MIRKO GAROFALO

The role and distribution of nominalized clauses:
A view from Icelandic

1. Introduction

Nominalized clauses (a term adopted from Roussou (1991)) are clausal arguments or complements preceded by a) a personal pronoun (see (1)) which generally also plays the role of the expletive (see e.g. Sudhoff 2016 for German, Ruys 2010 for Dutch, Kaltenböck 2003 for English), or b) a determiner, like a definite article (see (2)) or a demonstrative pronoun (see (3)):

(1)   I hate **it** that I must go out of town with my parents this weekend

(2)   **To**   oti efighe apodhiknii tin   enohi tis
        the.NOM that left-3sg proves   the.ACC guilt her
        'The fact that she left proves her guilt' (Roussou 1991)   [M. GREEK]

(3)   **In**   ke Maryam raft ma'alum e
        this.NOM that Maryam left clear   is
        'It is clear that Maryam left' (Hartman 2012)   [PERSIAN]

The existence of nominalized clauses in clausal complementation entails the occurrence of two different types of embedded clauses, i.e. nominalized and bare clausal arguments. However, their distribution differs, since the determiner appears to be at times mandatory, optional or even ungrammatical. We can observe such variation in languages like Icelandic:

(4) a. Ég fagna [*(**því**) að ég skuli hafa hætt við flugið]
       I rejoice that.DAT that I   shall have cancelled flight.the
       'I'm happy that I cancelled my flight'

    b. Ég harma [(**það**)   að hafa sært þig]
       I regret that.ACC to have hurt you
       'I regret hurting you'

    c. Ég ætla   [(***það**) að fara í   bíó   á morgun]
       I intend that   to go   in cinema tomorrow
       'I intend to go to the cinema tomorrow'

The preliminary data we have just observed here raise two important questions. Firstly, considering the fact that bare clausal arguments already exist in languages that feature clausal nominalization and can be selected by certain lexical items like, for instance, the verb *harma* ('regret') in (4b), it is not quite clear what the need is for a nominalizer in the first place. Secondly, the very fact that nominalized clauses are at times mandatory, optional or ungrammatical suggests the existence of a grammatical pattern that needs to be interpreted. By observing empirical data from Icelandic, I will provide in this study a plausible answer to these issues. I propose that clausal nominalization is a phenomenon limited to DP positions which mainly occurs in order to check formal features (D, phi and case features) which otherwise

could not be checked directly by a bare clausal argument, which is a CP (cf. Rosenbaum 1967; see in particular Garofalo 2020). Nominalization can also occur without any need for feature checking if the clausal argument surfaces in a DP position (which explains pronominal optionality), but speakers tend to prefer bare clausal arguments in these instances since they are less costly. The complex distribution of nominalized and bare clausal arguments across syntactic positions is the result of the restrictions applied by formal features like lexical case on bare clausal arguments (as in (4a)) and the ones applied on DPs in non-DP positions, which affect nominalized clauses as well (as in (4c)).

This analysis will be structured as follows. In the second section, I will briefly discuss the most relevant hypotheses that aim to explain clausal nominalization in the linguistic literature and, at the same time, we will see how Icelandic is not consistent with them. In the third one, I will present the methodology to collect the empirical data needed to observe the phenomenon across syntactic positions since the main hypotheses proposed in the syntactic literature are problematic. In the fourth, I will present and discuss the data collected. In the fifth, I will explain in more detail the purpose and distribution of clausal nominalization in light of the empirical data. In the sixth, the main relevant conclusions of this analysis will be presented.

2. Main hypotheses in contrast with Icelandic

Among the hypotheses presented in the linguistic literature that aim to explain the purpose of clausal nominalization, three are the most common. The first one is what I call here the *Semantic Hypothesis* (cf. Kiparsky and Kiparsky 1971; Höskuldur Thráinsson 1979; Kastner 2015 among others), according to which the meaning of the matrix verb selecting the clause has an impact on the structure of the clausal argument selected. Especially in the case of Icelandic, Höskuldur Thráinsson (1979:chapter 4) proposes that true factives like *fagna* ('rejoice') select more structurally complex arguments, i.e. nominalized clauses as in (4a), while non-factives like *halda* ('think') tend to opt for a non-nominalized structure. The various categorizations of matrix verbs differ from account to account, but the main idea is that certain categories require a more complex argument on the basis of their meaning while other ones do not, which causes the nominalizer to be sometimes mandatory or optional.

The second hypothesis is the *Case-Marking Hypothesis* (see Knyazev 2016; Roussou 1991 among others), which correlates the use of the nominalizer to a) case assignment or b) case visibility. Regarding option (a), it is posited, on the basis of the so-called *Case Resistance Principle* (Stowell 1981), that a clausal argument, which is analyzed as a CP, cannot be assigned case because it already contains case-assigning features. Therefore, the determiner that nominalizes the clause intervenes to check case on its behalf (Roussou 1991). As for option (b), it is proposed (Knyazev 2016) that clausal arguments are often DPs independently of the presence of a determiner, which entails that, if a nominalizer must surface, it is because case needs to be visible. The two alternatives within the Case-Marking Hypothesis described here also analyze examples of clausal subjects like in (2) differently. The case assignment proposal suggests that pronominal obligatoriness depends on the fact that structural case must be assigned, while it is claimed in the case visibility account that nominalization is obligatory because the subject D-head must be visible (based on Landau 2007). The former option has been challenged by Anton Karl Ingason (2018), who proposes against CRP that a CP can undergo direct case-marking due to the fact that floating quantifiers associated to clausal arguments are assigned case in Icelandic.

The third common hypothesis is what I call the *DP Selection Hypothesis* (see e.g. Borsley and Kornfilt 1999; Kornfilt and Whitman 2011), based on which clausal

arguments are considered verbal categories and must be nominalized (i.e. embedded into a DP shell) when a verb takes a nominal category as complement. Nominalization is a phenomenon that can be observed at multiple structural levels, that is, for instance TP or vP, as with gerunds in English, where the verb is still able to select a verbal complement but, at the same time, it can be modified by a possessive like a noun:

(5)   His **playing** the violin was quite remarkable

According to the DP Selection Hypothesis, something similar occurs at the CP level where the verbal category is embedded into a functional nominal category (which explains the presence of a determiner) in order to become a suitable argument of a lexical item that makes a requirement for a nominal category.

Now, it is surprising to see that Icelandic appears to be problematic for all of these hypotheses. First of all, let us take into account the fact that it has been shown that nominalized and bare clausal arguments in Icelandic are DPs and CPs respectively (Garofalo 2020). In particular, this difference becomes evident when we extract an item from a clausal argument. As observed by Höskuldur Thráinsson (1979), Anton Karl Ingason (2018) and Wood (2012), extraction is only allowed when the pronoun does not surface, which suggests a) that a DP layer is only projected when *það* is present, b) that this DP layer is a barrier for extraction and c) that a non-nominalized clause is not embedded into a DP projection since it does not block extraction:

(6)   Ólaf$_i$ ákvað hún (***það**)   að heimsækja t$_i$
          Ólafur.ACC decided she that.ACC to visit
          'She decided to visit Ólafur'

Now, assuming this structural distinction, we can see, for instance, that the Case-Marking Hypothesis is problematic. The determiner *það* ('that')[1] in Icelandic, in fact, is optional with clausal subjects in first position, although it is mandatory after the finite verb. This undermines the idea that clausal subjects must be preceded by a determiner with structural case assignment. Even if we considered all clausal subjects as DPs following the case visibility proposal, it is interesting to see that the D-head does not need to be phonetically realized in (7a) in first position in Icelandic. In other words, case-marking by itself is unable to capture the contrast presented in (7):

(7)   a. [(**Það**)   að læra á hljóðfæri] er eins og að læra að hjóla
         that.NOM to learn to instrument is like   to learn to ride a bike
         'Learning to play an instrument is like learning to ride a bike'

      b. Er [*(**það**) að læra á hljóðfæri] eins og að læra að hjóla?
         is that.NOM to learn on instrument like   to learn to ride a bike
         'Is learning to play an instrument like learning to ride a bike?'

The alternative hypothesis presented by Anton Karl Ingason (2018), which considers clausal complements able to undergo direct case-marking in Icelandic (challenging CRP), is also problematic, since the existence of obligatory nominalizers (as in (4a)) indicates that a bare clausal argument is not enough to check case directly.

---

[1] See Garofalo (2020) for various arguments that *það* before clausal arguments is a demonstrative pronoun and not a personal pronoun.

In contrast to the Semantic Hypothesis, it is quite puzzling to see that, when we take a look at accusative clausal objects in Icelandic, every possible structural difference of clausal arguments due to the meaning of the matrix verb is neutralized, as shown in (8) (a similar situation appears to occur in nominative objects, see section 4.2). I am not aware of any clear exception to this trend:

(8) a. Ég harma [(**það**) að hafa sært þig]
         I   regret that.ACC to have hurt you
         'I regret hurting you'

     b. Karl samþykkir [(**það**) að einhver annar stjórni fundinum]
        Karl accepts   that.ACC that someone else   leads   meeting.the
        'Karl accepts the fact that someone else chairs the meeting'

     c. Ég held [(**það**)   að kjöt almennt sé skaðlegt fyrir heilsuna]
        I   think that.ACC that meat generally is harmful for   health.the
        'I think meat as a whole is harmful to people's health'

It is important to keep in mind that there are some restrictions on clausal nominalization which are applied to those instances where DPs cannot surface as complements of the verb as in (9) and (10), which suggests that the ability of DP selection is an underlying condition to allow clausal nominalization:[2]

(9) a. *Ég skipaði honum **þetta**   **verkefni**
         I   ordered him.DAT this.ACC assignment.ACC
         Literally: 'I ordered him this assignment'

     b. Ég skipaði honum (***það**)   að klára   þetta verkefni
        I   ordered him.DAT that.ACC to complete this   assignment
        'I ordered him to complete this assignment'

(10) a. *María hjálpaði honum **flutningana**
         María helped   him.DAT moving out.ACC
         Intended: 'María helped him to move out'

     b. María hjálpaði honum   (***það**) að flytja
        María helped   him.DAT that.ACC to move out
        'María helped him to move out'

The accusative pattern observed in (8), however, contrasts with dative and genitive objects in Icelandic, where the nominalizer tends to be mandatory or preferable, although according to Höskuldur Thráinsson (1979) there are exceptions like (11c) with the verb *spá* ('predict'). In other words, the Icelandic distribution of clausal arguments based on case is quite problematic if we consider the Semantic Hypothesis alone:

(11) a. Ég fagna [*(**því**)   að ég skuli hafa hætt við   flugið]
         I   rejoice that.DAT that I   shall have cancelled flight.the
         'I am happy that I cancelled my flight'

---

[2] The examples in (9) and (10) are ungrammatical even if the nominalizer or the DP tested in (9a) and (10a) have dative or genitive case.

b. Jónas krefst   [*(**þess**) að við sendum skýrsluna strax]
   Jónas demands that.GEN that we send   report.the immediately
   'Jónas demands that we send the report immediately'

c. Jón spáði   [(**því**) að Gísli myndi sigra]
   Jón predicted that.DAT that Gísli would win
   'Jón predicted that Gísli would win'

Lastly, as for the DP Selection Hypothesis, there is an important question one can raise based on the facts we have observed so far. If Garofalo's (2020) structural distinction between nominalized and bare clauses is correct, we can definitely explain pronominal obligatoriness through the DP Selection Hypothesis by proposing that a lexical item might require a nominal category instead of a verbal one. However, it does not clearly explain pronominal optionality, as it is not immediately clear why a clausal argument has to be embedded into a DP projection if it is already acceptable as a CP complement. Similarly, also the other hypotheses can account for mandatory pronouns, as one could attribute the presence of *það* either to case assignment or to the need for a matrix verb to have a clausal complement with a more complex structure, but they are unable to account for pronominal optionality. At this point, it is necessary to gather more empirical data in order to explore the properties of optional clausal nominalization. At the same time, however, we also need to collect data on mandatory pronouns so that we can empirically compare the characteristics of both instances of clausal nominalization and develop a more valid hypothesis to explain the distribution of *það*.

3. Methodology

In order to find out useful empirical data on optional and mandatory pronouns in Icelandic, I decided to investigate two aspects of nominalized and non-nominalized clauses. In general terms, following Labov (1966), when we analyze a syntactic variable like the one we are discussing, we are supposed to gather tokens of the variants involved from a specific source and compare their frequency. However, depending on which type of data we are observing, we can attempt to investigate the degree of acceptability of the relevant variants on one hand, and the degree of preferability of a syntactic variant on the other hand. From naturalistic data like corpora, we have the chance to compare actual tokens and assess the degree of preferability of a variant, as a speaker cannot use more than one variant at the same time in a specific utterance. Questionnaires can help us to assess how acceptable both variants are. These two perspectives allow us to possibly find recurrent patterns across syntactic positions where a certain variant is preferable even if both variants are acceptable.

Consequently, in the case of Icelandic, I opted for looking for naturalistic data in the so-called *Icelandic Gigaword Corpus* or *Risamálheild* (henceforth RMH; see Steingrímsson 2019), which currently contains about 2,5 billion words grammatically marked. The types of clauses I investigated specifically are *that*-clauses, infinitives and indirect questions. After that, I administered a set of short questionnaires to native speakers of Icelandic. I gathered different groups of participants online (at least 20 people in each group) through social media as well as *Ugla*, the main student portal of the University of Iceland. Every questionnaire was made up of 25 sentences (17–18 of them were fillers) that the participants had to judge on a scale from 1 (unnatural) to 7 (natural). Every questionnaire was shared in two different versions, A and B. Considering the fact that the two variants of clausal nominalization differ only by the presence or absence of *það*, it would be very easy

for a participant to spot the difference and figure out the objective of the questionnaire, which would lead to biases in their judgments. Therefore, the two variants of a certain token with a clausal argument (i.e. nominalized and bare) were separated and one put in version A and the other one in version B, so that none of the participants saw both variants of a token, but just one. Moreover, the order of tokens in version B was backwards compared to the one in A, in order to guarantee that the order of items did not affect the judgments (cf. Höskuldur Þráinsson et al. 2013). Many tokens of clausal arguments were built upon the naturalistic data from RMH to ensure they could sound as natural as possible to speakers. The results on variants of the same sentence have been compared by calculating the weighted average of judgments on both variants of the same token. I considered acceptable those examples that got a minimum average of 4.44 (I follow here Pimentel's (2019) suggestion on intervals in a 7-point Likert scale).

Some items, however, have been more difficult to test in questionnaires (see section 4.4 in particular), which compelled me to investigate the matter further in interviews with various speakers, who would judge both variants orally.

The data I gathered are from all argument and complement positions, but due to space requirements for this article, I will focus specifically on subjects, direct and indirect objects, and nominal predicates, which are already enough to have a clear picture of the pattern of clausal nominalization in Icelandic.

4. Data

4.1 Subjects

Here are the results on nominative subjects from RMH, both when the clausal argument is in first position and when it surfaces after the finite verb. Note that I did not look for examples where the clausal argument is extraposed, as it is unclear whether *það* can be considered a nominalizer or an expletive when it surfaces in first position. For the sake of consistency in both data sets and also due to the difficulties in finding examples of extraposition in RMH, I only gathered tokens where *það* and the sentential subject occupy together the same position:

(12)

Position	Valid	-það	+það	%það
First	9485	7054	2431	**25.63**
After verb	71	5	66	**92.96**

The table in (12) shows both data on nominative clausal arguments in first position and after the finite verb, with and without nominalization (as shown in the columns +*það* and -*það*). The frequency of *það* expressed in percentage (see the last column) indicates that it is best to have a bare clausal argument in first position, while *það* is highly preferable after the finite verb. By contrast, if we take a look at the data from questionnaires, we get the following results:

(13)

Sentence (subject in first position)	-það	+það
(Það) að lifa í núvitund snýst um að njóta augnabliksins 'Living in mindfulness is about enjoying the moment'	6.41	6.08
(Það) að vörurnar séu ekki enn komnar skiptir engu máli 'It doesn't matter that the products have not arrived yet'	4.51	5.34
(Það) hvort hann komi eða ekki skiptir engu máli 'It doesn't matter whether he comes or not'	4.20	4.91

(14)

Sentence (subject after the finite verb)	-það	+það
Er (það) að læra á hljóðfæri eins og að læra að hjóla? 'Is learning to play an instrument like learning to ride a bike?'	3.64	5.67

The results in (13) show that both variants are acceptable in first position (apart from bare embedded questions where the average is below the minimum threshold), but not if the subject surfaces after the finite verb (see (14)). In other words, while the bare variant is more preferable in first position, the nominalized variant is mandatory after the finite verb.

4.2 Direct objects
As for clausal direct objects, I looked for examples in RMH where the clausal argument is selected by any of the verbs listed in (15). Some of these verbs have been selected due to the fact that Höskuldur Thráinsson (1979) tested them in order to analyze the role of *það*. These verbs are all the dative-assigning ones as well as *harma*. The verbs which are shown in bold have also been taken into account in the questionnaires:

(15)   Nom: *líka* ('like'), **leiðast** ('be tired of')
         Acc: **harma** ('regret'), **gagnrýna** ('criticize'), *opinbera* ('disclose')
         Dat: **spá** ('predict'), **fagna** ('rejoice'), *fresta* ('postpone')
         Gen: **krefjast** ('demand'), **sakna** ('miss'), *iðrast* ('regret, repent')

Here are the results on nominative and accusative clausal objects from RMH:

(16)

Verb	Valid	-það	+það	%það
líka + N	246	212	34	**13.82**
leiðast + N	577	553	24	**4.16**
gagnrýna +A	4241	3739	502	**11.84**
harma + A	2492	1897	595	**23.88**
opinbera + A	462	364	98	**21.21**

The data in (16) suggest that it is preferable to drop *það* before nominative and accusative clausal objects. By contrast, and somewhat surprisingly, dative and genitive clausal objects are preferably nominalized, also with *spá* ('predict') and *fresta* ('postpone'), which Höskuldur Thráinsson (1979) associated with an optional pronoun as in (11c):

(17)

Verb	Valid	-það	+það	%það
spá + D	12865	1045	11820	**91.88**
fagna + D	8926	157	8769	**98.24**
fresta + D	1602	154	1448	**90.39**
krefjast + G	32593	278	32315	**99.15**
iðrast + G	209	9	200	**95.69**
sakna + G	5228	124	5104	**97.63**

Let us take a look at the results from questionnaires. As for nominative and accusative clausal direct objects, the weighted averages are all acceptable, which suggests that, even if the bare variant is more preferable as shown in (16), it is still allowed to nominalize the clause:

(18)

Sentence (nominative and accusative clausal objects)	-það	+það
Ýmsir hafa gagnrýnt (það) að flóttafólki sé vísað úr landi 'Many have criticized the fact that refugees are deported'	6.00	5.89
Ég gagnrýni (það) hvernig þið öfluðuð ykkur þessara upplýsinga 'I criticize how you got this information'	5.04	4.66
Við hörmum (það) að stofnunin vilji hætta við verkefnið 'We are sorry that the institute wants to cancel the project'	5.74	5.58
Þeim leiddist (það) að ég væri alltaf að stoppa og taka myndir 'They were tired of the fact that I was always stopping and taking photos'	4.66	5.56
Mér líkar (það) að bókin sé myndskreytt 'I like the fact that the book is decorated with pictures'	5.10	5.00

As for dative and genitive clausal objects, we can see that the bare variant is degraded (as the weighted average is below the minimum threshold), while the nominalized variant is acceptable (differently from Höskuldur Thráinsson's (1979) observations):

(19)

Sentence (dative and genitive clausal objects)	-það	+það
Einar spáir (því) að liðið komist í úrslitaleikinn 'Einar predicts that the team will reach the final'	4.15	6.18
Ég þori ekki að spá (því) hvað verður um bækur í framtíðinni 'I don't dare to predict what will happen to books in the future'	3.96	4.69
Sindri fagnar (því) að bændur hafi fengið lækkun á raforkuverði 'Sindri is happy that the farmers got a reduction of the price of electricity'	2.73	5.26
Þóra var farin að sakna (þess) að búa á Íslandi 'Þóra started missing living in Iceland'	2.81	6.24
Ragnheiður krafðist (þess) að Árni bæði starfsmennina afsökunar 'Ragnheiður demanded that Árni apologizes to the staff'	3.76	6.30

In sum, it appears that, in instances of structural case, nominalization is optional but it is still more preferable to have a bare variant (consistently with nominative subjects in first position). By contrast, in instances of lexical case, the nominalized variant is generally mandatory.

4.3 Nominal predicates
Nominal predicates are a special category in the domain of clausal nominalization due to the fact that the nominalizer is assigned the same gender and number features of the subject, as shown in the following example:

(20)   Afleiðingin   er **sú**   að   við skuldum meira núna
consequence.the.F-SG is that.F-SG that we owe   more now
'The consequence is that we owe more money now'

Considering that the pronoun can, at this point, show up in six different forms in agreement with the gender and the number of the subject in nominative case, I chose and tested tokens where nominal predicates come after six specific subjects, as shown here below (note that these subjects generally surface with a clitic article in constructions like (20)):

(21)

Subject	Features	Valid	-það	+það	%það
galli ('drawback')	M. Sg.	325	97	228	**70.15**
afleiðing ('consequence')	F. Sg.	659	123	536	**81.34**
vandamál ('problem')	N. Sg.	721	678	43	**5.96**
kostir ('advantages')	M. Pl.	39	19	20	**51.28**
niðurstöður ('results')	F. Pl.	54	20	34	**62.96**
rök ('arguments')	N. Pl.	142	32	110	**77.46**

The data in (21) are crucial information, because whenever the pronoun is assigned *non-default* features (masculine, feminine and plural), the frequency of clausal nominalization increases, suggesting that nominalization is preferable. Only when *það* keeps its default properties (neuter singular), nominalization loses its preferability.

Let us take a look at the data from questionnaires. In this case, I considered enough to test only three out of the six subjects in (21), as the majority of these nouns apart from *vandamál* and *kostir* tend to exhibit similar results:

(22)

Sentence (nominal predicates)	-það	+það
Afleiðingin er (sú) að Bandaríkin eru í alvarlegri skuldakreppu 'The consequence is that the USA is in a serious debt crisis'	5.38	6.07
Kostirnir eru (þeir) að kerfið er ódýrt og einfalt í notkun 'The pros are that the system is cheap and easy-to-use'	6.65	6.38
Vandamálið er (það) að fólk vill ekki láta bólusetja sig 'The problem is that people do not want to be vaccinated'	5.88	5.44

Nominalization is also optional in this case, as it was for nominative subjects in first position and nominative and accusative objects.

4.4 Indirect objects
Clausal indirect objects are quite rare and more difficult to construct. As one could expect, it was impossible to find any naturalistic data in RMH. However, I still had the chance of testing the acceptability of clausal nominalization in questionnaires. I opted for two verbs in this case: *veita* ('provide') for dative case and *svipta* ('deprive') for accusative case. In order to construct the examples presented in (23), I asked some native speakers (who are also students in linguistics) for assistance.

They suggested that extraposing the clause while leaving the pronoun *in situ* could appear more acceptable to speakers. Therefore, only in this case, in order to avoid that neither variant was considered acceptable, I applied extraposition to the sentences in (23). The results are partially satisfying, as shown in the following table:

(23)	Sentence (clausal indirect objects)	-það	+það
	*Jón veitti (**því**) enga athygli að Sara væri að gráta* 'Jón paid no attention to the fact that Sara was crying'	1.86	6.33
	*Vopnahléið svipti (**það**) öllum hetjuljóma að hermenn fórnuðu lífi sínu* 'The armistice deprived the fact that the soldiers sacrificed their life of all its heroism'	2.10	2.57

The data on dative case indicate that *það* is mandatory, but the ones on *svipta* did not return significant results, which led me to investigate the matter further. Just to be sure, I tested the same sentence once more, but without extraposition. The results were problematic as well:

(24)	Sentence (clausal indirect objects)	-það	+það
	*Vopnahléið svipti (**það**) að hermenn fórnuðu lífi sínu öllum hetjuljóma* 'The armistice deprived the fact that the soldiers sacrificed their life of all its heroism'	1.09	1.42

Therefore, I resorted to some oral interviews with native speakers. In the interviews, all the speakers involved (*n* = 17) mostly considered the nominalized variant grammatical in the case of *svipta,* although, unexpectedly, they often reported that the non-extraposed nominalized variant was the best one. All this contrasts with the results here above and also indicates that participants to online questionnaires found difficult to parse the sentence with *svipta* while reading it (which is understandable since these constructions are rare). Be that as it may, the data from interviews tell us that *það* is mandatory independently from the case that is assigned. This is quite suspicious, especially because this is the first instance where we see that *það* is also mandatory with accusative case assignment. How can we explain this? Some interesting facts from passive constructions can help us to answer this question. There is an important contrast between accusative and dative case. If the indirect object is passivized, the pronoun remains obligatory with dative case, but not accusative (see Garofalo 2020; the judgments in (25) are based on the oral interviews as well):

(25) a. [*(**Því**) að Sara væri að gráta] var engin   athygli   veitt
      that.DAT that Sara was to cry   was no.NOM attention.NOM paid
      'No attention was paid to the fact that Sara was crying'

    b. [(**Það**)   að hermenn fórnuðu lífi sínu var svipt   öllum hetjuljóma
      that.NOM that soldiers sacrificed life their was deprived all heroism.DAT
      'The fact that the soldiers sacrificed their life was deprived of all heroism'

In other words, we are observing in accusative clausal indirect objects the same issue we observed with clausal subjects before and after the finite verb. Pronominal optionality in first position when the indirect object is passivized and moved to Spec,C suggests that obligatoriness *in situ* is not due to case assignment, but rather

to another property common to subjects and indirect objects. We are going to discuss this in the next section.

5. Discussion

On the basis of the empirical data, we can find some consistent patterns. First and foremost, the distribution of clausal nominalization is mainly built (but not exclusively) upon case assignment. Whenever lexical case is assigned, the pronoun is generally mandatory. The same cannot be said of structural case, where the pronoun is optional, but it is still best to opt for a bare clausal argument. This suggests that, while lexical case corresponds to formal features to check in syntax, structural case might not correspond to any formal feature at all. This resonates to much extent with the so-called *Dependent Case Theory* (see e.g. Wood 2011; also Marantz 2000), where accusative and nominative are actually the morphological representation respectively of dependent case (which is based on relations between DPs that have not been assigned lexical case) and unmarked case.

The data from nominal predicates corroborate this view since non-default features in gender and number are correlated to a higher frequency of *það* in naturalistic data. In other words, whenever a clausal argument needs to have a non-default feature checked, whether in relation to case, gender or number, nominalization tends to be the most preferable option, if not the only one available (as in lexical case).

In relation to subjects and clausal indirect objects, a possible explanation of pronominal obligatoriness lies in D-features. By standard assumptions in Minimalist Theory (see Chomsky 1993 and much subsequent work), the D-feature in T can only be checked by the DP that lands in subject position. If we consider the facts from Icelandic we observed so far on subjects and indirect objects and their interesting similarities, it is plausible that Spec,T and Spec,Appl (if we follow the *Applicative Head Approach* in Pylkkänen (2000)) are two positions where a DP is required as only this category can check D-features (which also entails that Appl should have a D-feature as well). If a clausal argument can escape those positions and reach the position before the finite verb, it can still surface as a bare CP if no case feature has to be checked (see (25)).

As for the optionality issue, the most plausible explanation at this point is that, since accusative and nominative case, which are generally associated to pronominal optionality, do not correspond to any feature to check, nominalization must be systematically allowed by virtue of the fact that the clausal argument is in a DP position. In other words, the underlying condition that restricts nominalization in non-DP positions is the same one that opens the way for nominalization to surface in DP positions. However, since there are no formal case features to check, it is still best to drop the pronoun, most probably because it is less costly from a structural perspective.

Lastly, from a cross-linguistic perspective, it is plausible that also clausal subjects in other languages like Modern Greek and Persian (see (2) and (3)) must be nominalized by virtue of D-feature checking and not by case assignment as proposed by Roussou (1991) or D-head visibility in subjects as in Knyazev (2016). The problem here is that these languages are not V2 languages like Icelandic and, therefore, it is not possible to see there a different distribution of nominalized and bare clausal arguments before and after the finite verb.

6. Conclusions

On the basis of the empirical data, the role of nominalizer *það* in Icelandic is checking formal case, gender and number features on behalf of the clausal argument. Only optionally, it is possible to nominalize the clause independently of feature checking, but this is plausibly due to the fact that merging in a DP position is an

underlying condition which allows clausal arguments to be systematically allowed to surface as DPs. Structural case, neuter and singular appear to not correspond to formal features to check, which leads the speaker to prefer a bare clausal argument since it is structurally less costly, while lexical case and non-default gender and number features (masculine, feminine and plural) correspond to formal features to check, which trigger clausal nominalization or make it more preferable (as in nominal predicates). Whenever a clausal argument surfaces or lands in a position which is linked to D-feature checking, nominalization is obligatory, independently of case, gender or number.

REFERENCES

Borsley, Robert D. and Jaklin Kornfilt. (1999). Mixed extended projections. *Syntax and Semantics*, *32*, 101–131. https://doi.org/10.1163/9781849500098_006

Chomsky, Noam. (1993). A minimalist program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser (Eds.), *The view from building 20* (pp. 1–52). MIT Press.

Garofalo, Mirko. (2020). The case of Icelandic clausal arguments. *Working Papers in Scandinavian Syntax*, *104*, 1–28.

Hartman, Jeremy. (2012). *Varieties of clausal complementation*. PhD thesis, MIT. http://hdl.handle.net/1721.1/77800

Ingason, Anton Karl. (2018). Icelandic case-marked CP. *Canadian Journal of Linguistics*, *63*(3), 444−453. https://doi.org/10.1017/cnj.2017.55

Kaltenböck, Gunther. (2003). On the syntactic and semantic status of anticipatory *it*. *English Language & Linguistics*, *7*(2), 235–255. https://doi.org/10.1017/S1360674303001096

Kastner, Itamar. (2015). Factivity mirrors interpretation: The selectional requirements of presuppositional verbs. *Lingua*, *164*, 156–188. https://doi.org/10.1016/j.lingua.2015.06.004

Kiparsky, Paul and Carol Kiparsky. (1971). Fact. In Danny Steinberg and Leon Jacobovits (Eds.), *Semantics* (pp. 345–369). Cambridge University Press.

Knyazev, Mikhail. (2016). *Licensing clausal complements: The case of Russian čto-clauses*. Doctoral dissertation, Utrecht University. https://www.lotpublications.nl/Documents/418_fulltext.pdf

Kornfilt, Jaklin and John Whitman. (2011). Introduction: Nominalizations in syntactic theory. *Lingua*, *121*(7), 1160–1163. https://doi.org/10.1016/j.lingua.2011.01.006

Labov, William. (1966). The linguistic variable as a structural unit. *Washington Linguistics Review*, *3*, 4–22.

Landau, Idan. (2007). EPP extensions. *Linguistic Inquiry*, *38*(3), 485–523. https://doi.org/10.1162/ling.2007.38.3.485

Marantz, Alec. (2000). Case and licensing. In Eric J. Reuland (Ed.), *Arguments and case: Explaining Burzio's Generalization* (pp. 11–29). John Benjamins.

Pimentel, Jonald L. (2019). Some biases in Likert scaling usage and its correction. *International Journal of Science: Basic and Applied Research* (*IJSBAR*), *45*(1), 183–191.

Pylkkänen, Liina. (2000). What applicative heads apply to. *University of Pennsylvania Working Papers in Linguistics*, *7*(1), 197–210.

Rosenbaum, Peter Steven. (1967). *The grammar of English predicate complement structures*. MIT Press. https://dspace.mit.edu/bitstream/handle/1721.1/16391/03190468-MIT.pdf

Roussou, Anna. (1991). Nominalized clauses in the syntax of Modern Greek. *UCL Working Papers in Linguistics*, *3*, 77–100.

Ruys, Eddy G. (2010). Expletive selection and CP arguments in Dutch. *The Journal of Comparative Germanic Linguistics*, *13*(2), 141–178. https://doi.org/10.1007/s10828-010-9037-5

Steingrímsson, Steinþór. (2019). Risamálheildin. *Orð og tunga*, *21*, 159–168. https://doi.org/10.33112/ordogtunga.21.9

Stowell, Timothy Angus. (1981). *Origins of phrase structure*. Doctoral dissertation, MIT. http://dspace.mit.edu/handle/1721.1/15626

Sudhoff, Stefan. (2016). Correlates of object clauses in German and Dutch. In Werner Frey, André Meinunger and Kerstin Schwabe (Eds.), *Inner-sentential propositional proforms: Syntactic properties and interpretative effects* (pp. 23–48). John Benjamins Publishing.

Thráinsson, Höskuldur. (1979). *On complementation in Icelandic*. Garland.

Thráinsson, Höskuldur, Ásgrímur Angantýsson and Einar Freyr Sigurðsson (Eds.). (2013). *Tilbrigði í íslenskri setningagerð: Markmið, aðferðir og efniviður*. Málvísindastofnun Háskóla Íslands.

Wood, Jim. (2011). Icelandic let-causatives and case. *Working Papers in Scandinavian Syntax*, *87*, 1–52.

Wood, Jim. (2012). Against the movement theory of control: Another argument from Icelandic. *Linguistic Inquiry*, *43*(1), 322–330. https://www.jstor.org/stable/pdf/23256616.pdf

ÚTDRÁTTUR

'The role and distribution of nominalized clauses: A view from Icelandic'

Keywords: Icelandic syntax, nominalizers, complement clauses, case marking, determiner phrases

Nafnyrtar fallsetningar eru algengt fyrirbæri þvert á tungumál þar sem ákveðniorð eða persónufornafn kemur fremst í fallsetningu. Í íslensku má sjá flókna dreifingu nafnyrtra og ónafnyrtra fallsetninga þar sem fornafnið getur verið skyldubundið, valfrjálst eða ótækt. Tilraunagögn segja okkur að nafnyrðing eigi sér stað ef fallsetning þarf að gáta fall-, kyn- eða töluþætti eða jafnvel ákveðniþátt. Ákveðniþættir og orðasafnsfall valda skyldubundinni nafnyrðingu (sem í óbeinum andlögum og í frumlagssæti á eftir persónubeygðu sögninni), og nafnyrðing verður æskileg ef fallsetningin þarf að gáta kyn- eða töluþætti (eins og í sagnfyllingum). Valfrelsi í notkun fornafnsins má sjá í frumlagssetningum í nefnifalli og andlags-setningum í nefnifalli og þolfalli og það stafar af því að fallsetningin kemur fram í ákveðniliðarstöðu þótt best sé að sleppa setningafornafninu í þessum tilfellum. Þetta bendir til þess að setningafornafn sé ótækt í stöðum þar sem ákveðniliður getur aldrei komið fram. Mynstrið sem kemur í ljós sýnir að nefnifall, þolfall, hvorugkyn og eintala eru ekki formlegir fallþættir (en það útskýrir af hverju fornafninu er sleppt), ólíkt orðasafnsfalli eða mörkuðum kyn- og töluþáttum.

Mirko Garofalo
Háskóla Íslands
mig@hi.is