Case error corrections for noun phrases containing deverbal attributive nouns in Greenlandic

Judithe Denbæk
Oqaasileriffik / The Language Secretariat of Greenland
judithe@oqaasileriffik.gl

Abstract

This paper presents very early findings using Constraint Grammar (CG) in semantic annotation of a specific type of noun phrases in Greenlandic (Kalaallisut), in which the attributive noun is a nominalized predicative verbal stem. The annotation is used in a grammar checker pipeline for the purpose of making case error correction suggestions.

1 Introduction

The paper presents initial ideas on using semantic valency information in a grammar checker pipeline for Greenlandic, with no significant findings yet. As such, the purpose of this contribution to the workshop *Constraint Grammar and Finite State NLP – Rule-based and hybrid methods and tools for user communities at NoDaLiDa 2025* is mainly an invitation to a fruitful discussion and for the author to gain valuable input from experienced peers. The work is carried out as part of Oqaasileriffik's[1] implementation process of a grammar checker pipeline for Greenlandic. The grammar checker pipeline is developed by Giellatekno[2].

Greenlandic only has 3 major word classes (nouns, verbs and particles) with each their subclasses and lacks adjective as a word class on its own. The formal distinction between the word classes is by their obligatory inflection. Particles have no inflection. Nouns can be inflected in two grammatical cases (*absolutive* (ABS) or *relative*[3] (REL)) and six oblique cases (*ablative* (ABL), *vialis*[4] (VIA), *aeqvalis*[5] (AEQ), *instrumental* (INS), *terminalis*[6] (TRM), *locative* (LOK)). Verbs can be transitive or intransitive and have 4 superordinate moods (*indicative* (IND), *interrogative* (INT), *imperative* (IMP), *optative* (OPT)) and 4 subordinate moods (*causative* (CAU), *conditional* (CON), *contemporative* (CONT), *participial* (PAR)).

The basic word-formation is as follows, where the elements in parentheses are optional and *encl* stands for *enclitic particle*:

```
stem + (suffix(es)) + inflection + (encl)
```

The suffixes have 4 classes: nominalizers, verbalizers, verbal and nominal. The noun class can change several times between the stem and the inflection. Enclitic particles can attach to any word class. Although Greenlandic doesn't have adjectives, adjectival meaning can be translated. There is a class of both genuine nouns, nominal suffixes, and nominalized predicative verbal stems that more frequently function as attributive nouns. These are morphologically or syntactically placed to the right of their stem or head noun. In a simplified glossing style, each type is exemplified below:

adjectival noun

(1) illu qorsuk

illu.SG.ABS qorsuk.SG.ABS
house.SG.ABS green.SG.ABS

'green house'

adjectival nominal suffix

(2) illunnguaq

illu-nnguaq.SG.ABS
house-small.SG.ABS

'small house'

[1] https://oqaasileriffik.gl/

[2] https://giellalt.github.io/proof/gramcheck/GrammarCheckerDocumentation.html

[3] also called *ergative* by some authors

[4] also called *perlative* by some authors

[5] also called *equative* by some authors

[6] also called *allative* by some authors

Proceedings of the 9th Workshop on Constraint Grammar and Finite State NLP, pages 50–58

nominalized predicative verbal stem as adjectival noun

(3) meeraq nuannaartoq

meeraq.SG.ABS nuannaar-toq.SG.ABS
child.SG.ABS be.happy-PTCP.SG.ABS

'happy child'

For a list of the particular abbreviations used in this paper, the reader is referred to the end of the paper. The interpretation of the diminutive suffix in (2) as being adjectival might not be completely adequate, since the assumption is based on the glossing of it as *'small'* in english and would have been *'lille'* in danish. This example might not have been included if english or danish instead formed the diminutive sense derivationally. Whether the PTCP example in (3) gets interpreted as either adjectival or as relative sentence-like probably depends on the speaker of the language and whether or not the form is lexicalized. Other types of verbal stems can be formed with PTCP, it is not reserved for predicative stems only, and as interesting as it is to explore what types of stems behave more substantival than the predicative stems when PTCP is formed, this subject falls outside the scope of the current paper. The attributive noun, as is shown in example (3), is placed after it's head noun. It is in case and number agreement with the head noun, but in recent years there has been a development in which this structure more or less gets treated in a compound-like manner by speakers of the language, wherein the word order is still the same, but the head noun is inflected in a functionally bare absolutive case and only the attributive noun is inflected in the appropriate case, depending on the function of the noun phrase as a whole. Although this can objectively be viewed as a natural development for the language, until there is consensus among the language users for it's acceptance, for the time being this paper will view them as grammatical errors to be given correction suggestions in a grammar checker.

The development of a grammar checker for Greenlandic is in its infancy, and the infrastructure is developed and made available for adaptations to Greenlandic by Giellatekno[7]. For an introduction to Giellatekno infrastructure, see (Moshagen et al. 2013), and for the description of Giellatekno's grammar checker module, see (Wiechetek et. al. 2019).

[7] https://giellalt.github.io/proof/gramcheck/GrammarCheckerDocumentation.html

2 Predicative verbal stems

The verbal stems that are termed predicative verbal stems in this paper are semantically portmanteau. They contain both a copular verbal element and an adjectival meaning. This combination of meanings is borne by a single stem that may or may not be comprised of more than one morpheme that in turn might or might not be lexicalized (this subject needs it's own investigation). In their verbal form, provided that the stem is not transitivized, they predicate the subject of the sentence. Take a look at the example (3) from the previous section without the PTCP and with indicative inflection:

(4) meeraq nuannaarpoq

meeraq.SG.ABS nuannaar-IND-3SG
child.SG.ABS be.happy-IND.3SG

'the child is happy'

One would think that an extension of the morphology (i.e. the PTCP form) of a word would add to its semantics, but in this case at least cross-linguistically speaking there seems to be a reduction, perhaps as would be expected with a nominal participle that semantically denotes a resulting state of the stem to which it is attached. The copular element seems to be wiped out and we are left with the semantics of the adjectival part. The copular element can resurface by word formation, for instance, when there is added adverbial-like suffix(es), such as negation, between the predicative stem and the PTCP. In such cases, the PTCP often corresponds to relative sentences in other languages:

(5) meeraq nuannaanngitsoq

meeraq.SG.ABS nuannaa-nngit-soq.SG.ABS
child.SG.ABS be.happy-NEG-PTCP.SG.ABS

'a child who is not happy'

3 Semantic classification for Greenlandic

The semantic classification used in Oqaasileriffik's semantically annotated lexical database, Katersat[8], is based on, or rather, inspired by the system developed in The Danish FrameNet project[9], developed by the University of Southern Denmark and GrammarSoft Aps[10].

[8] https://oqaasileriffik.gl/en/dict/
[9] https://FrameNet.dk
[10] https://grammarsoft.com/

For the predicative verb stems in Greenlandic, the general semantic frame is called "be_attribute, Be" which in some cases has been combined with an additional secondary semantic tag for adjectives. The semantic tags for adjectives are in square brackets starting with a "j". The tag for the word *happy* would be <jpsych>, which is defined as *"psychological, feeling, intellectual, inherent character trait"* and is attributed to nouns that are semantically classified as human, acts and semantics: <H>, <act> and <sem> (Bick, 2019).

In this paper, a combination of verbal frame and prototypes for adjectives will be used in tagging predicative verbal stems. The use of adjectival prototype tags is in purely semantic sense, since the adjectival words in question actually belong to nouns. In PTCP word-formations, this combination of verbal frame and adjectival semantic prototype will be split in purely adjectival prototype and verbal stem frame. By keeping the verbal stem frame tag, the *"pertainym"*[11] relation is informed.

4 Semantic annotation for predicative verbal stems in Greenlandic

Giellatekno grammar checkers already use semantic valency information in their grammar checker pipelines (Wiechetek, 2017). Oqaasileriffik has a semantic module for Oqaasileriffik's analyzer that combines FST[12] analysis with the semantically annotated lexical database, the aforementioned Katersat. This module is not implemented or implementable in the grammar checker for Greenlandic as of December 2024. For this reason for the purpose of this presentation, the annotations are done manually in a CG file within the pipeline for the grammar checker. Example (3) and (4) are annotated in this manner, with output as shown in CG-3 IDE[13]:

input string: (3) meeraq nuannaartoq
CG-3 IDE output:

```
"<meeraq>"
    "meeraq" N Abs Sg <H> §TH
    ADD:68 ADD:92
"<nuannaartoq>"
    "nuannaar" Gram/IV TUQ Der/vn N
    Abs Sg <STEM:f:be_attribute_jpsych>
    <jpsych> <p:<H>_N_Prop>
    ADD:72 ADD:85 SUBSTITUTE:95
    SUBSTITUTE:96
    "nuannaar" Gram/IV V Par 3Sg
    <f:be_attribute_jpsych>
    <§TH_@SUBJ_<H>_N_Abs>
    ADD:72 ADD:85
```

input string: (4) meeraq nuannaarpoq
CG-3 IDE output:

```
"<meeraq>"
    "meeraq" N Abs Sg <H> §TH
    ADD:68 ADD:92
"<nuannaarpoq>"
    "nuannaar" Gram/IV V Ind 3Sg
    <f:be_attribute_jpsych>
    <§TH_@SUBJ_<H>_N_Abs>
    ADD:72 ADD:85
```

Grammarsoft Aps uses <fn:..> as tags for semantic frames, where *"fn"* probably stands for FrameNet. The examples in this paper are just <f:..>, where *"f"* stands for frame.

The CG rules in question are as follows.

definitions:

```
LIST human = "meeraq" ;
LIST objectmarking =
(/[1-4][SP][gl]O/r) ;
LIST subjectmarking =
(/[1-4][SP][gl]/r) ;
SET ITR_pers = subjectmarking
- objectmarking ;
LIST frame = /<f:.*>/r ;
LIST jsem = /<j.*>/r ;
LIST noun_number = (N Sg) (N Pl) ;
LIST-TAGS +=
<f:be_attribute_jpsych>
<H>
<jpsych>
<§TH_@SUBJ_<H>_N_Abs>
<p:<H>_N_Prop> ;
LIST be_attribute_jpsych =
"ajuallap" "pilluar" "nuannaar"
"uumila" ;
```

Both "objectmarking" and "subjectmarking" are defined in order for the rules to specify an intransitive verb ITR_pers. Transitive verbs are easier to capture in rules by simply saying any person (1-4) in any number (Sg or Pl) with the letter for object O at the end, whereas capturing a verbal form by it's subject inflection will work on both transitive and intransitive inflection, so an intransitive verb is defined as subject marking and no object marking. In order to be able to control where the semantic tags go in the cohort, both "any frame tag" and "any adjective prototype tag" are defined within regular expressions. Regular expressions are formulated in slashes ending in the letter "r"[14], /<f:.*>/r means any tag (square brackets) for semantic frames (starts with an f followed

[11] https://aclanthology.org/W19-0406.pdf
[12] Finite State Transducer
[13] https://edu.visl.dk/cg3.html
[14] https://edu.visl.dk/cg3/chunked/tags.html#regex-icase

by a colon) followed by any number of any character ".*". The LIST-TAGS += enables definition of multiple tags.

annotations for input: "nuannaarpoq" in
(4) meeraq nuannaarpoq

```
ADD:73 <f:be_attribute_jpsych>
AFTER ITR_pers OR noun_number
be_attribute_jpsych ;
ADD:85 <§TH_@SUBJ_<H>_N_Abs>
AFTER frame <f:be_attribute_jpsych> ;
```

Whether or not to include keywords such as TARGET and IF is a question of personal taste or idea of readability. The formal expression of the ADD rule can be read in the Constraint Grammar Manual[15] (Didriksen, 2010). In the example above (ADD:73), the keyword TARGET would have been right before be_attribute_jpsych, which is the definition for different stems, including the stem "nuannaar" in example (4). The predicative verb in the example is intransitive, and the rule adds the tag <f:be_attribute_jpsych> after the person and number inflection, whether it is a predicative verb or a nominalized predicative verbal stem expressing a psychological state (be_attribute_jpsych). The rule ADD:85 is an attempt at defining the frame template for the frame that is added at ADD:73, and the tag is placed after the frame tag. The frame template is a formulation of dependency slots. We expect that the syntactic function of the subject @SUBJ has the semantic role §TH[16] (theme), and that it is of a semantic prototype human <H>, which is a noun in absolutive case <§TH_@SUBJ_<H>_N_Abs>.

tagging for input: "meeraq" in
(4) meeraq nuannaarpoq

```
ADD:69 <H> human | Pron
+ (/[12][SP][gl]/r) ;
ADD:90 §TH <H> + Abs
+ (/\([SP][gl]\)/r)
(*1 <f:be_attribute_jpsych>
+ <§TH_@SUBJ_<H>_N_Abs>
LINK 0 (VSTR:/^[1-4]$1$/r)
- objectmarking) ;
```

The rule ADD:69 adds the tag <H> to the list called "human", where the noun "meeraq" is defined. The rule ADD:90 adds the semantic role tag §TH to the noun that is tagged as human <H>, which is in absolutive case. The absolutive case is the case that is used for the subject of an intransitive verb. The number is formulated as being in agreement with the (person and) number that the verb of the sentence is inflected for. The agreement is expressed in a variable string[17]. The absolutive can be singular or plural:

```
(/\([SP][gl]\)/r)
```

the regular expression group match is in escaped parentheses, and is matched with the variable string match $1 in:

```
(VSTR:/^[1-4]$1$/r)
```

One could consider adding the semantic role tag @SC (subject predicative/complement)[18], in case the tag is to be used cross-linguistically, e.g. for MT[19] purposes.

tagging for input: "nuannaartoq" in
(3) meeraq nuannaartoq

```
SUBSTITUTE:95 (/<f:.*_j.*>/r)
(<STEM:$1$2>v <$2>v) TARGET
(<\(f:be_attribute_\)\(j.*\)>r)
+ TUQ + noun_number ;
SUBSTITUTE:96 <§TH_@SUBJ_<H>_N_Abs>
<p:<H>_N_Prop> AFTER jsem <jpsych> ;
```

The rule SUBSTITUTE:95 splits the tag in two, where <f:be_attribute_jpsych> is split in <STEM:f:be_attribute_jpsych> and <jpsych>, when the predicative verbal stem is in PTCP form. SUBSTITUTE:96 changes the frame template with specification of head type, which in this case is a human prototype noun or proper noun.

5 Quantitive predicative stem with an intensifier

The following example will be of another type of predicative verbal stem, namely <jquant> *"quantitive"* that combines with countable head nouns. Countable nouns are in this paper tagged with <+countable>[20]. This combination of <jquant> and <+countable> will be used to demonstrate case error correction.

In this noun phrase type, there is a free adverbial particle between the head noun and the attributive noun. This structure is chosen for the case correction, because the context provides us with means to find potential errors within Oqaasileriffik's corpus[21] more easily than if we were to

[15] https://edu.visl.dk/cg3/chunked/rules.html#add

[16] https://edu.visl.dk/tagset_cg_all.pdf

[17] https://edu.visl.dk/cg3/chunked/tags.html#variable-strings

[18] https://edu.visl.dk/tagset_cg_all.pdf

[19] Machine Translation

[20] square brackets might have been preferable as feature tags for the sake of distinction from semantic tags

[21] https://oqaasileriffik.gl/en/langtech/corpus/

search for two adjacent nouns where the first is in absolutive case and the next one is in another form of case. The adverbial particle between the head noun and the attributive noun is *taama* (*"so"*, *"as"*, *"that"*), and it is in this structure inseparable with the deverbal suffix -tigə- that is suffixed between the predicative verbal stem and the PTCP:

(6) *ukiut taama amerlatigisuni inuusuttunik atuartitsisarsimavoq

ukioq.PL.ABS taama
year.PL.ABS that
amerla-tigi-su.PL.LOK
be.many-that-PTCP.PL.LOK
inuusuttoq-PL.INS
young.people-PL.INS
atuartit-si-sar-sima-IND-3SG
teach-HTR-HAB-PST-IND-3SG

'(he/she) has been teaching young people for that many years'

When the predicative verbal stem has an adjectival sense, the combination of *taama* and -tigə- is in equative sense (i.e. *as beautiful as*, *as clever as*). In example (6) above, when the predicative verbal stem in question (*amerla-*) is in quantitive sense (be.many), the combination of the adverbial *taama* and the suffix -tigə- instead functions as an intensifier (i.e. *that many*).

The word in need of case error correction in example (6) is *ukiut*, which is in absolutive case. The case needs to be corrected to locative, *ukiuni*, since the noun phrase is meant to be an adverbial phrase in locative case, as indicated by the case of the attributive noun.

The correction is handled in the grammar checker file:

```
WITH Abs + <+countable> + $$NUMERUS
(1 ("taama") LINK 1 TIGE + <jquant> +
$$NUMERUS)
{
ADD:msyn-abs-taama-case
&msyn-abs-taama-case (*) ;

COPY:msyn-abs-taama-case
(VSTR:$1 &SUGGEST) EXCEPT
(Abs &msyn-abs-taama-case)
BEFORE NUMERUS (*)
(2 (/^\(Rel|Abl|Via|Aeq|Ins
|Trm|Lok\)$/r)) ;

ADDRELATION:msyn-abs-taama-case
($2 RIGHT) (*) TO (jC1 (*)) ;
} ;
```

The rules are grouped with the WITH rule (see Swanson et al., 2023). We want to capture what is on the same line as the keyword WITH, which is a countable noun in absolutive case, and which is in the same number $$NUMERUS as another word which is defined in the context. The context is defined on the next line in parenthesis. The word immediately to the right of the head in absolutive case is "taama", and the second word to the right contains the suffix TIGE, has the adjectival semantic prototype tag <jquant> and has the same number inflection as the head in absolutive case. The 3 rules within the curly brackets apply to the context that is defined before the curly brackets. The ADD rule adds a grammar checker tag starting with an ampersand, &msyn-abs-taama-case, the asterisk in the parenthesis refers to the noun in absolutive case that is specified in the same line as the operator WITH. The COPY rule duplicates the reading with the &msyn-abs-taama-case tag, replacing this tag with another tag &SUGGEST that is used in the grammar checker pipeline to suggest corrections. Furthermore the duplicated reading replaces the absolutive case with whatever the case that the attributive noun in PTCP is in, formulated with a variable string.

The suggestion output in the grammar checker is as follows:

error:
ukiut taama amerlatigisuni
correction:
ukiuni taama amerlatigisuni

The feedbacks provided to the user are as replicated below:

```
There should be case (and number)
agreement in noun phrases

The case ending of "ukiut"
needs to be the same as
the case ending of "amerlatigisuni"
```

Specifying combinatorial requirements for a lexical element according to it's semantic features in this manner is practical in determining whether or not for example the noun in absolutive case in question actually is correct, by additionally defining semantic restrictions on the verb of the sentence. The frame for the verb in the example in question is "teach", and the form is in intransitive. The transitive form of the verbal stem that denotes "teach" is detransitivized with the so-called half-transitive[22] suffix, which means that the syntactic subject of the verb with the semantic role of agent

[22] also called antipassive by some authors

is in absolutive case, and what would have been the object of the transitive form is in instrumental case (oblique object in instrumental case). In our particular example, the disambiguation is straightforward in that there has to be person and number agreement between subject and inflection for subject in the verb. The noun in question in our example *ukiut* is in plural, whereas the verb is inflected for a subject in singular. Greenlandic is a pro-drop language, and the subject in the example is implicit. If however the inflection for the subject in the verb had been in plural, morphological information alone would not be sufficient to determine if the noun is semantically compatible as the subject of the sentence, and we would have to rely on ad hoc list definition for what the particular verbal stem could have as a subject.

Our rules need an addition to their contextual requirements in the file containing semantic annotation rules:

```
LIST hum_hum =
(/"atuar"\ Gram/.V\ TIP\ Der/vv\ Gram/../l)
<f:teach> ;
SET nonhum = (<.*>r) - (<H.+>r) ;
ADD <f:teach> AFTER objectmarking
OR ITR_pers hum_hum ;
ADD
<$AG_@SUBJ_<H>_N_Rel-$BEN_@OBJ_<H>_N_Abs>
AFTER frame <f:teach> + objectmarking ;
ADD
<$AG_@SUBJ_<H>_N_Abs-$BEN_@OBJ_<H>_N_Ins>
AFTER frame <f:teach> + ITR_pers ;
```

in the grammar checker file we add:

```
(NEGATE 0 <H> + (VSTR:$1)
LINK *1 V + hum_hum
+ ITR_pers +
(/^[34]\([SP][gl]\)$/r))
```

In short, the verb of the sentence (teach) is a human to human verb (both subject and object are human). We don't wan't to make corrections in cases where the absolutive human noun might be in number agreement with the subject inflection of the intransitive verb, because the noun in absolutive case then could be the subject of the sentence.

6 Exceptions

Oqaasileriffik's corpus has a total of 1.443 sentences that contains the phrasetype

```
Abs + taama
+ [predicative_verb_stem+TIGE+PTCP
+non-absolutive case]
```

A quick glance at these examples shows plenty of instances where the first part in the absolutive case needs no correction. In those instances, there is in most cases a clear violation of semantic combinatorial restrictions between the head noun and the attributive noun, indicating that the noun in absolutive is not to be corrected, but perhaps more likely is the subject of an intransitive subordinate or superordinate verb in the sentence. The examples below start with the Greenlandic sentence, where the noun phrase in focus is in boldface and translated to English below the sentence. The incompatibility between the semantic prototype of the head noun before the particle *taama* and the attributive noun after *taama* are shown below in English, and the translation of the whole phrases is also provided. Case is indicated in square brackets. These square brackets in the following examples have no function and are not used as feature tags:

(7) Taamaattumik **Inughuit taama sakkortutigisumik**
isummersinnaasoqalersimappata tamanna tupigineqassanngilluinnarpoq

Inughuit that harsh(ly)

"Therefore it absolutely shouldn't be a surprise to anyone if someone among the **Inughuit** *has acquired the ability to express their opinions* **that harshly***"*

```
<H> [ABS] "taama" <jdegree> [INS]
```

(8) Niaqornaveeqqutitik qallersaassuatillu peeriarlugit tamanut tamaanga igiinnarsimavaat,
kinaluunniimmi taama kiatsigisumi
meqqulualissuarmik qallersaateqarluni sininnaviannginmat.

nobody (in) that warm

"They had removed their helmets and their thick outerwear and spread them all over the place, **nobody** *could sleep* **in (a place) that warm** *wearing outerwear with down (in it)"*

```
<H> [Abs] "taama" <jtemp> [LOK]
```

this particular word-form to denote "warm" is not compatible for metaphoric or literal use to describe humans in Greenlandic. This ambiguity can be resolved with a domain tag for weather specific adjectives <Dweather>[23]

[23] https://edu.visl.dk/semantic_prototypes_adj.pdf

(9) Taamani sikorsuaqarnera pillugu **umiarsuit taama sivisutigisumik** uninngapput Nuup umiarsualiviani.

ships that long

"Back then, the **ships** *were docked in the harbor for* **that long**, *because of storis (a floating mass of closely crowded icebergs and floes[24])"*

```
<Vwater> [Abs] "taama" <jtime> [INS]
```

this particular word-form to denote "long" is for temporal use only in Greenlandic

(10) **Apeqqutissaralu taama inuusutsigisumut** angivallaassagaluarnersoq neriuppunga silassorissunnguugavit akisinnaassagit

and my question (to) that young

*"***And** *I hope that since you are so clever, that you can tell whether or not my* **question** *would be too big* **for such a young (person)***"*

```
<sem-s> [Abs] + taama + <jage> [TRM]
```

(11) niaqqi qattornup qulaagut nuisippaa, **kinguninngualu taama sukkatigisumik** tarrisillugu

right after that quick(ly)

"he/she/it let his/her/its head appear at the top of the hill, and made it disappear **as quickly** *(as it had appeared)* **right afterwards***"*

```
<temp> [Abs] + taama + <jspeed> [INS]
```

(12) piginnittuanulli **kusaginninnini taama sukkatigisumik** takutippaa

her/his admiration that quick(ly)

"but he/she showed **his/her admiration** *to the owner* **that quickly***"*

```
<f-psych> [Abs] + taama + <jspeed> [INS]
```

(13) **pajugutaali taama sukkatigisumik** utertinneqarpoq

his/her gift that quick(ly)

"but **his/her gift** *was returned* **that quickly***"*

```
<cc> [Abs] + taama + <jspeed> [INS]
```

The instrumental case is used for 3 main functions in Greenlandic: oblique object (when the main verb is halftransitive[25]), modifier to an incorporated[26] object, or manner adverbial. The latter category is interesting in our examples, since all the instrumental cases that are used in the examples are used adverbially. These examples are perhaps comparable to danish so-called t-adverbials, that are formed by adding a -t on adjectives [27], and are categorized in 3 different types: *manner*, *time* and *degree*. Notice that this is consistent with our examples: (7) <jdegree> (9) <jtime>, and (11) + (12) + (13) falls into the broader category *manner*, namely <jspeed>.

At the moment, Oqaasileriffik's corpus is not available in semantically annotated version. The next step in this pilot project would be to extract those total of 1.443 sentences from corpus and have them parsed with their semantic annotations, manually identify case errors in the head noun, formulate further semantically compatibility restrictions, and make case error corrections in the grammar checker that later can be used for accuracy testing purposes.

7 Concluding remarks

The pilot project described in this paper is unfortunately at its very early beginning and, for this reason, the paper has not been able to show any statistics for accuracy testing. What the findings show is that there is a crucial need to define semantic compatibility restrictions where morphological information does not suffice. This would prevent false positives. Furthermore, the frame information used for predicative verbs needs to be made more specific by adding information about the prototype for the attributive meaning that the nominalized version in PTCP would have, in order to use them to describe their combinatorial restrictions in an accurate manner.

[24] *https://www.merriam-webster.com/dictionary/storis?fbclid=IwZXh0bgNhZW0CMTEAAR3zaIs8MNtCW4Ja-D8viUx-_tIZLNSUs6ENW5JKyL-2sh71vf1KZXZv2i0_aem_AYFjJKn7WSz_3CP1QLnqig*

[25] also called *antipassive* by some authors

[26] also called *inderived* by some authors

[27] https://sproget.dk/typiske-problemer/adverbielt-t-biords-t/

8 List of abbreviations, semantic prototype labels and semantic frames

ABL	ablative case
ABS	absolutive case
AEQ	aequalis case[28]
CAU	causative mood
CON	conditional mood
CONT	contemporative mood
HAB	habitual aspect
HTR	halftransitive suffix
IMP	imperative mood
IND	indicative mood
INS	instrumental case
INT	interrogative mood
LOK	locative case
NEG	negation
OPT	optative mood
PAR	participial mood
PL	plural
PST	past[29]
PTCP	nominal participle[30]
REL	relative case
SG	singular
TRM	terminalis case[31]
VIA	vialis case[32]
3SG	third person singular
<cc>	object countable
<+countable>	countable noun
<f-psych>	feature psychological
<H>	human
<jage>	age
<jdegree>	degree, intensity
<jspeed>	speed
<jtemp>	temperature
<jtime>	time
<sem-s>	semiotic, speech
<temp>	temporal
<Vwater>	vehicle water

[28] also called *equative* by some authors

[29] In the form of a suffix. Tense is not a grammatical category in Greenlandic.

[30] also called *active participle* by some authors

[31] also called *terminative* or *allative* by some authors

[32] also called *perlative* by some authors

Acknowledgements

I would like to extend my gratitude to the welcome reception I have been given, when visiting Giellatekno during late 2023 and 2024, Trond Trosterud (Professor of Sámi computational linguistics, leader of the research group Giellatekno. Department of Language and Culture) who first introduced me to Giellatekno's grammar checker (among many other things), Linda Wiechetek (Senioringeniør. Department of Language and Culture) who spent a great deal of her time to give me a thorough introduction, Kevin Brubeck Unhammer (Trigram AS) for taking time in giving a helping hand with grammar checker CG-related issues whenever needed, Sjur Nørstebø Moshagen (Sjefsingeniør / leiar for Divvun-gruppa Institutt for språk og kultur) and Flammie Pirinen (Engineer. Department of Language and Culture) for technical support whenever needed, last but not least Tino Didriksen who not only tirelessly answers any CG-related questions that I have, but also has been assisting in the whole process of implementing the grammar checker pipeline to Greenlandic.

9 Bibliography

Daniel Swanson, Tino Didriksen, and Francis M. Tyers. 2023. WITH Context: Adding Rule-Grouping to VISL CG-3. In Proceedings of the NoDaLiDa 2023 Workshop on Constraint Grammar - Methods, Tools and Applications, pages 10–14, Tórshavn, Faroe Islands. Association of Computational Linguistics.

Eckhard Bick. 2019. A Semantic Ontology of Danish Adjectives. IN: Simon Dobnik, Stergios Chatzikyriakidis, Vera Demberg,

(eds.), *Proceedings of the 13th International Conference on Computational Semantics – Long Papers*. May 2019. Gothenburg, Sweden. Association for Computational Linguistics. pp. 71-78.
https://aclanthology.org/W19-0406/

Kenneth R. Beesley, Lauri Karttunen. 2003. Finite State Morphology. CSLI Studies in Computational Linguistics. CSLI Publications. Linda Wiechetek. 2017. "When grammar can't be trusted – Valency and semantic categories in North Sámi syntactic analysis and error detection". A dissertation for the degree of Philosophiae Doctor. Department of Language and Culture (ISK). December 2017.

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.

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

Tino Didriksen. 2010. Constraint Grammar Manual. 3rd version of the CG formalism variant. GrammarSoft ApS, Denmark.
http://visl.sdu.dk/cg3/vislcg3.pdf