The Meaning of Case:  
Morphosyntactic Bootstrapping and Icelandic Datives

Iris Nowenstein, Sigríður Sigurjónsdóttir, Charles Yang,  
Anton Karl Ingason, and Joel Wallenberg

1. Introduction

Do children use the same resources to learn verb meaning across languages? One approach to language acquisition in which universality has been extensively debated is the syntactic bootstrapping hypothesis, which proposes that children use the argument structure of a verb as a cue to its meaning (Landau & Gleitman 1985, Gleitman 1990, Naigles et al. 1993). In recent years, the extent to which verbal morphology and morphosyntax can be informative of verb semantics has been the subject of cross-linguistic research, with one of the primary questions being whether possibly (syntactic) universal cues have an advantage over language-specific (morphological) ones (e.g. Lidz et al. 2003, Göksun et al. 2008, Matsuo et al. 2012, Trueswell et al. 2012 and Leischner et al. 2016).

Using corpora and experimental acquisition data from Icelandic, a language with almost no argument-drop and rich case morphology, we provide qualified support for a morphosyntactic bootstrapping account that does not exclusively rely on universal cues, since a learning model detects the available systematic mappings of form and meaning (Yang 2016). In specific contexts, we argue that morphology can be as salient as the number of arguments. Additionally, we argue that experimental comprehension results show the necessary basis for the well-documented productivity of the Icelandic non-default dative (Maling 2002, Svenonius 2002, Jónsson and Eythórsson 2005, Ingason 2010 and Barðdal 2011 i.a.). Specifically, we show that non-default subject case marking rules can be accounted for with Yang’s (2016) Tolerance Principle (TP).

Lidz et al. (2003), based on ideas of universal syntax-semantics mapping, argued that children initially rely on argument number and ignore morphological form to bootstrap verb meaning, even when the morphology provides stronger cues. This has been challenged from various perspectives, one of them being typological evidence against the universality of argument structure cues (Brown & Bowerman 2008).

Still, even work on argument-drop languages such as Japanese and Turkish reveals that children use syntactic frames as cues – in addition to e.g. case

* Corresponding author: Iris Nowenstein, University of Iceland, ien1@hi.is. We thank the audience at the 44th BUCLD for useful feedback. We are also very grateful to all the participants in our experiments.

morphology (Göksun et al. 2008 and Matsuo et al. 2012). Furthermore, research on German (Leischner et al. 2016) shows that children rely less on the number of arguments and more on case when word order is highly flexible. But what about languages that do not drop arguments and have a relatively rigid word order (like English) but still have a rich morphological case system (like Turkish)? Icelandic is such a language, with robust semantically driven dative productivity in subject and object case, and also well-documented links between case and lexical semantics (e.g. Jónsson 1997–1998, Maling 2002, Svenonius 2002 and Barðdal 2008).

2. Background
2.1. Icelandic datives, variation and Yang’s Tolerance Principle (2016)

As seen in (1), Icelandic has semantic minimal pairs where word order is not informative but case, either in subject or object position, is the differentiating factor:

(1) a. Hún klóraði köttinn/kettinum
     she.NOM scratched the.cat.ACC/DAT
     ‘She scratched the cat.’ (ACC=malefactive, DAT=beneffective)

b. Hún skaut boltann/boltanum
     she.NOM shot the.ball.ACC/DAT
     ‘She shot the ball.’ (NOM=patient, DAT=theme)

c. Skrímslið/skrímslinu er kalt
     the.monster.NOM/DAT is cold
     ‘The monster is/feels cold.’ (NOM=theme, DAT=experiencer)

In object position (1a and 1b), dative case can e.g. indicate a benefactive\(^1\) object rather than a malefactive one or a movement theme rather than a patient. In (1a), the accusative case frame indicates a reading where the cat is hurt by the scratching, while the dative indicates a pleasant scratch and, unlike the accusative, requires the object to be animate. In (1b), the accusative case frame indicates an event where the ball is shot with e.g. an arrow or a bullet, while the dative case frame yields an interpretation in which the ball itself necessarily moves somewhere. In subject position (1c), case alternations can, for example, differentiate between theme and experiencer subjects. In (1c), the nominative subject case frame has no animacy requirement and indicates that the monster is cold to the touch, without it necessarily feeling the cold. The dative case frame,

\(^1\) In this context, the terms benefactive and malefactive are used to differentiate between positive and negative effects on the patient/theme. Under some analyses, a more appropriate term would be (affected) experiencer (see discussion in e.g. Bosse, Bruening and Yamada 2012).
on the other hand, requires an animate argument and indicates that the monster feels cold without it necessarily being cold to the touch. Note that for subjects, this type of construction can only form a minimal pair when the subject is in the neuter singular and the agreement on the adjective therefore is syncretic to default agreement.

An extensive body of research addresses the relationship between case, argument structure and meaning in Icelandic (e.g. Yip, Maling and Jackendoff 1987, Jónsson 1997–1998, Maling 2002, Svenonius 2002, Barðdal 2001 and 2008 and 2011, H.Á Sigurðsson 2012, Wood 2015 and E.F. Sigurðsson 2017). Although the correspondence between case and argument role is far from being straightforward or categorical, some broad generalizations, arguably relevant from the standpoint of acquisition, can be extracted. In general, the default (structural) case marking is nominative for subjects and accusative for objects. The default case marking has few semantic restrictions, but agents are always nominative (so non-nominative always corresponds to non-agent), with the exception of the appropriately named exceptional case marking (ECM) contexts. Additionally, proto-typical patients tend to be accusative.

Although the default case frame is the most prominent pattern of verbal argument structure (types and tokens), Icelandic shows non-default case marking with subjects and objects (as shown in (1)). In the present paper we focus on dative, the most frequent non-default case of verbal arguments, well-known for its robust productivity in Modern Icelandic (Andrews 1976, Svavarsdóttir 1982, Maling 2002, Barðdal 2001 and 2008, Svenonius 2002 and Jónsson and Eythórsson 2005 i.a.).

In a corpus study of Modern Icelandic texts, Barðdal (2001) found that 69.4% of object tokens were accusative, 25% were dative, 3% nominative and 2.6% genitive. When it comes to type frequency, Jónsson’s (2005) representative lists indicate that 135 of 477 (28.3%) monotransitive verbs require a dative object, 60.6% take an accusative object, 7.1% appear both with accusative and dative objects and 4% appear with genitive objects. For ditransitives, the majority of verbs, 72.4%, require the indirect object to be dative. This has prompted analyses in which dative case on indirect objects, largely recipients or goals, is considered predictable and labeled semantic/thematic/inherent as opposed to less predictable idiosyncratic/quirky case (e.g. Zaenen, Maling and Thráinsson 1985, Jónsson 2003, Woolford 2006 and E.F. Sigurðsson 2017). When it comes to monotransitives, detailed overviews (see Maling 2002 in particular) show that although no categorical semantic motivation for case marking can be found, some patterns emerge. In object position, animate dative arguments tend to be beneficiaries while inanimate dative arguments are often the themes of movement verbs.

Dative subjects are far less common than dative objects. In the corpus study mentioned earlier, Barðdal (2001) found that 93–94% of subject tokens were nominative. If we take a closer look at non-nominative subjects only and exclude verbs requiring nominative subjects, we can see that dative is the most common case. In Jónsson’s (2005) list, 301 verbs/predicates with non-nominative subjects
are listed, 81.1% of them being dative. As we will show in section 2.2, very few of those verbs appear in child-directed speech, with dative subjects appearing more frequently than accusative ones. As has been mentioned, all non-nominative subjects are non-agents. Most of them are experiencers of psych verbs and require animacy (1c and 2a), but some are theme subjects as shown in (2b).

(2) a. Mig/mér langar í epli.
   me.NOM/DAT longs for apple (DAT → innovation)
   ‘I want an apple’

   b. Báttinn/báturinn rak á land.
   the.boat.ACC/NOM drifted to land (NOM → innovation)
   ‘The boat drifted to shore’

In (2), we additionally exemplify well-documented patterns of variation in subject case marking, with Dative Substitution (DS) in (2a) and Nominative Substitution (NS) in (2b). Extensive research, including numerous surveys on teenagers’ and adults’ case marking with the verbs in question, show that this variation in subject case, both inter-speaker and intra-speaker, is widespread (Svavarðsdóttir 1982, Jónsson 2003, Jónsson & Eythórsson 2005, Thráinnsson 2013 and Nowensteinn 2017). Within approaches which link case to theta roles (e.g. Yip, Maling & Jackendoff 1987, Jónsson 2003 and Woolford 2006), dative subjects have been associated with experiencers and not themes. Jónsson and Eythórsson (2005) and others have based their explanations on the differences between NS and DS on such assumptions, arguing that oblique subject case with motion or change of state verbs (theme verbs, e.g. (2b)) in Icelandic is always unpredictable. Dative subjects of experiencer verbs (e.g. (2a)) are on the other hand assumed to be more regular, and their productivity is therefore apparent in DS.

Although the subjects of theme verbs, just like the subjects of experiencer verbs, are originally both accusative and dative, it has been noted that their dative fails to attract the accusative in the same way that dative experiencer subjects do. Instead, both accusative and dative are replaced by nominative (Jónsson 2003 and Jónsson & Eythórsson 2005, but see Guðmundsdóttir et al. 2019). On the other hand, the productivity of the non-default dative in the context of experiencer subjects is undisputed. This difference in the directionality of variation between themes and experiencers is also predicted in Yang’s (2016) analysis, with the application of the Tolerance Principle. In simple terms, the Tolerance Principle (Yang 2016) accounts for the productivity of rules or patterns. A rule is productive if the maximum number of exceptions (e) under the Tolerance Principle threshold ($\theta_N$), calculated with the natural log of the total number of types, is not exceeded: $e \leq \theta_N = N/\ln N$. By applying the Tolerance Principle to the number of attested oblique subject verbs shown in Table 1, Yang (2016) predicts that the dative subject verbs are not numerous enough for productivity in the case of themes, while they are in the case of experiencers.
Table 1: Application of the Tolerance Principle (Yang 2016:165) to Icelandic subject case variation.

<table>
<thead>
<tr>
<th></th>
<th>ACC</th>
<th>DAT</th>
<th>Total (N)</th>
<th>( \theta_N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theme</td>
<td>14</td>
<td>19</td>
<td>33</td>
<td>9</td>
</tr>
<tr>
<td>Experiencers</td>
<td>37</td>
<td>227</td>
<td>264</td>
<td>47</td>
</tr>
</tbody>
</table>

As shown in Table 1, the accusative theme subject verbs exceed the computed threshold while the accusative experiencer subject verbs do not. Dative should therefore be productive for non-nominative experiencer subjects but not for theme subjects.

Although the role of mechanisms such as overgeneralization and leveling have regularly been brought up in the context of Icelandic subject case variation (e.g. Jónsson and Eythórsson 2005), only very recent accounts include modelling of this dative productivity in the context of language acquisition (Yang 2016, Nowenstein 2017 and Guðmundsdóttir et al. 2019). Additionally, the acquisition of Icelandic case marking in general remains largely unexplored.

2.2. Dative acquisition and productivity

Previous research on the acquisition of case in Icelandic has focused on production (Sigurðardóttir 2002), showing that children acquire the default case marking first, that is nominative on subjects and accusative on subjects. This happens early in acquisition, with the two-word stage already including case marked arguments (Sigurjónsdóttir 2005). Dative objects appear from around age two, and dative/non-nominative subjects appear last, around age three. Overgeneralizations are in line with this developmental path, as well as research on the acquisition of case in other Germanic languages (Schütze 1997, Eisenbeiss et al. 2006, Schmitz 2006), with nominative on subjects and accusative on objects being frequently overgeneralized. Additionally, children acquiring Icelandic overgeneralize the non-default dative (Nowenstein 2019) in subject and object position – unlike German children (Schmitz 2006). In object case, dative arguments can appear instead of benefactive prepositional phrases or accusative objects with verbs such as lesa (‘read’), kyssa (‘kiss’) and hug (‘knúsa’), and less frequently, accusative themes of motion verbs such as færa (‘move’) are replaced with the dative. Additionally, children overgeneralize dative subjects with predicates such as vera sveitt(ur) (‘be sweaty’), and show Dative Substitution patterns as shown in (2a). Therefore, it can be argued that the dative productivity discussed in the context of adult variation in the previous section is also found in child Icelandic.
To explore this further, we use Icelandic child language corpora (Strömqvist et al. 1995, Einarsdóttir 2018 and Sigurjónsdóttir corpus, unpublished) and attempt to extend Yang’s (2016) analysis of Dative Substitution. With a compiled corpus of approximately 500,000 words, we extracted all experiencer verbs with non-nominative subjects and a frequency >1 (note that no non-nominative subjects were themes), as shown in Figure 1 and discussed below.

![Figure 1: Verbs with experiencer subjects in the compiled child language corpus, by frequency and subject case marking.](image)

We find 17 verbs with dative subjects and 7 verbs with originally accusative subjects (but note that they often appear in the dative in adult production), yielding 24 verbs. We calculate the threshold for productivity using Yang’s Tolerance Principle (2016): $e \leq \theta_N = N/\ln N = 24/\ln(24) = 8$. Since the number of verbs with accusative subjects is below the threshold, this indicates that dative should be productive for non-nominative experiencer subjects under the Tolerance Principle. In further work, we aim to conduct similar analyses for datives in object position.

Now that the productivity of datives in Icelandic language acquisition has been established and accounted for, we turn back to comprehension and case as a cue for verb meaning. Although work on Icelandic datives has been completely focused on production data, the robust semantically conditioned productivity which has been reported has implications for comprehension. Specifically, we argue that a sensitivity to the semantic distribution of case not only is a prerequisite for productivity, it also provides morphosyntactic cues to verb meaning. To investigate this possibility, we conducted the comprehension experiments reported on in section 3.
3. Case as a cue for verb meaning: Experiments

As has been mentioned, the robust semantically-conditioned case marking patterns, in addition to a relatively rigid word order and almost no argument-drop, make Icelandic a convenient test case in the context of (morpho)syntactic bootstrapping. In this section, we report on two picture-selection comprehension tasks, with choice of relevant thematic role as the outcome variable (see e.g. Yuan, Fisher and Snedeker 2012). The experiments were designed to test whether Icelandic speaking children use case as a cue for verb meaning. The first experiment is a minimal pair task with existing Icelandic verbs/predicates, targeting both object and subject case, while the second one is a novel (nonce) verb task targeting subject case only.

3.1. Minimal pair task

Minimal pair examples such as the ones discussed in section 1.2 (example 1) are particularly relevant in the context of morphosyntactic bootstrapping, since word order, a syntactic and possibly universal cue, is uninformative. Meanwhile, morphological case, a language-specific morphosyntactic cue, is the only differentiating factor. To explore whether children are able to use case to differentiate between semantic interpretations in such a context, we conducted an experiment with 48 preschoolers aged 2;4–6;4. Eight frequent Icelandic dative/non-dative sentence pairs were used, five object case pairs and three subject case pairs.

Figure 2: Example of the pictures used in the minimal pair task.

All participants were presented with the eight pairs, yielding a total of 16 stimuli presented with two possible interpretations in the forms of pictures. Figure 2 shows an example for the predicate *be cold*, where a dative subject indicates an experiencer reading (on the right) while the nominative indicates a theme (on the left), see discussion for example (1c).
3.2. Novel verb task

In addition to the minimal pair task, 146 children aged 2–13 participated in a novel verb task with a 2x2 design, manipulating subject case (nominative/dative) and argument number (transitive/intransitive). This makes it possible to quantify the effects of case and argument number respectively, as well as the interaction between the two variables.

![Figure 3: Pictures used in the novel verb task (response to stimuli).](image)

All participants were presented with two novel verb stimuli for each condition, eight sentences in total. Examples are shown in (3), with the transitive condition in (3a) and the intransitive one in (3b). Note that hann (‘he/him’) is syncretic and could therefore be either in the nominative or accusative.

(3) a. **Hesturinn/hestinum gorpar** hann
   the.horse.NOM/DAT gorps he/him
   ‘The horse gorps him’

   b. **Hesturinn/hestinum gorpar**
   the.horse.NOM/DAT gorps
   ‘The horse gorps’

---

2 This experiment took place within the MoLiCoDiLaCo project at the University of Iceland (PIs: Sigriður Sigurjónsdóttir and Eiríkur Rögnvaldsson), which was awarded a grant of excellence from the Icelandic Research Fund 2016-2019.
Participants were then asked to choose the appropriate picture out of six options (see Figure 3), three with a single argument and three with two arguments. The outcome measure was the choice of an experiencer (first option from the right in both rows).

3.3. Results

Figures 4 and 5 illustrate the results from the minimal pair task: Figure 4 shows the general results, while Figure 5 shows the results by younger (2–4 years old, \(N = 27\)) and older (5–6 years old, \(N = 21\)) participants.

**Figure 4:** Results from the minimal pair task (\(N = 48\)). Choice of thematic role by case.

The outcome variable in these figures is the choice of thematic role indicated by the dative case in adult Icelandic, namely experiencers in subject position and either beneficiaries or movement themes in object position. Starting with Figure 4, we can see that preschoolers learning Icelandic seem to be able to take advantage of some case cues when interpreting sentences where word order is not informative. The tendency is much stronger for (experiencer) subjects as compared to objects. Children associate dative subject case with an experiencer reading. With objects, case is more conclusive for beneficiaries.
Figure 5: Results from the minimal pair task, by age group.

Turning to Figure 5, we see that the subject case effect is stable across age groups. However, younger children seem to associate the dative with beneficiaries more strongly than older children. For movement themes, case does not seem to be interpreted as a differentiating factor across age. This distinction between beneficiaries and movement themes pairs well with the developmental facts discussed in section 2.2, where dative is overgeneralized earlier and more frequently with beneficiaries. Indeed, there are few reported instances of children overgeneralizing the dative with movement themes. As a contrast, the productivity of the dative for movement themes seems to be preserved in adult Icelandic. This appears in the case marking of many recently coined Icelandic verbs of movement, such as dánlóda (‘download’), which take a dative object. For new verbs with recipients/beneficiaries, on the other hand, dative does not seem to be productive (see discussion about the case marking of new Icelandic verbs in Barðdal 2008). We could therefore speculate that the association between dative and beneficiaries/recipients is only productive for a brief time in early acquisition.

To explore the results further, we conducted a nested comparison (Likelihood Ratio Test) of mixed effects logistic regression models (lme4 in R, Bates et al. 2015). Age (as a continuous variable, in months) was included in every model as well as random intercepts for participants and items. We found a significant improvement to the model fit by adding case ($\chi^2 (5) = 31, p < 0.001$), thematic role ($\chi^2 (6) = 19.7, p < 0.001$) and a case*syntactic position interaction ($\chi^2 (6) = 2.1, p < 0.001$).

Figures 6 and 7 illustrate the results from the novel verb task: Figure 6 shows the general results, while Figure 7 shows the results across four different age groups (2;0–4;11 years old: $N = 37$, 5;0–7;11 years old: $N = 50$, 8;0–10;11 years old: $N = 33$ and 11;00–13;11 years old: $N = 26$). Looking at Figure 6, the tendency to interpret dative subjects as experiencers seems to be confirmed in the novel verb task.
Using nested model comparison again, we see significant improvements to the fit of the model by adding transitivity ($\chi^2 (5) = 60.3$, $p < 0.001$) and case ($\chi^2 (6) = 56.3$, $p < 0.001$) but not transitivity*case to the model ($p = 0.386$) – so we do not observe an augmented case effect in the intransitive condition, for example. Both effects (case and transitivity) interact with age, as is illustrated in Figure 7. The effects of both variables are augmented as the children grow older. When correcting for age, case (estimate: -1.05, SE: 0.17) and argument number (estimate: -1.1, SE: 0.17) show comparable effect sizes.
4. Discussion

Putting the experimental results presented in section 3 in the context of the syntactic bootstrapping literature, we argue that the Icelandic data provides additional support for the claim that children can rely on language-specific, morphological cues to derive verb meaning (Göksun et al. 2008, Matsuo et al. 2012 and Leischner et al. 2016). Our results show that children acquiring Icelandic can, early on, use case to determine verb meaning when word order is uninformative. We furthermore show, by testing transitivity and case morphology with the same experiment, that case morphology can be as salient as the number of arguments in specific contexts. Both types of cues become more prominent with age. In general, Icelandic children’s use of case as a cue for verb meaning points towards a morphosyntactic bootstrapping account that does not exclusively rely on universal cues. Additionally, the properties of Icelandic allow us to show that the use of morphological cues does not rely on argument-drop or word order being less rigid.

It is important to bear in mind that the use of case cues varies across contexts in our experimental results. Dative on subjects, as opposed to objects, seems to be much more salient. This could be interpreted in various ways, one of them being that the low frequency of non-nominative subjects in general, paired with the generalization that non-nominative corresponds to non-agents, provides a more accessible contrast to default case marking. In general, the properties of Icelandic dative subjects could prove important in research on the acquisition of psych verbs (e.g. Hartshorne et al. 2015 and Harrigan et al. 2016). Another interpretation could be related to effects of parsing on the availability of cues (Trueswell et al. 2012). It could then be argued that the different positions of subjects and objects play a role, with non-nominative on a subject forcing an earlier reinterpretation.

The fact that the use of case cues is context-specific is not problematic if we assume a learning model that detects the available systematic mappings of form and meaning (Yang 2016). We argue that the same learning model, through the Tolerance Principle, predicts Dative Substitution in Icelandic experiencer subjects. This is an extension of Yang’s (2016) work on the well-documented variation in Icelandic subject case, and shows how generalizations which guide verb learning are also relevant in production. The results are therefore relevant for work on productivity and rule formation in child language: Icelandic dative productivity provides a scenario where exceptions are structured and non-default patterns can be generalized systematically based on distinctions which arise in comprehension.

References


Schmitz, Katrin. 2006. Indirect objects and dative case in monolingual German and bilingual German/Romance language acquisition. In Hole, Daniel, Meinunger, André & Abraham, Werner. (eds.): *Datives and other cases: Between argument structure and event structure*, pp. 239–268.


