Case variation: viruses and star wars

Sigurðsson, Halldor Armann

Published in:
Nordic Journal of Linguistics

DOI:
10.1017/S033258651300005X

2012

Citation for published version (APA):
Case variation: viruses and star wars

Halldór Ármann Sigurðsson

This article discusses morphological case variation, arguing that individual cases are not syntactic objects or features but PF interpretations of a range of different underlying syntactic relations. Nevertheless, it turns out that case variation can, to a large extent, be analyzed in terms of only two atomic ‘ingredients:’ event licensing of NPs and PF marking of the licensing relation (where marking is analyzed in terms of Chomskyan case stars). Ergative is a Voice/AG*-case, whereas accusative is a v*-case, licensed under c-command by Voice/AG (ergative and accusative marking thus being two sides of the same coin). Individual cases in case-expanding morphological/PF case systems, it is argued, behave like viruses, striving to expand beyond their original ‘reasonable’ domain.

Keywords: case, case expansion, case growth, case star augmentation, case star deletion, case variation, externalization, internal vs. external language

1. INTRODUCTION

Case is perhaps the most intriguing and also the most widely discussed of all grammatical categories. It is fair to say that case, more than any other category of language, highlights the fact that our understanding of language externalization and variation is limited. Many languages do not, in fact, have any overt case marking. The figures in (1) are drawn from Iggesen 2011b.

(1)  NUMBER OF CASES IN 266 LANGUAGES
     a. No morphological case marking       100 languages
     b. 2 cases                             23 languages
     c. 3 cases                             9 languages
     d. 4 cases                             9 languages
     e. 5 cases                             12 languages
     f. 6-9 cases                           50 languages
     g. 10 or more cases                    24 languages
     h. Exclusively borderline case marking 24 languages

Iggesen’s classification applies to ‘regular’ or ‘general’ noun-case distinctions. Thus, English counts as a two-case language (Nom and Gen), rather than a three-case language even though it has a three-case distinction in parts of its pronominal system. Actually, it would be more logical to count only marked cases; that is, to refer to languages as zero-case languages (no marked

1 In the words of Iggesen (2011b): ‘The feature value exclusively borderline case-marking refers to languages which have overt marking only for concrete (or "peripheral", or “semantic”) case relations, such as locatives or instrumentals’ [Iggesen’s emphasis].
case), one-case languages (one marked case, two distinct terms), and so on. I will return to this issue in relation to nominative case in Section 2.

Blake (2001:156) suggested that there is a morphological case hierarchy or scale, such that languages ‘pick’ their cases in the preference order in (2).²

(2) Nom > Acc/Erg > Gen > Dat > Loc > Abl/Inst > others (Part, Com, Purp, Perl, …)

The following tendencies can thus be discerned:

(3) COMMON TYPOLOGICAL CASE GENERALIZATIONS

a. If a language has cases, one of them is morphologically and/or functionally unmarked in relation to the other(s) and is thus likely to be called ‘nominative’ (or ‘direct’) by linguists.

b. The second case (the first marked one) is likely to distinguish objects (Acc) from subjects or to specifically distinguish agentive subjects (Erg) from other core arguments.³

c. A third case is likely to be an adnominal ‘possessor case’ (Gen).

d. A fourth case is likely to be an ‘additional core argument case’ (Dat, etc.).

e. A fifth case is likely to mark spatial relations (Loc, etc.).

f. Additional cases make more fine grained distinctions between arguments or NPs in general (Abl, Inst, Part, Com, etc.).

Presumably, these are third-factor effects in the sense of Chomsky 2005: natural options allowed for by underspecification of Universal Grammar (UG).⁴

Blake (2001:156ff) discusses a number of variously rich case systems, including the following ones (disregarding the vocative and some vestigial cases as well as complex locative case systems):

---

² I adopt the case abbreviations used in Blake 2001, including Nom(inative), Acc(usative), Erg(ative), Dat(ive), Gen(itive), Voc(ative), Abl(ative), Inst(rumental), Loc(ative), Ill(ative), Ade(ssive), All(ative), Part(itive), Com(itative), Purp(osive), Perl(ative), Obl(ique).

³ Blake (2001) refers to absolutive case as nominative. I adopt that terminology.

⁴ Chomsky (2005:6): ‘3. Principles not specific to the faculty of language’. The idea I am adopting is that cases have the general function of making overt distinctions between arguments and/or NPs (Sigurðsson 2006b), and that such a distinctive or classifying capacity is not confined to UG, the specifically linguistic department of the human mind/brain—a reasonable assumption. As for abstract Case, I refer to Sigurðsson 2012, where it is extensively argued that it is entirely unrelated to morphological case and must instead be analyzed as phi-licensing. On underspecification, see Roberts & Roussou 2003, Biberauer et al. 2009, Holmberg 2010.
EXAMPLE LANGUAGES: CASES:

(4)  a. Tamil  Nom  Acc  Gen  Dat  Loc  Abl  Inst  Com
   b. Classical Armenian  Nom  Acc  Gen  Dat  Loc  Abl  Inst
   c. Turkish  Nom  Acc  Gen  Dat  Loc  Abl
   d. Russian, Czech  Nom  Acc  Gen  Dat  Loc  Inst
   e. Latin  Nom  Acc  Gen  Dat
   f. Ancient Greek, German  Nom  Acc  Gen  Dat
   g. Modern Greek, Arabic  Nom  Acc  Gen

However, this gives an idealized picture of order in a complex ‘case world.’ First, many languages do not ‘fit’ into the Blake hierarchy. Faroese, for instance, is losing Gen but keeping Nom, Acc, Dat (Thráinsson et al. 2004:248ff.), which is the opposite of the development in Greek. Gaelic has Dat and Gen but no Acc, Hungarian has no Gen but Dat, Inst, etc., and Finnish has a general Gen but no inflectional Acc except for pronouns, also having Part, ‘instead of’ the higher ranked Dat, Loc, Abl and Inst. For more examples of ‘deviant’ systems, see Blake 2001, Malchukov and Spencer 2009a, 2009b. Obviously, also, notions like ‘dative case’ are not usually ‘constant’ across languages with variably rich case inventories: ‘the dative’ is something different in Faroese, Turkish, and Tamil.

Second, languages with identical case inventories can show various splits and asymmetries across grammatical categories, such as Person, Tense and Aspect, or across lexical categories, such as nouns and pronouns (see Iggesen 2011a). A familiar split or asymmetry of this sort is seen in a number of European languages, including English, with a general two-case system (Nom, Gen) for nouns (and NPs) and a three-case system for central parts of their pronominal systems (Nom, Acc, Gen). Even ‘simple’ systems of this sort show considerable complexities and cross-linguistic variation (see Quinn 2005 on English). And there are many more types of split case systems, including split ergative systems of various sorts.

Third, languages with the same number of cases can distribute their cases quite differently across semantic and grammatical roles – call this CASE MISMATCHES. Case mismatches are most easily demonstrated for two-case languages, where various patterns can be discerned (see, e.g., Blake 2001, Arkadiev 2009, Malchukov and Spencer 2009b, Iggesen 2011a). Thus the oblique (or the ‘marked’/‘indirect’) case in nominative-oblique or morphologically unmarked-marked systems may mark various relations, as sketched in (5) for a few languages (disregarding ergative systems, which add another dimension to this case mismatching picture).5

5 The abbreviations in (5): A = subjects of transitive predicates; S = subjects of intransitive predicates (unergative, unaccusative); DO = direct objects (usually denoted as P in the typological literature); IO = indirect objects (commonly either covered by P or ignored in the typological literature); Poss = possessor.

A caveat is in order here: Typological studies do not usually aim at exhaustive descriptions and analyses of case systems, and many such studies also use varying and ill-defined notions, so the presentation in (5) has many lacunas (e.g., commonly lacking information on Poss, IO and numerous other functions, including nominal
Developing some understanding of all this variation is a major challenge. As suggested by the presentation in (3) above, it presumably follows from general third-factor effects and principles. That is, case is not provided by or part of Universal Grammar (UG), a conclusion supported by numerous facts, for example extensive case mismatches and also the simple fact that having no case marking at all is highly common, not only in ‘usual’ languages, but even more so in creoles and sign languages (see Sandler and Lillo-Martin 2006:23; see also Meir 2003).\(^6\) It is thus unsurprising that UG based approaches to case variation have never been successfully developed.

Saying that language variation follows from underspecification of UG and third-factor effects makes a very general statement. It does not release linguistics from developing some coherent understanding of variation. As in any other scientific endeavour, developing such understanding requires not only broad overviews but also detailed analyses of specific data. In the following I will look into some of the details of Icelandic case assignment, comparing it with case assignment in some other languages, above all German, Icelandic and German having an identical case inventory: Nom, Acc, Gen, Dat. Even though the Icelandic/German four-case system is an offspring of the Proto-Indo-European eight-case system (Nom, Acc, Gen, Dat, Loc, Inst, Abl, Voc), I will be focusing on case growth, setting case decline aside.\(^7\)

Section 2 discusses the nature of nominative case and its correlation with verb agreement, Section 3 discusses the non-nominative argument cases and argument case growth (case star augmentation), Section 4 describes non-core case of various sorts, Section 5 describes and analyzes case expansion from one subsystem of case marking to another, and Section 6 concludes the paper.

The following journey is part of a much longer voyage (Sigurðsson 2004, 2006a, 2006b, 2011b, 2011c, 2012, etc.), aiming at showing that internal syntax is even smaller or narrower and that the externalization component of language is thus much broader and richer (‘more syntactic’) than commonly assumed in generative approaches. As it turns out, however, case

---

\(^{6}\) Even though creoles have more morphology than has often been assumed, see de Graf 2001.

\(^{7}\) See Barðdal and Kulikov 2009 on some of the PF observable processes typically involved in case decline.
variation can, to a large extent, be analyzed in terms of only two atomic ‘ingredients’: event licensing of NPs and variable PF marking of the licensing relation.

2. NOMINATIVE

All languages ‘begin life’ without any case marking, a conjecture that is not only reasonable on general conceptual grounds but also strongly supported by the fact that all known sign languages are caseless or next to caseless, sign languages generally being very young (see, e.g., Sandler & Lillo-Martin 2006). If so, original or non-borrowed case systems plausibly always come into being by a simple split in two, with a marked or an oblique case and an unmarked one, commonly referred to as ‘nominative’, some such systems further developing or growing by adding further cases over time (and then potentially declining, then even expanding again in new directions, see Kulikov 2009). As mentioned above, it would actually be more logical to count only marked cases, languages, then, generally being ‘born’ as zero-case languages, then going through a case stage with one marked case, with either a marked core case (Acc or Erg) or a marked borderline case (Loc, Inst, etc.), then through a stage with two marked cases (e.g., Acc and Gen, Acc and Loc), and so on. On this view, nominative is always a non-case.\(^8\)

However, unlike German and most other modern European case languages, but like Latin, Icelandic has a split marked nominative system, in the morphological sense (rather different, it seems, from the African-systems described in König 2009). That is, it has a morphologically marked nominative form in certain declensions but an unmarked nominative form in other declensions. Icelandic has many different noun declensions—a comprehensive description minimally requires around 70 distinct paradigms (even if much morphophonological variation is disregarded).\(^9\) Simplifying, we can say that the so-called strong masculine declensions usually have marked nominative forms in the singular (mostly in -ur), while most other declensions have unmarked or not clearly marked singular nominatives. A few of the most regular patterns are exemplified in (6); the strong-weak dichotomy is a formal one (pure vowel endings in all weak singular forms), with no semantic or syntactic import or correlates.

---

\(^8\) Many others have argued that Nom is ‘no case’ in some sense, including many generative researchers (see for example Markman 2009:402ff, Ashbury 2010:18).

(6) **CENTRAL ICELANDIC NOUN DECLENSIONS**

<table>
<thead>
<tr>
<th></th>
<th><strong>Masculine nouns</strong></th>
<th><strong>Feminine nouns</strong></th>
<th><strong>Neuter nouns</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Strong 1</strong></td>
<td><strong>Strong 2</strong></td>
<td><strong>Weak</strong></td>
</tr>
<tr>
<td>SINGULAR</td>
<td>‘horse’</td>
<td>‘way’</td>
<td>‘pen’</td>
</tr>
<tr>
<td>Nom</td>
<td>hest-ur</td>
<td>veg-ur</td>
<td>penn-i</td>
</tr>
<tr>
<td>Acc</td>
<td>hest</td>
<td>veg</td>
<td>penn-a</td>
</tr>
<tr>
<td>Dat</td>
<td>hest-i</td>
<td>veg-i</td>
<td>penn-a</td>
</tr>
<tr>
<td>Gen</td>
<td>hest-s</td>
<td>veg-ar</td>
<td>penn-a</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th><strong>Masculine nouns</strong></th>
<th><strong>Feminine nouns</strong></th>
<th><strong>Neuter nouns</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Strong 1</strong></td>
<td><strong>Strong 2</strong></td>
<td><strong>Weak</strong></td>
</tr>
<tr>
<td>PLURAL</td>
<td>hest-ar</td>
<td>veg-ir</td>
<td>penn-ar</td>
</tr>
<tr>
<td>Nom</td>
<td>hest-a</td>
<td>veg-i</td>
<td>penn-a</td>
</tr>
<tr>
<td>Acc</td>
<td>hest-um</td>
<td>veg-um</td>
<td>penn-um</td>
</tr>
<tr>
<td>Dat</td>
<td>hest-a</td>
<td>veg-a</td>
<td>penn-a</td>
</tr>
</tbody>
</table>

This is the traditional analysis of the case endings. Alternatively, one could analyze the Nom.Sg forms of all weak nouns as the basic stem form, other forms being derived by vowel truncation, deleting the first of two unstressed vowels, a process that is in any case operative elsewhere in the language (/penn-ur, /vika-r/ > penn-a, vik-a, etc.). Be that as it may, it is clear that Nom is marked to a variable degree, depending on both number and declension class.\(^\text{10}\) Despite this morphological split, all nominatives show parallel behavior with respect to syntactic distribution and agreement processes. That is, the marking of marked nominative forms is not syntactically triggered, instead being a purely morphological (PF) phenomenon.

Case researchers commonly make no clear distinctions between morphological case terminology, functional case terminology and alignment case terminology. Morphologically, there is a distinction to be drawn between marked and unmarked case forms (Dixon 1979). Functionally, there is a distinction between general/common and less general/common, often referred to as marked vs. unmarked (or default) as well (see Schütze 2001). In contrast, traditional case terms, Nom, Acc, Erg, etc., are commonly used as alignment terms, indicating which case usually aligns with which thematic role (Agent, Patient, etc.) or grammatical function (subject, direct object, etc.). This is simply illustrated in (7).

(7) **Morphological distinctions:** unmarked, marked (e.g., by some case suffix)
**Functional distinctions:** general (unmarked/default), less general (marked)
**Alignment distinctions:** subjects (agentive, nonagentive), direct objects, etc.

\(^{10}\) Alternatively, the -ur and the -i markers in the nominative singular masculine might perhaps be analysed as markers of gender, number or even of inflectional class rather than of nominative case (which is of course compatible with a no-case approach to the nominative, see Neeleman & Weerman 1999, Weerman & Evers-Vermeul 2002).
Commonly, the term ‘nominative’ is used to designate 1) the morphologically unmarked case form, 2) the functionally general or default case, and 3) the case of subjects (in non-ergative systems), but there are obviously many mismatches between morphological, functional and alignment mappings.

Mainstream generative approaches to case (Chomsky 1980, etc.) take it that Nom is a syntactically active case, ‘assigned a value under agreement’ with the Tense head (T) in finite clauses (Chomsky 2001:6). The assumption that Nom is the ‘responsibility’ of the finite, ϕ-complete, T has commonly been taken to account for finite verb agreement in regular Nom-Acc languages, such as English and German. However, as we have seen, it is more coherent to assume that Nom in Nom-Acc systems is simply the unmarked opposite of a marked Acc or Obl case. If so, verb agreement (in case + agreement languages) enters non-null-subject (PF) grammars as an extra alignment strategy, for example along the path sketched in (8)\(^\text{11}\).

(8) **Development of Some Nom-Acc Systems (in Non-Null-Subject Languages):**

<table>
<thead>
<tr>
<th></th>
<th>0 case, no agreement:</th>
<th>1 marked case, no agreement:</th>
<th>1 marked case + agreement:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>NP1 T NP2</td>
<td>NP1 T NP2Acc</td>
<td>NP1 T(_\text{AGR/NP1}) NP2Acc</td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As suggested by this presentation, agreement is commonly redundant in non-null-subject systems. When such a system develops marked nominatives, a number of patterns may emerge, including patterns with two marked cases. Consider the Icelandic examples in (9) (\(\text{AGR} = \text{overt agreement}\)).

(9) a. **Ólafur** sér **Guðrúnú.** 2 marked cases + AGR

\(\text{Ólafur} . \text{NOM} \text{ sees.3SG} \text{ Guðrúnú} . \text{ACC}\)

‘Ólaf sees Guðrún.’

b. **Ólafur** sér **Ólaf.** 1 marked case (Nom) + AGR

\(\text{Ólafur} . \text{NOM} \text{ sees.3SG} \text{ Ólaf}\)

‘Ólaf sees Ólaf.’

c. **Guðrún** sér **Ólaf.** 0 marked case + AGR

\(\text{Guðrún} \text{ sees.3SG} \text{ Ólaf}\)

‘Guðrún sees Ólaf.’

d. **Guðrún** sér **Guðrúnú.** 1 marked case (Acc) + AGR

\(\text{Guðrún} \text{ sees.3SG} \text{ Guðrúnú} . \text{ACC}\)

\(\text{AGR} = \text{overt agreement}\)

\(\text{null subject grammars of roughly the Italian/Latin type have presumably always been more common than (referential) non-null-subject grammars of roughly the German/French/English type, but that does not affect the point I am making, so I put this aside.}\)

---

\(\text{null subject grammars of roughly the Italian/Latin type have presumably always been more common than (referential) non-null-subject grammars of roughly the German/French/English type, but that does not affect the point I am making, so I put this aside.}\)
‘Gudrun sees Gudrun.’

Here, the verb form does not yield any clear alignment information (as both the subject and the object are in the third person singular), but it does in many other cases, for example when the subject and the object are in different numbers. In many such cases, nominative alignment (with A/S) is doubly marked, with a marked Nom + verb agreement. If the object carries a marked Acc, this yields a threefold alignment marking (as in (9a)), where single marking (only marked Acc, only marked Nom, or only verb agreement) would be sufficient to disambiguate the NP-predicate relations.

Nom is a non-case.\(^\text{12}\) That is, whenever grammar gives no specific case instructions to the morphological case component, the NP in question will show up in Nom, regardless of the overt shape of the nominative elements expressed. Finite verb agreement, in turn, arises in morphology whenever the verb successfully probes a non-cased NP (be it morphologically marked or not), given, of course, that the language in question has finite verb agreement. Thus, we have to sharply distinguish between grammar instructions leading to case marking and case realization or externalization in PF. As we will see, actually, case morphology often behaves like a virus in morphological systems like the Icelandic one, without any clear semantic or syntactic import or correlates. Marked nominatives are just one example of such ‘PF viruses.’\(^\text{13}\)

\(^{12}\) Schütze (2001) argues that Acc is a default non-case in the English pronominal case system (see also the discussion in Weerman & Evers-Vermeulen 2002, Quinn 2005, Parrot 2007, 2009). I develop an analysis where Acc in English (and Danish) is actively assigned in (PF) grammar (see also Sigurðsson 2006b). If Nom is in some sense the marked or non-default case in English, then verb agreement is with the non-default case, which would be highly exceptional for accusative systems (see Baker 2008). However, that may not be the right perspective. That is, the notions ‘nominative’ and ‘accusative’ are perhaps misnomers for English and Danish. Regardless of perspective, any analysis must account for the fact that Burzio’s Generalization applies to English in much the same manner as to, e.g., German, and it must also accommodate or allow for not only the Icelandic/German and the English/Danish types of systems but also the various mixed types described in Sigurðsson 2006b (including historical stages of English). I cannot discuss this any further here, only pointing out that the overt distribution of case forms and the grammar mechanism that derives or licenses these forms must be kept apart. The analysis developed here is fully compatible with the standard view (Chomsky 1980, etc.) that the English v/p ‘accusative’ is an assigned or an actively ‘decided’ case (albeit in PF, not in narrow syntax).

\(^{13}\) PF case spreading of a different type is seen in various case agreement phenomena. One widely discussed phenomenon of this sort is case marking of Icelandic PRO in examples like she asked Olaf.ACC to go just alone.ACC to the party (see Sigurðsson 2008 and the references there). Case marking of PRO has no semantic effects, arguably involving PF percolation of an abstract morphological case feature under control, as sketched in (i).

\[(i)\text{ [CP ... Olaf.ACC ... [CP ... PRO ... alone.ACC ... ]]}\]
\[\uparrow \text{Control} \uparrow \text{Agree(ment)}\]

PF percolation (also seen in, e.g., tense agreement in Sequence of Tenses contexts) behaves like a reflex of sorts, utilizing syntactic correlations (Control/Agree) as a gateway to transmit a morphological feature in a top-down
3. THE NON-NOMINATIVE ARGUMENT CASES

Blake’s hierarchy in (2) and the related generalizations in (3) would seem to suggest that the first cases to get marked in the historical development of case systems are the core cases: First Acc or Erg, then Gen, then Dat, and so on. However, as the hierarchy is mostly based on observations on established ‘dinosauric’ case languages (with perhaps a 100,000 years old history), it is unclear whether it reflects usual historical developments of case systems (for some observations and speculations on the development of case systems, see Blake 2001, Kulikov 2009). It is quite possible that a language develops some borderline or peripheral cases before developing Acc or Erg; 24 of the 266 languages in Iggesen 2011b have exclusively borderline case marking, including Plains Cree ‘whose only case-inflecting device is the locative suffix –ehk’ (Iggesen 2011b). For expository purposes, however, I will assume that Blake’s hierarchy and the generalizations in (3) reflect at least a common case growth pattern over time.

In ‘Derivation by phase’, Chomsky (2001) suggests that Acc in regular accusative systems is the responsibility of $\phi$-complete v, designated as v*. In contrast, he analyzes defective v as not assigning or licensing any case value, thereby rendering the underlying object in defective vP-types accessible to Nom (as in canonical passives, unaccusatives, and anticausatives). As illustrated in (10), this yields Burzio’s Generalization (for English), that is to say, the Acc-to-Nom conversion typical of defective predicate types.

<table>
<thead>
<tr>
<th>Case</th>
<th>Description</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>We sank them.</td>
<td>transitive v*-V: NOM – ACCi</td>
</tr>
<tr>
<td>b.</td>
<td>They were sunk.</td>
<td>passive v-V: NOMi</td>
</tr>
<tr>
<td>c.</td>
<td>They sank.</td>
<td>unaccusative v-V: NOMi</td>
</tr>
</tbody>
</table>

CASE STAR AUGMENTATION, then, yielding v* in addition to plain v, is involved when a language develops from a no case language to an accusative language, distinguishing Acc NPs from non-cased (Nom) NPs. This development is a PF change, not applying in caseless languages. In a language where it has taken place, active Acc assignment in the v-system is based on the structure in (11), whereas the corresponding unaccusative and passive structures yield Nom, as sketched in (12); the arrows connecting v*-V and NP1 indicate a matching relation that gets

---

14 In contrast to these authors, however, I will not be discussing the various PF-origins of case markers.

15 There is substantial evidence that the case star notation and $\phi$-completeness must be kept apart, but I will not discuss this here (instead referring to the extensive discussion in Sigurðsson 2012).
interpreted in terms of non-nominative case in PF morphology, here Acc (as indicated on the right hand side).16

\[(11) \quad \ldots [\text{TP} \ldots \text{Voice}/\text{AG} \ldots [\text{NP}_2 \ldots \text{v}^*-\text{V} \text{ NP}_1 \ldots]]] \quad \rightarrow \quad \text{NP}_1/\text{ACC} \text{ in PF} \]
\[
\uparrow \quad \uparrow
\]
\[
\rightarrow \quad \text{NP}_2/\text{NOM} \text{ in PF}
\]

\[(12) \quad \begin{align*}
a. \quad & \ldots [\text{TP} \ldots \text{Voice}/\text{PASS} \ldots [\text{v} \text{-V} \text{ NP}_1 \ldots]]] \quad \rightarrow \quad \text{NP}_1/\text{NOM} \text{ in PF} \\
b. \quad & \ldots [\text{TP} \ldots \text{Voice}/\text{EXPL} \ldots [\text{v} \text{-V} \text{ NP}_1 \ldots]]] \quad \rightarrow \quad \text{NP}_1/\text{NOM} \text{ in PF}
\end{align*}
\]

I assume that all predicates are embedded under some Voice head, passives being embedded under Voice/\text{PASS}, while unaccusatives as well as anticausatives are embedded under expletive Voice, Voice/\text{EXPL} (even though unaccusatives and anticausatives have different vP-internal structures). Agentive or active Voice, Voice/\text{AG}, in turn, introduces the subject in active transitive structures, such as (11) (Kratzer 1996 and much related work).

In ergative systems, PF case star augmentation applies to Voice/\text{AG}, yielding a case licensing Voice*/AG, as sketched in (13).

\[(13) \quad \ldots [\text{TP} \ldots \text{Voice}*/\text{AG} \ldots [\text{NP}_2 \ldots \text{v}-\text{V} \text{ NP}_1 \ldots]]] \quad \rightarrow \quad \text{NP}_2/\text{ERG} \text{ in PF} \]
\[
\uparrow \quad \uparrow
\]
\[
\rightarrow \quad \text{NP}_1/\text{NOM} \text{ in PF}
\]

Ergativity and accusativity are thus two sides of the same coin: while ergative case is directly licensed by Voice*/AG, accusative is licensed by a v* under c-command of Voice/AG.17

The case decision for a non-nominative NP can be described in three steps. First, syntax transfers structures such as (11) and (13) to deep PF, without any case stars (hence all case marking is out of semantic/syntactic reach, even when it has semantic correlates, see Section 5). Thus, the syntactic message or information to PF is purely structural (and ‘lexical’), without any inflectional instructions: [Voice (↔ NP) ↔ v/V ↔ NP] type structures, where the arrows denote matching relations (and where V and NP contain abstract ‘lexical’ roots). This first (plausibly universal) step is not shown in (11)–(13). Second, language-specific PF assigns a particular type of case star to one or more licensing heads (v, Voice, etc.), depending on the syntactic structure (and on lexically related case star selection, see Section 5). This yields the input structures in (11) and (13) (i.e., the structures to the left of the horizontal arrows). Third, PF interprets the particular case star as one of the morphological cases of the language in question, again depending on the transferred syntactic information (and on deep PF properties, most importantly

---

16 There is a (very relevant) matching correlation with Voice as well, not indicated in these sketchy structures.

17 Voice and v are members of the same extended V-projection or head-chain (Voice-v-V); the highest member or the ‘head’ of the chain is the case licenser in ergative systems, while it is the lowest member or the ‘foot’ of the chain that is the case licenser in accusative systems.
lexical categories). The right hand horizontal arrows in (11)–(13) indicate this third and most shallow step in this case-marking description. Notice that I say ‘description.’ It seems reasonable to assume that steps two and three are instantaneously intertwined or just a single step in the actual morphological derivation.

On this approach, case is absent from Universal Grammar (and hence the previously mentioned fact that 100 out of 266 languages in Iggesen 2011b lack case marking is unsurprising). Notice that the approach is fully compatible with the Inclusiveness Condition (obviously, the Inclusiveness Condition does not apply to PF, see Chomsky 1995:228, 381, n. 10). These claims and assumptions will become more transparent and understandable as we proceed.

In addition to either Acc or Erg, languages with two marked cases (three-case languages) most commonly have adnominal Gen or benefactive/recipient Dat. These cases do not usually involve Voice or v, the core heads of the verbal system. Rather, the canonical Gen is adnominal, and indirect object Dat is commonly taken to be introduced by an applicative head, Appl (Marantz 1993, McGinnis 2001, Cuervo 2003, Pylkkänen 2008, Schäfer 2008, Wood 2012b). The Appl enriched structure of an ordinary ditransitive clause like Icelandic Hún gaf mér bókina ‘she.NOM gave me.DAT the book.ACC’ is illustrated in (14).

\[
(14) \quad \text{.. [TP .. Voice/Ag .. [ NP_3 .. Appl* [ NP_2 v*-V NP_1 ..] → NP_1/ACC in PF}\n\]
\[
\quad \uparrow \uparrow \uparrow \uparrow \quad \uparrow \uparrow \uparrow \quad \text{NP_2/DAT in PF}\n\]
\[
\quad \text{NP_3/NOM in PF}\n\]

Arguments, then, are EVENT LICENSED by specialized heads (Pylkkänen 2008, Sigurðsson 2012): agentive or active subjects by Voice/Ag, indirect objects by Appl, and direct objects by v-V. The TP structures of regular transitive and ditransitive clauses are sketched in (15) and (16).
DO, IO and SU are just convenient expository labels (instead of NP₁, NP₂, NP₃ or θ₁, θ₂, θ₃). The derivation of (16) starts out by introducing DO (NP₁/θ₁) and event licensing it by v-V, then introducing IO and event licensing it by Appl, then introducing SU and event licensing it by Voice/AG.¹⁸

In caseless languages none of the event licensing heads are augmented by a case star, whereas languages like German, Icelandic and Turkish have both case licensing v* heads and Appl* heads (licensing Acc vs. Dat). Ergative systems, in turn, have a case star augmented Voice*/AG, as discussed above. The generalization that emerges from these observations is stated in (17).

(17)  a. NPs are event licensed by heads: Voice, Appl, v, n, a, p, …¹⁹
    b. Event licensers of NPs may be case star augmented
    c. The augmented case star may be variably marked: *, *+, *++
    d. Nom = a non-case (regardless of its overt marking)

¹⁸ By the Inclusiveness Condition (Chomsky 1995:228) there are no ‘bar-levels in the sense of X-bar theory’. Thus, arguments that have been traditionally assumed to occupy specifier positions in in the sense of X-bar theory are ‘adjoined’ to maximal categories and subsequently event licensed by a specialized head (see Sigurðsson 2012 for a more detailed account). I do not distinguish between high and low applicatives (the putative distinction is immaterial for my purposes).

¹⁹ Functional categories are not stored in Universal Grammar, instead being materialized in individual I-languages, perhaps universally so (which is not a contradiction, see Sigurðsson 2011c). They need not have any exponents in E-language (but each of the C-, T-, v-subsystems typically, perhaps universally, has at least one exponent).
The plus notation on case stars may be interpreted as ‘more marked’, $*^+$ thus being more marked than $*$, $*^{++}$, in turn, being still more marked than $*^{+}$. As we will see, a morphological case may be variably marked depending on which structure it is licensed in; that is to say, the markedness of some particular case, say Dat, depends on whether it is an Appl-case (unmarked as compared to Acc), a v-case (more marked than Acc), and so on. As event licenseurs and case stars are the basic ‘ingredients’ in case systems, I refer to (17) as the Case Ingredients Generalization.\footnote{More than three degrees of markedness is unusual, perhaps non-existent, in morphological systems. However, as markedness is a general 3rd factor phenomenon, it is not clear whether there are any principled linguistic limits to markedness. The notion of markedness is widely adopted in linguistics but it is notoriously moot (in all approaches to language that I am aware of) and I do not have a sophisticated theory of markedness, beyond the reasonable but largely common sense approach in Bresnan 2001. For my purposes a morphological value is the more marked the more peripheral it is within a given domain (where ‘domain’ loosely refers to some particular structure or configuration, such as [v-V NP]). In general this means that a value or a form is the more marked the more specified information is required to trigger it.}

It does not come as a surprise, of course, that argument-introducing heads may also license some overt case marking. In the often cited words of Blake (2001:1), case is ‘a system of marking dependent nouns for the type of relationship they bear to their head.’ The variable markedness generalization in (17c) is less expected, but it is corroborated by many facts, one being that indirect objects are variably marked both within and across languages. Thus, while Icelandic has more than 280 ditransitive verbs that take a dative indirect object (mostly Nom-Dat-Acc), it also has around 60 ditransitive verbs that take an accusative indirect object (mostly Nom-Acc-Dat verbs (Jónsson 2000, Thráinsson 2007:173). Two examples are given in (18); as indicated, the ‘first’ object is optional, as indirect objects commonly are.

\begin{align*}
\text{(18) a. } & \text{Þeir leyndu } \text{(mig)} \text{ sannleikanum.} \\
& \text{they.NOM concealed me.ACC truth.the.DAT} \\
& \text{‘They concealed the truth (from me).’} \\
\text{b. } & \text{Þeir rændu } \text{(hana)} \text{ töskunni.} \\
& \text{they.NOM robbed her.ACC bag.the.DAT} \\
& \text{‘They robbed the bag (from her).’}
\end{align*}

As suggested by these examples accusative indirect objects commonly resemble Latin ablatives of separation. Presumably, they came into being as Acc ‘invaded’ the Appl-system from the v-system (perhaps in tandem with or as a consequence of the Germanic ‘Abl-death’). As these ‘ablative accusatives’ are clearly more marked in the Icelandic Appl-system than are regular dative indirect objects, we need to distinguish between Appl*, licensing dative indirect objects and Appl*+, licensing the ‘ablative accusatives.’ Languages like English and Swedish (see Holmberg & Platzack 1995:197), with only accusative indirect objects, have only one Appl head,\footnote{For somewhat reminiscent (but also quite different) approaches, see Caha 2009 and Pesetsky 2010.}
Appl*, whereas Gaelic has lost its Nom-Acc distinction, thus having only nominative indirect objects, licensed by plain Appl. Thus, ‘marked’ and ‘more marked’ are relative notions, depending not only on event licensers but also on languages.22

Variable case marking also applies to direct objects in many languages (see, e.g., Malchukov & de Swart 2009); thus, Russian has Acc, Gen, Dat, Inst direct objects (see Richardson 2007, Bailyn 2012), German has Acc, Gen, Dat direct objects, Icelandic has Acc, Gen, Dat, Nom direct objects, and so on.23 Some of this variation is language-internally regular to some (variable) extent, but it is typically unpredictable cross-linguistically. Thus, ‘Maling (1996) [an unpublished work] contains a list of more than 750 [Icelandic] verbs which in at least one sense occur with a dative object ... [whereas the] corresponding number of verbs for German is approximately 140, and, for Russian fewer than 60’ (Maling 2002:31). There is nothing (synchronically) in the case systems of these languages that would lead one to expect these dramatic differences – recall that Icelandic and German have identical 4 case inventories: Nom, Acc, Gen, Dat.

The Icelandic verbs in (19) take a dative direct object:

(19) Some Dat-taking Icelandic verbs

ausa ‘scoop,’ beina ‘direct,’ bjarga ‘rescue,’ hjóða ‘invite,’ bylta ‘overturn,’ dreifa ‘spread,’ fagna ‘welcome,’ fleygja ‘throw away,’ fleyta ‘float,’ fljúga ‘fly’ (e.g. an aeroplain), giftast ‘marry,’ gleyma ‘forget,’ heilsa ‘greet,’ hella ‘pour,’ henda ‘throw, throw away,’ hjúkra ‘nurse,’ hlífa ‘protect, spare,’ launa ‘pay, reward,’ misþyrma ‘torture,’ ráða ‘decide,’ riða ‘ride’ (e.g. a horse), róa ‘row,’ sigla ‘sail,’ snúa ‘turn,’ stjórna ‘control, govern, rule,’ sökkva ‘sink,’ tortíma ‘exterminate,’ yta ‘push, shift,’ þjóna ‘serve,’ póknast ‘please’

Corresponding verbs in German all take an accusative object (auslößeln ‘scoop,’ retten ‘rescue,’ einladen ‘invite,’ etc.).

The Icelandic verbs in (20) all take a genitive direct object (some of them can alternatively take a PP complement).

(20) Some Gen-taking Icelandic verbs

---

22 For expository ease, one could use notations like $v^{ACC}$ and Appl$^{acc}$, p$^{acc}$ etc., but then one would need to account separately for markedness (e.g., the fact that Acc is less marked in the $v$-system than in other subsystems). Nothing crucial hinges on which notation one opts for (they are in any case PF notations), but, as argued in previous work (including Sigurðsson 2012), the star notation has virtues not shared by other notations.

23 Some direct object datives might be analysed as Appl datives rather than as $v$-datives (see the discussion in McFadden 2004), but that would not alter the basic fact that many direct object datives are $v$-datives (Wood 2012b:131ff.). In addition, of course, there are $v$-genitives, $v$-instrumentals, etc.
Corresponding verbs in German (bitten ‘ask for,’ erwähnen ‘mention,’ etc.) take either an accusative object or a PP complement.

The examples in (21) and (22) illustrate this Icelandic-German dichotomy with respect to direct object case marking.

(21) a. Hún kastaði boltanum/*boltann.  
   *she threw* ball.the.DAT/*ACC*  
   *Icelandic Dat/*Acc*  
   b. Sie hat den Ball/*dem Ball geworfen.  
   *she has* the ball.ACC/*DAT* thrown  
   ‘She threw the ball.’

(22) a. Hún saknaði hans/*hann.  
   *she missed* him.GEN/*ACC*  
   *Icelandic Gen/*Acc*  
   b. Sie hat ihn/*seiner vermisst.  
   *she has* him.ACC/*GEN* missed  
   ‘She missed him.’

Importantly, there are NO semantic differences between Icelandic Dat/Gen objects in examples like (21a) and (22a) and corresponding Acc objects in German examples like (21b) and (22b).

Case mismatches of this sort are not only found across related languages but also within individual languages. Consider the Icelandic direct object case marking contrasts in (23)–(24) (showing just a few samples of such contrasts); the verb pairs either have similar (sometimes identical) or opposite meanings.

<table>
<thead>
<tr>
<th>DATATIVE:</th>
<th>ACCUSATIVE:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(23) a.</td>
<td>(24) a.</td>
</tr>
<tr>
<td>bifa ‘(slightly) move’</td>
<td>hreyfa ‘move’</td>
</tr>
<tr>
<td>bjarga ‘rescue’</td>
<td>lifga (við) ‘revive’</td>
</tr>
<tr>
<td>fylgja ‘follow, accompany’</td>
<td>elta ‘follow, chase’</td>
</tr>
<tr>
<td>hjálp ‘help’</td>
<td>aðstöða ‘assist’</td>
</tr>
<tr>
<td>hjúkra ‘nurse’</td>
<td>lækna ‘cure’</td>
</tr>
<tr>
<td>hrósa ‘praise’</td>
<td>lofa ‘praise’</td>
</tr>
<tr>
<td>koma ‘(manage to) move’</td>
<td>flytja ‘move’</td>
</tr>
<tr>
<td>ljúka ‘finish’</td>
<td>klára ‘finish (up)’</td>
</tr>
</tbody>
</table>
These facts highlight two important albeit often neglected or ‘downgraded’ aspects of case variation: First, it is cross-linguistically unpredictable, even for closely related languages with identical case inventories; that is, there are no straightforward universal generalizations or ‘truths’ behind case variation. Second, also language internally there are case irregularities and mismatches; that is, even internally to individual languages variable case marking cannot be fully described in terms of regularities. This is in fact made quite obvious by historical sporadic case marking changes in otherwise stable case systems. Icelandic witnesses many such minor shifts in case use without any concomitant semantic changes (see Barðdal 2001, Jónsson & Eythórsson 2005, Thráinsson 2007). When a verb starts, say, taking a dative object instead of an accusative one, without any meaning shift or system shift, then it cannot be the case that both historical stages (V-Acc vs. V-Dat) represent ‘the regular system.’

Languages are of course full of forms that correlate with some semantics. Such form-semantics correlations are typically significant but incomplete and variably regular and cannot be explained as a result of a direct interpretative connection between PF morphology and semantics (a widely held misconception). Rather, they stem from the fact that semantic interpretation and morphological interpretation is read off at the interfaces from a common underlying syntactic structure (on the minimalist single cycle approach to the derivation up to transfer). Thus, there are indirect but no direct or causal correlations between semantics and morphology.
‘[C]ommunication is a more-or-less affair, in which the speaker produces external events and hearers seek to match them as best they can to their own internal resources’ (Berwick & Chomsky 2011:40). Understanding external language (E-language) is possible because language users interpret or decode it in relation to their own internal language (I-language), and not because of any exact mappings between I- and E-language. Seemingly meaningful morphemes and items like, say, English –s, -ed, and apple, are GRAMMATICAL COINS: By using them, the speaker ‘pays’ for accessing and activating someone else’s I-language, ‘buying’ understanding or reaction. The coins are valid currency, not because they are part of anybody’s I-language but because they relate to and activate I-languages: You need a pen or an apple and you pay with a coin because it gets you what you desire (and not because the coin IS a pen or an apple). This understanding does not, of course, release linguists from the duty—or deprive them of the joy—of carefully analyzing form-semantics correlations in individual constructions and languages, but it ought to save them from the pitfall of analyzing such (more or less inaccurate) relations as being part of I-language.

Let me emphasize that developing an analysis of exactly how some grammatical coin, like, say, a dative case marker (or ‘simply’ a lexical item like pen or apple), can gain currency in some linguistic community is a worthy undertaking. However, the complexity of the issue cannot be overstated. Understanding it requires a theory of at least three subsystems, each of which is largely beyond present day limits of science. First, it requires a coherent understanding of the structure and nature of I-language, the linguistic thought system (there biolinguistics actually seems to be making some encouraging progress, see, e.g., Berwick & Chomsky 2011, Hinzen 2012). Second, it requires an analysis of how one individual I-language can relate to another I-language via the grammar of external language (E-language) and body expressions (acoustic, facial, manual, etc.; cf. the double transfer approach in Sigurðsson 2011c). Third, it also requires a theory of how E-grammar and body expressions can establish themselves as a part of the ‘E-language contract’ in a community.

Regardless of how we deal with these fundamental issues, the study of language and thought is subject to the dilemma of Plato’s cave allegory: We can only access and study I-language via E-language, but E-language is merely a fluctuating shadow of I-language, the relation between the two being fundamentally and inescapably non-isomorphic.\footnote{For a discussion of this, see Sigurðsson 2011c and related work; see also Hinzen 2012.} It follows that a dative case marker, for instance, cannot really be a marker of dative case, paradoxical and distressful as that may seem. Not any more than a coin is a pen or an apple.

4 BEYOND THE V/N/APPL-SYSTEMS
The Icelandic cases are found widely outside of the core argument system. The overview in (25)-(28) is a very brief non-exhaustive sketch of the distribution of the non-nominative cases beyond core arguments (disregarding dislocated NPs and other case agreement phenomena).

(25) **ACCUSATIVE:**
   a. Complements of certain prepositions
   b. Temporal and spatial adverbial NPs

(26) **DATIVE:**
   a. Complements of most prepositions (including agentive NPs in *af- ‘by’ phrases)
   b. A few adnominal NPs (*/she looked in eyes me/; /*to defence me/ = ‘to my defence’)
   c. Complements of certain adjectives (*/she was me kind/*)
   d. Comparative and iterative/frequentative adverbial NPs

(27) **GENITIVE:**
   a. Complements of some prepositions
   b. Most adnominal NPs, reflecting an array of semantic/syntactic relations
   c. Complements of a handful of adjectives (*/she was not worthy his/ = ‘worthy of him’)
   d. Some partitive NPs (*/most theirl = ‘most of them’)
   e. NPs in various other adverbial or adverbial-like functions (see Kress 1982:228ff)

Only a few illustrative examples follow. The accusative adverbial NPs in (28a–b) are temporal, while the one in (28c) is spatial.26 As indicated to the right, other cases are excluded.

(28) a. Hún kom **annan júlí.**  
    *öðrum júlí.DAT, etc.
    *she came second July.ACC
    ‘She came the second of July.’

   b. Hún var **mánuð á Íslandi.**
    *mánuður.NOM, etc.
    *she was month.ACC in Iceland
    ‘She stayed for a month in Iceland.’

   c. Hún **keyrði alla leið.**
    *allrar leiðar.GEN, etc.
    *she drove all way.ACC
    ‘She drove all way.’

---

25 That is, like most other case languages, Icelandic has many adnominal genitives, morphologically homogeneous but syntactically heterogeneous.

26 Adverbial accusatives of this sort measure out a span, temporal or spatial, getting a universal reading (reminiscent of that of universal perfects), indicating that the eventuality of the predicate holds for all subintervals of the temporal or spatial span in question. This ‘measuring property’ is not, as a rule, shared by other accusatives (but see Svenonius 2002 for a different interpretation).
The datives in (29) are iterative/frequentative, while the ones in (30) are comparative.

(29) a. Hann söng *fjógurra sinna.GEN, etc.
    he sang song.the four times.DAT
    ‘He sang the song four times.’

    b. Hann spilar einu sinni í viku.
    he plays one time.DAT in week
    ‘He plays once a week.’

(30) a. Hún stóð hans.GEN, etc.
    she stood him.DAT further
    ‘She was more able than him.’

    b. Hún var þrem ár.ára.ACC, etc.
    she was three years.DAT older than he
    ‘She was three years older than him.’

The adjectival complements in (31) are dative (not uncommon) and genitive (rare).

(31) a. Hann var barnanna.GEN, etc.
    he was children.the.DAT kind
    ‘He was kind to the children.’

    b. Hann var verðlaunum.DAT, etc.
    he was worthy prices.the.GEN
    ‘He was worthy of the price(s).’

These few examples are only meant to give an initial idea about the nature of some of the Icelandic borderline case phenomena. Similar (but only partly identical or overlapping) facts are found in many other case languages. I will not develop any detailed analysis of these facts here – they are lexically restricted with idiomatic or semi-idiomatic properties. What these facts serve to show, however, is that morphological case can behave very much like a virus, penetrating every possible corner of a language, not tolerating a single NP without some morphological case marking. As previously mentioned, marked nominatives, case agreement and case percolation into PRO infinitives also illustrate this tendency to case mark nominal categories even when they are not subject to any independent case instructions.

This virus-like behavior of morphological case is perhaps most clearly demonstrated by adpositional case. Icelandic prepositions fall into four classes with regard to case marking; simple prepositions are listed in (32), with only their most central translations.27

27 There are also numerous complex prepositions or combinations of adverbial elements and propositions, but these always include one of the simplex basic prepositions in (32). Usually, the case assigned depends on the last element in such combinations, innan ár ‘from within,’ upp ár ‘out of, up from’ for instance assigning dative, as the simplex ár ‘out of, from.’ The exception is the –an adverbs/prepositions in (32b2), taking accusative instead of genitive when combined with fyrir ‘of:’ fyrir sunnan landið ‘south of the country.ACC,’ vs. sunnan landsins ‘south of the
All prepositions containing the string /um/ take an accusative complement and any single-word preposition/adverb containing /an/ assigns genitive. Beyond these bleeding phonological relations, dative is the unmarked, general prepositional case.

The variation in (32d) is to an extent semantically predictable and similar phenomena are found in other Indo-European languages (see Libert 2002). Generally, however, the case licensing properties of prepositions are not decided by their semantics. Thus að ‘towards, up to’ and af ‘off, from’ have roughly opposite meanings, and so do gegn ‘against’ and handa ‘for’ as well as, e.g., frá ‘from’ and hjá ‘at, with, beside,’ and yet, all these prepositions obligatorily take a dative complement. Similarly auk ‘in addition to’ and án ‘without’ have roughly opposite meanings, both nevertheless taking genitive complements (án variably took Acc, Dat or Gen in Old Norse).

Again, a comparison between Icelandic and German is instructive (inasmuch as such a comparison is possible and plausible). There are a number of roughly synonymous prepositions that assign the same case in both languages, such as dative-assigning German bei and Icelandic hjá, both meaning ‘by, at,’ accusative assigning German durch and Icelandic gegnum meaning ‘through,’ and so on. However, there are also many prepositional case mismatches between the languages. Recall that they have identical case inventories, and, yet, we find differences like the ones listed in (33).
The case literature contains numerous suggestions that inherently case-marked NPs are actually part of larger structures, PPs or KPs (with a silent K(ase) head that is responsible for the case assignment, functioning much like an overt adposition). See for instance Emonds 1987, Bittner & Hale 1996, Bayer et al. 2001, McFadden 2004, Asbury 2010 for a number of different versions of this basic idea. Demotion (from a case marked NP to an adpositional complement) and patterns like Dative Shift (give the book to X, vs. give X the book) are commonly taken to lend support to this view.

There are four basic combinations of case markers and adpositions, as simply sketched in (34), where NP₀ and NP_C stand for a non-case marked NP vs. case-marked NP, and where P denotes ‘adposition.’

(34)  a.  NP₀    English:       They stayed a month.
       b.  NP_C  Icelandic:  Þeir voru mánuð/mánuður.
              they were month.ACC/NOM (i.e., ‘stayed a month’)
       c.  P + NP₀    English:   I bought three bottles of water.
       d.  P + NP_C  Icelandic: Ég keypti þrjár flöskur af vatni/vatn.
              I bought three bottles of water.DAT/NOM

Some so-called case markers can be successfully analyzed as adpositions in some constructions in some languages; this seems for instance to be true of various locative markers in Finno-Ugric languages, for instance the inessive and the terminative markers in the Hungarian examples in (35) (from Asbury 2010:5).

(35)  a.  Géza olvas a kert-ben.
        Geza reads the garden.INESS
        ‘Geza is reading in the garden.’
       b.  Kovácséknál voltam két hétfög.
        Kovavacs stayed.ISG two week-TERM
        ‘I stayed with the Kovácses for two weeks.’
However, even though it seems profitable to analyze the Hungarian elements *ben* and *ig* in (35a, b) as postpositions rather than genuine case makers, and to analyze the adverbial NPs in constructions like the ones in (34a–b) as PPs or KPs with a zero head (see McFadden 2004, inter alia), there is no gain in analyzing inherently case marked NPs in the Latin/Icelandic/German type of languages as always being PPs or KPs. If, for example, plain datives (indirect objects, direct objects, quirky subjects, etc.) are analyzed as K + NP, we are forced to either assume an empty K even in the presence of an overt P, as in (34d), or to come up with some plausible account of why some K heads must and some must not be spelled out as overt Ps.

Adverbial NPs like the ones in (34a–b) do have similar distributional and semantic properties as certain PPs, but inherently case marked objects commonly have properties that are quite distinct from those of PPs. Thus, Icelandic dative and genitive objects raise to subject in numerous construction types (passives, etc.), whereas full-fledged PPs never do, as illustrated in (36) vs. (37).

(36) a. Hún hélt bókinni.
   *she held book.*
   ‘She kept the book.’

b. Var bókinni haldið?
   *was book.*
   ‘Was the book kept? / Did they keep the book?’

(37) a. Hún hélt á bókinni.
   *she held on book.*
   ‘She held/carried the book.’

b. *Var á bókinni haldið?
   *was on book.*
   ‘Was the book held/carried?’

c. *Var bókinni haldið á?
   *was* book.
   ‘Was the book held on?’

d. Var haldið á bókinni?
   *was held on book.*
   ‘Was the book held/carried?’

As suggested in Sigurðsson 2011a, an overt preposition arguably blocks T_Φ from probing prepositional objects (in Icelandic). That is: Prepositions introduce both structural and semantic information that is otherwise absent.

On a GENERAL PP or KP analysis of inherently case marked NPs, one would presumably expect the P + NP_C type in (34d) to be exceptional in case languages, as compared to the P + NP_0 type in (34c). The opposite is true: Typological research suggests that most morphological case
languages do indeed have case-marked complements of adpositions, and, conversely, that non-case marked complements of adpositions are highly exceptional in case languages (see Libert 2002).

Observations of this sort indicate that morphological marking, even in the simplest and most regular morphological systems, is an unnecessary ‘quirk’ that languages really can do without—suggesting, in turn, that morphological variation is a sociobiological phenomenon rather than a strictly linguistic one. In other words, morphological marking is plausibly not part of or ‘predicted’ by the narrow innate and internal language faculty (Universal Grammar), instead arising in the (PF) externalization process (controlled and affected by biological and social factors that are not specific to the language faculty).  

5. VIRUSES AND STAR WARS

Recall from Section 3 that case star augmentation, yielding v* in addition to plain v and Voice/AG* in addition to plain Voice/AG is involved when a language develops from a no case language to either an accusative or an ergative language, distinguishing Acc or Erg NPs from non-case d (Nom) NPs. Case star augmentation (and case star deletion) is a general phenomenon in case systems.

Recall also the Case Ingredients Generalization in (17), repeated here.

(17) a. NPs are event licensed by heads: Voice, Appl, v, n, a, p, ...
   b. Event licensors of NPs may be case star augmented
   c. The augmented case star may be variably marked: *,*, *
   d. Nom = syntactically a non-case (regardless of its overt marking)

Accordingly, little v, for instance, comes in several flavors in individual case languages, as v, v*, v*,v*,v*++ (Sigurðsson 2012). Thus, the v-case system in nominative-accusative/genitive/dative languages can be described as in (38), where the arrow reads as ‘yields’ (in PF morphology).

---

28 A common misconception is that this conclusion entails that language variation is generally unlimited or unpredictable. It entails, instead, that language variation largely follows from interaction of the 3rd factor with the 1st factor of the genetic endowment and the 2nd factor of experience (cf. Berwick & Chomsky 2011)–it cannot plausibly follow from Universal Grammar as such if it is indeed universal and innate, hence invariable, nor can it be fully explained by the 2nd factor (or else variation could not have arisen in the first place; see the discussion in Sigurðsson 2004, 2011b, 2011c). If we had complete knowledge and understanding of the 3rd factor and its interaction with the 1st and the 2nd factors, we would presumably be able to predict much of the attested variation of the languages of the world, including visible and tactile sign languages. We have a long way to go, but I am optimistic that biolinguistics is (very) slowly moving into the right direction.

29 As mentioned before, Icelandic, of course, arrived at its 4 case system by decline from the 8 case system of Proto-Indo-European. Another way of arriving at (38) is by gradual case growth along the lines sketched below.
The case star approach has one thing in common with KP approaches, namely that it explicitly states that a cased NP is licensed in a more marked or informative structure than is a non-cased (Nom) NP, the more so the more oblique the case. Importantly, however, the augmented markedness or complexity is morphological and not syntactic. That is, SYNTACTICALLY, Icelandic is just like caseless languages, Chinese, Thai, Vietnamese, etc., in having only a general object-licensing v, the v*-flavors being activated in the externalization process, as (abstract) language-specific PF interpretations. In other words, as discussed in Section 3, case instructions (e.g., the ones in (38)) are PF externalization instructions (or interpretations).

Presumably, both Dat and Gen always ‘begin life’ as non-v cases, Gen as an adnominal n*-case and Dat as an Appl*-case. Subsequent introduction of both Dat and Gen into the v-system can be seen as the result of two ‘virus invasions’ or ‘star attacks.’ Gen is commonly more oblique or peripheral than Dat within the v-system (cf. Haspelmath & Michaelis 2008), thus being licensed by a more marked v type head than both the other v-cases, as indicated in (38). That suggests, in turn, that Dat invaded the v-system before Gen did. Thus, nominative-accusative / n-genitive / Appl-dative / v-dative / v-genitive case systems might (for instance) come into being along the lines sketched in (39).

(39) a. \(v^* v \rightarrow \text{Gen} \) yielding the Nom-Acc distinction in the v-system
b. \(n^* n \rightarrow \text{Dat} \) introducing Gen into the n-system
c. \(\text{Appl} > \text{Appl*} \) introducing Dat into the Appl-system
d. \(v, v^* v > v, v^* v^* \) adding Dat to the v-system
e. \(v, v^*, v^* v > v, v^*, v^* v^* \) adding Gen to the v-system

Similarly, Dat can invade the n-system, yielding adnominal datives, such datives being more marked than Gen within the n-system, thus licensed by n*. Also, as mentioned in Section 3, Acc has invaded the Appl-system from the v-system (presumably at least partly as a consequence of the loss of ablative in the prehistory of Germanic), thus being more marked than Dat within the Appl-system (licensed by Appl*), even though it is less marked than Dat within the v-system. In addition, both the adjectival (a) and the prepositional (p) subsystems,

---

30 German (and similar Icelandic) *Ich sah ihm in die Augen* ‘I looked him.DAT in the eyes’, etc. Notice that I am not claiming that these datives are in any kind of a free variation relation to the more central n-system genitive.

31 As pointed out to me by Jim Wood, passive Acc-to-Nom conversion applies to Appl accusatives, indicating that
discussed in Section 4, have developed case licensing a- and p-heads. Thus, (much of) the Icelandic case system can be described as in (40) (where Nom is disregarded).

(40)   *  *+  *++  
  a. v-system    Acc  Dat  Gen  
  b. Appl-system  Dat  Acc  
  c. n-system     Gen  Dat  
  d. p-system33  Dat  Acc  Gen  
  e. a-system    Dat  Gen  

The five subsystems thus make use of only three ‘common’ cases to mark 12 distinct relations. Other languages sometimes take different routes, introducing specific cases for (or limiting them to) some of the subsystems (‘dative’ for instance being confined to prepositional complements in Scottish Gaelic). Subsystem-specificity of this sort can yield systems with many more cases than the Icelandic one (an interesting factor that is nevertheless largely disregarded here). A parallel sketch for English is given in (41).

(41)   *  
  a. v-system    Acc  
  b. Appl-system  Acc  
  c. n-system     Gen  
  d. p-system    Acc  

The cases are evidently not unitary building blocks. That is, applicative dative is not the ‘same’ case as verbal dative, adnominal genitive is not the ‘same’ case as verbal genitive, and so on. Not surprisingly, the traditional view that individual cases are primitives yields recalcitrant puzzles. Thus, it has been a well-known but mysterious fact (see Holmberg 1991:143; see also Thráinsson 2007:173ff. and the references there) that Icelandic has genitive direct objects but no genitive indirect objects. The simple account or description under the present approach is that Gen has invaded the v-system as opposed to the Appl-system. Another fact illustrating that the cases are case star deletion is not as straightforward as the analysis in Sigurðsson 2011a, 2012 would seem to suggest. However, I must put this aside here. – Also, other historical developments than the ones described here are of course possible. My ambition is not to reconstruct the case history of any particular language, but rather to outline how a four case system can develop from a non-case system.

32 On little p as distinct from lexical Ps, see Svenonius 2007.
33 Recall that Dat is the unmarked p-case. I assume that Gen is more marked than Acc as a p-case.
34 As pointed out by a reviewer, this is a description rather than a principled explanation. There is no such explanation nor should there be, as the cases are not syntactic elements. If there was any principled syntactic account of all the variation facts discussed here, we would also have to develop syntactic ‘counterexplanations,’ for example of the fact that most case languages neither have direct nor indirect objects that are inflectionally genitive.
not unitary is that Appl-datives differ from v-datives in being retained in the anticausative -st-construction (see below).

So-called ‘inherent’ case marking of objects is structural rather than lexical (pace Woolford (2006) and many others). That is, the case is not the ‘direct responsibility’ of the lexical V root, but of the particular v type that combines with V to make up a full-fledged verb (and similar considerations apply to the other lexical categories). The v-selection of V, however, is commonly idiomatic or ‘fossilized,’ such that a particular V may only select one particular v-type, seemingly yielding a ‘lexical’ case marking relation. As one would expect, though, many Vs may select more than one v type. Thus, to mention only one type of such variation, some Icelandic verbs make a distinction between dative objects (commonly benefactive) and accusative objects (commonly affected), yielding minimal pairs like ‘comb the child.DAT’ vs. ‘comb the hair.ACC’ (for more observations of this sort for Icelandic, see Barðdal 2001, Maling 2002, Svenonius 2002, Jónsson 2003, 2005, Thráinsson 2007; more generally, Malchukov & de Swart 2009 and the references there).

Actually, the so-called ‘inherent’ or ‘lexical’ cases may be erased in certain Voices. Thus, the expletive Voice head in the anticausative -st-construction, triggers deletion of v-datives, as opposed to Appl-datives. This is illustrated in (42)–(43).

(42)  a. Við lokuðum gluggum. Active Nom-Dat
    we closed windows.the.DAT
    b. Gluggarnir lokuðust. Anticausative Nom
    windows.the.NOM closed-ST

(43)  a. Pétur bauð mér vinnu. Active Nom-Dat-Acc
    Peter offered me.DAT job.ACC
    b. Mér bauðst vinnu. Anticausative Dat-Acc
    me.DAT offered-ST job.NOM

The generalization behind these facts is that Voice_{expl} erases all v-case stars as opposed to the Appl-case star, the latter triggering applicative dative marking (for a more detailed discussion of these and related facts, see Sigurðsson 2011a, 2012, Wood 2012b).

The ‘final’ case of an argument, thus, is not only decided by the initial case licenser, such as v* and v*+, but also by Voice type, a fact which shows, again, that the individual morphological cases are not simple unitary building blocks, neither across languages nor language internally. Rather, they are morphological entities that interpret a variety of underlying syntactic structures, with extensive (but not unlimited) variation both internal to and across languages.
6 CONCLUDING REMARKS

The Blake hierarchy in (2) represents a long tradition in the study of language, where overt markers are taken to be largely unitary and uniform building elements, both language-externally and across languages, each with canonical functional properties. This way of thinking of easily observable variation in some surface form as being linked in some more or less one-to-one like fashion to a deeper and a more general system has been widely abandoned in the natural sciences, but it is still the prevailing view in linguistics. Call it the ANIMISTIC VIEW (avoiding the term ‘realism’): it is animistic in the sense that it ascribes some deeper sense to observable objects, such as case markers.

If the animistic view of case could be upheld—even only some very weak version of it—we would expect a parametric approach to case variation to be basically successful. However, no such approach has ever been developed. Symptomatically, Baker (2001) mentions only one putative case parameter in his ambitious attempt to construct a ‘periodic table of languages,’ the ‘ergative case parameter.’

However, recall the Case Ingredients Generalization in (17), repeated here.

(17) a. NPs are event licensed by heads: Voice, Appl, v, n, a, p, …
   b. Event licensors of NPs may be case star augmented
   c. The augmented case star may be variably marked: *, *, *
   d. Nom = syntactically a non-case (regardless of its overt marking)

In view of the extensive case variation observed within and across languages, the ingredients of case systems are amazingly limited: only a number of event licensing heads and a few variably marked case stars. These atomic elements readily lend themselves to relatively simple ‘parametric’ statements, as we already saw in Section 3 in relation to ergativity and accusativity, the former arising as a consequence of case star augmentation of Voice/AG itself, the latter as a consequence of case star augmentation of little v c-commanded by Voice/AG. Similarly, as we saw, case star augmentation of Appl yields specific case marking of indirect objects, invasion of n*-case into the v-system yields a grammar with some genitive objects, and so on.35

However, statements of this sort are not UG anchored. Markedness is a general third-factor phenomenon and syntactic heads are arguably materialized in individual I-language (out of

---

35 Actually, the issues are more complex than this would seem to indicate. First, Voice and v seem to be the only head types that are full-fledged independent event licensors (other event licensors must enter a matching relation with either v or Voice, it seems, but I must put this aside here). Second, as mentioned above, I am largely disregarding the fact that some languages develop subsystem-specific cases instead of tolerating extensive case invasions/overlaps of the Icelandic sort.
general ‘concept material’) rather than pre-stored in UG (Sigurðsson 2011b, 2011c). That is, the statements in question are third-factor externalization generalizations and not UG parameters in the traditional Principles and Parameters sense. In view of the fact that UG parametric approaches cannot be upheld (Berwick & Chomsky 2011, Boeckx 2011, Sigurðsson 2004, 2011b, 2011c, 2012), it is interesting, even uplifting, that language variation, nevertheless, can be systematically analyzed. It raises some hope that the incredibly complex externalization component of language can be successfully studied.

ACKNOWLEDGMENTS

This article is a ‘follow up’ of sorts to the paper I presented at the Brussels Case at the interfaces conference (BCGL 5) in December 2010 (see Sigurðsson 2012). I would like to thank the organizers of the conference for their hospitality and the audience for helpful comments and suggestions. Many thanks also to Thórhallur Eythórsson, Martin Ringmar, Jim Wood, and, not least, Jeff Parrot for valuable comments and discussions. Thanks also to three reviewers for critical remarks that helped me clarify some moot issues and improve my presentation.

REFERENCES


Pesetsky, David. 2010. Russian case morphology and the syntactic categories. Ms., MIT.


