Analogical transfer and analogical levelling in the West Greenlandic verb paradigm

Robert Schikowski

Keywords: West Greenlandic, verbal person markers, analogy, analogical transfer, analogical levelling, building block metaphor

Abstract

As an Eskimo language, West Greenlandic has a rich verbal morphology: The prototypical verb is marked for one of eight moods and subjective (intransitive verb) or agentive and objective (transitive verb), respectively. The mood markers constrain the occurrence of following person markers, which are in turn dependent on each other and frequently amalgamated into portmanteaus, leading to a synchronically highly irregular system. In this paper I introduce some recent analogies in this system which are not yet described in the literature, specifically analogical transfer in the monopersonal contemporative mood and analogical levelling in the complex subordinate moods (causative, habitual and conditional) and the interrogative. In the conclusion I make some observations regarding which kind of language model the data support and which not.

0. Introduction

West Greenlandic or Kalaallisut is spoken by some 50,000 people in Greenland and Denmark (www.ethnologue.com, www.statgreen.gl). Its genetical background is Inuit > Eskimo > Eskaleut. To make the present paper easier to read, it should be mentioned that alignment is basically ergative (although there are clear accusative traits in many places) and that the language is polysynthetic. The verbal system is built from eight (suffixed) moods, which are properly divided into superordinate moods (indicative IND, interrogative INTR, imperative-optative OPT) and subordinate moods (contemporative CONT, participial PART, causative CAUS, habitual HABT, conditional COND), expressing superordinate and subordinate relations, respectively. The present time interrogative is for the most part identical to the indicative and will therefore not be given special consideration below except where necessary. Following the mood markers are the person markers for subjective (S) in intransitive and for agentive (A) and objective (O) in transitive forms. The mood markers place constraints on the occurrence of person markers, which in turn are in many cases fused into one morpheme. For this reason not a single person marker is the same through all moods and all bipersonal constellations, although there are many tendencies. It is the secondary aim of this paper to show that speakers are making use of these tendencies. The evidence I present are some recent analogies in the person marker system, the description of which constitutes the primary aim.

The descriptive part is based on the corpus Qassutit (QAS, about 1.8 mio. words) which has been extracted from the Greenlandic internet using the perl tool BootCaT (http://sslmit.unibo.it/~baroni/bootcat.html).
In addition to the abbreviations introduced en passant above I use the following (widely used ones not separately listed):

- **AP** antipassive
- **EQU** equative case
- **FUT** glosses for three different suffixes whose meaning is close to future tense
- **NOMZ** nominalizer
- **OBL** oblique case (cover term for locative, allative, ablative, perlative and equative)
- **PP** passive participle
- **RES** resultative
- **4** the traditional fourth person; close to a reflexive
- **>** signifies the relationship between A and O (e.g. 3SG>2SG = 3SG:A:2SG:O) or possessor and possessum (e.g. 1SG>SG:ABS = 1SG possessor, SG:ABS possessum)

Glossing is done according to the Leipzig Glossing Rules (www.eva.mpg.de/lingua/files/morpheme.html). To represent the morphophonological behavior of suffixes, the symbols + (additive), - (subtractive) and ± (mixed type) are used.

### 1. Analogical transfer in the contemporative

#### 1.1 Structure of the paradigm

The contemporative belongs to the subordinate moods. It is the only mood which is for the largest part monopersonal: with the exception of 1PL:A and 2PL:A, only S/O is marked. Prototypically the contemporative is used in cases where S/A of the superordinate and the subordinate clause are coreferential, so usually no ambiguities arise from its being underspecified.

The mood marker for the contemporative is +llu-. Below I have inserted the paradigm. This and all following paradigms are taken from Fortescue 1982 and rearranged for better viewability.

<table>
<thead>
<tr>
<th>A</th>
<th>S/O</th>
<th>1SG</th>
<th>2SG</th>
<th>3SG</th>
<th>4SG</th>
<th>1PL</th>
<th>2PL</th>
<th>3PL</th>
<th>4PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ø (ITR)</td>
<td>-nga</td>
<td>-tit</td>
<td>-</td>
<td>-ni</td>
<td>-ta</td>
<td>-si</td>
<td>---</td>
<td>---</td>
<td>-tik</td>
</tr>
<tr>
<td>1SG</td>
<td>---</td>
<td>-tit</td>
<td>-gu</td>
<td>-ni</td>
<td>---</td>
<td>-si</td>
<td>-git</td>
<td>-tik</td>
<td></td>
</tr>
<tr>
<td>2SG</td>
<td>-nga</td>
<td>---</td>
<td>-gu</td>
<td>-ni</td>
<td>-ta</td>
<td>---</td>
<td>-git</td>
<td>-tik</td>
<td></td>
</tr>
<tr>
<td>4SG</td>
<td>-nga</td>
<td>-tit</td>
<td>-gu</td>
<td>-ni</td>
<td>-ta</td>
<td>-si</td>
<td>-git</td>
<td>-tik</td>
<td></td>
</tr>
<tr>
<td>1PL</td>
<td>---</td>
<td>-tit</td>
<td>-tigu</td>
<td>-ni</td>
<td>-</td>
<td>-si</td>
<td>-tigit</td>
<td>-tik</td>
<td></td>
</tr>
<tr>
<td>2PL</td>
<td>-singa</td>
<td>---</td>
<td>-siuk</td>
<td>-ni</td>
<td>-tigut</td>
<td>---</td>
<td>-sigit</td>
<td>-tik</td>
<td></td>
</tr>
<tr>
<td>4PL</td>
<td>-nga</td>
<td>-tit</td>
<td>-gu</td>
<td>-ni</td>
<td>-ta</td>
<td>-si</td>
<td>-git</td>
<td>-tik</td>
<td></td>
</tr>
</tbody>
</table>

The bipersonal forms of 1PL:A and 2PL:A must have been optional at an earlier time, since Kleinschmidt, the writer of the first grammar, notes: “Der Infinit. [i.e. CONT] kommt zwar gewöhnlich nicht mit andern, als den dort aufgeführten
Suffixen vor [i.e. only with O-suffixes], doch kann er, wenn die deutlichkeit des
ausdrucks es erfordert, an deren statt in der 1sten u. 2ten pers. die endungen
des Cjnct. u. Sbjnct. [i.e. CAUS and COND] annehmen, nämlich das, was nach
dem vocal des moduszeichens steht: vsiuk, vtigo etc.; es geschieht dies aber
selten anders, als in der mehrh. des subjects (ihr, wir), und auch da noch nicht oft (...)
(Kleinschmidt 1851:53). Kleinschmidt’s hypothesis about the origin of the
bipersonal suffixes is however doubtable because the contemporative has
single consonants where the causal and the conditional have geminates (e.g. vs
and vt correspond to modern ss and ts), and there is no obvious reason why
there should have been a split if the suffixes had been identical from the
beginning.
The markers might have soaked in from other moods (OPT and INTR contain
identical forms). The evidence for this is however weak because these two
moods are a minority among the other moods and stand in no formal or
functional relationship to CONT. It might as well be that the forms are old.

1.2 Analogical transfer
In addition to the these forms, three other bipersonal forms are attested in QAS
which have not been described yet. These are -ingga [4PL>1SG], -ssi
[1SG>2PL] and -tsigit [1PL>3PL]. Each of these forms is only attested once,
which may be taken as evidence for the recency of their origin. I list them all
below.

(1) (...) inequunartumik atserusussimavaannga,
iniquinar+Su±m+ik atsir-rusuC+sim+Va-annga
cute-PART-SG-INST give.a.name.to-feel.like-RES-IND:TR-3PL>1SG

aaliangersimappullu nunatsiini atimik
aaliangir+sima+pput=lu nuna-tsiC+ni atiq±m+ik
decide-RES-IND:ITR:3PL=and land-1PL>SG:OBL-LOC name -SG-INST

atugaavallaangitusumik atserususullunga.
atur-gaq-u+Vallaar-nngit+Su±m+ik atsir-rusuC+llu-nnga
use-PP-be-too.much-not-PART-SG-INST give.a.name.to-feel.like-CONT-4PL>1SG

‘They wanted to give me a cute name, and they decided to give me a
name not used to often in our country’ [QAS 26905]

(2) Neriuppunga takoqqikkumaarlussi!
niriuc+V+u-nga taku-qqiC+Jum+ar+llu-ssi
hope-IND-ITR-1SG see-again-FUT-CONT-1SG>2PL

‘I hope I’ll see you again some day.’ [QAS 52910]

(3) (...) pegatigiiffinnut saaffiginneriarluni aningaasanik
piqatigiivik+t+ut saavvigiC+niC-riar+llu-ni aningaasaq+t+ik
association-PL-ALL turn.to-AP-shortly-CONT-4SG money-PL-INST

1 ‘The contemporative usually does not cooccur with other suffixes than the ones cited above;
however, if required for clearness of expression, it can take the endings of the 1
and 2 person of the causative and the conditional instead, that is, what comes after the vowel of the
mood marker: vsiuk, vtigo etc.; this does rarely happen except in the plural of the subject (you,
we), and even there it is not frequent.’
Robert Schikowski

katersiriarluni oqallinneq allatsinniarli taava
katirsi-riar+llu-ni oqall+C+niq allat+tit+niar+li taava
collect-Shortly CONT-4SG discuss-NOMZ write-let-FUT-3SG:S then

immitsinnut nanissuugut suut
immi-tsiC+nut naniC+Su-u+V+u-gut su-ut
self-1PL>SG:OBL-ALL cave.in-PART-be-IND-ITR-1PL what-PL

tunngaviginerlutsigit.
tunngavik-gi+nir+llu-tsigit
basis-have.as-POT-CONT-1PL>3PL

‘He should appeal to associations, collect money and participate in the discussion, then, taking these things (his achievements) as a basis, we may cave in to ourselves.’ [QAS 6952]

The appearance of these forms is all the more surprising as the contemporative is basically monopersonal. This means that they do not stem from analogical levelling within the paradigm or with corresponding forms in other paradigms but that they are a true case of analogical transfer.

-nnga for 3PL>1SG can be found in almost all paradigms. IND and PART have the slightly different form -anna. Consequently -nnga must have been transferred from one of the other moods, that is, OPT or CAUS-HABT-COND.

-ssi appears in all paradigms. Its origin in the CONT paradigm cannot be determined.

-tsigit is marking the constellation 1PL>2SG in IND, OPT and PART. However, as this interpretation is impossible in (3), the form can only be taken over from CAUS-HABT-COND, where it designs 1PL>3PL.

This shows that at least some person markers must be identified with each other across mood paradigm boundaries.

2. Analogical levelling in the complex subordinate moods and in the interrogative

2.1 Structure of the paradigms

CAUS, HABT and COND can be grouped together as a subgroup within the subordinate moods on the grounds of various formal and functional commonalities (‘komplexe Nebensatzmodi’, Schikowski 2006:62). In particular, their person markers are identical for the most part.

Below is the paradigm of the causative, which is representative of the two other moods. The mood marker is +mma- for the 3SG/PL:S/A and +Ga- for all other persons.
Analogy in the West Greenlandic Verb Paradigm

The person markers for the other two complex subordinate moods are +Gaanga- for the habitual (all persons) and +ppa- (3) / +Gu- (other persons) for the conditional. The forms of the 4S/A beginning with -m- in CAUS and HABT begin with -n- in COND.

In addition we need to consider the paradigm of the interrogative, which is largely identical to the indicative, but still maintains some special forms (there were more dedicated forms at an earlier time, cf. Kleinschmidt 1853). The general mood marker is +Vi-. However, the 3SG/PL:S shows the irregular forms +Va and +ppat, where mood and person markers are fused. Below the paradigm is shown. The forms printed in bold letters are the ones different from IND. All other forms are taken from IND and do not take the INTR mood marker.

### 2.2 Analogical levelling

Considering comparative evidence from the other Eskimo languages, (one set of) person markers for 2SG:O and 1PL:O can be reconstructed as *-ten und *-tikut. The -s- in -sit [4SG>2SG] and -sigut [4SG/PL>1PL] in CAUS-HABT-COND and -sigut [2SG/PL>1PL] in INTR arose through progressive palatalization after old -i- (t > s / i_V). However for all of these, forms with a -t- instead of the -s- are attested in QAS, e.g. in:
Robert Schikowski

(4) **Illersuinngikkutta** **kiap** **Illersussagamisigut?**
illsrsur-i-ningit+Gu-tta ki-ap illersur-ssa+Ga-misigut
protect-AP-not-COND-1PL:S who-ERG protect-FUT:1-CAUS-4SG>1PL

‘If we don’t protect ourselves, who will protect us?’
[QAS 58174]

(5) (...) **taava aappartaavit** **paasigunitit** **taamatut**
taava aappaq+taaq-vit paasi+Gu-nilit ta-imat+tut
then partner-new-2SG>SG:ERG understand-COND-4SG>2SG ANA-so-EQU

piumasatit eqqortinniarqarsinnaavai (...) pi+Juma-saq-tit iqqur+tit+niar+sarit+sinnaa+V+a-i
DMY-want-PP-2SG>PL:ABS come.true-let-FUT2-HAB-can-IND-TR-3SG>3PL

‘Then, if your new partner understands you, that way he can make your
wishes come true.’ [QAS 93351]

(7) **Heeey, qaqugu tikeraarniarpitigut?!**
heeeey qaqugu tikiraar+niar+Vi-tigut
heeeey when.FUT visit.on.a.trip-FUT

‘Heeey, when will you come visit us?’ [QAS25298]

Only 4 forms out of 17 attested in QAS contain the ‘correct’ -s-. For INTR, the
correct forms are not attested at all. The distribution of the forms is as follows.

<table>
<thead>
<tr>
<th></th>
<th>4SG&gt;2SG</th>
<th>4SG&gt;1PL</th>
<th>4PL&gt;1PL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAUS</strong></td>
<td>+Ga-misit 0</td>
<td>+Ga-misit 2</td>
<td>+Ga-misigut 2</td>
</tr>
<tr>
<td></td>
<td>+Ga-mitit 4</td>
<td>+Ga-mitit 2</td>
<td>+Ga-mitigut 1</td>
</tr>
<tr>
<td><strong>HABT</strong></td>
<td>+Gaanga-misit 0</td>
<td>+Gaanga-misigut 0</td>
<td>+Gaanga-misigut 0</td>
</tr>
<tr>
<td></td>
<td>+Gaanga-mitit 0</td>
<td>+Gaanga-mitigut 0</td>
<td>+Gaanga-mitigut 1</td>
</tr>
<tr>
<td><strong>COND</strong></td>
<td>+Gu-nisit 0</td>
<td>+Gu-nisigut 0</td>
<td>+Gu-nisigut 0</td>
</tr>
<tr>
<td></td>
<td>+Gu-nitit 1</td>
<td>+Gu-nitigut 1</td>
<td>+Gu-nitigut 0</td>
</tr>
</tbody>
</table>

|                | 5 | 5 | 4 | 14 |

<table>
<thead>
<tr>
<th></th>
<th>2SG&gt;1PL</th>
<th>2PL&gt;1PL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTR</strong></td>
<td>+Vi-sigut 0</td>
<td>+Vi-sigut 0</td>
</tr>
<tr>
<td></td>
<td>+Vi-tigut 2</td>
<td>+Vi-tigut 1</td>
</tr>
</tbody>
</table>

The form -tit for the 2SG:O is the regular (that is, documented) form in the
following moods, in combination with the following A:

IND (3SG, +V+a-atit, 38)
CONT (all A, +llu-tit, 4262)
PART (3SG, +Ga-atit, ~4)
CAUS (3SG/PL, +mma-tit, ~6)
HABT (3SG/PL, +Gaanga-tit, 0)
COND (3SG/PL, +ppa-tit, 9)

2 The reason for the frequency of +llu-tit is that it marks five different constellations, including
the frequently used intransitive 2SG:S.
In contrast, stable -sit only occurs in OPT (3SG/PL, +li-sit, 2). -tigut for the 1PL:O is regularly found in:

IND (3SG/PL, +V+a-atigut, 90)
OPT (2SG/PL, +tigut, frequency not separable)
CONT (2PL, +llu-tigut, 2)
PART (3SG/PL, +Ga-atigut, ~31)
CAUS (3SG/PL, +mma-tigut, 22)
HABT (3SG/PL, +Gaanga-tigut, 1)
COND (3SG/PL, +ppa-tigut, 8)

Considering the stable -s- in OPT -sit it would seem plausible that the -s- in +li-sigut- [OPT-3SG>1PL] should be stable, too; however, the form is not attested in QAS.

From the evidence above it becomes clear that the regular s-forms are by far outnumbered by the regular t-forms both in terms of type frequency (that is, in this case, frequency of a form-meaning conjunction across mood paradigms: 2:9 for -sit/-tit and 4:13 for -sigut/-tigut) and token frequency (2:~300 for -sit/-tit, 0:~150 for -sigut/-tigut). It would seem that the pressure for analogical levelling is therefore quite high; however, if it were, the OPT forms should be levelled, too (though it might be the case that they are just too rare to be attested in their levelled form in a relatively small corpus as QAS). For the time being, nothing can be said about the reason for the rise of the analogical forms.

It is interesting that the analogical levelling may not be said to be motivated by similarity of the concerned forms with other bipersonal constellations as a whole, because in none of the paradigms containing regular -tit and -tigut do these follow the 4SG/PL:A (2SG/PL>1PL for INTR). If the analogical forms are not similar to the whole of any other person marker constellation, this means that the patterns after which they are formed must be segmented synchronically by speakers. Even more interestingly, of the patterns listed above only IND and PART allow for a complete analysis, the remains of subtracting -tit and -tigut constituting a proper A marker -a-, which is present throughout the whole 3SG:A row for both moods. In contrast, A is amalgamated into the mood marker in CAUS and COND and unmarked in HABT and OPT.

3. Some general conclusions

In the following I would like to consider the meaning of what has been said above for a general theory of language. Which kind of language model do the findings from this paper support, and why?

First I would like to address what is so bewildering about the West Greenlandic verbal system (and any other similarly chaotic morphological system) at first sight. As we have seen above, mood and person markers are intimately connected. There are very general tendencies concerning the distribution of person markers (e.g. all moods which use -gu/-uk for the 3SG:O also use -git for the 3PL:O) as well as total idiosyncracies (e.g. the use of -ma for 2SG>1SG in IND as opposed to -mma in PART and CAUS-HABT-COND). Only in a few
cases can the variation be interpreted as allophony. Furthermore, the person markers themselves are very often inseparable from each other, as in +V+a-\textit{kkit} [IND-TR-1SG>2SG] or +li-\textit{ssuk} [OPT-3PL>3SG]. Now if we assume that forms like -\textit{kkit} or even +V+a+kkit are stored as a whole, we are confronted with the problem of what to do with the better segmentable forms like +\textit{Gu-ni nga} [COND-4SG:A-1SG:O]. The problem is that despite their (linguistic!) segmentability, their segments are never fully coextensive with the meaning ascribed to them because their applicability is restricted.

Linguists traditionally feel uneasy with the thought that forms lacking a certain minimal degree of mobility should be meaningful. For example, the sound sequence [st] in English \textit{stiff} is bound to be considered meaningless on structuralistic grounds, because it only gains meaning in connection with the rest of the word: [st] uttered without context tells us nothing, as is the case for the rest of the word, [\textit{if}]. Furthermore it is not possible to insert other forms in the slot of [st] in order to change the meaning of the whole in a regular way. So it would seem like [st] is meaningless: It is never experienced as mobile by speakers and therefore not accessible as an element on its own. But this is only true for the side of production. On the side of perception, speakers experience a statistical connection between the occurrence of [st] and a very general radial concept containing ideas such as straightness, erectness and rigidity, which is also present, for instance, in \textit{stick, stave, stab, stubble, stare} and \textit{steady}. The uneasiness with this way of thinking has its ultimate source in the widespread metaphor of language as something like a game of building blocks. In this view of language partial analyzability of forms is not allowed, for what can be interpreted (that is, what can be segmented) has got to be laid out for being interpreted (that is, it has got to be assembled).

Now back to West Greenlandic person markers: If we analyze +\textit{Vaatit} as +V+\textit{a-a-tit}, we postulate meaningful units +V- [IND], +a- [TR], -a- [3SG:A] and -\textit{tit} [2SG:O]. But all these units are extremely limited in their mobility: For example, the forms +V+a+kkit [IND-TR-1SG>2SG], +V+a-tsigit [IND-TR-1PL>2SG] and +V+a-atsit [IND-TR-3PL>2SG], which likewise contain 2SG:O semantically, do not contain -\textit{tit}. If we take the rigid building block metaphor as our starting point, we have to choose between two absurd options: If we insist that +\textit{Vaatit} is +V+\textit{a-a-tit} (and for the time being ignore possible differences in the ordering of A and O), we would have to say that +V- is suffixed as soon as the meaning of IND becomes salient, subsequently followed by +a-, -a- and -\textit{tit} as soon as TR, 3SG:A and 2SG:O become salient, respectively. It is only if TR determines the use of +a- that can we say that +a- has the meaning TR. Consequently we are forced to have +a- in the underlying forms of all transitive forms. These underlying forms, most of which never have seen the light of day, have then to be rewritten by a mass of idiosyncratic rules stating, for example, that something like +\textit{Ga-a-p+nga} [CAUS-TR-2SG:A-1SG:O] > +\textit{Gamma}. This seems immediately undesirable.

The other option is to assume that the form of the marker is only fixed after all the information necessarily contained in a verbal ending is known. In other words, it is only when we know that what we want to say is IND and TR and contains a 3SG:A and a 2SG:O that we can decide about the form of the ending. But then +\textit{Vaatit} is clearly no longer assembled, because it comes up as a whole. In this case we would have to let go of the analysis +V+a-a-tit as a
synchronic phenomenon and instead have to assume something like +Vaatit [IND:TR:3SG>2SG]. This is equally awkward but would seem feasible if it were not for the analogies described above: Person markers clearly cannot exert influence on other markers if they are not analyzed as such - in the case of the new contemporative forms, they have at least to be separated from the mood marker +llu-, and in the case of the complex subordinate moods and the interrogative, the person markers themselves have to be kept apart, too. So in the end this extreme is no good either.

By the way I doubt that any linguist is actually thinking in such extreme terms (at least I hope so). What I want to say is just that if we try to think about the consequences of some seemingly inoffensive linguistic commonplaces (which is not usually done) we easily get into extremes.

Sure enough the answer to the dilemma described above is that we should abandon the rigid building block metaphor. It may seem like linguistic common sense has for long arrived at this standpoint - for example, in speaking of allomorphs we are implicitly assuming that the rigid building block metaphor is of no descriptive use. My claim is that the meaning of such implicit assumptions has until now too seldom been explored in-depth. Of course this paper cannot be the place of such an exploration - the evidence comes from just one language, and the space is limited. Instead I would like to append a very rough sketch of how I would like to think about extreme morphological interdependence and analogies of the type described above.

If we abandon the building block metaphor, we have to accept that there must be something like partial analyzability: In some ways a form is analyzable, and in other ways it is not. One very simple way to enable partial analyzability is to separate the processing of a form from its production: Even if +Vaatit should be monolithic on the side of production because it can only be formed after all its semantics components are known it can easily be segmented on the processing side, where each of its parts can be related to separate meanings also experienced in other forms. Most theories usually do not have problems with uniting several forms to one (that is, with assigning one meaning to several forms), in contrast to splitting up one form into several (that is, producing several alloforms from one meaning) without properly describing the conditions of this process.

Partial analysability could even be reintroduced to the processing side if we conceptualize the assembling of a form not as a serial process, but as a parallel one: As various semantic components gain saliency during the formulating of an utterance, the probability of certain forms rises; e.g. the presence of a 3SG:A activates all the forms statistically connected to the corresponding semantic component. This activation state is then subsequently modified by the stimuli of the remaining components until finally the strongest form emerges as the complete marker. One possible way for analogy to arise in such a model would be that frequent forms which are easilier activatable because of their high access rates eventually override the activation value of old forms.

I would be glad if this paper could become one tiny piece of evidence for the ones who are taking on the building of a theory heading for this direction.
References

www.ethnologue.com (SIL Ethnologue)
www.oqaasileriffik.gl (Greenlandic Language Council)
www.statgreen.gl (Statistics Greenland)