

An RRG Parser for Middle Constructions in Faroese

SIGMUND VESTERGAARD – B00108126

Institute of Technology Blanchardstown

In this article we describe a parser for Middle Constructions in Faroese built on Role and Reference Grammar (RRG).

Keywords: RRG, Faroese, Middle Constructions, Parsing

1. INTRODUCTION

In this paper we describe the architecture and implementation of a Role and Reference Grammar (RRG) parser for middle constructions in Faroese. The parser will be able to take a so-called "middle" sentence and tell us whether it is grammatical or not, and if it is, show an RRG tree for it.

In the first section below we will briefly describe the morphology and syntax of Faroese in general. The description is based on (Thráinsson et al. 2012). We will then move on to describe middle constructions in particular in the second section. After this we will describe the design of the lexicon in section 3 and the parser itself in section 4. We will finish off in sections 5 and 6 by testing the parser.

2. FAROESE MORPHOLOGY AND SYNTAX

Faroese is derived from the Norse language of the primarily Norwegian settlers who moved to the islands in the ninth century. Due to a lack of Viking Age and medieval sources little is known about the development of Faroese into the language we know today. Nothing of substance was written in Faroese until the 1770s, by which time most the features of the modern language must have developed.

2.1. *Nominal categories*

Like Icelandic, Faroese has three grammatical genders; masculine, feminine, and neuter. Faroese nouns inflect for gender, but also number (singular/plural) and case (nominative/accusative/dative/genitive). The inflectional endings vary according to gender. Definiteness of nouns is indicated by a suffixed article.

Just like nouns, adjectives also inflect for number, case, and gender. Additionally they also inflect for degree (positive/comparative/superlative). Adjectives typically have two forms of inflections - strong or weak - depending on the definiteness of the noun phrase they are a part of.

Finally, articles, pronouns, the cardinal numbers 1 to 3, and the ordinal numbers, inflect for number, case, and gender.

Faroese noun phrases show extensive agreement, including number agreement between nouns and the adjectives that modify them. This holds for both attributive and predicative adjectives. We illustrate this using the noun *drongur* 'boy' with the adjective *klókur* 'smart' (Thráinsson et al. 2012: 61):

- (1) (a) ein klókur drongur klókir drongir
 a smart boy.SG smart boy.PL
 'a smart boy' 'smart boys'
 (b) drongurin er klókur dreingirnir eru klókir
 boy-DET is smart boy-DET,PL is.PL smart.PL
 'the boy is smart' 'the boys are smart'

The gender of nouns is reflected in the different forms of the personal pronouns used to refer to them and the gender of adjectives and articles used to modify said nouns. We illustrate this with the following example:

- (2) Hetta er ein klókur drongur
 this is a.M smart.M boy.M
 Hann. er klókur
 he.M is smart.M

Hetta er ein klók genta
 this is a.F smart.F girl.F

Hon er klók
 she.F is smart.F

Hetta er eitt klókt barn
 this is a.N smart.N child.N

Tað er klókt
 it.N is smart.N

Old Norse and older Faroes had four morphologically distinctive cases - nominative, accusative, dative, and genitive - but only the three first are productive in modern spoken Faroese (modern *written* Faroese still retains the genitive to a degree). We illustrate this with a small example:

- (3) Gentan svav
 girl-DET.NOM sleep.PST
 'The girl slept'

Eg sá gentuna
 I see.PST girl-DET.ACC

'I saw the girl'

Hetta er hundurin hjá gentuni
This is dog-DET with girl-DET.DAT

'This is the girl's dog'

Despite this a genitive form can be produced for nouns and personal pronouns, but less so for adjectives. The genitive form of personal pronouns is widely used while the genitive form of many nouns is found in fixed expressions and as the first part of certain compounds, but it is uncertain whether speakers intuitively interpret these forms as genitive (Thráinsson et al. 2012: 62).

Instead of genitive, modern spoken Faroese prefers prepositional constructions involving a dative form of the noun as illustrated in these examples:

- (4) Her eru húsini hjá einum ríkum manni
here is.PL house-DET.PL with a.DAT rich.DAT man.DAT
'Here is a rich man's house/home'

Kettlingurin hjá kettuni hjá mær er vakur
kitten-DET with cat-DET.DAT with I.DAT is beautiful

'My cat's kitten is beautiful'

Above we saw examples with the preposition *hjá* 'with', but Faroese speakers also use other prepositions with the dative, depending on the semantic function. We'll look at some examples (Thráinsson et al. 2012):

- (5) takið á húsinum motorurin í bilinum
roof-DET.NOM.N on house-DET.DAT.N motor-DET.NOM.M in car-DET.DAT.M
'the roof of the house' 'the car's engine'

abbi at dreinginum aldurin á kirkjuni
grandfather to boy-DET.M.DAT age-DETON church-DET.DAT

'the boy's grandfather' 'the age of the church'

halin á kúnni tenninar í hundinum
tail-DETON cow-DET.DAT tooth-DET.PL in dog-DET.DAT

'the cow's tail' 'the dog's teeth'

høvdið á mær eyguni í honum
head-DETON I.DAT eyes-DET in he.DAT

'my head' 'his eyes'

With nouns denoting family relationships an accusative form is normally used instead of genitive or a prepositional phrase as illustrated here:

- (6) pápi dreingin mamma gentuna beiggi Jógvan
 father boy-DET.ACC mum girl-DET.ACC brother Jógvan.ACC
 'the boy's father' 'the girl's mother' 'Jógvan's brother'

As mentioned above, adjectives can be grouped into two categories: *strong* adjectives and *weak* adjectives. Which category they belong to depends on the definiteness of the noun phrase they form a part of. The general rule is that the adjective takes the weak form if the noun phrase is definite, and the strong form if the noun phrase is indefinite. We illustrate this with two examples (Thráinsson et al. 2012: 65):

- (7) (a) Hetta er ein stórir bilur og ein lítill
 this is a big.NOM.SG.M car.NOM.SG.M and a small.NOM.SG.F
 bók.
 book.NOM.SG.F
 'This is a big car and a small book'
- (b) Hetta er tann stóri bilurin og tann
 this is the big.NOM.SG.M car-DET.NOM.SG.M and the
 lítla bókin.
 small-DET.NOM.SG.F book
 'This is the big car and the small book'

Finally, most adjectives can be inflected for degree by adding the suffixes *-(a)r* and *-(a)st* in comparative and superlative, respectively. Indeclinable adjectives express difference in degree by using the auxiliary verbs *meiri* 'more' (comparative) and *mest* 'most' (superlative). We'll finish by showing a couple of examples of this:

- (8) (a) gulur gul-a-ri gul-ast-ur
 yellow yellower yellowest
- (b) hóskandi meiri hóskandi mest hóskandi
 appropriate more appropriate most appropriate

2.2. Verbal categories

Faroese verbs are inflected by person, number, and tense, with the two following characteristics with respect to person inflection (Thráinsson et al. 2012: 67):

1. Faroese verbs do not show any person distinctions in the plural and regular (weak).
2. Faroese verbs do not show any person distinctions neither in the singular nor in the past tense.

This can be illustrated with the following example:

- (9) (a) eg kalli tú kallar hann/hon/tað kallar
 I call.1SG.PRS you call.2SG.PRS he/she/it call.3SG.PRS

'I call' 'you call' 'he/she/it calls'

- (b) vit kalla tit kalla teir/tær/tey kalla
 we call.1PL.PRS you.PL call.2PL.PRS they.M/F/N call.3PL.PRS

'we call' 'you call' 'they call'

- (c) eg kallaði tú kallaði hann/hon/tað kallaði
 I call.1SG.PST you call.2SG.PST he/she/it call.3SG.PST

'I called' 'you called' 'he/she/it called'

Faroese has two distinct imperative forms, plural and singular, as illustrated here:

- (10) Gev/gevið hesum manninum gætur!
 give.SG/PL this man-DET.3SGDAT attention
 'Give attention to this man!'

The singular *gev* would be used if addressing one person, and the plural *gevið* if more than one person is being addressed. There was no distinctive plural imperative in Old Norse, where the 2PL indicative had this role, but in Faroese there is a distinction between the default finite forms and the imperative forms. We illustrate this below, where the non-imperative forms are referred to as *indicative*, although it is uncertain that one can speak of indicative in Faroese, because there is no productive contrasting subjunctive in Faroese (Thráinsson et al. 2012: 67-68):

- (11) (a) Tú fert til hús.
 you.2SG go.2SG.PRS.IND to house
 'You go home'
 (b) Far til hús!
 go.2SG.IMP to house
 'Go home!'
 (c) Tit fara til hús.
 you.2PL go.2PL.PRS.IND to house
 'You go home'
 (d) Farið til hús!
 go.2PL.IMP to house
 'Go home!'

We mentioned above that the subjunctive is not productive in Faroese anymore, and it should be added that only a few relic forms exist in main clauses in relatively fixed expressions and in religious language. And where they exist they almost exclusively express *optative modality*, as illustrated below (Thráinsson et al. 2012: 68):

- (12) (a) Jesus fylgir tær
 Jesus follow.3SG.IND you

- 'Jesus is with you'
- (b) Jesus fylgi tær
 Jesus follow.3SG.SUBJ you
 'Jesus be with you'
- (c) Gud signar Føroyar
 God bless.3PL.IND Faroes
 'God blesses the Faroes'
- (d) Gud signi Føroyar
 God bless.3PL.SUBJ Faroes
 'God bless the Faroes'

Examples of other, relatively fixed optative forms, which aren't religious expressions, are:

- (13) (a) Hann leingi livi!
 He long live.3SG.SUBJ
 'Long live he!'
- (b) Gævi at tað skjótt varð heystfrí!
 Give.3SG.PST.SUBJ that it soon become.SG.PST autumn break
 'I wish we had autumn break soon!'
- (c) Hevði tað nú bara gingist henni væl.
 Have.3SG.PST.SUBJ it now just go.SG.PST her well
 'I wish things would go well for her'

Of these the first one uses the present subjunctive while the two others use the past subjunctive (or what was the past subjunctive in older Faroese).

Typically the passive is formed with the auxiliary verbs *verða* 'be, become' and *blíva* 'be, become'. The participle agree in case, gender, and number with a nominative subject, and the agent is more frequently left out than in English. If the agent is included, it is with the auxiliary *av* 'by', which takes a dative form. We illustrate this with some examples (Thráinsson et al. 2012: 69):

- (14) (a) Hann kysti hana
 he.3SG.NOM kiss.SG.PST her.3SGACC
 'He kissed her'
- (b) Hon varð/bleiv kyst (av honum)
 she.3SG.NOM be.SG.PST.AUX kiss.NOM.SG.F.PST.PTCP (by him.DAT)
 'She was kissed by him'
- (15) (a) Hon kysti teir
 she.3SGNOM kiss.SG.PST.IND them.ACC.PL.M
 'She kissed them'

- (b) Teir vórðu/blivu kystir (av henni)
 they.3PL.NOM.M be.PL.PST.AUX kiss.NOM.PL.M (by her.DAT)
 'They were kissed by her'
- (16) (a) Teir smurdu hann av
 they smear he.3SG.ACC off
 'They beat him up'
- (b) Hann varð/bleiv avsmurdur
 he.3SG.NOM be offsmear.NOM.SG.M.PST.PTCP
 'He was beaten up'

In Faroese it is frequently possible to form so-called *-st*-forms, or *middle* forms, by adding the suffix *-st* to various inflectional forms of the verb. The meaning of the middle forms varies widely in Faroese, but the usages most frequently mentioned in discussions of the middle forms are reflexive, reciprocal, or passive. Here we give some examples, but we will discuss the middle forms in more detail in the next section:

- (17) (a) Eg settist niður
 I sit down
 'I sat down' (reflexive meaning)
- (b) Teir berjast altíð
 they fight always
 'They always fight' (reciprocal meaning)
- (c) Oyggjin kallast Nólsoy
 island-DET call Nólsoy
 'The island is called Nólsoy' (passive meaning)

Perfect tense is either formed with the auxiliary *hava* 'have' and the supine (SGN of PSTPTCP) of the main verb, or it is formed with the auxiliary *vera* 'be' and the inflected and agreeing past participle. *Hava* is used with all transitive verbs and most intransitive verbs (Thráinsson et al. 2012: 72):

- (18) (a) Hon hevur lisið bókina.
 she have.3SG.PST read.SUP book-DET.ACC
 'She has read the book'
- (b) Teir hava sovið leingi.
 they have.3PL.PST sleep.SUP long
 'They have slept for long'
- (c) Hann hevur verið ríkur.
 he have.3SG.PST be.SUP rich
 'He has been rich'
- (d) Hann er vorðin ríkur.
 he.NOM.SG.M is become.NOM.SG.M rich

'He has become rich'

Past perfect is formed with past tense of the relevant auxiliary (*vera/havaas* mentioned above), and the perfect passive is formed with the auxiliary *vera* 'be', not *hava* 'have'. We illustrate this with a couple of examples:

- (19) (a) Hann hevði verið ríkur.
 he have.PST be.SUP rich.NOM.M
 'He had been rich'
- (b) Hann var vorðin ríkur.
 he be.3SG.PST become.PST.PTCP rich.NOM.M
 'He had become rich'
- (20) Hann er/*hefur ofta vorðin/blivin avsmurdur.
 he be.3SG.PRS/*have often become.PST.PTCP off-smear.NOM.M
 'He has often been beaten up'

The indicative-subjunctive distinction is not productive in Faroese and past subjunctive forms generally do not exist. Past subjunctive was commonly used in Old Norse (and still is in Modern Icelandic) to indicate a counterfactual or hypothetical situation. The regular past tense can have this function in Faroese, but the meaning of such forms are typically ambiguous. We illustrate this with a few examples:

- (21) (a) Eg gjørði tað fegin.
 I do.PST it gladly
 'I did it gladly.' or 'I would gladly do it.'
- (b) Hann hevði dripið hundin.
 he have.PST kill.PST.PTCP dog-DET.ACC.M
 'He had killed the dog.' or 'He would have killed the dog.'
- (c) Hann tók bókina.
 he take.PST book-DET.NOM.F
 'He took the book.' or 'He would gladly take the book if...'
- (d) Eg hevði fegin gjørt tað, um eg fekk pengar fyri
 I have.PST gladly do.PST.PTCP it if I get.PST money for
 tað.
 it
 'I would gladly have done it if I was paid for it.'
- (e) Hann drap hundin, um hann fekk hendur á honum.
 he kill.PST dog-DET if he get.PST hand.NOM.PL. on it
 'He would kill the dog if he got his hands on it.'
- (f) Hann hevði tikið bókina frá mær, um hann
 he have.PST take.PST.PTCP book-DET.NOM from me if he
 hevði sæð meg lisið í henni.
 have.PST see.PST.PTCP me read.PST.PTCP in it

'He would have taken the book from me if he had seen me reading it.'

The default word order in Faroese is subject-verb-object (SVO) or subject-auxiliary-main verb-object (SAVO), both in main clauses and embedded ones. We look at some examples (Thráinsson et al. 2012: 236):

- (22) (a) Jógvan las bókina.
Jógvan read.PST book-DET.ACC.SG.F
'Jógvan read the book.'
- (b) Jógvan hevur lisið bókina.
Jógvan have.PST read.PST.PTCP book-DET.ACC.SG.F
'Jógvan has read the book.'
- (c) Eg haldi, at Jógvan hevur lisið bókina.
I think.PRS that Jógvan have.PST read.PST.PTCP book-DET
'I think that Jógvan has read the book.'

As a rule, the indirect object precedes the direct object and typically appears in the dative, although indirect objects in the accusative form also appear. Lets look at some examples:

- (23) (a) Turið gav Hjalmar nógvar bókur.
Turið.NOM give.PST Hjalmar.DAT many book.NOM.PL
'Turið gave Hjalmar many books.'
- (b) Eg spurdi, um Zakaris seldi Eivindi tann gamla bilin.
I ask if Zakaris.NOM sell.PST Eivind.DAT the old car-DET.ACC
'I asked if Zakaris sold the old car to Eivind.'
- (c) Hon lærði meg niðurlagið.
she teach.PST me.ACCrefrain-DET.ACC
'She taught me the refrain.'

Should we move an object or a prepositional phrase, or some other non-subject, to the front of a sentence, as is done in Topicalisation, the finite verb shows up in second place followed by the subject, i.e. Faroese is a "verb-second" (V2) language like the other Germanic languages except English. We take a look at some examples (Thráinsson et al. 2012: 238-239):

- (24) (a) Hesa bókina hevur Jógvan lisið.
this book-DET.ACC have.PST Jógvan.NOM read.PST.PTCP
'This book has Jógvan read.'
- (b) Tann gamla bilin seldi Zakaris Eivindi.
the old car-DET.ACC sell.3.SG.PST Zakaris.NOM Eivind.DAT.
'The old car Zakaris sold to Eivind.'

- (c) *Jóannes haldi eg eigur hesa bókina.*
 Jóannes.NOM think I own.3SG.PRS this book-DET.ACC
 'Jóannes, I think, owns this book.'

As is typical for modern Germanic languages, adjectives precede the noun they modify:

- (25) (a) *ein vøkur genta, ein bláur bilur*
 a beautiful.NOM.F girl.NOM.F, a blue.NOM.M car.NOM.M
 'a beautiful girl', 'a blue car'
 (b) *Hann kom súkklandi á einari gamlari súkklu.*
 he come.PST ride.PRS.PTCP on an old.DAT.F bicycle.DAT.F
 'He came riding on an old bicycle.'

When it comes to adverbs we can broadly distinguish between three basic adverbial positions (Thráinsson et al. 2012: 241):

1. the medial position, following the finite verb
2. the verb phrase position, following a possible object and other elements of the verb phrase
3. the modifying position, when the adverb is modifying an adjective and other adverbs

We illustrate this with a few examples:

- (26) (a) *Tey hava ikki/ivaleyst/jú/aldri lisið bókina.*
 they have not/undoubtedly/actually/never.ADV read book-DET.ACC
 'They have not/undoubtedly/actually/never read the book.'
 (b) *Tey hava lisið bókina tá/har/væl og virðiliga.*
 they have read book-DET.ACC then/there/well and thoroughly.ADV
 'They have read the book then/there/well and thoroughly.'
 (c) *Tey hava lisið hesa ógvuliga longu bókina sera*
 they have read this extremely.ADV long book-DET.ACC very.ADV
væl.
 well
 'They have read this extremely long book very well.'

Let's end by looking at the case of subject, object, and indirect object. The regular *subject* case in Faroese is *nominative*, as we already have seen in previous examples:

- (27) (a) *Hann skrivur.*
 he.NOM.SG write.3SG
 'He writes.'

- (b) *Hon* *arbeiðir.*
 she.NOM.SG work.3SG
 'She works.'
- (c) *Børnini* *spæla.*
 child-DET.NOM.PL play.3PL
 'The children play'
- (d) *Vit* *settu niður epli* *í gjár.*
 we.NOM.PL put down potato.PL yesterday
 'We planted potatoes yesterday.'

Before we continue, it should be mentioned that some verbs take non-nominative subjects in modern Faroese, but we won't look further into that here.

Looking at *direct object* case we can say that the *accusative* is the default case, as we have already seen in previous examples:

- (28) (a) *Hon keypti bókina.*
 she buy.PST book-DET.ACC
 'She bought the book.'
- (b) *Hann seldi telduna.*
 he sell.PST computer-DET.ACC
 'He sold the computer.'

It should be mentioned, however, that a number of verbs take a direct object in dative case, which we also touched upon earlier, but we will restrict ourselves here to saying that among the verbs that take a direct object in the dative are verbs of helping, ordering, praising, thanking, welcoming, etc., leading one to associate the dative objects with thematic roles such as recipients and experiencers. The semantic verb classes and thematic roles we have mentioned are roughly the same as the ones taking a dative object in Icelandic, although dative objects are getting more rare in modern Faroese while still being prevalent in Icelandic (Thráinsson et al. 2012: 257-258).

At last we have the *indirect object*, which takes the *dative* as its default case, as we have seen in previous examples. Among the verbs taking an indirect object are verbs meaning 'sell', 'lend', 'give', 'send', etc. Let's look at three examples:

- (29) (a) *Hann beyð henni starv.*
 he offer.PST she.DAT job.ACC
 'He offered her a job.'
- (b) *Tey finga sær bil.*
 they get.PST themselves.DAT car.ACC
 'They got themselves a car.'
- (c) *Fyrigev honum syndir hansara.*
 forgive.IMP him.DAT sin.PL.ACC he.GEN
 'Forgive him his sins.'

We notice from the examples that these verbs, *bjóða* 'offer', *fáa* 'get', *fyrigeva* 'forgive', are examples of verbs that take two objects, one dative (indirect) object and one accusative (direct) object. We call these verbs *ditransitive*. This dative–accusative pattern is also the most common case marking pattern of ditransitive verbs in Icelandic, but there we also have many other patterns, such as dative–dative, dative–genitive, accusative–dative, accusative–genitive, and accusative–accusative, most of which are not found in Faroese (Thránsson et al. 2012: 262–263).

3. MIDDLE CONSTRUCTIONS IN FAROESE

Above we briefly touched on middle forms in Faroese, but since they are the main subject of this paper we need to look deeper into them.

We mentioned that the middle constructions are formed by adding *-st* to various inflectional forms of the verb, but let's now look further at the origin and semantics of the middle constructions.

The origin of the *-st* suffix is considered to be the reflexive pronoun *sik* in Old Norse, which corresponds to the reflexive *seg* in modern Faroese. Thus Old Norse *setja sik* became *setjast* 'sit down', possibly via the intermediate form *setjask*. In modern Faroese this evolved to *setast* (the Old Norse verb *setja* became *seta* in modern Faroese) (Thránsson et al. 2012: 277).

As we mentioned in the previous section, the middle forms in various can have a wide variety of meanings, i.e. the semantics varies according to what verb is used. We now mention the various semantic meanings with accompanying glossed examples (we add *+st* to the gloss to indicate an *-st*-form) (Thránsson et al. 2012: 277–278).

Reflexive meaning

- (30) (a) Eg settist niður.
I sit+st.PST down
'I sat down.'
- (b) Tey vandust skjótt við hitan.
They get-used+st.PST soon with heat-DEF
'They soon got used to the heat.'

Reciprocal meaning

- (31) (a) Maðurin og konan heilsaðust.
Man-DEF and woman-DEF greet+st.PST
'The man and the woman greeted each other.'

- (b) Tey bítast og klórast.
 They bit+st.PRS and scratch+st.PRS
 'They bite and scratch each other.' (i.e. 'They fight.')

Modal meaning

- (32) (a) Ikki slepst uppafur uttan hjálp.
 not get+st.PRS upp again without help
 'One cannot get back up without help.'
- (b) Ikki kemst uttanum, at málið hevur týdning.
 not get+st.PRS around that language-DEF have.PRS importance
 'One cannot get around the fact that language is important.'

Middle (or passive) meaning

- (33) (a) Íbúð ynskist til leigu.
 apartment wish+st.PRS to rent
 'Apartment for rent is sought.'
- (b) Hon hoyrdist syngja langa leið.
 she hear+st.PST sing long way
 'She was heard singing from a distance.'

4. IMPLEMENTATION OF THE LEXICON

5. IMPLEMENTATION OF THE PARSER

6. ANALYSIS OF TEST DATA

7. TESTING OF THE PARSER

REFERENCES

Thráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen & Zakaris Svabo Hansen. 2012.
Faroese – an overview and reference grammar. Faroe University Press and Linguistic Institute,
 University of Iceland 2nd edn.

*Author's address: Institute of Technology Blanchardstown, Blanchardstown Road
 North, Dublin 15, Ireland
 B00108126@student.itb.ie*