been posted in September nineteen eighty-eight and was signed by a retired schoolteacher from Guildford in Surrey, writing on behalf of Amnesty International to plead for the release of a blind Somali preacher who’d been imprisoned for his religious beliefs. (Hasselgård 2004)

The paragraph in (30a) is the start of a story, but the story proper only sets out in (30b) with the introduction of the “discovered letters”. This then becomes the theme that is picked up subsequently in (30c). The informative presupposition *it*-cleft in English, then, functions as a transition between two types of paragraphs, it helps in structuring the text.

While the main function of the *it*-cleft in English undoubtedly is that of expressing stressed focus, it appears that English too, like Swedish, Norwegian, and Chechen, uses the *it*-cleft as a text structuring device. The difference between the languages appears to be one of quantity: English uses the *it*-cleft for text-structuring infrequently, Norwegian and Swedish more so, and Chechen stands out, because it *only* uses the *it*-cleft for this text-structuring purpose.

5. Discussion and conclusions

Chechen, a North-East Caucasian language, contains *it*-clefts with a number of typologically interesting characteristics: (a) they *it*-clefts are predicational (as opposed to specificational) and (b) of the informative-presupposition type, (c) the clefted constituent grammatically always is an NP, but (d) the only licit ones are temporal adjuncts. The function of the *it*-cleft in Chechen is exclusively that of text-structuring: setting out a story or providing a clear transition to a new paragraph.

This paper introduces and adopts a potentially cross-linguistically applicable method of determining whether a construction can be labeled as *it*-cleft or not. The definition that is proposed here takes a language’s copula construction as a starting point. For a construction to be accepted as an *it*-cleft, the copula construction must be extended with a relative clause, the relativized argument or adjunct must coindex with the predicative argument of the main copula clause, and the pronominal subject of the copula clause must satisfy strict criteria. The *it*-cleft definition presented in this paper should make it possible for researchers into other languages to identify *it*-cleft constructions in an objective way, and this opens the path for a broader investigation into the forms and functions associated with this construction typologically.

The Chechen construction argued to be an *it*-cleft satisfies all the criteria that are proposed for an *it*-cleft, and it introduces variation in a number of typologically interesting areas: (a) the expletive pronominal subject is either left unexpressed or is conveyed by a demonstrative pronoun, (b) the clefted constituent necessarily is an NP with an attributive element, but it may optionally contain a generic head noun *xaan* 'time', which comes in addition to a more informative head noun (such as *butt* 'month', *sho* 'year', *sahwt* 'hour'), and (c) the order of the relative clause with respect to the main copula clause varies between the canonical model, where the relative clause follows the copula clause, and the 'reversed' order, where the relative clause precedes the main clause.

The function of the *it*-cleft in Chechen is derived by looking at examples in their wider context, which shows that it is used for text-structuring. The same function of text-structuring also appears to be present in a large part of the *it*-clefts found in Norwegian and Swedish, but English too uses informative presupposition *it*-clefts for text-structuring.

References

- Almirza, M. 2005. Our interview -- on the path of development. In *Dajmuoxk*. 54.
- Arsanukaev, Wabdulla Magomedovich. 2008. *Noxchiin literatura: 9-chu klassana uchebnik*. [Chechen literature: primer for grade 9.] Grozny: Abat.
- Baerman, Matthew, Brown, D. P. and Corbett, Greville G. 2002. Case Syncretism in and out of Indo-European. Proceedings of the CLS 37: The main session. Papers from the 37th meeting of the Chicago Linguistics Society, Chicago.
- Ball, Catherine N. 1991. The historical development of the *it*-cleft. Ph.D. dissertation, University of Pennsylvania.
- Berg, Helma van den. 2005. The East Caucasian language family. *Lingua* 115(1): 147-190.
- Collins, Peter C. 1991. *Cleft and pseudo-cleft constructions in English*. London and New York: Routledge.
- Cowie, Jim. 2011. Personal communication. Email: 21/mar/2011.
- Declerck, Renaat. 1983. Predicational clefts. *Lingua* 61: 9-45.
- Delin, Judy. 1992. Properties of *it*-cleft presupposition. *Journal of Semantics* 9: 179-196.
- Ezhaev, Umalat Xozovich and Ezhaeva, R. 2007. *Noxchiin literatura: 9 klass uchebnik-xrestomati, shoolgha daaq'a*. [Chechen literature: primer and reader for grade 9, part two.] Grozny: GUP "Knizhni izdatelstvo".
- Faarlund, Jan Terje, Lie, Svein and Vannebo, Kjell Ivar. 1997. *Norsk referansegrammatikk [Norwegian reference grammar]*. Oslo: Universitetsforlaget.
- Gómez-González, María Á. 2000. *The theme-topic interface. Evidence from English*. Amsterdam and Philadelphia: John Benjamins.
- Good, Jeff. 2003. Clause combining in Chechen. *Studies in language*. 27(1): 113.

- Gundel, Jeanette K. 2002. Information structure and the use of cleft sentences in English and Norwegian. In *Information structure in a cross-linguistic perspective*. 39, Hilde Hasselgård, Stig Johansson, Bergljot Behrens and Cathrine Fabricius-Hansen (eds). Amsterdam, New York: Rodopi.
- Hasselgård, Hilde. 2004. Adverbials in it-cleft constructions. In *Advances in corpus linguistics*. K. Aijmer and B. Altenberg (eds), 195-211. Amsterdam & New York: Rodopi.
- Hasselgård, Hilde. 2010. *Adjunct adverbials in English*. Cambridge University Press.
- Hedberg, Nancy. 1988. The discourse function of cleft sentences in spoken discourse. In *Linguistic Society of America Meeting*. New Orleans, Louisiana.
- Hedberg, Nancy. 1990. Discourse pragmatics and cleft sentences in English. Ph.D. thesis dissertation, University of Minnesota.
- Higgins, Francis Roger. 1979. *The pseudo-cleft construction in English*. New York: Garland Publishers.
- Holisky, Dee Ann. 1987. The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua* 71(1): 103-132.
- Hukari, Thomas E. and Levine, Robert D. 1995. Adjunct extraction. *Journal of Linguistics* 31: 195-226.
- Jespersen, Otto. 1937. *Analytic syntax*. London: Allen and Unwin.
- Johansson, Mats. 2001. Clefts in contrast: a contrastive study of it clefts and wh clefts in English and Swedish texts and translations. *Linguistics* 39(3): 547-582.
- Johansson, Mats. 2002. Clefts in English and Swedish: a contrastive study of IT clefts and WH clefts in original texts and translations. Ph.D. thesis dissertation, Lund University.
- Komen, Erwin R. 2015. *The Nijmegen Parsed Corpus of Modern Chechen (NPCMC)* Nijmegen, Netherlands: Radboud University Nijmegen, <http://erwinkomen.ruhosting.nl/che/crp>.
- Komen, Erwin R. 2007. Focus in Chechen. Master's Thesis dissertation, Leiden University.
- Komen, Erwin R. 2009. *Corpus Studio manual* Nijmegen: Radboud University Nijmegen. http://erwinkomen.ruhosting.nl/software/CorpusStudio/CrpStu_Manual.pdf
- Komen, Erwin R. 2010. Overriding negative concord. *Linguistics in Amsterdam* 3(2): 1-20. <http://www.linguisticsinamsterdam.nl/cgi/t/text/get-pdf?idno=m0302a02>
- Komen, Erwin R. 2013. *Finding focus: a study of the historical development of focus in English*. Utrecht: LOT.
- Kudrinskii, Maksim. 2014. The Chechen enclitic =a as an additive focus particle. M.A. thesis dissertation, Radboud University Nijmegen.
- Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39(3): 463-516.
- Nichols, Johanna. 1994. Chechen. In *The indigenous languages of the Caucasus, vol. 4, The North East Caucasian languages*. Riek Smeets (ed), 1-77. Delmar: Caravan books.
- Nichols, Johanna. 2007. *An all-ASCII Latin practical orthography for Ingush* <http://linguistics.berkeley.edu/~ingush/orthography.html#Practical>
- Nichols, Johanna. 2011. *Ingush grammar*. Berkeley, CA: University of California Press.

- Nunaev, Said-Hwamzat. 1991. *Jilbazan yemar*. [The era of the devil.] Grozny: Kniga.
- Patten, Amanda. 2010. Cleft sentences, construction grammar and grammaticalization. Ph.D. thesis dissertation, The University of Edinburgh.
- Patten, Amanda. 2012. *The English it-Cleft a Constructional Account and a Diachronic Investigation*. Berlin: De Gruyter.
- Plaster, Keith, Polinsky, Maria and Harizanov, Boris. 2013. Noun classes grow on trees. *Language typology and historical contingency: In honor of Johanna Nichols* 104: 153.
- Prince, Ellen. 1978. A comparison of *it*-clefts and WH-clefts in discourse. *Language* 54: 883-906.
- Prince, Ellen F. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. In *The limits of syntax*. Peter W. Culicover and Louise McNally (eds). New York: Academic press.
- Quirk, Randolph, Greenbaum, S., Leech, G. and Svartvik, J. 1972. *A grammar of contemporary English*. London: Longman.
- Schachter, Paul. 1973. Focus and relativization. *Language* 49(1): 19-46.
- Zacharski, Ron and Cowie, Jim. 2011. *Chechen parallel bilingual and monolingual corpus* <http://guidetodatamining.com/appendices/corpora/>