Arguments and Information Management in Inuktitut

Elke Nowak

(1) Introduction

Research on a variety of structurally different languages suggests that information is assigned to grammatical form in way of preferred representations of arguments. These preferences can be captured by four interacting constraints which are based on the analysis of spoken and written discourse. These constraints represent measurable discourse preferences: pragmatically unmarked utterances seem to follow them blindly and widely. Consequently, the preferences motivating these constraints seem to represent the default structuring of discourse in immediate relation to elementary grammatical form. Discourse is no longer viewed as acting upon grammatical form, but as being ‘grammatical’ itself.

For grammar a quantity constraint holds, limiting the optimal number of lexical arguments, as opposed to pronominal ones to one. With respect to grammatical roles it can be stated that external arguments, subjects, of transitive structures, are rather not represented lexically.

i. One lexical argument constraint: Avoid more than one lexical argument per clause.

ii. Lexical A constraint: Avoid lexical arguments in A-position, i.e. as external arguments of transitive clauses.

These constraints are met by matching constraints on the pragmatic side, again concerning quantity and role: the optimal number of arguments representing new information is limited to one. As optimal locus for given information the external argument/subject of a transitive structure is identified.

iii. One new argument constraint: Avoid more than one new argument per clause.


The Preferred Argument Structure Hypothesis (PAS) tries to establish a substantial and universal correlation between elementary discourse patterns and grammatical coding. It was first developed as a contribution to the debate on ergativity. DuBois (1987) proposed that ergative marking and grouping is best to be understood as coding of a discourse pattern: ergative languages code the structural positions alike where “new information” is most often represented. This is accomplished by default case marking, the absolutive, which is often zero marked. Nominative-accusative languages single it out by accusative marking. In ergative languages carried on information, the “red thread” of discourse, is specially marked, by ergative case, again in relation to its preferred locus of representation. Since protagonists of carried on information tend to be animate or even human rather than inanimate or abstract, it is likely for them to be active participants, ergates. What emerged as explanation for ergative coding is of course
not restricted to this phenomenon. It seems to be the case that languages of different genetic affiliation and of clearly different structure prefer the syntactic position of direct object or the sole argument of an intransitive structure for introducing new information. The subject or external argument of a transitive structure, on the other hand, is the preferred locus for carrying on an already introduced or accessible topic/theme, i.e. “given information”. Since first and second person protagonists are immediately accessible in discourse, i.e. given, it is the third person arguments which are crucial. Further, new information is introduced preferably by lexical mentions, referring expressions, while carried on information is characteristically represented by pronominal or zero anaphora.

Any claim concerning discourse patterns and the preferred distribution of discourse roles in relation to grammatical form must be understood as being “soft” in nature – its violation does not render an utterance ungrammatical. Such claims cannot be based on introspection or judgements of grammaticality. If any such correlation exists, it must be tracked down in the most common ways of how utterances are construed: where, in what grammatical position, new information as compared to “given”, and carried on information, most likely is packaged - spontaneously, in the course of immediate utterance. At the same time, the question as to how given or new information usually is presented has to be answered: is it represented as lexical mention or in any other, non-lexical form available in the respective language. Since PAS claims to capture a substantial and universal correlation between elementary discourse patterns and grammatical coding, its application to polysynthetic languages qualifies as promising test case. In a language like Inuktitut given and new information should be assigned to identifiable grammatical constituents with clear speaker preferences – just as in any other language. In the following I will argue that information distribution in Inuktitut is directed towards preferred sites, too, but these differ from those identified by the present version of PAS. Since the syntactic pivot positions\(^1\) identified as reference points in PAS lack descriptive adequacy in Inuktitut, it must be reconsidered what possibly qualifies as target position of arguments. Due to the polysynthetic nature of Inuktitut, it must be reconsidered what constitutes a “lexical mention” of an argument.

(2) Inuktitut
In Inuktitut, as in other polysynthetic languages, arguments are represented morphologically on verbal complexes, i.e. by affixes often called ‘pronominal arguments’. Verbal roots or complex verbal stems must be specified for their arguments, there are no infinite forms. The pronominal arguments are not cliticized pronouns, but bound morphemes. They resemble inflection in that their “pronominal” content is fused with other grammatical information such as number, valence, and mood. In transitive inflection two arguments are expressed but morphologically fused, and consequently it is impossible

\(^1\) These positions are traditionally labelled S, A and O. S indicating the sole argument of an intransitive sentence, while A and O stand for external and internal argument respectively.
to establish any kind of immediate (structural) asymmetry between these two arguments. In the case of third person arguments, these may be lexically specified by constituents outside the verbal complex, which are then cross referenced by ergative and/or absolutive case respectively.

(1)

\[
\begin{align*}
\text{piqatiga} & \quad \text{ikumalirijuq} \\
\text{piqati} & \quad \text{-ga} & \quad \text{ikuma} & \quad \text{-liri-} & \quad \text{-juq}
\end{align*}
\]

companion -1s.poss.abs engine²-operate- -3s.itr

‘my companion was working the engine’

First and second person pronouns cannot be employed to specify arguments. Third person pronouns are better considered demonstratives and additionally exhibit a wide range of spatial meaning.

While first and second person arguments are never expressed lexically, the reluctance of speakers of Inuktitut to lexically specify third person arguments is amazing. In spoken discourse the number of lexically represented arguments as compared to morphologically represented arguments is suspiciously low. Research on a large corpus of Inuktitut child language (Allen and Schröder 2003) suggests near avoidance of lexical mentions: only 5.1% of the arguments were represented lexically.

…”only 7.8% of referring expressions in the Inuktitut corpus are represented lexically (5.1% of the arguments and 77.6% of the obliques)…” (Allen and Schröder 2003:312)

The results presented here support this impression, although not to such an extreme. They are based on a very small corpus representing an entirely different genre. Yet, when compared to Allen and Schröder’s findings, but especially in comparison to research on other languages, they seem to shed some light on the matter of “lexical mentions” and the distribution of information in Inuktitut.

The corpus is based on a narrative told by Armand Tagurnaaq, edited by Alexina Kublu and Mick Mallon. As opposed to child language, anchored in situation and context, it represents a genre of high information pressure. It is directed towards an unknown audience, with a stage to be set, the protagonists to be introduced. It can be assumed that the narrator did his very best with respect to style and elaborate, ‘good’ expression. As a consequence, an unusually high degree of explicitness can be expected, with no relief from situation and context. The corpus comprises a total of seventy-nine constituents, i.e. expression units which are separated by a blank space in writing. Of these, thirteen or 16.5% are prepositional/ adverbial; they will not be considered here, although they seem to represent a major source of new information. The focus will be on arguments proper. Forty-two or 53.1% of the constituents are verbal complexes. Of these, eighteen are transitive, i.e. specified for two arguments; twenty-four are intransitive, i.e. specified for one argument, the total of arguments

² ikuma literally translates as ‘fire’; the plural ikumat is lexicalized as ‘engine’. Bare plurals are lost with incorporation, but see (3), below.
amounting to sixty. The remaining constituents are particles such as tagva, deictic demonstratives (taingna), exclamations (atii) and conjunctions (amma).

None of the verbal constituents is simple, i.e. a verbal root inflected for its argument(s), but all are complex and comprise a considerable amount of synthesis, including incorporation.

Twelve constituents are lexical specifications of arguments, which amounts to 20% of the sixty arguments manifested in the verbal complexes. In the vast majority of cases arguments manifest as part of the verbal complex, as pronominal arguments. In order to identify the two arguments represented by transitive inflection, resort to case marking assigned to possible lexical specifications has to be made. It must be emphasized that talking of ‘ergative’ and ‘absolutive’ arguments exclusively serves this purpose. An ‘ergative argument’ is the one which may be lexically specified by a then ergative marked constituent. An ‘absolutive argument’ on the other hand may be lexically specified by an absolutive, zero-marked constituent. With respect to the ‘One lexical argument constraint’ and the ‘Lexical A constraint’, it is the lexically specified arguments which are of interest, although the lexical specifications do not have argument status themselves. For transitive utterances the above introduced constraints allow these predictions: the ergative arguments most likely represent afore mentioned information; they are less likely to be specified lexically. New information is more likely to be introduced by the absolutive arguments in transitive as well as intransitive utterances. Lexically specified arguments are more likely to be absolutive arguments.

(3) Arguments

The ‘One lexical argument constraint’ is immediately met by the data: there is not a single case in which both arguments of a transitive verbal complex are lexically represented. There is not a single case in which a lexical mention of an argument is repeated as such.

Since it is impossible to have two ergative or two absolutive arguments in a transitive utterance, or an ergative argument in an intransitive utterance, the thirty-six arguments of the eighteen transitive verbal complexes can be neatly separated into eighteen absolutive and eighteen ergative arguments with the potential to be lexically specified.

Of these, altogether seven arguments are lexically specified (19.4%): two ergative arguments are specified by lexical mentions marked ergative, which amounts to 11.1%; five absolutive arguments are specified by lexical mentions marked absolutive, which amounts to 27.7%.

Of the twenty-four intransitive verbal complexes, the single absolutive argument is lexically specified five times, which amounts to 20.8%.

These lexical mentions are marked absolutive.

As was to be expected, these results are much less dramatic than the ones presented by Allen and Schröder, but are still far removed from results from other languages. In his 1987 study DuBois reports for Sakapultec Maya adult narratives a total of 44.2% lexical referring expressions, as opposed to 20% in the present study, and 5% in Allen’s
and Schröder’s. Of these Sakapultec lexical referring expressions, 48.1% represented lexical S, as compared to 20.8% intransitive absolutes; 45.9% represented lexical O, as compared to 27.7% transitive absolutes in the Inuktitut corpus discussed here. The dramatic difference to other languages is underlined by the studies by Kumpf (2003) on English teacher discourse, Clancy (2003) on Korean child language and England and Martin (2003) on five Maya languages, were lexical representations of S and O range up to 90%. Figures for lexical A, comparable to ergative arguments in Inuktitut, are low in all studies: in Sakapultec Maya 6.1% are represented lexically (DuBois 1987), in the study on five Maya languages the figures for lexical A range between 4% and 11%, English teacher discourse exhibits 8% of lexical A, while Korean child language seems to employ most lexical As, namely 17%.

Inuktitut child language is not only characterized by a very low rate of lexical mentions of arguments (5%), but also by a low degree of transitivity (Allen and Schröder 2003: 312). The fact that in the present study 42.9% of the verbal complexes are transitive, as compared to 27.4% in child language can be attributed to adult language and to genre. Nevertheless, the fact remains that lexical mentions of arguments are not very popular in Inuktitut. Even an increase in lexical mentions of arguments from 5% in child language to 20% in elaborated adult language of the present study is considerably lower than in any other language investigated. Speakers of Inuktitut seem to strangely avoid lexical mentions. But this is not quite so. There is a designated place for lexical arguments – they are incorporated.

(4) Incorporated arguments
Restrictions on incorporation in Inuktitut are very liberal; verbal affixes incorporate bare roots as in (2), but also inflected nominal complexes as in (3), or multiply derived ones as in (4). Pronouns and particles, simple and complex, may be incorporated as well, as can bee seen with (5) and (6).

(2) aiviqsiliramnuk
 aiviq -si- -liq- -ramnuk
 walrus -come.across- -begin- -1d.caus.itr
 we two suddenly come across a walrus

(3) illutinnuaqHunga
 illu –tinnut –aq- -tunga
 house -1p.poss.term -move- -1s.itr
 “I arrived (at our) home” (= I went home)

---

See the appendix. These studies were not harmonized, but investigated adult as well as child language and different genres. Interesting are very low figures for lexical O in the Maya languages Mam (6%) and Q’anjob’al (21%).

Examples (1), (2), (4) and (5) are taken from the corpus investigated here; (3) and (6) are from my field notes.
Do incorporated items qualify as arguments? To answer this question it is helpful to point out some properties of incorporated lexical items, which set them apart from lexical items involved in word formation processes in languages like German, or English.

Incorporated lexical items are strictly obligatory and governed by the incorporating verbal affix as much as the pronominal arguments are. They are never co-referent with these. In example (2) aiviq ‘walrus’, is incorporated by the verbal affix -si- which roughly corresponds to ‘come across’. The resulting verbal complex aiviqsi- is modified aspectually by verbal -liq- and completed by the inflectional ending -ramnuk. -ramnuk indicates a single (intransitive) argument of the first person dual in the relational (complement) mood causalis. It is evident that any kind of internal co-reference between the incorporated lexical item and the pronominal argument must be excluded.

Incorporated items need not be “bare” in the sense of excluding grammatical modification, as is demonstrated with (3). Since synthesis is a strictly binary process in Inuktitut, the incorporated lexical mention is illutinnut “to our house”, an adverbial nominal constituent inflected for possession, number and the directional case terminalis. In (4), tuluaqtigijavuk represents the only way to create a ditransitive structure in Inuktitut. -javuk represents two pronominal arguments, namely a first person dual ‘we two’ in relation to a third person singular ‘him/her/it’. None of these arguments refers to the incorporated tuluaqti ‘the gorer’. It is the third person singular of transitive –javuk, which is lexically specified by aiviq, a remarkable fact in itself.

Incorporated lexemes are frequently picked up, referred to and further elaborated outside the synthetic complex. The walrus introduced in (2)
is modified by three successive constituents, elaborating on its unpleasant and dangerous properties, discussed in detail in Nowak (2006). Martha Angugiaq Ungalaaq (1985: 71) begins her life story by (7)

(7) taimaguq maqruungnik ukiuqarliqlungaguuq...
taima -guuq maqruuk -nik ukiuq -qaq -liq --lunga -guuq
part -narrat. two-obj.d winter-have--progr--1s.vpart.itr -narrat.
“it is said, when I was two years (winters) old…”

Extensions of incorporated items are linked to the synthetic complex by a (case-) marker, the objective, which indicates number and, possibly, possession of the incorporated item. In (7) maqruuk ‘two’ must be marked by the objective dual, quantifying the incorporated ukiuq ‘winter’ of the following verbal complex. Nominal constituents marked objective may not be incorporated. In that they equal absolutive and ergative constituents and are set apart from the adverbial cases. Last but not least: Incorporated lexical items may be referential as in (3). Incorporated lexical items qualify as arguments. They are strictly obligatory, they can be referential and may be externally modified, quantified and specified. Since verbal affixes do not have roots as counterparts, a whole range of mostly very basic predications force incorporation 7.

If we take incorporated lexical constituents into account, the statistics with respect to lexical mentions is very much improved. Seven nominals, plus one indefinite pronoun, plus one particle, can be added to the twelve lexical mentions of pronominal arguments. This amounts to an increase of 58,3% or even 75% if all instances are counted. Allen and Schröder mention 166 incorporated lexical items for their corpus (2003:327), a fairly large increase, too. They do not consider them.

As can be seen with examples (2) and (4), both taken from the corpus, the pronominal arguments of verbal complexes based on incorporated lexical items may be specified lexically. But since these two cases are the only instances in the corpus, such a possibility does not seem to be frequently used. In both cases an absolutive argument is lexically specified. In terms of grammaticality, the co-occurrence with lexical ergative is by no means excluded, but such a case does not occur in the corpus.

Reconsidering the facts, it is tempting to distinguish two kinds of arguments for Inuktitut: those being represented morphologically, as pronominal arguments. They are strictly obligatory and complete a verbal complex. These arguments may be lexically specified, which is comparatively rare. Lexical referring expressions are frequently incorporated, incorporation must happen with affixal verbs. To talk of “internal arguments” with respect to these incorporated items is not just a nice metaphor. As it seems, Inuktitut exhibits a distinct argument structure with respect to how and where arguments are represented.

7 For further discussion see Nowak 2004.
(5) Information management

Finally, the distribution of given vs. new information with respect to arguments and with respect to lexical mentions of arguments and incorporated arguments must be considered.

What is given information, what is new information? With respect to the genre of the corpus it seems to be justified to cut the problem short and apply a simple rule “if a lexical item has been mentioned before, it is given, if not, it is new”. Such a perspective includes incorporated items, but not the morphological arguments. It does not pay attention to the distance between the first and next mention. Under such a perspective, four out of five lexical mentions of a single intransitive argument are new, but all absolutive arguments of transitives are given. Setting aside the incorporated indefinite pronoun and particle, of seven incorporated lexical items five are new, and two are given.

(8) Distribution of given and new information - lexical mentions

<table>
<thead>
<tr>
<th>Type</th>
<th>Total</th>
<th>New</th>
<th>Given</th>
</tr>
</thead>
<tbody>
<tr>
<td>lexical absolutive itr</td>
<td>5</td>
<td>3(4)</td>
<td>2(1)</td>
</tr>
<tr>
<td>lexical absolutive tr</td>
<td>5</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>lexical ergative tr</td>
<td>2</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>incorporated nominal items</td>
<td>7</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

With respect to the pronominal arguments, a tendency may be stated, too. Transitive verbal complexes are not favoured for introducing new information, be it lexical mentions, and be it the pronominal arguments. Only in one of eighteen transitive cases the absolutive argument represents new information; seven of the twenty-four intransitives provide new information, 29.2%. But the best result clearly is with incorporated items: 5 out of 7 amounts to 71.4%.

(9) Distribution of given and new information - pronominal arguments

<table>
<thead>
<tr>
<th>Type</th>
<th>Total</th>
<th>Given</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>transitive verbal complexes</td>
<td>18</td>
<td>94.4%</td>
<td>5.6%</td>
</tr>
<tr>
<td>both arguments represent</td>
<td>17</td>
<td>94.4%</td>
<td></td>
</tr>
<tr>
<td>the absolutive argument</td>
<td>1</td>
<td></td>
<td>5.6%</td>
</tr>
<tr>
<td>represents new information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>intransitive verbal complexes</td>
<td>24</td>
<td>70.8%</td>
<td>29.2%</td>
</tr>
<tr>
<td>given information</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>new information</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Taking into account the small size of the corpus investigated here, and its peculiar genre, a claim concerning the distribution of new information may be taken as tendency, at the best. What can be stated is that incorporated lexical items seem to be accessed for new as well as given information, as are the absolutive arguments of transitive and intransitive verbal complexes. Ergative arguments represent to 100% given information. Absolutive arguments and incorporated arguments group together insofar as they are target positions for introducing new information as well as for carrying on given information. What sets
them apart is lexicality: Incorporation represents the established domain for lexical representation of arguments.

(6) Concluding remarks
Considering the constraints of PAS once again, it can be stated that the “Given A constraint” is met. None of the ergative arguments provides new information. The obligatory distinction of 3rd and 4th person in relational inflection, clearly specifying anaphoric co-reference is not just a disambiguating device. With transitive utterances, it is always the ergative argument which serves as point of reference. Applying the strict rule again, which identifies “new information” as not immediately afore mentioned, 4th person differentiation might be interpreted as indication of “new information”, not just as “switch in reference”.

Since Inuktitut possesses the distinct argument structure described above, the options of how to accommodate new information are potentially increased. In my corpus, however, there are only two instances to be found, where a pronominal argument of a verbal complex based on an incorporated item is lexically extended. Only one of them, given as example (1), here repeated as (10), qualifies for the introduction of more then one new argument, one being represented by the incorporated lexical item ikuma(t) ‘engine’, the other by the lexical specification of an intransitive argument, piqatiga, ‘my companion’.

(10)
piqatiga
piqati -ga
companion -1s.poss.abs
ikumalirijuq
ikuma -liri- -juq
engine -operate- -3s.itr
‘my companion was working the engine’

Consequently, the “One new argument constraint” is observed, too. Keeping in mind that the predictions made by PAS do not aim at grammaticality, but at preferences, it can be concluded that both discourse related constraints are supported by the Inuktitut data. But further investigation is dearly needed, also with respect to other polysynthetic languages. As for the constraints related to grammatical form, both were in need for re-interpretation with respect to how arguments are represented in Inuktitut: morphologically, as affixes and as incorporated lexical items. But these alternative manifestations of grammatical form do not render the grammatical constraints inapplicable. As has been shown, the “One lexical argument constraint” as well as the “Lexical A constraint” hold with respect to the synthetic complex. In Inuktitut, information structure manifests within the domain of the polysynthetic word.
(7) Appendix

(1) Inuktitut corpus
aiviq tuluaqtigijavuk by Armand Tagurnaaq

(1.1)
Total of constituents 79
Total of verbal complexes 42 53.1%

transitive verbal complexes (specified for two arguments) 18 42.9%
intransitive verbal complexes (specified for a single argument) 24 57.1%

Lexical specifications of pronominal arguments: 12 15.2%

prepositional/adverbial constituents 13 16.5%

(1.2)
Arguments represented by affixes (pronominal arguments)
total 60
lexical extensions of these arguments 12 20%

Lexical mentions with ergative indexing -up 2 11.1%
Lexical mentions, transitive, -O 5 27.7%
Lexical mentions, intransitive, -O 5 20.8%

(2) Distribution of given and new information in the Inuktitut corpus

(2.1) pronominal arguments

transitive verbal complexes total: 18
both arguments represent given information: 17 94.4%
absolutive argument represents new information 1 5.6%

intransitive verbal complexes total: 24
given information 17 70.8%
new information 7 29.2%

(2.2) lexical mentions

incorporated items total: 7 new: 5 given: 2
lexical absolutive itr total: 5 new: 3(4) given: (2)1.
lexical absolutive tr total: 5 new: (1) given: 5
(3) Data from other languages

(3.1) Sakapulteco adult narratives  
 lexical A 6.1%
 lexical S 48.1%
 lexical O 45.9%

(3.2) English teacher discourse  
 lexical A 8%
 lexical S 52%
 lexical O 60%

(3.3) Korean child language  
 lexical A: 17%
 lexical S 39%, 32%
 lexical O 44%, 51%

(3.4) five Maya languages: Sakapulteco, Mam, Tektiteko, Mocho, Q’anjob’al (England and Martin 2003: 140)
Roles occupied by core argument lexical NP’s in clauses with one lexical argument, as % of total one-argument clauses with lexical NP’s

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>Ma</th>
<th>T</th>
<th>Mo</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>lexical A</td>
<td>5%, 6%, 11%, 6%, 4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lexical S</td>
<td>58%, 89%, 56%, 58%, 74%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lexical O</td>
<td>37%, 6%, 32%, 35%, 21%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(8) Abbreviations

The hyphen-minus indicates open morpheme boundaries which need to satisfied.

1,2,3,4  first, second, third, forth person
s.d.p,  singular, dual, plural
tr  transitive
itr  intransitive
caus  causalis mood
vpart  verbal participle
poss  possessive
abs  absolutive
erg  ergative
obj  objective
term  terminalis
neg  negation
narrat  narrative
prog  progressive
part  particle
(9) References


Angugatiaq Ungalaaq, Martha 1985. Recollections. Arviat: Inuit Cultural Institute


Tagurnaaq, Armand 1999. aiviq tuluaqtigijavuk. In Intermediate Inuktitut. Mallon, Mick and Alexina Kublu. Ittukuluuk Language Programs. 4-1, 4-2, 4-3. Iqaluit: Nunavut Arctic College

© Elke Nowak, 19-Apr-07.