Correcting well-known interference errors – Towards a L2 grammar checker for Inari Saami

Trond Trosterud
The Arctic University of Norway
`trond.trosterud`
`@uit.no`

Marja-Liisa Olthuis
Oulu University
`marja-liisa.olthuis`
`@oulu.fi`

Linda Wiechetek
The Arctic University of Norway
`linda.wiechetek`
`@uit.no`

Abstract

We present *GramDivvun*, the first Inari Saami grammar checker for L2 users. The grammar checker is an important tool in the revitalisation of the language, in particular for strengthening the literary language. As the Inari Saami language community needs language tools predominantly for language learners, the focus is on grammatical interference errors made by (mostly Finnish-speaking) learners. Six of these errors are featured in the first version of the grammar checker. For non-proofread text written by inexperienced writers, precision is good, 73%. With experienced text and proofread text, alarms are rare but precision considerably lower, 19.5 % on average, but varying considerably between the error types. The paper discusses reasons for this variation. Future plans are improving results by means of increased testing, especially for complex sentences, and eventually also including more error types.

1 Introduction

Knowing what a language community needs is the basis for creating meaningful writing tools. In the Inari Saami case, most speakers at work age have learnt the language as adults, and now they are also taking great responsibility for creating modern writing culture, with linguistic help of a handful of native speakers. The next step is to write more and thereby becoming more proficient writers. In order to facilitate this, there is a need for writing tools for L2 users. According to feedback from L2 writers, they need a grammar checker to correct their own texts. The focus on the subsequent revision process can then be on errors outside the scope of both spell checker and grammar checker.

Modern language technology tools, in particular a spellchecker and a morphologically-aware e-dictionary, have been introduced to Inari Saami writers, and these tools are in active use. Based on Inari Saami proofreading experience, the most common type of errors is syntactic interference errors copying Finnish syntax, as described in chapter 3.

In order to help writers correct such errors, we have built a grammar checker for Inari Saami, *GramDivvun*, with L2 writers as its primary target group. The whole grammar checker is freely available on our web page[1] and can be integrated in Google Docs and Microsoft Word as described on the web page. In this article we shall investigate whether *GramDivvun* for some specific constructions is able to change the grammatical structure in L2 writers' Inari Saami text from an underlying Finnish grammar to a correct Inari Saami one. At the same time*GramDivvun* should also be usable as an L1 speakers' grammar checker. In order to do this, we made a rule-based grammar checker for Inari Saami with the same technical framework as the ones for North Saami (Wiechetek et al., 2019) and Lule Saami (Mikkelsen et al., 2022), but where the focus was not on L1 but on L2 users.

Section 2 presents the language community and the technical background for the programs chosen. Section 3 discusses grammatical differences between Inari Saami and Finnish (the native language of most Inari Saami writers) and section 4 presents the grammar checker program. Then section 5 presents how the grammar checker deals with errors and evaluates its performance. Finally, in section 6 comes a conclusion.

2 Background

2.1 The language community

The Inari Saami language has been strongly revitalised since 1986. The focus of the revitalisation movement has initially been on oracy. The

---

[1] `https://divvun.no/korrektur/gramcheck.html`

language has been revitalised by language nest activities and in schools as language of instruction (Olthuis et al., 2013). Also two intensive project-based language learning study years have been organised for working people. The CASLE programme took place in 2009-2010 and trained primarily teachers and nursery school teachers, but also journalists, clergy and civil servants. (Olthuis et al., 2013) The Agile University project trained teachers for all three Saami languages in Finland in 2019–2022: North, Inari and Skolt Saami. (Mattanen et al., 2023)

Especially during recent years, writing culture and written domains have been strengthened. Since 2015, writing tools have been developed by Divvun/Giellatekno at the Arctic university of Norway, UiT. That same year, a project for creating 100 writers for the Inari Saami language was born during the Inari Saami machine translation project in Tromsø, see (Morottaja et al., 2018) pp. 62-63 for a presentation. Since then, the language has gotten more visibility as a literary language. Now it has a communal magazine *Anarâš* (since 1987) and a youth magazine *Loostâš* (since 2020) which are published up to four times a year, see (Anarâškielâ servi, 2023). Furthermore, it has a brand-new e-newspaper, Anarâš aavis² published since March 2023 by the Inari Saami association *Anarâškielâ servi*³.

The association has the intention to publish 100 children's and youth books TODO SOURCE, in the absence of Inari Saami language reading materials for these generations.

Text revision in the above mentioned writing domains shows the differences in the writing skills of the authors. The same observation has been done by (Morottaja et al., 2018) pp. 62. Some L2 writers are having difficulties with syntactic and grammatical structures. *GramDivvun* is mainly made for L2 writers, but should also avoid false alarms for L1 users.

The language community counts only a handful of L1 writers. In this group the speakers are mainly either elderly people or children after a strong language revitalisation movement. Elderly speakers have never learnt to write in their own language, and therefore, if they write something, they rather tend to use their own personal orthographies, being a less fruitful basis for developing spell-checkers

(cf. (Morottaja et al., 2018)). The linguistic competence of L1-speakers is good, so the L2 speakers/writers can profit from their language skills in common. The students have the elderly speakers as language masters, in order to learn the daily spoken language fluently.

We argue that offering proofreading and writing support helps to increase the number of publications and motivates to write creatively and translate literature. The needed experience in reading and writing will come with time.

2.2 Technical background

The Inari Saami grammar checker and all its modules are part of a multilingual infrastructure *GiellaLT*, which includes 130 languages altogether.

The technological implementation of the grammar checker is based on finite-state automata for morphological analysis (Beesley and Karttunen, 2003; Lindén et al., 2013; Pirinen and Lindén, 2014) and constraint grammar for syntactic and semantic as well as other sentence-level processing. Constraint Grammar is a rule-based formalism for writing grammars that disambiguate and syntactically label text. It was initially presented by Fred Karlsson (Karlsson, 1990; Karlsson et al., 1995), we use the free open source implementation VISLCG-3 (Bick and Didriksen, 2015; Didriksen, 2016). The Inari Saami morphological analyser and lexicon is included in the *GiellaLT* infrastructure (Moshagen et al., 2013) and is publicly available⁴.

The grammar checker is built on a pipeline of modules: we process the input text with morphological analysers and disambiguate and then apply grammar rules on the disambiguated sentences, as described above. The grammar checker takes this input and sends it to a number of other modules, the core of which are several Constraint Grammar modules for tokenisation disambiguation, morphosyntactic disambiguation and a module for error detection and correction. The full modular structure is described in Wiechetek (2019).

2.3 Earlier research

Inari Saami is a language with agglutinative morphology combined with a rich array of stem changing processes, as shown in (Olthuis, 2000) and (Valtonen et al., 2022).

Work on Inari Saami language technology

started out with a project on machine translation ((Antonsen et al., 2017)). This work also gave rise to a spellchecker ((Morottaja et al., 2018)).

L2 writers are known to make errors based upon Finnish interference, but there has so far not been published systematic research on the topic.

3 Grammatical error types

3.1 A typology of errors

Based upon our earlier experience, we assume that L2 speakers have only minor problems with orthography. Their major challenges are related to syntax (mainly interference from Finnish), morphosyntax and morphophonology.

As Inari Saami proofreading has shown, one of the most common error types are interference errors, as well as grammatical errors due to inflectional forms (especially case forms) being similar to each other.

Syntax errors are mainly interference errors, copying the Finnish syntax unchanged into Inari Saami. The syntactic structure of Finnish and Inari Saami are quite similar, but still different enough to give rise to a well-known set of interference errors. The main focus of the present grammar checker is upon these errors. At the same time, both L1 and L2 writers need the program to be robust enough to not give too many false alarms.

Given that the Inari Saami language community is small (appr. 450 speakers and even fewer writers), we have a very limited amount of written text at our disposal. We will still investigate to what extent it is possible to draw conclusions from it.

The following section shows a number of real-world error examples that have served as a basis for our error typology.

3.2 Object marking errors

In transitive sentences, the totality object in plural is in the nominative-like accusative in Finnish (1), whereas the object case is only accusative in Inari Saami (2):

(1)   Minä ostin   kirjat.
        I.NOM buy.1PL.PRT book.PL.ACC;NOM
        'I bought the books.'

(2)   Mun ostim   kiirjijd.
        I.NOM buy.1PL.PRT book.PL.ACC

(3)   *Mun ostim   kirjeh.
        I.NOM buy.1PL.PRT book.PL.NOM

The typical interference error for L2 writers is to use plural nominative also in Inari Saami, instead of the accusative *kirjeh*, pro *kiirjijd*, as in (3). Furthermore, if the totality in Inari Saami needs to be stressed, it should be given by adding an attribut, like *puoh kiirjijd* 'all books' or *taid kiirjijd* 'those books'. In Inari Saami the use of accusative like in example 4 gives a perfect counterpart for the Finnish partial object

(4)   Minä ostin   kirjoja.
        I.NOM buy.1PL.PRT book.PL.ACC;PAR
        'I bought some books.'

In negative sentences, the object case marking is again different: Where Finnish negative objects are in the Partitive-like accusative (5), the Inari Saami ones are not sensitive to negation, and occur in the Accusative (6).

(5)   Minä en   ostanut   kirjaa.
        INom Neg.1Sg buy.PRFPTC book.SG.PAR
        'I did not buy any book.'

(6)   Mun jiem   uástám   kirje.
        I.Nom Neg.1Sg buy.PrfPtc book.Sg.Acc
        'I did not buy any book.'

These types of sentences mainly do not cause any troubles for the L2 speaker.

In Finnish, the object can express completion (by using genitive) and incompletion (by using partitive) of a process, cf. (7) vs. (8):

(7)   Minä luen   kirjan.
        I.NOM read.1PL.PRS book.PL.GEN
        'I read the book.'

(8)   Luen   kirjaa.
        read.1PL.PRS book.PL.PAR
        'I am reading the/a book.'

The content in (8) should in Inari Saami be expressed with the present continuous, as in (9):

(9)   Mun lam   luhâmin
        I.NOM read.1sg.prt.cont.1sg book.sg.par
        kirje.

        'I am reading a/the book.'

The different ways the two languages express incompleted action seems to cause problems for the L2 speakers.

Also the object of an imperative verb is often erroneously realised as plural nominative, modeled after Finnish *Osta sukset!* (10), instead of the correct Accusative ((11)).

(10) *Uásti   saveheh!
      buy.ɪᴍᴘʀᴛ ski.ᴘʟ.ɴᴏᴍ
      'Buy skis!'

(11) Uásti   savehijd!
      buy.ɪᴍᴘʀᴛ ski.ᴘʟ.ᴀᴄᴄ
      'Buy skis!'

3.3 Existential clauses and the habitive construction

The agreement pattern in existential clauses and habitive constructions shows several differences between Finnish and Inari Saami.

Firstly, interference occurs in E(xistential)-subject marking, for example where the Finnish plural partitive ((12)) is erroneously realised as accusative in prohibitions, like in the example (13), as compared to correct (14). In Finnish, the verb is in singular , whereas in Inari Saami the verb agrees with the E-subject.

(12) Minulla ei   ole
      I.ʟᴏᴄ   ɴᴇɢ.3ꜱɢ/ɴᴇɢ.3ᴘʟ be.CᴏɴNᴇɢ
      ystäviä
      friend.Pʟ.Aᴄᴄ/Pᴀʀ
      'I have no friends'

(13) *Must ij/iä   lah
      I.ʟᴏᴄ ɴᴇɢ.3ꜱɢ/ɴᴇɢ.3ᴘʟ be.ᴄᴏɴɴᴇɢ
      ustevijd
      friend.ᴘʟ.ᴀᴄᴄ
      'I have no friends'

(14) Must iä   lah   usteveh
      I.ʟᴏᴄ ɴᴇɢ.3ᴘʟ be.ᴄᴏɴɴᴇɢ friend.ᴘʟ.ɴᴏᴍ
      'I have no friends'

The same E-subject congruence also applies for the affirmative clauses, in both clause types, with (15) being the Finnish pattern, (16) the interference error (with accusative representing partitive and (17) being the correct Inari Saami form:

(15) Pihalla on   koiria.
      yard.ᴀᴅᴇ be.3ꜱɢ.ᴘʀꜱ dog.ᴘʟ.ᴘᴀʀ
      'There are dogs in the yard.'

(16) *Šiljoost lii   pennuid.
      yard.Lᴏᴄ be.3ꜱɢ.ᴘʀꜱ dog.ᴘʟ.ᴀᴄᴄ

(17) Šiljoost láá   pennuuh.
      yard.ʟᴏᴄ be.3ᴘʟ.ᴘʀꜱ dog.ᴘʟ.ɴᴏᴍ
      'There are dogs in the yard.'

Despite Inari Saami having a partitive case, it behaves different from its Finnish counterpart, and the case used during interference from Finnish in existential sentences is the accusative plural, *pennuid*. Also, the plural verbform *láá* is often re-placed with the singular *lii*, as shown in (16). Some writers note the verb congruence but forget the e-subject in the accusative, though.

The grammar checker will target those errors.

4 The grammar checker

The Inari Saami grammar checker is built on handwritten Constraint Grammar rules. The grammar checker module uses mainly two syntactic rule types, *ADD* for adding error labels, and *COPY* for creating correct morphological strings that are then generated by the morphological *FST* generator. In this version of *GramDivvun*, we use flat syntactic structures, including valencies and semantic categories. There is an option for including dependencies if the specific error type requires it. However this has not been the case for the rules implemented for Inari Saami yet.

The simplified ADD-rule in the box below adds an error tag (&msyn-obj-plnom-placc) to a plural nominative noun if it has a left context with a transitive finite verb that is preceded by a nominative. We further exclude (with the *0C* condition) the possibility that the target noun has other readings than nominative plural. In addition we exclude that there is a third person plural verb in agreement with the noun to its right.

The COPY-rule on the other hand replaces the nominative tag with an accusative tag. In addition it removes the error tag and replaces it with the label &SUGGEST marking that this line is a correction of the error.

```
ADD (&msyn-obj-plnom-placc) TARGET N
IF
(*-1 VFIN + TV LINK -1 Nom)
(0C N + Pl + Nom)
(NOT *1 V + Pl3 BARRIER NOT-ADV);


COPY (Acc &SUGGEST)
EXCEPT (Nom &msyn-obj-plnom-placc)
TARGET (N Pl &msyn-obj-plnom-placc);
```

The present version of the grammar checker[5] contains 160 Constraint Grammar rules (dated 16.03.2023) that map error labels onto word forms, for 88 different error types. In this article we focus on a smaller selection of the most frequent error types that get corrected reliably without too many

---

[5]The grammar checker is documented at https://giellalt.github.io/lang-smn/tools-grammarcheckers-grammarchecker.cg3.html, with link to the source code at the end of the document.

false alarms. The focus is on releasing a preliminary tool that can be tested by users.

5 Evaluation

5.1 Test setup

For texts written by the target audience (language learners) we hand-picked uncorrected early versions of Wikipedia, written by L2 users, cf. 5.2.

We also looked at how the grammar checker copes with texts written by more experienced writers. For that, we evaluated some 1,27 million words, and got 227 relevant alarms. These alarms are evaluated in section 5.3 below.

5.2 Evaluation of the L2 results

Table 1 illustrates the results. The false positives are unsuccessfull corrections. However, 7 of the 9 instances are successfull error detections.

	interference corp
TP	24
FN	95
FP	9
Precision	72.73%
Recall	20.17%

Table 1: Evaluation of the Inari Saami grammar checker

The quality is measured using basic precision, recall and f_1 scores, such that recall $R = \frac{t_p}{t_p+f_n}$, precision $P = \frac{t_p}{t_p+f_p}$ and f_1 score as harmonic mean of the two: $F_1 = 2\frac{P \times R}{P+R}$, where t_p is a count of true positives, f_p false positives, t_n true negatives and f_n false negatives.

The qualitative evaluation of the results is shown in table 1. Looking at some of the examples, we se in ex. (18) a case error in the subject *čuoigâmkammuuh* (nominative plural should be accusative). This case error disrupts the agreement between the presumable plural subject (which should be an object) and the third person singular verb *koolgâi*. The grammar checker finds that there is an error in the sentence, but instead of fixing the case error it suggests an error in the verb form *koolgâi*. Even though this counts as a false alarm, *GramDivvun* is successful in error detection in general.

(18)   Máttáátteijee čielgij,   maht
        teacher.SG.NOM explain.3SG.PAST, what
        čuoigâmkammuuh koolgâi
        skishoe.PL.NOM   should.3SG.PAST

kiddiđ   saveháid.
fasten.INF ski.PL.ILL
'The teacher explained how the ski shoes should be fastened to the ski.'

Also in the next sentence, the verb *koolgâi* is corrected by GramDivvun instead of the nominative plural noun *oppâkirjeh*.

(19)   Talle oppâkirjeh   koolgâi
        then textbook.PL.NOM should.3SG.PAST
        jurgâliđ   suomâkielâst sämikielân.
        translate.INF Finnish.SG.LOC Saami.ESS
        'Then the text books needed to be translated from Finnish into Saami'

Here, case errors as in *kuobbâreh*, which should be *kuobbârijd* (Acc Pl), and *tábáhtusâid* (N Pl Acc corrected to nominative msyn-extsubj-acc-nom), which should be *tábáhtusah*, are identified correctly.

(20)   Nubeh tobdeh   kuobbâreh
        others identify.3PL mushroom.NOM.PL
        ivneest.
        colour.LOC
        'Others identify mushrooms by their colour.'

(21)   Mist láá   eenâb-uv tábáhtusâid.
        I.LOC be.3SG.PRES more   event.PL.ACC
        'We have even more events than that.'

5.3 Presicion evaluation of proofread, published texts

In order to evaluate precision we ran a test on a larger corpus of blogs, news and science texts. These texts were proofread and published. The size of test corpus was appr. 1.27 million words. The total number of alarms for the relevant error types was 169. The result was as shown in table 2[6]. As can be seen, the result of this second evaluation deviates strongly from the first test.

Error type	TP	FP	Precision
Ext. verb Sg3 → Pl3	9	2	81.0 %
Ext. verb Pl3 → Sg3	15	43	25.9 %
E-subj acc → nom	5	45	10.0 %
E-subj gen → nom	4	46	8.0 %
Overall precision	33	136	19.5 %

Table 2: Evaluation of Inari Saami *GramDivvun* on a corpus of news and science texts (N=1266071)

[6]TP = true positives, FP = false positives, Precision = TP/(TP+FP)

The best result (81 %) was provided by the rule set correcting singular existential verbs from sg3 into pl3. For the false alarms of the type Sg3 → Pl3, the rule failed to identify the clause boundary between singular *lii* and plural *virgeomâhááh*.

(22) Olgoštem muuneeldestim
discrimination prevention.N.SG.GEN
tááhust lii
perspective.N.SG.LOC be.V.PRS.PL3
tehálâš, et virgeomâhááh já
important, that official.N.PL.NOM and
palvâlusfälleeh ovdedeh
servant.N.PL.NOM promote.v.prs.pl3
toimâidis oovtviärdásâžžân já
activity.N.ACC.PXPL3 equalN.ESS
pasteh tarvaniđ meid olgošteijee tooimân.

'From the point of view of preventing discrimination, it is important that public authorities and service providers promote equality and are able to tackle discriminatory behaviour.'

For the opposite correction, Pl3 → Sg3, the results were somewhat worse. It seems advanced writers have several ways of expressing plural referents, ways that are not captured by the grammar checker.

In (23), the problem was a wrongly disambiguated *anarâškielâ*. The word could be either nominative or genitive, but since it was disambiguated as nominative, the grammar checker errouneously corrected the plural verb *láá* to singular *lii*, despite the subsequent plural form *uáppeeh*.

(23) škoovlâst láá uđđâ
school.N.SG.LOC be.V.PL3.PRES new
anarâškielâ uáppeeh..
Inari.Saami.N.SG.GEN pupil.N.PL.NOM
'In the school, there are new Inari Saami pupils'

In (24), the problem is again an error of disambiguation. The time expression *manuppaje* has errouneously lost its accusative analysis, and as a percieved nominative singular it thus blocks for the reference to the correct subject *čielgiittâsah*.

(24) Taan kalenderist láá kevttum
this calendar be.V.PL3.PRES used
mánuppaje kuobbârist
month mushroom.N.PL.LOC
siämmáálágán čielgiittâsah ko
similar explanationN.PL.NOM as
oovdeld mainâšum
before mentioned

kuobârkirjeest-uv.
mushroom.bookN.SG.LOC.FOC
'This calendar uses the same explanations of the mushroom of the month as the fungus book mentioned above.'

The authors of the larger test corpus often used complex noun phrases, such as the coordination NP in (25), or the apposition *maailmist* following the quantor phrase in (26).

(25) Čielgâsávt stuárráámus uási Island
clearly largest part Iceland
tuálvumist láá kyeli
income.N.SG.LOC be.V.PRS.PL3 fish.N.SG.NOM
já kyelipyevtittâsah.
and fish.product.N.PL.NOM
'The largest part of Icelands income clearly consists of fish and fish products.'

(26) [m]aŋgâ kielâ maailmist
many language.N.SG.GEN world.N.SG.LOC
láá lappum
be.V.PRS.PL3 disappear.V.PTC
'Many languages of the world have disappeared'

Removing *maailmist* from the quantor phrase would have removed the false alarm. The improvement needed is thus to take such appositions into account.

A recurrent phenomenon in the experienced writers' corpus was the use of third person plural pro-drop, like in (27):

(27) Tommittáá tast láá
so.much that.ADV.LOC be.V.PRS.PL3
sárnum já čáállám sehe sämi
talk.V.PRFPRC and write.V.PRFPRC both Sámi
já lädimediast.
and Finnish.media.N.SG.LOC
'So much about this people have talked and written both in Sámi and Finnish media.'

This phenomenon is not common in Finnish and has thus not been present in the learner corpus on which the grammar checker development has been focused.

Topicalised e-subjects like the one shown in (28) were rarely encountered in the learner corpus and thus not within the range of the grammar checker rule set. In the reference corpus there were several instances of it.

(28) Kommemainâseh-uv sust
ghost.story.N.PL.NOM.FOC s/he.PRON.SG.LOC

láá,   veikkâ ij   tain
be**v.prs.pl3**, although not.v.sg3 that
poollâđ   taarbâš.
fear.v.inf need.v.conneg
'Ghost stories s/he has, even though one
does not need to be afraid of them'

The grammar checker also has not taken into account listing of referents following a colon. Cf. (29).

(29)   Sii lasseen   ive 2018
       they in.addition.to year 2018
       kielâpiervâlijn   láá
       language.nest.n.pl.loc be.v.prs.pl3
       porgâm: Tarja Passi, Tiina Lehmuslehti,
       work.v.ptc: T.P.,   T.L.,
       Minna Lampinen, ...
       M.L.,   ...
       'In addition to (the before mentioned), in the year 2018 the following people have worked in the language nests: T.P., T.L., M.L., ...'

List formatting such as the one in (29) was not encountered in the learner corpus, hence the grammar checker did not take them into account.

The poorest results were found for the errors correcting e-subject case errors. Here, a recurrent problem was complex sentences with several NPs following the existential sentence proper. The grammar checker targeted simple sentence constructions and failed to identify barriers to prevent it from errouneously flagging the right dislocated NPs as subject errors. (30) is a case in point.

(30)   [A]narâškielâ várás iä   lah
       Inari.Saami   for   notv.pl3 be.v.conneg
       kielâtotkei   virgeh,   já
       linguist.n.pl.gen position.n.pl.nom, and
       nuuvt jieškote-uv totkee   ferttee kavnâđ
       thus   each   researcher must   find
       vuovijd   maht ruttâdiđ jieijâs
       way.n.pl.acc that finance ones.own
       projektijd já   pargoid.
       projects   and works
       'There are no permanent linguist positions for Inari Saami, and thus each and every researcher must find a way to finance his or her projects and work.'

6   Conclusion and future perspectives

We have written a basic Inari Saami grammar checker (*GramDivvun*). Here we evaluate a subset of if, containing rules for the 5 most common error types of L2 users. Based on our evaluation of a test corpus of learner texts, the subset of L2 error rules presented in section 3 has a fairly good precision of 73% This corpus had a high number of interference errors, and the grammar checker was able to identify them quite reliably. Compared to this, the precision in a larger reference corpus written by more experienced writers was lower.

The grammar checker focuses on interference errors from Finnish. The experience from the present study indicates that the focus for some time to come should be upon improving precision for the rules discussed in the present paper, with a focus on recall being the next step. Investigating other error types in Inari Saami text and correcting them we leave for further research.

7   Bibliographical References

References

Anarâškielâ servi. 2023.   Anarâškielâ servi publications.   `https://anaraskielaservi.fi/Almostitmeh/`.

Lene Antonsen, Ciprian Gerstenberger, Maja Kappfjell, Sandra Nystø Rahka, Marja-Liisa Olthuis, Trond Trosterud, and Francis M. Tyers. 2017.   Machine translation with North Saami as a pivot language. In *Proceedings of the 21st Nordic Conference on Computational Linguistics, NoDaLiDa, 22–24 May 2017, Gothenburg, Sweden*, volume 29 of *NEALT Proceedings Series*, pages 123–131. Linköping University Electronic Press.

Kenneth R Beesley and Lauri Karttunen. 2003. Finite-state morphology: Xerox tools and techniques. *CSLI, Stanford*.

Eckhard Bick and Tino Didriksen. 2015. CG-3 — beyond classical constraint grammar. In *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)*, pages 31–39, Vilnius, Lithuania. Linköping University Electronic Press, Sweden.

Tino Didriksen. 2016. *Constraint Grammar Manual: 3rd version of the CG formalism variant*. GrammarSoft ApS, Denmark.

Fred Karlsson. 1990. Constraint grammar as a framework for parsing running text. In *COLING '90 Proceedings of the 13th conference on Computational linguistics*, volume 3, pages 168–173, Helsinki.

Fred Karlsson, Atro Voutilainen, Juha Heikkilä, and Arto Anttila. 1995.   *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*. Mouton de Gruyter, Berlin.

Krister Lindén, Erik Axelson, Senka Drobac, Sam Hardwick, Juha Kuokkala, Jyrki Niemi, Tommi A

Pirinen, and Miikka Silfverberg. 2013. Hfst—a system for creating nlp tools. In *International workshop on systems and frameworks for computational morphology*, pages 53–71. Springer.

Tiina Mattanen, Marja-Liisa Olthuis, Anni-Siiri Länsman, and Sari Harmoinen. 2023. Máŋggahámat skuvlen oassečoavddusin sámegielat ávnnasoahpaheaddjiid váilumii. *Dutkansearvvi dieđalaš áigečála*, 7(1):6–30.

Inga Lill Sigga Mikkelsen, Linda Wiechetek, and Flammie A Pirinen. 2022. Reusing a multi-lingual setup to bootstrap a grammar checker for a very low resource language without data. In *Proceedings of the Fifth Workshop on the Use of Computational Methods in the Study of Endangered Languages*, pages 149–158, Dublin, Ireland. Association for Computational Linguistics.

Petter Morottaja, Marja-Liisa Olthuis, Trond Trosterud, and Lene Antonsen. 2018. Anarâškielâ tivvoomohjelm – kielâ- já ortografiafeeilâi kuorrâm tivvoomohjelmáin. *Dutkansearvvi dieđalaš áigečála*, 1(2):63–84.

Sjur N. Moshagen, Tommi A. Pirinen, and Trond Trosterud. 2013. Building an open-source development infrastructure for language technology projects. In *NODALIDA*.

Marja-Liisa Olthuis. 2000. *Kielâoppâ. Inarinsaamen kielioppi*. Sämitigge, Inari.

Marja-Liisa Olthuis, Suvi Kivelä, and Tove Skutnabb-Kangas. 2013. *Revitalising Indigenous Languages. How to recreate a lost Generation?* Multilingual Matters, Bristol.

Tommi Pirinen and Krister Lindén. 2014. State-of-the-art in weighted finite-state spell-checking. In *Computational Linguistics and Intelligent Text Processing : 15th International Conference on Intelligent Text Processing and Computational Linguistics, CICLing 2014, Kathmandu, Nepal, April 6-12, 2014, Proceedings, Part II*, volume 2, pages 519–532, Berlin Heidelberg. Springer-Verlag.

Taarna Valtonen, Jussi Ylikoski, and Luobbal Sámmol Sámmol Ánte (Ante Aikio). 2022. 178Aanaar (Inari) Saami. In *The Oxford Guide to the Uralic Languages*. Oxford University Press.

Linda Wiechetek, Sjur Nørstebø Moshagen, Børre Gaup, and Thomas Omma. 2019. Many shades of grammar checking – launching a constraint grammar tool for North Sámi. In *Proceedings of the NoDaLiDa 2019 Workshop on Constraint Grammar - Methods, Tools and Applications*, NEALT Proceedings Series 33:8, pages 35–44.