



**UNIVERSITY  
OF ICELAND**

# Building yourself a variable case system

The acquisition of Icelandic datives

**Iris Edda Nowenstein**

Dissertation towards the degree of Doctor of Philosophy

2023

**Faculty of Icelandic and Comparative Cultural Studies**

Íslensku- og menningardeild Háskóla Íslands  
hefur metið ritgerð þessa hæfa til varnar  
við doktorspróf í íslenskri málfræði

Reykjavík, 24. febrúar 2023  
Gauti Kristmannsson  
deildarforseti

The faculty of Icelandic and Comparative Cultural Studies  
at the University of Iceland has declared this dissertation  
eligible for defense leading to a Ph.D. degree in Icelandic Linguistics

Doctoral Committee:  
Sigríður Sigurjónsdóttir, supervisor  
Anton Karl Ingason  
Joel C. Wallenberg  
Charles Yang

*Building yourself a variable case system – The acquisition of Icelandic datives*

© Iris Edda Nowenstein  
Reykjavík 2023

Dissertation for a doctoral degree at the University of Iceland. All rights reserved.  
No part of this publication may be reproduced in any form without written permission of the author.

ORCID 0000-0003-1945-163X

# Abstract

The Icelandic case marking system has for decades proved an important testing ground for theories on case and productivity. This is in part because Icelandic has the rare property of marking argument function redundantly, with both rich morphological case marking and a relatively rigid word order (Kiparsky 1997, Fedzechkina et al. 2016). Additionally, the Icelandic (non-default) dative is surprisingly robust and productive, both in object and subject case (e.g. Andrews 1976, Thráinsson 1979, Svavarsdóttir 1982, Maling 2002, Jónsson & Eythórsson 2005 and Barðdal 2008). The productivity of the Icelandic dative does appear in typologically expected semantic contexts, namely with experiencer and recipient arguments, but also with e.g. themes of motion verbs (Maling 2002, Jónsson 2003, Barðdal 2008). Despite these well-known associations between morphological dative case and specific semantics, the relationship between case and meaning in Icelandic is still a matter of debate. Although various patterns have been identified (agents are always nominative, patients are typically accusative and indirect objects (recipients) most frequently dative), linguists have pointed out that the correlations are not exact and exceptions can be found (e.g. Maling 2002, Thráinsson 2007, H.Á. Sigurðsson 2012 and Wood 2015). Conflicting views can therefore be discerned in previous work: Case productivity is assumed to be semantically conditioned while the relationship between case and meaning is rejected on the basis of being too approximative. But do patterns (rules of language) always have to be absolute and without exceptions?

In this thesis I argue that from the standpoint of language acquisition, which has been lacking in research on Icelandic case marking, rules do not need to be exceptionless to be discovered by children and thus become part of the grammar. If the distributional evidence for case-semantics associations is present in children's language environment, they will learn these patterns as long as the number of exceptions is tolerable (Yang 2016). This hypothesis is supported by the results of this thesis, which builds on experimental data from 148 children aged 2–13 as well as incremental child language corpus analyses. Broadly, the results show that children associate the dative more with experiencers, recipients and motion themes in comparison to other thematic roles,

despite known exceptions. The association between the dative and applied arguments (experiencers and recipients) is formed earlier than the association with themes of motion verbs. The corpus analysis furthermore shows that these mappings of form and meaning can be derived from the distributional properties of the input, which can also predict the well-studied patterns of dative productivity and case marking variation in Icelandic. I argue that the learnability perspective is crucial to the understanding of language variation and change, as well as providing valuable insights into theoretical approaches to case.

However, the main contributions of the thesis lie in the field of language acquisition, where the properties of the Icelandic case marking system also have the potential to become important testing grounds for various fundamental questions. The results of the dissertation have implications within three distinct but related topics in language acquisition research: (1) (Morpho)syntactic bootstrapping, (2) Rule formation and productivity and (3) Variation and change in acquisition.

First, I 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, Trueswell et al. 2012 and Leischner et al. 2016). The results show that children acquiring Icelandic can, early on, use case to determine verb meaning when word order is uninformative. I furthermore show, by testing transitivity and case morphology within the same experiment, that case morphology can be as salient as the number of arguments in specific contexts. Second, the data call for a model of rule formation which accounts for nested non-default productivity and does not hinge on frequency to determine gradient overgeneralization, since patterns which are not statistically dominant (the dative) can become productive and multiple rules can apply to a context simultaneously. I argue this is most compatible with rule-based accounts which assume that rules are derived from distributional information in the input (Yang 2002, Albright & Hayes 2003, O'Donnell 2015 and Yang 2016). Third, I show how such an account, a combination of Yang's Tolerance Principle (2016) and his Variational Model of language acquisition (2002), can also account for the acquisition of the Icelandic case marking system in well-known scenarios of language variation and change, and that a case should be made for stochastic exceptions in addition to stochastic rules. I argue that the underlying mechanisms in language acquisition, namely a constant search for

productive rules while tracking exceptions, and both convergence with and divergence from the language environment, shape the directionality and dynamics of language variation and change.

# Útdráttur

Íslenska fallmörkunarkerfið hefur um áratugaskeið reynst mikilvægur prófsteinn á formlegar kenningar um fall, virkni málfræðireglna og breytileika í máli, m.a. vegna þess að íslenska hefur þann sjaldgæfa eiginleika að merkja hlutverk rökliða bæði með nokkuð fastri orðaröð og ríkulegri fallbeygingu (Kiparsky 1997 og Fedzechkina o.fl. 2016). Þar að auki telst íslenska þágufallið bæði óvenju algengt og virkt miðað við það sem þekktist í skyldum málum og birtist þessi virkni bæði í falli frumlaga og andlaga (sjá t.d. Andrews 1976, Höskuld Þráinsson 1979, Ástu Svavarsdóttur 1982, Maling 2002, Jóhannes Gísla Jónsson og Þórhall Eypórsson 2003 og Jóhönnu Barðdal 2008). Virkni íslenska þágufallsins kemur bæði fram í samhengi sem er vel þekkt í tungumálum heimsins, þ.e. með rökliðum sem teljast til skynjenda og þiggjenda, og einnig í óvenjulegra samhengi eins og á þemum hreyfisagna. Þannig er viðurkennt að virkni þágufalls í íslensku fer að einhverju leyti eftir merkingarlegum þáttum, þar sem þágufall birtist oftast með skynjendum, þiggjendum og þemum hreyfisagna (sjá t.d. Jóhannes Gísla Jónsson 2003 og Jóhönnu Barðdal 2008), en á sama tíma er umdeilt að hversu miklu leyti fall felur í sér merkingu. Bent er á að þrátt fyrir ýmis mynstur í dreifingu falls og merkingarhlutverka (t.a.m. eru gerendur alltaf í nefnifalli, þolendur iðulega í þolfalli og óbein andlög (viðtakendur) í þágufalli) sé samband falls og merkingar ekki áreiðanlegt og ýmsar undantekningar á mynstrunum séu til staðar (sjá t.d. Maling 2002, Höskuld Þráinsson 2007, Halldór Ármann Sigurðsson 2012 og Wood 2015). Þannig togast í raun á andstæð sjónarmið í fræðunum: Gert er ráð fyrir því að virkni falls fari eftir merkingarlegum þáttum en um leið er tengslum falls og merkingar hafnað á grundvelli óáreiðanleika og undantekninga. En þurfa mynstur (málfræðireglur) að vera algild til að teljast raunveruleg?

Í ritgerðinni eru færð rök fyrir því að út frá sjónarhóli máltökunnar, sem hefur skort í umræðu um íslenskt fall hingað til, þurfi reglur ekki að vera algildar til þess að verða til í málkerfi barna og þ.a.l. fullorðinna. Ef börn uppgötva mynstur í dreifingu falls og merkingar í málumhverfi sínu tileinka þau sér þau mynstur svo lengi sem undantekningar fara ekki yfir ákveðin þolmörk eða virkniþröskuld (Yang 2016). Rannsóknargögnin sem niðurstöður þessarar ritgerðar byggja á eru annars vegar tilraunagögn frá 148 börnum á aldrinum tveggja til þrettán ára, sem safnað var að mestu innan öndvegisverkefnisins

„Greining á málfræðilegum afleiðingum stafræns málsambýlis“ (verkefnisstjórar: Sigríður Sigurjónsdóttir og Eiríkur Rögnvaldsson), og hins vegar greiningar á algengustu sögnunum í barnamálgögnum sem að mestu var safnað af Jóhönnu T. Einarsdóttur og Sigríði Sigurjónsdóttur. Niðurstöður sýna á heildina litið að börn tengja þágufall frekar við skynjendur, þiggjendur og hreyfipemu en við önnur merkingarhlutverk þrátt fyrir undantekningar. Tenging þágufalls við skynjendur og þiggjendur kemur auk þess fram fyrir en tengingin við hreyfipemu. Niðurstöður greininga á barnamálgögnum sýna síðan að börn geta áttað sig á þessum tengslum forms og merkingar út frá vísbendingum í málumhverfinu. Í ritgerðinni eru þessar niðurstöður ræddar og tengdar áður nefndri fræðilegri umræðu um merkingu falls og færð rök fyrir því að í formlegum kenningum um fall þurfi frekar að taka mið af því hvernig fallmörkunarkerfið byggist upp í máltöku en að einblína á stöku undantekningar frá þekktum mynstrum.

Helsta framlag þessarar doktorsritgerðar er þó innan máltökufræða. Sýnt er fram á að þágufall í íslensku, og íslenska fallmörkunarkerfið almennt, getur einnig svarað ýmsum grundvallarspurningum um máltöku barna. Niðurstöður ritgerðarinnar fela þannig í sér nýja þekkingu innan þriggja ólíkra en skyldra sviða máltökufræða: (1) notkunar forms til að leiða út merkingu í máltöku (e. *(syntactic) bootstrapping*), (2) reglumyndunar barna og virkni málfræðireglna þeirra og (3) tileinkunar barna á tilbrigðum í máli og tengslunum á milli máltöku barna og málbreytinga. Uppbygging ritgerðarinnar fylgir þessum þremur sviðum og hver kafli, fyrir utan þann fyrsta sem er almenn kynning á falli í íslensku, inniheldur sérstakar tilraunir og niðurstöður þar sem afmörkuðum spurningum er svarað.

Í öðrum kafla er meginspurningin sú hvort börn nýti fallmörkun til þess að leiða út merkingu sagna, en þekkt hefur verið um árabíl (Landau og Gleitman 1985) að börn geta nýtt setningagerð til þess að draga ályktanir um merkingu orða. Hins vegar er álitamál í hversu miklum mæli breytur eins og fallmörkun gagnast að þessu leyti. Niðurstöður kaflans benda þó til þess, í samræmi við rannsóknir á tyrknesku, japönsku, tagalog og þýsku (Göksun o.fl. 2008, Matsuo o.fl. 2012, Trueswell o.fl. 2012 og Leischner o.fl. 2016), að börn geti notað fall sem vísbendingu um merkingu sagna. Þannig geta íslenskumælandi börn t.a.m. snemma notað fall til þess að leiða út merkingu sagna þegar setningagerðin felur ekki í sér neinar viðbótarupplýsingar um merkingu. Í þriðja og fjórða kafla er áhersla lögð á reglumyndun og virkni í máltöku barna og athugað hvort og þá

hvernig fall tengist merkingu í svokölluðum framköllunarprófum (e. elicited production) á falli. Niðurstöður þessara kafla, þar sem skýr merkingarlega skilyrt virkni þágufallsins kemur fram, kalla á líkön um reglumyndun í máli barna sem geta gert grein fyrir virkni markaðra mynstra og reiða sig ekki eingöngu á tíðni til þess. Færð eru rök fyrir því að reglumiðuð (e. *rule-based*) líkön sem byggja á tölfræðilegri dreifingu ílagsins séu best til þess fallin (t.d. Yang 2002, Albright og Hayes 2003, O'Donnell 2015 og Yang 2016), bæði þegar gert er ráð fyrir máltöku með og án tilbrigða í málumhverfinu. Í fimmta kafla er sjónum nefnilega beint að tileinkun fallmörkunar þegar umfangsmikill innri breytileiki er staðar eins og í tilviki þágufallshneigðar. Sýnt er fram á að breytileiki í því hvort þágufallshneigð kemur fram í máli einstaklings eða ekki er skilyrtur af m.a. persónu og tölu frumlagsins og því hvort samfall beygingarmynnda kemur fram í beygingarmynstri þess eða ekki. Samfall nefnifalls og þolfalls, þar sem ekki kemur eins skýrt fram að um aukafallsfrumlag er að ræða og í samfalli aukafalla, hefur í för með sér aukna notkun þágufalls. Athugað er hvort umfang þágufallshneigðar foreldra spái fyrir um tíðni þágufallshneigðar hjá börnum þeirra en svo reynist ekki vera. Börn sýna almennt meiri þágufallshneigð en fullorðnir en tileinka sér þó sömu málfræðilegu skilyrðingu á breytileikanum og birtist í máli foreldra þeirra, sem er í samræmi við nýlegar rannsóknir á því hvernig börn tileinka sér breytileika í máli (t.d. Hall og Maddeaux 2020 og Repetti-Ludlow og MacKenzie 2022). Færð eru rök fyrir því að þau ferli sem eru undirliggjandi í máltöku barna, annars vegar stöðug leit þeirra að virkum reglum og skráning á undantekningum frá þeim og hins vegar bæði sam- og sundurleitni með málumhverfinu, séu ráðandi öfl í þróun tilbrigða og málbreytinga.

Í sjötta kafla eru niðurstöður ritgerðarinnar loks fléttaðar saman í eina heild þar sem sýnt er fram á að íslenskumælandi börn byggja sér breytilegt fallmörkunarkerfi með umfangsmikilli þágufallsvirkni þar sem merkingarlegir þættir skipta höfuðmáli. Þetta ferli mótar tilbrigði í máli og hefur áhrif á þróun málbreytinga, og ætti auk þess að mynda grunninn að formlegum kenningum um fallmörkun – í takt við þá staðreynd að málkerfi fullorðinna byggir á máltöku barna.

# Table of Contents

<b>List of Figures</b>	12
<b>List of Tables</b>	15
<b>Acknowledgments</b>	17
<b>Introduction</b>	21
<b>1. Icelandic Datives</b>	29
1.1 The Icelandic case marking system at a glance	29
1.1.1 On the surface	29
1.1.2 Distribution and status	33
1.1.3 Semantics	38
1.2 Acquisition and productivity	43
1.3 Variation and change	48
1.3.1 Dative Substitution (DS)	48
1.3.2 Nominative Substitution and relationship to DS	53
<b>2. The Meaning of Case</b>	56
2.1 Morphosyntactic bootstrapping	56
2.1.1 Main questions and result summary	61
2.2 Methods	63
2.2.1 Minimal pair task	63
2.2.2 Novel verb task	66
2.3 Results	68
2.3.1 Minimal pair task	69
2.3.2 Novel verb task	72
2.4 Discussion	76
<b>3. Producing non-default case</b>	78
3.1 Rule formation and productivity	78
3.1.1 Main questions and result summary	81
3.2 Methods	83
3.2.1 Existing verbs	83
3.2.2 Novel verbs	87
3.3 Results	89
3.3.1 Existing verbs	90
3.3.1.1 Subject case	90
3.3.1.2 Object case: Accusative	94

3.3.1.3 Object case: Dative	98
3.3.1.4 Object case: Genitive	101
3.3.1.5 A note on borrowings	105
3.3.2 Novel verbs	107
3.3.2.1 Subject case	107
3.3.2.2 Object case	110
3.3.2.3 Typical and atypical contexts for the dative	112
3.3.2.4 Predicting productivity at the individual level	115
3.4 Discussion	117
<b>4. Dative Derivability: A Case for Exceptions</b>	<b>120</b>
4.1 Learning from distributional information	120
4.1.1 Main questions and results	125
4.2 Methods	125
4.2.1 Thematic roles	128
4.2.1.1 Agents	128
4.2.1.2 Experiencers	129
4.2.1.3 Patients	130
4.2.1.4 Themes	132
4.2.1.5 Goals	133
4.2.2 Patterns and hypothesis formation	134
4.3 Results	136
4.3.1 Overview	136
4.3.2 Datives in production	146
4.3.2.1 Form	146
4.3.2.2 Meaning	152
4.3.2.3 Form and meaning	163
4.3.2.4 A note on adjectival predicates and ditransitives	167
4.3.3 Datives in comprehension	168
4.3.3.1 Form	168
4.3.3.2 Meaning	171
4.3.3.3 Form+ and meaning	175
4.4 Discussion	178
<b>5. Acquiring and shaping case marking variation</b>	<b>183</b>
5.1 Inconsistent input	184
5.1.1 Main questions and results	187
5.2 Methods	187
5.3 Results	189

5.3.1 Parental variant rate	189
5.3.2 Conditioning factors	192
5.3.2.1 Person-Specific Retention	192
5.3.2.2 Syncretism effects	194
5.4 Discussion	196
<b>6. Synthesis</b>	202
6.1 Case has meaning and children know it	202
6.2 Non-defaults can be productive	205
6.3 Case marking rules can be derived from the input	207
6.4 Acquisition specializes variation	211
6.5 Future work	213
6.5.1 The Tolerance Principle as an adaptation	214
<b>References</b>	217
<b>Appendix A: Comprehension task stimuli</b>	237
<b>Appendix B: Production task stimuli</b>	243
<b>Appendix C: Variation task stimuli</b>	247
<b>Appendix D: Modeling results</b>	250
<b>Appendix E: IceCASE corpus</b>	260

# List of Figures

<b>2.1</b> Minimal pair task stimuli, subject condition: <i>Vera kaldur/kalt</i> ‘be cold’.	<b>66</b>
<b>2.2</b> Minimal pair task stimuli, object condition: Klóra ‘scratch’.	<b>66</b>
<b>2.3</b> Minimal pair task stimuli, object condition: Skjóta ‘shoot’.	<b>67</b>
<b>2.4</b> Pictures used in the novel verb task, subject condition (response to stimuli).	<b>68</b>
<b>2.5</b> Pictures used in the novel verb task, object condition (response to stimuli).	<b>69</b>
<b>2.6</b> Results from the minimal pair task. Proportional choice of (dative) thematic role interpretation by case.	<b>70</b>
<b>2.7</b> Results from the minimal pair task, by age group. Proportional choice of (dative) thematic role interpretation by case.	<b>71</b>
<b>2.8</b> Results from the novel verb task. Choice of experiencer subject by case and transitivity.	<b>73</b>
<b>2.9</b> Results from the subject condition of the novel verb task, by age group.	<b>74</b>
<b>2.10</b> Results from the novel verb task. Choice of motion theme by case.	<b>75</b>
<b>2.11</b> Results from the object condition of the novel verb task, by age group.	<b>76</b>
<b>3.1</b> Stimuli for the novel and semi-novel predicates in the forced-choice production task.	<b>89</b>
<b>3.2</b> Forced-choice production task with existing verbs, historical subject case.	<b>91</b>
<b>3.3</b> Forced-choice production task with existing verbs by age group, historical subject case.	<b>92</b>
<b>3.4</b> Forced-choice production task with existing verbs, subject case.	<b>93</b>
<b>3.5</b> Forced-choice production task with existing verbs by age group, subject case.	<b>94</b>
<b>3.6</b> Forced-choice production task with existing verbs, historical accusative object case.	<b>96</b>
<b>3.7</b> Forced-choice production task with existing verbs by age group, historical accusative object case.	<b>96</b>
<b>3.8</b> Forced-choice production task with existing verbs, object case with historically accusative verbs.	<b>97</b>
<b>3.9</b> Forced-choice production task with existing verbs by age group, object case with historically accusative verbs.	<b>98</b>

<b>3.10</b> Forced-choice production task with existing verbs, historical dative object case.	<b>100</b>
<b>3.11</b> Forced-choice production task with existing verbs by age group, historical dative object case.	<b>100</b>
<b>3.12</b> Forced-choice production task with existing verbs, object case with historically dative verbs.	<b>101</b>
<b>3.13</b> Forced-choice production task with existing verbs by age group, object case with historically dative verbs.	<b>101</b>
<b>3.14</b> Forced-choice production task with existing verbs, historical genitive object case.	<b>103</b>
<b>3.15</b> Forced-choice production task with existing verbs by age group, historical genitive object case.	<b>103</b>
<b>3.16</b> Forced-choice production task with existing verbs by age group, historical genitive object case.	<b>104</b>
<b>3.17</b> Forced-choice production task with existing verbs by age group, object case with historically genitive verbs.	<b>105</b>
<b>3.18</b> Forced-choice production task with novel verbs, subject condition.	<b>109</b>
<b>3.19</b> Forced-choice production task with novel verbs by age group, subject condition.	<b>110</b>
<b>3.20</b> Forced-choice production task with novel verbs, object condition.	<b>111</b>
<b>3.21</b> Forced-choice production task with novel verb by age group, object condition.	<b>112</b>
<b>3.22</b> Forced-choice production task with novel verbs by age group, subject and object condition.	<b>113</b>
<b>3.23</b> Associations between the dative and typical and atypical semantic contexts, by syntactic role. Comprehension results on the left (mean experiencer/theme rate) and production results (mean dative rate) on the right.	<b>114</b>
<b>3.24</b> Associations between the dative and typical and atypical semantic contexts, by syntactic role. Production results by age group.	<b>115</b>
<b>3.25</b> Dative productivity and expressive vocabulary, linear model fitted (95% confidence) and age specified.	<b>118</b>
<b>4.1</b> Distribution of analyzed verbs in the IceCASE corpus by case frame.	<b>138</b>

<b>4.2</b> Distribution of analyzed verbs in the IceCASE corpus, presence of a dative argument.	<b>139</b>
<b>4.3</b> Proportion of different case frames in the IceCASE corpus, top 50 most frequent verbs.	<b>140</b>
<b>4.4</b> Proportion of different case frames in the IceCASE corpus, top 150 most frequent verbs.	<b>141</b>
<b>4.5</b> Proportion of different case frames in the IceCASE corpus, top 300 most frequent verbs.	<b>142</b>
<b>4.6</b> Proportion of different case frames in the IceCASE corpus, all verbs (410).	<b>143</b>
<b>4.7</b> Distribution of analyzed verbs in the IceCASE corpus, by presence of a dative argument. Reflexives included.	<b>144</b>
<b>4.8</b> Distribution of the top 50 most frequent verbs in the IceCASE corpus by case frame. Icelandic verbs above and translation below.	<b>145</b>
<b>4.9</b> Distribution of the top 50 most frequent verbs in the IceCASE corpus by presence of a dative argument, reflexives included in the lower frame.	<b>146</b>
<b>5.1</b> Rate of original case retention in child-caregiver dyads, by age group.	<b>191</b>
<b>5.2</b> Dative selection/acceptance in the first and third person singular, by group (children and caregivers).	<b>193</b>
<b>5.3</b> Dative selection in the first and third person singular, by age group (children).	<b>194</b>
<b>5.4</b> Dative selection/accusative rejection with/without syncretism on the subject, by group (children and caregivers).	<b>196</b>
<b>5.5</b> Dative selection with/without syncretism on the subject, by age group (children).	<b>197</b>

# List of Tables

<b>1</b> Examples for possible syncretism patterns in Icelandic noun declension. Syncretism shown with bold lettering and color patterns.	<b>33</b>
<b>2</b> Distribution of object case with transitive verbs, based on Jónsson's (2005) representative lists for Modern Icelandic.	<b>34</b>
<b>3</b> Nominative-accusative syncretism and lack of syncretism in Icelandic subject case and possible relationship to Dative Substitution rate.	<b>52</b>
<b>4</b> Syncretism patterns in the Icelandic pronominal system.	<b>53</b>
<b>5</b> Application of the Tolerance Principle (Yang 2016, p. 165) to Icelandic subject case variation.	<b>55</b>
<b>6</b> Linear regression predicting dative productivity.	<b>117</b>
<b>7</b> The maximum number of exceptions for a productive rule over N items.	<b>124</b>
<b>8</b> Argument case by syntactic role, the IceCASE corpus.	<b>148</b>
<b>9</b> Evaluation of the [+subject → +nominative] rule in the IceCASE corpus.	<b>148</b>
<b>10</b> Evaluation of the [-subject → -nominative] rule in the IceCASE corpus.	<b>149</b>
<b>11</b> Evaluation of the [-subject → +accusative] rule in the IceCASE corpus.	<b>150</b>
<b>12</b> Evaluation of the [+direct object → +accusative] rule in the IceCASE corpus.	<b>151</b>
<b>13</b> Evaluation of the [-subject, -direct object → +dative] rule in the IceCASE corpus.	<b>152</b>
<b>14</b> Argument case by thematic role, the IceCASE corpus.	<b>154</b>
<b>15</b> Evaluation of the [-agent → -nominative] rule in the IceCASE corpus.	<b>155</b>
<b>16</b> Evaluation of the [-agent → +accusative] rule in the IceCASE corpus.	<b>156</b>
<b>17</b> Evaluation of the [+patient → +accusative] rule in the IceCASE corpus.	<b>156</b>
<b>18</b> Evaluation of the [-agent, -patient → +dative] rule in the IceCASE corpus.	<b>157</b>
<b>19</b> Evaluation of the [+goal, +experiencer → +dative] rule in the IceCASE corpus.	<b>158</b>
<b>20</b> Evaluation of the [+goal → +dative] rule in the IceCASE corpus.	<b>158</b>
<b>21</b> Evaluation of the [+goal → +dative] rule in the IceCASE corpus.	<b>159</b>
<b>22</b> Evaluation of the [+experiencer → +dative] rule in the IceCASE corpus, randomized samples.	<b>160</b>

<b>23</b> Evaluation of the [+obl. experiencer → +dative] rule in the IceCASE corpus.	<b>161</b>
<b>24</b> Evaluation of the [+theme → +accusative] rule in the IceCASE corpus.	<b>162</b>
<b>25</b> Evaluation of the [+theme, -motion → +accusative] rule in the IceCASE corpus.	<b>163</b>
<b>26</b> Evaluation of the [+theme, +motion → +dative] rule in the IceCASE corpus.	<b>163</b>
<b>27</b> Evaluation of the [+direct object, +patient → accusative] rule in the IceCASE corpus.	<b>165</b>
<b>28</b> Evaluation of the [+direct object, -patient → accusative] rule in the IceCASE corpus.	<b>165</b>
<b>29</b> Evaluation of the [+direct object, +theme → accusative] rule in the IceCASE corpus.	<b>166</b>
<b>30</b> Evaluation of the [+direct object, +theme, -motion → accusative] rule in the IceCASE corpus.	<b>167</b>
<b>31</b> Evaluation of the [+direct object, +theme, +motion → dative] rule in the IceCASE corpus.	<b>167</b>
<b>32</b> Evaluation of the [+nominative → +subject] rule in the IceCASE corpus.	<b>170</b>
<b>33</b> Evaluation of the [-nominative → -subject] rule in the IceCASE corpus.	<b>171</b>
<b>34</b> Evaluation of the [+accusative → +direct object] rule in the IceCASE corpus.	<b>171</b>
<b>35</b> Evaluation of the [+dative → -subject, -direct object] rule in the IceCASE corpus.	<b>172</b>
<b>36</b> Evaluation of the [+nominative → +agent] rule in the IceCASE corpus.	<b>173</b>
<b>37</b> Evaluation of the [+accusative → +patient] rule in the IceCASE corpus.	<b>174</b>
<b>38</b> Evaluation of the [+accusative → +patient OR +theme] rule in the IceCASE corpus.	<b>174</b>
<b>39</b> Evaluation of the [+dative → +goal OR +theme] rule in the IceCASE corpus.	<b>175</b>
<b>40</b> Evaluation of the [+dative → +goal] rule in the IceCASE corpus.	<b>175</b>
<b>41</b> Evaluation of the [+subject, +nominative → +agent] rule in the IceCASE corpus.	<b>176</b>
<b>42</b> Evaluation of the [+direct object, +dative → +theme] rule in the IceCASE corpus.	<b>178</b>
<b>43</b> Final rules assumed in the incremental corpus analyses with the TP. Rule productivity in yellow, lack of productivity in blue.	<b>181</b>

# Acknowledgments

Despite the individualistic values which dictate so many aspects of academia, I think that creating scientific knowledge is inherently, and perhaps increasingly so, a collaborative effort. Every new contribution builds on and creates a dialogue with past work, and at the same time a contribution's value is in part determined by others' engagement with it. For the past few years, I have very much felt how my own work has been enriched, and really just made possible, by the people who I have been fortunate enough to share this journey with in one way or another.

First and foremost, I want to thank Sigga (Sigríður Sigurjónsdóttir), my advisor. I owe so many of the opportunities I have been given to her. She has believed in me and trusted me to develop my own ideas while always pushing me towards more rigor and precision. One of the things I am most thankful to Sigga for, however, is the fact that her support extended beyond my work and to me as a whole person. She was a constant reminder that having a life outside academia, pursuing other passions and dedicating time to friends and family, was not only okay but just frankly necessary. This is something I wish every mentor would understand and encourage.

I was lucky enough to have three other brilliant people on my committee: Charles Yang, Anton Karl Ingason and Joel Wallenberg. Working with them has been tremendously rewarding. Throughout this project, and even while I was an undergrad, they have always been ready to engage with me and give me the time I needed to understand previous ideas and develop my own through (sometimes way too long) conversations, while asking me the hardest questions but taking my answers seriously (even when they were bad), the necessary ingredients for progress. I think what I am most thankful for is their willingness to give even my most unusual ideas a shot, which taught me to not be afraid of convention in the pursuit for knowledge.

So many others have been willing to discuss case, variation and/or acquisition with me for the whole decade I have been obsessed with these subjects, showing me new perspectives and inspiring me. Einar Freyr Sigurðsson has been incredibly generous in this respect, and I also want to thank Jóhannes Gísli Jónsson, Höskuldur Þráinsson, Jim Wood, Sigríður Mjöll Björnsdóttir, Matthew Whelpton, Joe Jalbert, Katie Schuler, John Trueswell, Joan Maling, Sigríður Magnúsdóttir, Heimir Freyr Viðarsson, Laurel

MacKenzie, Meredith Tamminga, John Grinstead, Silvina Montrul, Ásgrímur Angantýsson, Cherlon Ussery, Ari Páll Kristinsson and Eiríkur Rögnvaldsson, as well as the outstanding set of people I have met at various workshops and conferences when presenting my work (e.g. Rask, PLC, FWAV, DiGS, GALANA, GURT, TALV, BUCLD and NWAV). It has been invaluable to make friends on this journey, and even though I did not have a proper cohort at the University of Iceland, I shared being a PhD student in (Icelandic) linguistics with an amazing group of women: Takk fyrir að vera samferða mér, Sirrí, Þorbjörg, Dagbjört, Ingunn, Tinna, Sirrý, Binna, Steffi og Lilja!

For the first three years of my PhD, I got to work with the lovely MoLiCoDiLaCo team on questions of acquisition, variation and change in the context of language contact and preservation, tackling socially important issues which were prominent in Icelandic public discourse. It taught me a lot, and being part of the small community surrounding the research project was an absolute joy. Most of the experimental data this thesis is based on were collected within the project, so I owe a lot to all the researchers involved. Of course, doing language acquisition experiments is not possible without extremely generous parents or teachers and most importantly, children who are willing to put up with researchers like me. I am extremely grateful to the parents and children who participated in the MoLiCoDiLaCo testing sessions, as well as everyone who helped me gather additional data at the Klettaborg preschool. When it comes to the child language corpora, the data Jóhanna T. Einarsdóttir and Sigríður Sigurjónsdóttir shared made the whole difference, and Hinrik Hafsteinsson was instrumental in getting probably the messiest corpus in the history of Icelandic language technology in shape for the analyses I wanted to do.

Another pretty important component of doing a PhD is securing funding, and I was lucky enough to be awarded a doctoral grant from The University of Iceland Research Fund in 2019. I would never have been able to do that without the superhero of every PhD student in the Humanities at the University of Iceland: Margrét Guðmundsdóttir, who also happens to be an amazing linguist. Having her as an ally in the university's administration made an incredible difference for me through quite a few challenges. Þúsund þakkir fyrir allt!

Finally, I want to thank all the important people in my life outside academia for putting up with me through this crazy ride and filling it with happy moments and support when I needed it the most. These are all the friends I wish I had seen so much more of over the past few months, my incredibly funny and brilliant co-workers in my other life as an SLP at the National University Hospital and of course, my family. My two big little brothers have been the best, Thomas with his inspiringly good nature (despite being only 14!) and Nils with surprising (I'm kidding) common sense, wisdom and real talk at the most crucial moments. I can't even put into words how grateful I am for my parents. I would never have been able to get this done without them. My father taught me how to think and still asks me the type of questions which challenge me the most, and somehow he is always the one I call when my thoughts have gotten into the worst tangles. Gracias por todo, pa. My mother is truly a force of nature and gives everything she has to offer to the ones she loves. I owe so much of the best things in my life to her incredible support. Merci pour tout, ma maman. To my amazingly loving family-in-law: Thank you for always being there and making everything easier. And to my little family which got bigger while I worked on this thesis, Haukur, Elísabet and Ýmir: You are just the best. Thank you for reminding me every day what really matters in life, for being so patient (ok, maybe not Ýmir) and for giving me all this love. Dúbæ. These last few months have not been easy. In the most difficult moments, having Haukur by my side, cheering me on and always making me laugh (even when I mostly wanted to cry), made all the difference in the world. Takk fyrir allt, ástin mín.

**Last but not least: This thesis is dedicated to my two grandmothers, mi abuela and ma mamie, Nora and Liline.** They achieved incredible things in difficult times and gave so much back to their communities, big things and smaller ones. Both transmitted to me a love for conversation as well as reflection, and the importance of sharing good food and reading good books. Their love and generosity made me who I am, and I will always be inspired by their strength, wisdom and humility.

In general, I feel extremely privileged and grateful for having been able to pursue my passion for language science with this project. It is a strange feeling to be able to do what you love while the world is burning. Not only did I have access to all the education I aspired to receive, I also had the necessary support when I made the choice to carry a

child while working on my thesis. Here's to a world where everyone is given equal opportunities to thrive within the limits of our planet. This may sound naive and utopian, but I have always thought academia should be about finding realistic solutions to impossible problems.

# Introduction

In the broadest sense, this dissertation is a study of how children map form to meaning and use language productively, and how the mechanisms involved in language acquisition shape variation and change. The problem is approached from the perspective of the Icelandic case-marking system, with its notoriously robust non-default dative productivity and instances of well-documented variability. In this context, I focus on how children make use of distributional information to discover form-meaning mappings and case-marking rules, both in contexts where the input is consistent and in situations where it contains widespread stylistically and grammatically conditioned intra-speaker variation. I argue for an analysis in which children detect available systematic mappings of form and meaning and do not necessarily consistently favor possibly universal cues such as the number of arguments. Instead, the functionality of a cue in acquisition depends on context and salience, which has different implications for comprehension and production. When it comes to rule productivity, I emphasize that from a learnability standpoint, exceptions can be tolerated to a certain extent. I show that some of the broad generalizations between case, structure and semantics which have been discovered and subsequently discarded within formal approaches to case, on the basis of correlations not being exact, are in fact learnable from the input and acquired by children. Finally, this dissertation investigates how children acquire case marking rules when faced with widespread intra-speaker variation in the language community. I argue that the mechanisms underlying rule formation in acquisition can in part explain how variation can become conditioned and specialized while incrementation still occurs, and show that children do not match parental variant rate but replicate contrasts systematically instead.

This project, which builds on experimental data from 148 children aged 2–13 as well as incremental child language corpus analyses,<sup>1</sup> therefore has implications within

---

<sup>1</sup> All the experimental data (excluding the minimal pair task in Chapter 2) were collected within the research project “Modeling the Linguistic Consequences of Digital Language Contact”, directed by Sigríður Sigurjónsdóttir and Eiríkur Rögnvaldsson (MoLiCoDiLaCo, <https://molicodilaco.hi.is/>), which was financed by a Grant of Excellence from the Icelandic Research Fund from 2016 to 2019 (grant no. 162991). The child language corpus also contains data from the MoLiCoDiLaCo project but is mostly composed of data from the Icelandic CHILDES corpus (Strömquist et al. 1995), Sigurjónsdóttir’s (2008) longitudinal corpus and finally data from Einarsdóttir et al.’s (2019) corpus of language samples. In addition to being funded within the MoLiCoDiLaCo project, I received a grant from the UI doctoral grants fund.

three distinctive but related subjects of language acquisition research: (1) (Morpho)syntactic bootstrapping, (2) Rule formation and productivity and (3) Variation and change in acquisition. The organization of the dissertation follows from this partitioning, with each chapter including an independent study aside from Chapter 1, where I begin by providing a general descriptive background and overview of Icelandic datives and relevant previous research. I then proceed to morphosyntactic bootstrapping in acquisition and case as a cue for verbal semantics in comprehension (Chapter 2). The associations between case, meaning and syntactic role investigated in Chapter 2 form the basis for children's productive rule formation which is the subject of Chapters 3 and 4. In Chapter 5, I focus on how children acquire grammatically conditioned variation in subject case marking and how acquisition itself shapes change. Finally, Chapter 6 concludes with a general summary, discussion and possible directions for future work. The results from Chapters 2–5 are briefly summarized below.

In Chapter 2, I use two comprehension experiments, a minimal pair task (N = 48 preschoolers aged 2–6) and a forced-choice novel verb task (N = 146, children aged 2–13), to show that children form associations between case and meaning relatively early in acquisition, using case as a cue for verbal semantics. This is interesting in the context of the syntactic bootstrapping theory (originally proposed by Landau & Gleitman 1985), where it has been demonstrated that children use syntactic information to learn verb meaning. Broadly speaking, the results in Chapter 2 show that datives (in contrast with the nominative and accusative) are associated with experiencers in subject position and benefactives/goals and motion themes in object position. The strength of the association varies between age groups and contexts, with the largest effects appearing with experiencers in subject position, despite the low frequency and restricted semantics of non-nominative subjects in the input. The associations between the dative and goals and motion themes are more elusive, yielding less clear interpretation contrasts based on case. With novel verbs, the results show that in particular semantic contexts, morphological case, a language-specific cue,<sup>2</sup> can be as salient as the number of arguments, a potentially universal cue which has been instrumental in work arguing for a syntactic bootstrapping approach to language acquisition where universal cues are

---

<sup>2</sup> Language-specific in the sense that it is not universally (overtly) marked.

privileged (Lidz et al. 2003). Indeed, despite the fact that the universality of argument structure cues has been called into question (Bowerman & Brown 2008), support for the syntactic bootstrapping account has even been found in work with argument-drop languages such as Japanese and Turkish, where children use syntactic frames as cues – in addition to e.g. case morphology (Göksun et al. 2008 and Matsuo et al. 2012). In this respect, one hypothesis (Leischner et al. 2016) is that children rely less on the number of arguments and more on case when the syntax is less reliable (word order less rigid). The results in Chapter 2 suggest that this does not need to be the case. Under certain conditions, case and argument number can be equally valuable as cues.

The results of this thesis are therefore a contribution to an ongoing debate about the resources children use to map form to meaning, particularly in the context of psych verbs (Hartshorne et al. 2015 and 2016, White et al. 2017, Hacquard & Lidz 2018, Harrigan et al. 2019, Shablack et al. 2020), i.e. predicates describing psychological states, and their experiencers.<sup>3</sup> I argue that data from Icelandic, a language with very limited argument-drop and a relatively rigid word order (like English), but rich case morphology (like Turkish), can provide important insights in this context. The results from the experiments in Chapter 2 provide qualified support for a morphosyntactic bootstrapping account that does not exclusively rely on universal cues, and I therefore follow e.g. Göksun et al. (2008), Matsuo et al. (2012) and Leischner et al. (2016), in building on the insights from the syntactic bootstrapping literature without limiting the learning possibilities to argument number and order. The basic theory is the same: Form drives the acquisition of meaning and verbs can be grouped into semantic categories based on their structural behavior (see also Pinker 1989, Levin 1993 and Yang 2016). But other properties than word order and argument number can be relevant, and in languages like Icelandic, morphosyntactic bootstrapping, based on case, is a strong candidate. Since a learning model detects the available systematic mappings of form and meaning (Yang 2016), cue universality is not crucial to cue availability.

In Chapters 3 and 4, I show how these generalizations, which guide verb learning, are also relevant in production, further investigating the patterns which can be derived

---

<sup>3</sup> Note that terminology varies when it comes to psych/perception/attitude/mental state verbs. Here I use the term psych verb in the broadest sense, including all the categories just mentioned.

from the distribution of the dative and how they are learned. This is done with a production task featuring existing and novel verbs ( $N = 101$ , children aged 3–13), described in Chapter 4, and an incremental analysis of the arguments of verbs appearing with a frequency  $>5$  ( $N = 410$ ) in an approximately 750,000 word child language corpus (Einarsdóttir et al. 2019, Sigurjónsdóttir 2007, Strömqvist et al. 1995), described in Chapter 5. The arguments were syntactically and thematically annotated by hand, building on previously proposed generalizations on the associations between case, structural role and meaning. I refer to the resulting frequency list as the IceCASE (Icelandic Child Argument Structure Evidence) corpus. I then implemented a maximal approach to hypothesis formation, both in terms of the size of the set a rule applies to, rule specificity and the conceptual categories assumed (e.g. Rissman & Majid 2019). When a productive rule is not discovered, the set is further divided (Belth et al. 2021). Rule productivity was incrementally evaluated using the Tolerance Principle (Yang 2016), a parameter-free learning model which provides a threshold for the number of exceptions a productive rule can tolerate:

- (1) Let a rule  $R$  be defined over a set of  $N$  items.  $R$  is productive if and only if  $e$ , the number of items not supporting  $R$ , does not exceed  $\theta_N$ :

$$e \leq \theta_N = \frac{N}{\ln N}$$

The Tolerance Principle draws on the idea that productive rules must be applicable to a large enough number of candidates, without having too many exceptions. Additionally, the nature of the computation yields a higher proportion of tolerable exceptions in smaller sets. The corpus study therefore provides insights into the linking problem and the distributional information children could use to discover associations between argument form and role (e.g. Pearl & Sprouse 2021). It also revives the broad generalizations between case, structure and meaning which have been discussed by linguists (e.g. agents are always nominative, patients tend to be accusative and goals are usually dative) but have remained theoretically elusive, and their existence even contested (H.Á.

Sigurðsson 2012), on the basis of correlations not being exact (Maling 2002, Thráinsson 2007, Wood 2015, McFadden 2020).

I argue that from an acquisitional standpoint, which has been lacking in the debate on the semantics of case, generalizations do not need to be without exceptions to be discovered and learned. Exceptions to productive rules can be tolerated (Yang 2016) and even structured into subrules (see Belth et al. 2021 for a computational approach). By combining child language corpus analyses and experimental data, I show that broad associations between Icelandic case, structure and meaning, can in fact be discovered and learned by children and form the basis of the non-default dative productivity. Indeed, the results show that associations between Icelandic datives and specific semantic contexts are formed in language acquisition, despite known exceptions. The corpus analyses furthermore show that they are discoverable from the input and the novel verb experiments in Chapters 2 and 3 confirm their productivity. Dative is produced more frequently than the default nominative (subjects) or accusative (objects) when associated with experiencers, goals and themes of motion verbs. Furthermore, age patterns in production suggest that associations between the dative and syntactic role are established before semantic correlations. This precedence of form over meaning is also present in the corpus results and resonates with previous findings showing differences in the emergence of pattern recognition (e.g. Emond & Shi 2021) and semantically conditioned rule application (e.g. Schuler et al. 2016), where pattern recognition is attested sooner and possibly is a necessary condition for the discovery of form-meaning mappings (Yang 2020).

The use of the Tolerance Principle in the analysis of the child language corpus furthermore sheds light on various more specific puzzles relating to the productivity of Icelandic datives (e.g. the lack of dative subjects and indirect objects with novel verbs, see section 1.2). It becomes clear how complex patterns of productivity (and lack of productivity) can arise, both within structured exceptions (dative subjects, where nominative is the productive default) and subrule formation (dative objects, where the default accusative fails to reach productivity in various contexts), where a productive rule is not necessarily found. The abundance of productive non-default (dative) forms in the Icelandic case marking system therefore provides an interesting test case for approaches

to abstraction, rule formation, productivity and overgeneralizations in acquisition, and the findings of Chapter 3 and 4 are discussed in the context of alternative models of productivity in language acquisition (such as Pinker 1999, O'Donnell 2015, Blything et al. 2018, Ambridge et al. 2018, Goldberg 2019 and Ambridge 2020). I argue that the data presented in this thesis supports accounts where abstractions are derived from the input and productivity is binary.

In Chapter 5, I assess how general models of language acquisition account for the learning of variation and how acquisition in turn shapes variation. One of the fundamental questions within developmental sociolinguistics, and language acquisition research more broadly, has to do with children's reaction to variability in their input, or primary linguistic data (e.g. Labov 1989, Yang 2002, Hudson Kam & Newport 2005, Smith et al. 2009, Cournane & Pérez-Leroux 2020). Repetti-Ludlow and MacKenzie (2022) describe that at a glance, the literature points to a paradox: Children are both expected to diverge from and match their caregivers' speech. Despite children's tendency to go beyond the input, with e.g. overgeneralizations, we do expect children to learn their caregivers' dialect, and they have in fact been known to match the rates of variation found in their environment (Labov 1989, Smith et al. 2009, Johnson & White 2020). When a closer look is taken at the growing body of literature targeting children's acquisition of variation, it becomes clear that the paradox is in fact attested in reality and that both regularization and matching occur, but under different circumstances. Indeed, the nature of the developmental path can depend on a number of factors such as the learner's age and the amount and consistency of exposure to different dialects, but also the variable type (language domain), the complexity (or existence) of the conditioning factors and the social salience of the variable (e.g. Smith et al. 2009, Hendricks et al. 2018). In Chapter 5, the focus is placed on subject case variation and grammatical conditioning factors, where I argue that children neither regularize nor match their caregiver. Instead, they replicate the systematic contrasts they encounter. The main questions of the chapter are therefore whether children acquire the widespread intra-speaker variation found in the input, or regularize it, and whether they match their caregivers' variant rate. The associations between the dative and experiencer subjects in comprehension and production (Chapters 2 and 3) are put in the context of existing and well-documented change in progress, with

the corpus results in Chapter 4 additionally providing a context where variation is distributionally expected. To investigate how Dative Substitution (DS) is acquired and shaped by variation, DS was investigated in 101 children aged 3–13 and their caregivers (80 dyads) by using forced-choice tasks and grammaticality judgments across multiple items as a proxy for case use. The conditioning factors investigated, the Person-Specific Retention (Nowenstein & Ingason 2021) and less documented syncretism effects, are both instances of how change that can be traced back to leveling and regularization can subsequently lead to complex conditioning.

The results show that children do not regularize but instead acquire the widespread intra-speaker variation present in their caregivers' language, including the syncretism effects which have previously not been confirmed in Icelandic and consist of the dative being used more frequently if nominative-accusative syncretism is present in the inflectional paradigm. Interestingly, the syncretism effects are not present in the results from the youngest children, while the Person-Specific Retention appears in all age groups. Furthermore, the results show that when acquiring the grammatically conditioned case marking variation of Icelandic Dative Substitution, children do not match their parents' variant rate, as already mentioned. Instead, systematic contrast replication at the group and individual level is observed, or regularization within matched conditions. In other words, children acquire the conditioning patterns observed in their caregivers' results without producing the variants at the same rates. Their rate of DS is always higher, pointing towards incrementation. This is in line with results from previous work (e.g. Smith et al. 2007, Hall & Maddeaux 2020 and Repetti-Ludlow & MacKenzie 2022) and has implications for studies on the role of variant specialization (Wallenberg 2019) and the dynamics of variation in individuals (Tamminga et al. 2016) for the directionality of change. I furthermore argue that models of language acquisition need to be able to account for the emergence of productive intra-speaker variation patterns. In this context, I show that a combination of Yang's Variational Model of Language Acquisition (2002), where children can acquire competing rules, and the Tolerance Principle (2016) is a promising approach in that respect (as in Woods et al. 2021), with grammatical conditioning of variation being reframed in the context of productive rule discovery.

In Chapter 6, all the findings of the thesis are brought together and discussed generally. The main results are summarized and joined into a comprehensive account that describes how children learning Icelandic build themselves a variable case system with abundant non-default productivity, how acquisition shapes variation and change and how formal theories on case should build upon findings from acquisition, as adult grammars do.

# 1. Icelandic Datives

As Joan Maling (2002, p. 31) and others (e.g. Svenonius 2002, Thráinsson 2007, Barðdal 2008, Jónsson 2013, Wood 2015) have observed, “one of the striking features of Icelandic syntax is the frequency with which verbs seem to govern the dative case”. In this dissertation, I focus on how children approach this striking feature and learn to use it productively. But before turning to the question of how children approach Icelandic datives, it is necessary to understand how linguists have done so. The study of Icelandic datives and case in general spans decades of work from various different theoretical perspectives. When reviewing this work, it quickly becomes clear that the Icelandic case marking system is and has been informative for broader questions about the nature of case and its structural position, as well as questions about productivity, variation and change in language. One of the aims of this dissertation is to show that Icelandic case can also be informative about fundamental questions in the field of language acquisition, and how the learnability perspective is in turn important to the understanding of variation and change, as well as providing valuable insights into theoretical approaches to case. Therefore, the overview of the core data and previous research provided in this chapter will be done from a learnability standpoint. I start with a brief descriptive glance at the expression of the Icelandic case marking system on the surface before discussing relevant aspects of previous research on the distribution, structural status and semantics of the dative. I conclude with an overview of previous work on case marking in language acquisition and language variation and change.

## 1.1 The Icelandic case marking system at a glance

### 1.1.1 On the surface

Icelandic is characterized by a rich morphology despite a relatively rigid word order, a fact which in itself is a puzzle considering the cross-linguistic tendency to mark argument

function with either one or the other (Fedzechkina et al. 2016).<sup>4</sup> It is a nominative-accusative language with four morphological cases: nominative, accusative, dative and genitive (Thráinsson 1995). In contrast with most other Germanic languages, the morphological expression of case in Icelandic has remained robust (see McFadden 2020 for an overview). The case marking usually takes the form of inflectional suffixes which also can encode information about number, gender and inflectional class. As opposed to various other Germanic languages, which show impoverished or vestigial morphological case (such as English, Danish or Norwegian, see Parrott 2009), case distinctions are still apparent in all the major constituents of the Icelandic noun (or determiner) phrase. Since Icelandic has rich case agreement and nominal concord<sup>5</sup> as well, this can result in abundantly case-marked arguments such as in (2), where all three components of the object DP (the demonstrative pronoun, adjective and noun) appear with inflectional suffixes containing information about the case marking.

- (2) Við þurftum að ýta hin-um grá-a bíl-num.  
 We needed to push other.DAT gray.DAT car.the.DAT  
 ‘We had to push the other gray car.’

Still, the nature of this information differs. While the demonstrative pronoun and suffixed definite article both take forms which are unambiguously dative, the adjective takes a weak form which shows syncretism in all the oblique (or non-nominative) cases, as shown in the first inflectional paradigm in (3). To complicate things further, the adjective also can take a strong declension form if the noun is indefinite,<sup>6</sup> changing the inflectional paradigm significantly, as shown in the second paradigm in (4). With a weakly inflected adjective and a definite (singular) noun (*grái bíllinn*), the noun has different forms for each case but the adjective shows accusative-dative-genitive (oblique) syncretism, a common pattern in the Icelandic inflectional system. With a strongly inflected adjective and indefinite (singular) noun (*grár bíll*), on the other hand, the adjective differentiates

<sup>4</sup> See Kiparsky (1997) for the relationship between word order freedom and case morphology and Thráinsson (2007) for a discussion on the myth of a free Icelandic word order.

<sup>5</sup> See Norris (2014) for an extensive overview of nominal concord and Ingason & E.F. Sigurðsson (2017) as well as E.F. Sigurðsson (2017) for an account of nominal concord in Icelandic.

<sup>6</sup> See Pfaff (2017) for the marginal *grár bíllinn* (weak adjective declension but definite noun) order.

between the four cases while the noun shows accusative-dative syncretism, another common pattern in Icelandic.

(3)	case	definite noun	indefinite noun
	<b>nominative</b>	grá-i bÍl-l-inn	grá-r bÍl-l
	<b>accusative</b>	grá-a bÍl-inn	grá-an bÍl
	<b>dative</b>	grá-a bÍl-num	grá-um bÍl
	<b>genitive</b>	grá-a bÍl-s-ins	grá-s bÍl-s

Interestingly, both DPs in (3) are unambiguous throughout the paradigm as a whole, but this is not always the case. A similar paradigm with the DP *gláða barnið/glatt barn* ‘the/a happy baby’ would yield the same forms in the nominative and accusative for both the adjective and noun, as shown in (4).

(4)	case	definite noun	indefinite noun
	<b>nominative</b>	gláð-a barn-ið	glat-t barn
	<b>accusative</b>	gláð-a barn-ið	glat-t barn
	<b>dative</b>	gláð-a barn-i-nu	glöð-u barn-i
	<b>genitive</b>	gláð-a barn-s-ins	gláð-s barn-s

In DPs with only one constituent, the rate of syncretic/ambiguous form is of course inherently higher.<sup>7</sup> The most common syncretism patterns in Icelandic can be found in Table 1, including no syncretism with nouns such as *hestur* ‘horse’, with a strong masculine declension paradigm. In the present context, perhaps the most important aspect to note is the fact that nominative-dative syncretism never occurs unless there is also syncretism with the accusative, while nominative-accusative syncretism excluding

<sup>7</sup> This has never been quantified, but see Svavarsdóttir (1993) for the frequencies of different endings in a corpus of Modern Icelandic.

the dative and genitive is possible.<sup>8</sup> This means that to minimize ambiguity between nominative and non-nominative, the choice of dative over accusative provides a more salient contrast.<sup>9</sup>

**Table 1.** Examples for possible syncretism patterns in Icelandic noun declension. Syncretism shown with bold lettering and color patterns.

nominative	accusative	dative	genitive
tal-a 'number'	<b>töl-u</b>	<b>töl-u</b>	<b>töl-u</b>
<b>rós 'rose'</b>	<b>rós</b>	<b>rós</b>	rós-ar
<b>hús 'house'</b>	<b>hús</b>	hús-i	hús-s
drottning 'queen'	<b>drottning-u</b>	<b>drottning-u</b>	drottning-ar
hest-ur 'horse'	hest	hest-i	hest-s
<b>auga 'eye'</b>	<b>auga</b>	<b>auga</b>	<b>auga</b>

Although the complexities of Icelandic inflectional paradigms are not the topic of this dissertation, this brief description is relevant when considering the learnability of the case marking system, as it shows that despite a rich morphological system, syncretism is present and the case of arguments is therefore not always unambiguously expressed. For this reason, the acquisition of case marking is relevant within research on how children acquire categories which are expressed variably and sometimes ambiguously (e.g. Miller & Schmitt 2012).

<sup>8</sup> Table 1 does not include the pattern found exclusively in the words *kýr* 'cow' and *æf* 'sheep', which consists in nominative-genitive and accusative-dative syncretism.

<sup>9</sup> See Harðarson (2016) and Starke (2017) for an extensive discussion of syncretism in the context of case contiguity (Caha 2009).

### 1.1.2 Distribution and status

Regardless of how case is expressed, it is clear that in comparison to other Germanic languages, or even Slavic languages, Icelandic has an abundance of dative case marking (Maling 2002). This abundance is found both in subject and object case, where datives appear in addition to the most frequent (or default/structural) nominative for subjects and accusative for objects. Note that less frequent patterns of case marking also appear with Icelandic verbs, accusative subjects (e.g. with verbs such as *langa* ‘want’) and genitive objects for example (most commonly with *sakna* ‘miss’). 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 60.6% (289/477) of monotransitive verbs take an accusative object, 28.3% (135/477) require a dative object, 7.1% (34/477) appear both with accusative and dative objects and 4% (19/477) appear with genitive objects. For ditransitives, the majority of verbs, 72.4% (113/156), require the indirect object to be dative. The type frequency of verbs depending on object case is summarized in Table 2.

**Table 2.** Distribution of object case with transitive verbs, based on Jónsson’s (2005) representative lists for Modern Icelandic.<sup>10</sup>

Monotransitive: Accusative object	289
Monotransitive: Dative object	135
Monotransitive: Both ACC. and DAT. possible	34
Monotransitive: Genitive object	19
<b>Total monotransitives</b>	<b>477</b>

<sup>10</sup> With longer verb lists, a higher rate of the default patterns is expected.

Ditransitive <sup>11</sup> : DAT.-ACC. objects	87
Ditransitive: DAT.-DAT. objects	12
Ditransitive: DAT.-GEN. objects	14
Ditransitive: ACC.-ACC. objects	2
Ditransitive: ACC.-DAT. objects	34
Ditransitive: ACC.-GEN. objects	7
<b>Total ditransitives</b>	<b>156</b>
<b>Total verbs</b>	<b>633</b>

Dative subjects, on the other hand, 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 a closer look is taken at non-nominative (oblique) subjects only, and verbs requiring nominative subjects excluded, it can be seen that dative is the most common oblique subject case. In Jónsson's (2005) list, 301 verbs/predicates with non-nominative subjects are listed, 81.1% of them being dative. Of course, as is shown in Chapter 4, only a fragment of these predicates appear in child language, where the basis of the adult case marking system is formed.

Although oblique subjects have been argued to be present in a number of Indo-European languages (see e.g. Butt & King 2004 for Hindi/Urdu and Barðdal et al. 2012), the Icelandic oblique subjects have received an enormous amount of coverage in the linguistic literature over the past 45 years and played an important role in various formal approaches to case, some of them extremely influential (e.g. Andrews 1976, Thráinsson 1979, Zaenen et al. 1985, Yip et al. 1987, H.Á. Sigurðsson 1989, Marantz 1991, Harley 1995, Jónsson 1996, Kiparsky 1997, Baker 1997, Blake 2001, Schütze 2001, Svenonius 2002, McFadden 2004, Eythórsson & Barðdal 2005, Butt 2006, Woolford 2006, Legate 2008, Bobaljik 2008, Barðdal 2011, H.Á. Sigurðsson 2012, Preminger 2014, Anagnostopoulou & Sevdali 2015, Baker 2015, Wood 2015, E.F. Sigurðsson 2017,

<sup>11</sup> In contrast to Old Icelandic, the position of dative arguments is much more rigid in Modern Icelandic (see Viðarsson 2022), with the exception of so-called alternating predicates (Barðdal et al. 2014).

McFadden 2018, Schätzle 2018, Wood et al. 2020, Preminger 2020, Yuan 2022). This is in part due to the fact that contributions by Andrews (1976), Thráinsson (1979), Zaenen et al. (1985) and Yip et al. (1987) became known as the first conclusive demonstration of the subjecthood of non-nominative arguments in subject position, with Icelandic therefore differentiating itself from languages such as German, where comparable arguments did not fulfill subjecthood criteria.<sup>12</sup> More typologically unusual, as pointed out by Maling (2001, 2002), is the prominence and productivity of the dative marking on direct objects, particularly theme arguments of motion verbs (see example (5c)), which has received some (Svenonius 2002, Barðdal 2008, Jónsson 2005, 2013) but still much less attention than the oblique subjects. Examples of datives in subject and object position with frequent verbs can be found in (5).

- (5) a. **Mér** **finnst** gaman að baka brauð.  
 me.DAT finds fun to bake bread.ACC  
 'I like baking bread.'
- b. **Ég** **sagði þér** það.  
 I.NOM said you.DAT it.ACC  
 'I told you so.'
- c. **Þær** **náðu** **lestinni**.  
 they.NOM reached train.the.DAT  
 'They got the train.'
- d. Hann **fékk sér** ekki köku.  
 he.NOM got self.DAT not cake.ACC  
 'He didn't have cake/get himself cake.'

(5a) is an example of *finnst* 'find', the most frequent dative subject verb in the IceCASE corpus used in Chapter 4. It contains the morpheme *-st* which appears with a few other psych/experiencer raising verbs as well as various other types of constructions such as reciprocals and reflexives. Anderson (1990) and Wood (2015) provide comprehensive

---

<sup>12</sup> But see Barðdal & Eythórsson (2003) and Eythórsson & Barðdal (2005) for the view that oblique subjects are a common Germanic inheritance and that German has oblique subjects (and then Schätzle (2018) for a critique).

overviews of *-st* and Schätzle (2018) accounts for the diachronic lexicalization process of former middle verbs as stative experiencer predicates with a dative subject. (5b) has the most common ditransitive in the IceCASE corpus, *segja* ‘say, tell’, which appears with a dative indirect object. (5c) is an example of the dative direct object which has the highest frequency in the same corpus, with the verb *ná* ‘reach, get’, and finally, (5d) includes an obligatorily reflexive dative applicative argument representing an optional benefactive (Wood 2016) which appears with a few verbs, most frequently *fá* ‘get’.

Building on a long generative tradition (which can be traced back to Rouveret & Vergnaud (1980) and Chomsky (1981) and was developed for Icelandic by Yip et al. (1987)), recent theoretical accounts of case which build on Icelandic (Wood (2015) and E.F. Sigurðsson (2017) are representative in this respect, but see Schätzle (2018) and Galbraith (2018) for a Lexical-Functional Grammar approach), assume a distinction between structural and non-structural case, with the exception of Barðdal’s (2008, 2011) usage-based account of case in Icelandic. Nominative on subjects and accusative on objects, the ‘defaults’ which also are distributionally dominant, are then considered to be instances of structural case, while other case marking, and therefore all dative case marking, is non-structural. Although this has not always been the case, current criteria for this distinction are purely based on structure/form and not meaning. Structural case, which has always been considered independent of thematic roles (see Butt 2006), is considered by E.F. Sigurðsson (2017) to be assigned in syntax and then translated into either nominative or accusative in the morphology, “depending on whether the DP in question is the highest structurally case marked DP in its domain or not” (p. 24).

Under most approaches, non-structural case is then divided further into inherent/thematic/semantic case and idiosyncratic/quirky case (see Zaenen et al. 1985, Jónsson 2003 and Woolford 2006, in addition to Wood 2015 and E.F. Sigurðsson 2017). Semantic predictability was the original motivation for this distinction, with inherent/thematic/semantic case considered to be in direct relationship with the role of goal or experiencer (as in examples (5a–b)) while idiosyncratic/quirky case was considered truly lexical, as in the case of accusative subjects or direct dative objects as in (5c) (Jónsson 2003 and Woolford 2006). The most recent versions of this division into two types of non-structural case are also mostly structural in nature but still depend, at

least in part, on semantic predictability. In E.F. Sigurðsson (2017) approach, inherent case is assigned via a structure-building feature within e.g. an applicative phrase (5a, b and d), while quirky case is assigned via a probe feature, usually on Voice (placed on little *v* or *p* in Wood 2015), as would happen in (5c). This is partly motivated by the predictability of the case on applicatives, which are usually dative and consist of goals, experiencers and ‘affected’ roles such as benefactives and malefactives (Wood 2015), meaning they prototypically appear as animate entities. The (less established) predictability of dative themes of motion verbs is not incorporated into the inherent/quirky distinction but could be argued to still be structurally accounted for in Wood (2015) and E.F. Sigurðsson (2017), as these datives are related to either a *pP* or *PathP* as opposed to the verbal root.<sup>13</sup> These distinctions between structural and non-structural case are further reinforced in various contexts, such as preservation of non-structural case under different types of passivization (E.F. Sigurðsson (2017) for an overview and Barðdal (2011) for an account where the case preservation argument is questioned). Dative is occasionally not preserved under passivization, which under some analyses is taken to mean that dative can in fact be (partly) structural (Woolford 2006, Alexiadou et al. 2014).

In the next subchapter, I further illustrate how case is associated with meaning in Icelandic and speculate on the possible semantic case patterns and generalizations available to children. I focus on case-marking in canonical contexts and therefore exclude e.g. dative subjects in the passive. I do this based on evidence that children ‘filter’ out movement to some extent in the acquisition of argument structure (Perkins 2019). Note that a few very simple patterns in form have already emerged from this subchapter: Subjects are mostly nominative, direct objects are most frequently accusative and indirect objects (and optional benefactive reflexives, as in (5d)) are most often dative. From a learnability standpoint, these general but non-deterministic patterns which are available in the input could provide a starting point for the distributional acquisition of case marking. Additionally, the abundance of non-default (dative) forms in the Icelandic case marking system provides an interesting test case for various approaches to abstraction, rule

---

<sup>13</sup> Here it is worth noting that Icelandic has also featured prominently in the debate concerning the status of case as a syntactic or morphological phenomenon (e.g. Marantz 1991, Harley 1995, McFadden 2004, Caha 2009), with McFadden’s (2021) most recent account consisting of mapping the notion of case in the syntax on structural case and case in morphology on inherent case.

formation, productivity and overgeneralizations in acquisition (such as Pinker 1999, O'Donnell 2015, Yang 2016, Blything et al. 2018, Ambridge et al. 2018, Goldberg 2019, Ambridge 2020, Schuler et al. 2021), topics which have very much been dominated by research on the acquisition of English (e.g. the past tense, datives and the double object construction, a-adjectives and periphrastic causatives). A contrast from a related, morphologically rich language with relatively rigid word order and robust non-default patterns in case marking, Icelandic, could therefore provide valuable testing grounds. The semantic dimension to the case marking, described in subsection 1.1.3, arguably makes the contrast more complex and interesting.

### 1.1.3 Semantics

The examples in (6–8) illustrate this semantic dimension to case marking in Icelandic quite well, as they all include predicates in which case alternations are possible, creating minimal pairs in which the semantic characteristics of the dative become clear. Barðdal (2011) observes that this could be considered to be remnant differential object marking (where dative is used to mark e.g. animacy).

- (6) a. Stelpan        **skaut boltann/boltanum.**  
 girl.the.NOM shot ball.the.ACC/DAT  
 'The girl shot the ball.'
- b. Þau            **sópuðu**        **allt/öllu.**  
 they.NOM swept all.ACC/DAT  
 'They swept it all.'
- (7) a. Ég        **klóraði**        **hann/honum.**  
 I.NOM scratched him.ACC/DAT  
 'I scratched him.'
- b. Hann        **þvoði**        **það/því.**  
 he.NOM washed it.ACC/DAT  
 'He washed it.'

- (8) a. **Hún/henni**            **gengur**        vel.  
 she.NOM/DAT            goes/does    well.  
 'She is doing well/it is going well.'
- b. **Skrímlið/skrímslinu**        **er**        **kalt.**  
 monster.the.NOM/DAT        is        cold.  
 'The monster is cold.'

Examples (6a) and (6b) show an alternation between an affected patient and a motion theme. When shooting (*skjóta*) a ball in the accusative, the ball itself does not move and is the patient of the shot (with an arrow or a bullet for example), but when shooting a ball in the dative, the ball itself is the motion theme. When sweeping something in the accusative, the object refers to the surface being swept while sweeping something in the dative refers to the entity that is being swept away (e.g. dust). Both (6a) and (6b) would therefore be considered quirky in E.F. Sigurðsson's (2017) approach, with the dative being assigned via PathP or pP. Examples (7a) and (7b) show an alternation between a negatively affected patient and a benefactive (applicative). Scratching (*klóra*) in the accusative is interpreted as a (harmful) change-of-state for a patient which could be either inanimate (e.g. a couch) or animate (e.g. a person), with no animacy requirement, whereas dative scratches have (sentient) benefactives. Similarly, when inanimate entities or body parts and hair are washed, accusative can only be used, since the dative implies washing a sentient being. In the case of (7b), the neuter pronoun in the accusative could therefore refer to hair (neuter noun) but not a child (also neuter), in which case the dative would be appropriate. Finally, (8a) and (8b) contain alternations between themes and experiencers (also applicatives). When using the verb *ganga* 'go, walk', as in (8a), to describe the course of a study (feminine noun) led by a female researcher, a nominative pronoun would point to the study itself but a dative pronoun has an animacy requirement and therefore points to the researcher conducting the study. In (8b), the nominative has no animacy requirement and indicates that the monster is cold to the touch, without necessarily itself feeling the cold. The dative, on the other hand, requires a sentient 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 an apparent

minimal pair when the subject is in the neuter singular and the agreement on the adjective therefore is syncretic to the non-agreement present with non-nominative subjects.<sup>14</sup>

To summarize, the semantic minimal pairs in (6-8) show how the dative can indicate motion (perhaps licensed by a pP or PathP) or an applicative goal or experiencer meaning, where sentience is required. This is contrasted with nominative subject themes and accusative object patients, but it is important to note that dative case marking also appears on arguments which are neither motion themes nor applicatives. For example, various verbs of destruction and manner of killing appear with dative objects (such as *slátra* ‘slaughter’, *rústa* ‘ruin, destroy’ (which also appears with the accusative) and *splundra* ‘shatter’). Other examples which do not fit the broad semantic patterns just described include *gleyma* ‘forget’ and *loka* ‘close’, which have dative objects that are neither applicatives or motion themes. Icelandic therefore seems to have, in contrast with e.g. German, “truly direct object datives” (Wood 2015, p. 129) which are not derived from the semantic associations just described (or the structural distributions discussed in subsection 1.1.2).

These exceptions of course do not erase the patterns just described, which also have been highlighted in the literature (e.g. Van Valin 1991, Jónsson 1997–1998, 2000, 2003, 2005 and 2013, Maling 2002, Svenonius 2002, Thráinsson 2007, Barðdal 2008). These patterns, as is further discussed in Chapter 2, are particularly interesting in the context of how children can use form to acquire meaning and ongoing debates about the use of possibly universal cues (e.g. number of arguments) and/or language-specific cues (e.g. morphological case) in this respect (Lidz et al. 2003, Göksun et al. 2008, Matsuo et al. 2012, Leischner et al. 2016). In general, agents are consistently nominative, non-nominative subjects tend to be experiencers (but not all experiencers are non-nominative), patients are mostly accusative (but also appear in the dative), recipients (animate goals) are almost consistently dative and themes of motion verbs (ballistic motion in particular) are regularly dative. But as McFadden (2020) describes, despite the clear semantic relationship between e.g. dative and recipients/beneficiaries/goals in

---

<sup>14</sup> If a masculine noun such as *maður* ‘man’ would be used in (8b), the dative subject construction would contain the adjective in the neuter third person (*kalt*) while the nominative subject construction would yield gender (and number) agreement and the masculine form of the adjective (*kaldur*).

Germanic languages with a solid dative-accusative distinction in general, "the correlations between cases and semantic or thematic roles are notoriously approximate" (p. 295), with pairs of verbs with close semantics but different case frames being taken "to show that the relationship between cases and semantic/thematic roles cannot be direct" (p. 295). This is emphasized by e.g. H.Á. Sigurðsson (2012, p. 324), who uses Icelandic verb pairs such as *aðstoða* 'assist' (accusative object) and *hjálpa* 'help' (dative object) to argue against a relationship between case and semantics.

Maling (2002) furthermore states that although "the correlation between thematic role and morphological case is not one-to-one, there is clear evidence of semantic generalizations at work in the language" (p. 96). In a similar vein, Thráinsson (2007) observes that "although it is possible to find some relationship between thematic roles of arguments [...] and morphological case in Icelandic, the relationship is many-to-one and one-to-many" (p. 200). He adds that "the case-marking possibilities are to some extent restricted by the grammatical function (subject, object, indirect object) involved" (p. 200) and later that "no predictions can be made about the thematic role of a subject given its case" (p. 206). Similarly, as an argument against adopting Schäfer's (2008) account on Icelandic datives, Wood (2015) explains that "the semantic correlations are not exact" (p. 137) and in another instance states that "it remains true that one cannot predict from event semantics which case any given verb will assign, though there are some semantic regularities" (p. 135).<sup>15</sup> In the approach taken in this thesis, I make a case for exceptions: Semantic correlations do not need to be exact for the learner to discover and productively use the form-meaning mappings which have been observed by linguists.

To be clear, attempts have been made at formalizing case-semantics regularities under various approaches despite their dismissal in the aforementioned work. For datives in particular, Jónsson (2013) provides a more fine-grained analysis of the lexical semantics of dative motion themes (into e.g. emission verbs and verbs of ballistic motion as contrasted to accompanied and directed motion) and further defines the characteristics of non-nominative experiencers (Jónsson 2003) as compared to nominative experiencers, and Ingason (2010) sketches out an analysis where the subclass of

---

<sup>15</sup> Although the literature discussed here consists of recent approaches to case in Icelandic, the fact that the alignment of thematic roles and case/syntactic roles is notoriously approximate was highlighted much earlier (e.g. Rosen 1984).

experiencer subjects which experiences physical discomfort is associated with the accusative. More generally, in Barðdal's (2008, 2011) usage-based accounts, she proposes that the status of the nominative-accusative (nominative subjects and accusative objects) construction as a default for transitive predicates can be explained by the frequency (type and token) and high level schematicity of the case frame, as she demonstrates that the nominative-accusative construction has a much broader semantic range than the nominative-dative construction<sup>16</sup> and dismisses the dichotomy between structural and non-structural case. From a different perspective, recent accounts of case in Icelandic and Faroese within the LFG framework (Schätzle 2018 and Galbraith 2018), in continuation of Kiparsky's Linking Theory (1997, 2001) and Zaenen's Lexical Mapping Theory (1993, see also Kibort 2014), accentuate the role of event structure and lexical semantics in the licensing of case via binary abstract case features.

Schätzle's account (2018) builds in part on observations found in Svenonius' (2002) very influential aspectual account of Icelandic case,<sup>17</sup> where he argues that the difference between accusative and dative object case lies in the temporal alignment of the subevents encoded within the verb, with accusative appearing when the two subevents of a transitive verb are partly co-temporaneous, while dative occurs when the subevents can be separated. But as E.F. Sigurðsson (2017) notes, Svenonius' generalization, which is intended to hold across the data, accounts only for a subset of datives. E.F. Sigurðsson (2017) argues instead for a distinction between direct and indirect causation, following Kratzer's (2005) distinction between events of causing other events and events which lead directly to the caused event with no intermediate event. Thus, E.F. Sigurðsson (2017) generalization is not, contrary to Svenonius' (2002) generalization, assumed to hold across all the object case data.

The learnability approach taken here furthermore entails that the mappings formed by learners should be based on cognitively plausible event structure categories which are available early on. There is independent evidence for children's early sensitivity to agentivity (Rissman & Majid 2019), animacy (Opfer & Gelman 2011, see also Becker's

---

<sup>16</sup> See also Thráinsson's (2007) tables (p. 206 and 212) showing an overview of the semantic scope of different morphological subject and object case across thematic roles.

<sup>17</sup> Schätzle's implementation of the role of event structure in case marking still is rather based on Ramchand's (2008) event-decompositional framework of first-phase syntax.

(2014) overview) and (caused) motion (Kellman et al. 1987) but I have not found corresponding evidence for children's early sensitivity to the temporal alignment of subevents, with children in fact possibly acquiring various linguistic and cognitive properties of tense and aspect quite late (Sharon & Wynn 1998, Slabakova 2002, Van Hout 2005 and McCormack & Hoerl 2017, but see Wagner 2006 for the early acquisition of telicity). This might suggest that the correlation of aspect and case discussed by Svenonius (2002) does not entail causation, which does not exclude that the distinctions could come into play in the formation of specific subrules late in acquisition. This kind of evidence represents one way in which data from language acquisition and learning in general can inform formal case theory, which presumably aims to describe a system which originates in acquisition, and even provide criteria to evaluate them. But the main evidence which is imperative to take into account is of course data from the acquisition process of case itself, both the order of acquisition and the patterns which are abstracted, formed into productive rules and (over)generalized. Additionally, it is important to consider whether productivity in acquisition is maintained in adult language and appears with novel verbs and in known instances of variation, providing evidence for a continuity that would show how case marking in adult language is determined in acquisition. The acquisition of case marking rules in Icelandic, their productivity and the variation associated with them, are the topic of the next two subsections.

## 1.2 Acquisition and productivity

I begin by asking how children acquire the case marking system outlined in section 1.1 and what their developmental path can tell us about the nature of the system. Is the development of case in line with the distribution in form discussed in subsection 1.1.2, where subjects are mostly nominative, direct objects for the most part accusative and indirect objects dative? Do children acquire the most common distributional patterns first? What about the semantic patterns discussed in subsection 1.1.3? What rules could be considered productive based on the overgeneralizations found? Is this productivity maintained in adult language? Although these are the main questions of the current section, the acquisition of case in Icelandic is understudied, as has already been mentioned. Previous research, most prominently Sigurðardóttir's (2002) MA-thesis which

drew on experiments and corpora, has focused on production, showing that children acquire the structural/default/most frequent case marking first, that is nominative on subjects and accusative on objects. This happens early in acquisition, with the two-word stage already including case marked arguments (Sigurjónsdóttir 2005).<sup>18</sup> Dative objects appear from around age two, with indirect objects being acquired earlier, and dative/non-nominative subjects appearing 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 (direct and indirect) being frequently overgeneralized. This also fits quite well with the patterns described in subsection 1.1.2. But additionally, children acquiring Icelandic overgeneralize the dative (Sigurðardóttir 2002) in subject and object position – unlike German children (Clahsen et al. 1996, Schmitz 2006, Scherger 2018). In object case, dative arguments can appear instead of benefactive prepositional phrases or accusative objects with verbs such as *lesa* ‘read’, *kyssa* ‘kiss’ and *knúsa* ‘hug’ (see (9a, b and c)), and (seemingly a bit later) accusative motion themes of verbs such as *færa* ‘move’, *hrista* ‘shake’ and *grípa* ‘catch’ are replaced with the dative, see (9d, e and f). This can be seen in the examples in (9) which I have collected from various sources. Additionally, children have occasionally been reported to overgeneralize dative subjects with predicates such as *vera sveitt(ur)* ‘be sweaty’ (see experimental results from Sigurðardóttir (2002), based on an example reported by Gunnarsdóttir (1996)). Moreover, it has been established that children ‘overgeneralize’ the dative with psych verbs appearing with non-nominative subjects, but this could be interpreted as the acquisition of existing variation rather than overgeneralization (Nowenstein 2017).

- (9) a. Mamma                    **lesa**                    **mér.**  
          mommy.NOM            read.INF            me.DAT  
          ‘Mommy read to me.’  
          Child aged 2;4 (p.c.)

---

<sup>18</sup> But the arguments during the one-word stage mostly appear in the nominative.

- b. **Knúsa þér.**  
 hug.INF you.DAT  
 'Hug you.'  
 Child aged 3;6 (p.c.)
- c. Eva **kyssa mér.**  
 Eva.NOM kiss.INF me.DAT  
 'Eva kiss me.'  
 Child aged 2;0 (Sigurjónsdóttir longitudinal corpus (2007))
- d. Hún var að **færa bílnum mínum.**  
 she.NOM was to move.INF car.the.DAT my.DAT  
 'She just moved my car.'  
 Child aged 3;4 (Sigurjónsdóttir longitudinal corpus (2007))
- e. Já þú **grípa honum.**  
 yes you.NOM catch.INF him.DAT  
 'Yes you catch it.'  
 Child aged 3;0 (Einarsdóttir et al. corpus (2019))
- f. Maður á að **hrista fótunum.**  
 man.NOM ought to move.INF feet.the.DAT  
 'You should shake your feet.'  
 Child aged 7;6 (Einarsdóttir et al. corpus (2019))

If these examples of overgeneralization are interpreted in the context of the patterns described in section 1.1, it seems likely that (at least some) children acquiring Icelandic associate the dative with sentient benefactives/goals on the one hand (or positively affected patients for verbs like *kyssa* and *knúsa*, depending on the criteria for thematic roles) and with themes which undergo some kind of motion on the other hand. To explore the kind of productive rules (abstractions/patterns/schemas) which underlie overgeneralizations, the use of novel verb experimental paradigms is well established. Barðdal (2008) conducted such a study targeting Icelandic datives. 20 children aged 6–13 participated as well as 20 adults, producing sentences with five novel verbs which were presented along with existing Icelandic verbs with dative subjects and objects. Both groups produced datives (as well as the more frequent nominative for subjects and accusative for objects) with two experiencer subjects and one motion theme object. Only

one participant produced a dative object with a novel verb for which the dative had not been primed, which Barðdal interprets as an argument against semantic abstractions beyond the lexical level for dative subjects and objects (i.e. dative is assigned based on “synonymous verb-specific constructions” (p. 104)), while the nominative-accusative pattern is interpreted as a “high-level semantically-vacuous schematic construction” (p. 110). Such an approach does not account for the child-language overgeneralizations just described in (9), since the children produce dative despite the verb-specific accusative case marking found in the input. Indeed, this verb-specific approach to the dative also contrasts with the results from Barðdal’s (2008) more substantial study on recently borrowed verbs, where she assumes a higher level of schematicity (abstraction) for dative objects to account for the dative productivity she finds (36% of 107 borrowed verbs) which cannot be attributed to synonymous translational equivalents. These verbs are frequently associated with caused motion.

Jónsson and Thórarinsdóttir (2020) also use newly coined verbs as a way to assess the semantic predictability and productivity of the dative for object case in Icelandic, reaching the conclusion that a novel verb takes a dative direct object if it “(a) encodes motion of the object referent, or (b) has a translational substitute that takes a dative object” (p. 297). Despite varying approaches, it has therefore been well established that the dative is productive for direct objects of new verbs in Icelandic, particularly in the semantic context of caused motion. There is also consensus on the more restrictive nature of dative productivity as compared to the accusative for direct objects, and all accounts show that new verbs never appear with the much less frequent genitive objects. But what about the dative for indirect objects or subjects (in both contexts applicatives)? Interestingly, despite the notorious productivity of the dative in Icelandic and the frequency of dative for indirect objects, the goals of novel verbs of (technological) transfer like *ímeila* 'email' seem to appear more often as prepositional phrases rather than indirect objects, so (10a) would be preferred to (10b).

- (10) a.      Ég      **(i)meila**      **þetta/þessu** til      þín.  
           I.NOM (e)mail      this.ACC/DAT to      you.GEN  
           ‘I’ll (e-)mail this to you.’

- b.    ?Ég   **(i)meila**    **þér**            **þetta/þessu.**<sup>19</sup>  
       I.NOM (e)mail    you.DAT        this.ACC/DAT  
       'I'll (e-)mail you this.'

This is in line with results from a questionnaire conducted by Barðdal (2008), where the option with the indirect dative object received the lowest rating. Notice that there is variation between accusative and dative in the direct object case, with both cases being used by speakers, which also appears in Jónsson and Thórarinsdóttir's (2020) results. This possible lack of productivity of the double object construction in Icelandic presents an interesting contrast to its well-known productivity in English (see overview in Yang 2016) and creates an Icelandic twist to Baker's (1979) paradox, where children acquire the productivity of the double object construction despite known exceptions (e.g. *text someone an address* is fine but *donate the museum a painting* is not).

When it comes to dative subjects, the data on recent borrowings seems in line with the overgeneralizations in acquisition, dative is very rarely extended to the subjects of new verbs. To explain this, Barðdal (2008) points to their low frequency and narrow semantic restrictions. A recent example possibly illustrates this: the Icelandic version of Facebook makes use of the well established Icelandic dative subject taking verb *líka* 'like' as a translation for the English *like*. This has not been widely adopted by the speaker community, with the loan verb *læka* 'like on Facebook or other social media' emerging instead with a nominative subject. If the originally Icelandic *líka* 'like' is used, it also appears with a nominative subject (Ingason 2016). This is presumably related to the fact that liking something on social media often involves a more agentive meaning (with the 'hitting' of a button) than in the traditional use of the verb, therefore ruling out the use of an oblique subject. But even loan verbs with semantics which are very much compatible with known oblique subject verbs cannot seem to consistently appear in the dative. The only attested examples (Guðmundsdóttir et al. 2018) of oblique case appearing with a novel verb were found with the verb *kreiva/kreifa*, an adaptation of the English *crave* (11), but this verb predominantly appears in the nominative.

---

<sup>19</sup> Anton Karl Ingason has pointed out to me that the double-object construction potentially works better with a heavier direct object.

- (11) a. **mig**            **kreifar**            **í**            **ávexti.**  
 me.ACC            craves            in            fruit  
 'I crave fruit.'
- b.    **so**    **kreifar**            **honum**            **alltaf**            **í**            **eitthvað.**  
 then craves            him.DAT            always            in            something  
 'Then he always craves something.'

It is interesting that example (11a) shows an accusative subject in the first person singular, and (11b) shows a dative pronoun (third person singular) as a subject. This is in line with the so-called Person-Specific Retention in Icelandic subject case variation (Nowenstein & Ingason 2021) discussed in the next section, where accusative is more likely to appear with subjects in the first and second person. Indeed, even though new verbs generally do not appear with non-nominative subjects, Icelandic subject case variation is a context where the productivity of the dative is very well-documented. This variation is the topic of the next subsection.

## 1.3 Variation and change

As has been mentioned, the case marking system Icelandic-speaking children need to acquire displays variation. Indeed, variation in subject case is one of the most researched topics of morphosyntactic change in Insular Scandinavian (e.g. Jónsson & Eythórsson 2005). Variation in object case also exists but has not been extensively documented (but see e.g. Maling 2002 and Eythórsson & Thráinsson 2017) and is consequently less addressed within this project.

### 1.3.1 Dative Substitution (DS)

The most common form of variation in subject case in Icelandic is Dative Substitution (DS), also known as Dative Sickness (*'þágufallssýki'*) in the prescriptivist discourse. It is a relatively stable but stigmatized variant which spread at the end of the 19th century (but see Viðarsson 2022 for examples from Old Icelandic) and has been extensively studied

diachronically and in a series of large surveys (e.g. Svavarsdóttir 1982, Jónsson 1997–1998, Jónsson & Eythórsson 2003, Jónsson & Eythórsson 2005, Barðdal 2011, Thráinsson 2013, Nowenstein 2017 and Schätzle 2018). Studies furthermore indicate that children’s rate of DS is associated with their parents’ socio-economic status (Svavarsdóttir 1982 and Jónsson & Eythórsson 2005). The change consists in the dative being used instead of the original subject case for a number of psych verbs (12):

- (12) a. **Hana** langar í epli. → **Henni** langar í epli.  
her.ACC wants in apple. her.DAT wants in apple.  
‘She wants an apple.’
- b. **Ég** hlakka til. → **Mig/mér** hlakkar til.  
I.NOM look.forward to. me.ACC/DAT look.forward to.  
‘I look forward to (it).’

Most verbs with which DS occurs have traditionally accusative marked experiencer subjects, but two common DS verbs (*hlakka* ‘look forward to’ (see 12b) and *kvíða* ‘be anxious about’) also appear in the nominative, which is by far the most frequent case for subjects in Icelandic, as is outlined in subsection 1.1.2 (94% of tokens in Barðdal 2001). Although originally dative subjects outnumber accusative subjects, both are marked in contrast with the default (or structural) nominative. These oblique subjects have certain characteristics: In addition to not triggering subject-verb agreement (the verb is in the default third person singular form), non-nominative subjects cannot be agents. But even though all agents are nominative, this does not mean that all experiencer subjects are non-nominative. In a leveling situation, dative therefore also competes with the nominative default option. This appears clearly in children’s nominative overgeneralizations of oblique subjects mentioned earlier (Sigurðardóttir 2002, Nowenstein 2017) as well as in the coinage of new psych verbs, with only one attested example of non-nominative subject case marking with a novel verb as was mentioned in section 1.2 (Guðmundsdóttir et al. 2018). Nonetheless, children do acquire DS (Sigurðardóttir 2002, Nowenstein 2017) and associate dative subjects more than nominative ones with experiencers in novel verb tasks as is shown in Chapter 2 (Nowenstein et al. 2020).

The large-scale surveys mentioned above (Svavarsdóttir 1982, Jónsson & Eythórsson 2005 and Thráinsson 2013) furthermore indicate both incrementation and age-grading. The same four DS verbs were tested in a forced-choice task administered to 10–11-year-olds in 1980–1981 (Svavarsdóttir 1982, N = 202) and 2001 (Jónsson & Eythórsson 2005, N = 845) and to 14–15-year-olds in 2006–2007 (Svavarsdóttir 2013, Thráinsson 2013). For these four verbs, the rate of DS increases from 27.4% to 33.9% between the results of Svavarsdóttir (1982) and Jónsson & Eythórsson (2005) but the DS rate then drops to 20% in the study with 14–15-year-olds, which was administered 5 years later (see Svavarsdóttir 2013) with older participants. This is in line with the Hall & Maddeaux (2020) “two steps forward, one step back” approach in which children innovate and then retract (resulting in age-grading) but still contribute to incrementation.

More recently, it has been emphasized that intra-speaker variation in subject case marking is widespread and grammatically conditioned (Nowenstein 2012, 2014 and 2017, Svavarsdóttir 2013, Ingason 2015, Nowenstein & Ingason 2021). As is discussed in greater detail in Chapter 5, DS is not only specialized (Wallenberg 2019) along a stylistic dimension (s-conditioning) but also along an internal (grammatical) dimension (i-conditioning) (Tamminga et al. 2016). In Chapter 5, I focus on the acquisition of two instances of conditioning along the internal grammatical dimension: the Person-Specific Retention (PSR) and syncretism effects (see Ingason 2015 for prosodic constraints). The PSR is now well known and has been investigated in data from Icelandic adults and children as well as speakers of heritage North-American Icelandic (Nowenstein 2017). It involves dative being selected more often when the subject is in the third person in comparison to the first and second person (a number effect might be present as well as a smaller contrast between first and second person of the singular, see Nowenstein 2017 and Nowenstein & Ingason 2021). This has often been attributed to prescriptivism and self-correction but intra-speaker variation which is conditioned in this particular way can also be found in informal contexts and child language (Nowenstein 2017).<sup>20</sup> The second type of conditioning investigated here has not been confirmed before; the dative is

---

<sup>20</sup> It could additionally be interpreted in the context of an animacy hierarchy, with some languages treating first and second person as more animate than third person (Becker 2014, p. 64). This is particularly interesting considering the fact that DS verbs are psych verbs and therefore require their subjects to be animate and sentient.

hypothesized to be selected more often when the nominative and accusative are syncretic. This is shown in Table 3, with two inflectional paradigms which were used as stimuli in the study described in Chapter 5. *Stelpurnar* ‘the girls’ has syncretic nominative and accusative forms, while *strákarnir* ‘the boys’ has distinct forms for each case. Here it is important to note again that accusative-dative syncretism is common in Icelandic, but the nominative and dative never pattern together without the syncretism including accusative also, as already described in subsection 1.1.1. We could therefore interpret possible syncretism effects as a preference for contrast saliency when the input is variable, because the nominative-accusative syncretism entails a less clear signal for an oblique subject construction.

**Table 3.** Nominative-accusative syncretism and lack of syncretism in Icelandic subject case and possible relationship to Dative Substitution rate.

	<b>syncretic</b>	<b>non-syncretic</b>
<b>nominative</b>	<b>stelpur-nar</b> girls.the.NOM/ACC	strákar-nir boys.the.NOM
<b>accusative</b>	<b>stelpur-nar</b> girls.the.NOM/ACC	stráka-na boys.the.ACC
<b>dative</b>	stelpu-num girls.the.DAT	stráku-num boys.the.DAT
<b>translation</b>	the girls	the boys
<b>DS rate</b>	→ more DS	→ less DS

Under such an interpretation, the syncretism effects might be the result of processing pressure in variation,<sup>21</sup> while the PSR would be an example of how inconsistent input is regularized and mapped onto existing features. Still, the two types of conditioning might be intertwined, as the syncretism patterns in the Icelandic pronominal system (Table 4) match the PSR, with the first and second person singular showing no syncretism in the singular and only accusative-dative syncretism in the plural, while nominative-accusative syncretism appears in the third person pronouns.

**Table 4.** Syncretism patterns in the Icelandic pronominal system.

person and number	NOM.-ACC.-DAT. forms	syncretism
<b>1.p.sg.</b>	ég-mig-mér 'I/me'	none
<b>2.p.sg.</b>	þú-þig-þér 'you'	none
<b>3.p.sg.fem.</b>	hún-hana-henni 'she/her'	none
<b>3.p.sg.masc.</b>	<b>hann-hann-honum</b> 'he/him'	NOM.-ACC.
<b>3.p.sg.neut.</b>	<b>það-það- því</b> 'it'	NOM.-ACC.
<b>1.p.pl.</b>	við- <b>okkur-okkur</b> 'we/us'	ACC.-DAT
<b>2.p.pl.</b>	þið- <b>ykkur-ykkur</b> 'you'	ACC.-DAT
<b>3.p.pl.fem.</b>	<b>þær-þær</b> -þeim 'they/them' (feminine)	NOM.-ACC.
<b>3.p.pl.masc.</b>	þeir-þá-þeim 'they/them' (masculine)	none
<b>3.p.pl.neut</b>	<b>þau-þau</b> -þeim 'they/them' (neuter)	NOM.-ACC.

If syncretism could be a drive for leveling, where more saliently oblique forms (no syncretism with nominative) are preferred, this would reinforce the observed PSR and help establish it.

<sup>21</sup> In this context, Bornkessel-Schlesewsky et al.'s (2020) results on DS are interesting, since they show that "neurophysiological responses of native speakers already reflect projected changes that are not yet apparent in their overt behavior" (p. 1).

### 1.3.2 Nominative Substitution and relationship to DS

In addition to DS, the variation in Icelandic subject case also appears with the less discussed *Nominative Substitution* (NS) of theme subjects. NS comprises a change from an oblique subject case (accusative or dative) on themes (often inanimate) to nominative (Jónsson & Eythórsson 2005), see (13) in contrast to the DS in (12).

- (13) **Bátinn rak á land** → **Báturinn rak á land.**  
boat.the.ACC drifted to shore → boat.the.NOM drifted to shore  
'The boat drifted to shore.'

DS and NS have traditionally been viewed as an example of overgeneralization/leveling where productive, unmarked patterns are generalized at the expense of less productive, lexically specific and more marked patterns. Within approaches which link case to thematic roles, as discussed in subsection 1.1.3, patterns between datives and argument role in Icelandic have been used to argue for a division of non-structural case into inherent (predictable but not structural) and quirky/lexical (unpredictable) case (e.g. Yip et al. 1987, Jónsson 2003 and Woolford 2006). In this context, dative subjects are associated with experiencers and not themes. Jónsson & Eythórsson (2003, 2005) and others have based their explanations on the differences between NS and DS on such assumptions, arguing that oblique case on theme subjects (as in (13)) in Icelandic is always quirky, based on lack of predictability. Dative subjects of experiencer/psych verbs (as in (12)) are on the other hand assumed to be both more predictable and regular, and their productivity is apparent in DS. Still, dative is not the default structural case for experiencer subjects (whereas the nominative is). Instead, it can be considered to be the inherent case, too predictable to be quirky but regular enough to substitute quirky accusative case with experiencers. Previous ideas (e.g. Jónsson 2003, Jónsson & Eythórsson 2005) about changes in subject case marking contrasting DS and NS in this way are summarized below in (14):

- (14) a. **Nominative Substitution (themes)**  
Lexical ACC./DAT. (quirky) case → Structural case. **Dative is not productive.**  
b. **Dative Substitution (experiencers)**  
Lexical ACC. (quirky) case → Inherent case. **Dative is productive.**

Indeed, although the subjects of theme verbs, just like the subjects of experiencer verbs, are historically both accusative and dative, (14) represents the observation that dative fails to attract the accusative for themes in the same way that dative experiencer subjects do. In fact, it has been maintained that such patterns are impossible, since the dative fails to acquire the status of inherent case with theme verbs (Jónsson 2003, Jónsson & Eythórsson 2005). This difference in the directionality of change between themes and experiencers is furthermore predicted by the application of Yang's Tolerance Principle (2016). As has been described in the Introduction to this thesis (see example (1)), the Tolerance Principle accounts for the productivity of rules or patterns. A rule is productive if the maximum number of exceptions ( $\theta_N$ ) under the Tolerance Principle, calculated through the natural log of the total number of types ( $N/\ln(N)$  where N is the full set of items the rule could apply to), is not exceeded by the exceptions (e). Therefore, this is not a proportional threshold and smaller sets in fact allow for a bigger proportion of exceptions. By applying the Tolerance Principle to the number of attested oblique subject verbs shown in Table 5 (attested oblique subject verbs in Modern Icelandic), 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 5.** Application of the Tolerance Principle (Yang 2016, p. 165) to Icelandic subject case variation.

<b>role</b>	<b>accusative</b>	<b>dative</b>	<b>total (N)</b>	<b><math>\theta_N</math></b>
<b>theme</b>	14	19	33	9
<b>experiencers</b>	37	227	264	47

This is shown in Table 5, where the accusative theme subject verbs exceed the computed threshold while the accusative experiencer subject verbs do not. For themes, the accusative verbs represent too many exceptions for datives to be productive. Yang's (2016) analysis therefore yields the same conclusion Jónsson & Eythórsson (2005)

reached, although on different premises, where DS should not appear with theme subjects. Recent research indicates that this is consistent with the broad empirical picture, but that the patterns still are slightly more complicated, with some speakers showing non-nominative productivity for non-agents in general, particularly with verbs which tend to show subject case variation (Guðmundsdóttir et al. 2018). In this dissertation, the focus is placed on Dative Substitution exclusively, for the simple reason that Nominative Substitution verbs are rare and do not appear in child language. But how do children acquire the complex grammatical conditioning patterns outlined in this section and how does acquisition shape the directionality of the known changes in progress? Before addressing these questions in Chapter 5, and putting them in the context of developmental sociolinguistics (e.g. Labov 1989, Smith et al. 2009, Hendricks et al. 2018, Johnson & White 2020, Sneller & Newport 2019, Cournane & Pérez-Leroux 2020, Hall & Maddeaux 2020 and Repetti-Ludlow & MacKenzie 2022) as well as the dynamics of variation in individuals (Tamminga et al. 2016) and specialization (Wallenberg 2019), I begin by investigating the acquisition of case marking rules independent of variability in Chapters 2, 3 and 4.

## 2. The Meaning of Case

Based on previous research on the acquisition and productivity of datives in Icelandic, it seems reasonable to assume that learners associate case, at least in part, with semantics in Icelandic. In the current project, the nature of this relationship between the acquisition of verb meaning and case marking is considered a crucial factor in the discovery of productive patterns for comprehension and production. The learning of verbs is the first step in the hypothesis formation which leads to productive rules for the case marking of predicate arguments. But how is verb meaning acquired in general and in Icelandic in particular? Could the morphologically rich case marking system itself contribute to verb learning?<sup>22</sup>

### 2.1 Morphosyntactic bootstrapping

These questions are related to a field of language acquisition research in which universality has been extensively debated to answer, amongst other things, whether children use the same resources to learn verb meaning across languages. It is well established (Pinker 1984, 1989) that learning a verb is dependent on its semantic argument structure as well as its syntactic structure (but see arguments for item-based learning in e.g. Tomasello 2000). The syntactic bootstrapping hypothesis has played an important role in this context, as it proposes that children use the argument structure of a verb as a cue to its meaning (Landau & Gleitman 1985, Gleitman 1990, Naigles 1990). For example, it has been shown that learners as young as 15–19 months old associate novel verbs presented in a transitive frame, e.g. *the horse VERBs the dog*, more with a causative interpretation as compared to intransitive contexts such as *the horse VERBs* (Jin 2015).<sup>23</sup> 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 questions being whether possibly (syntactic) universal cues have an

---

<sup>22</sup> Part of the findings in this chapter were reported in Proceedings of the 44th annual BUCLD: Nowenstein et al. (2020).

<sup>23</sup> See also Arunachalam & Waxman (2010) for a contrast with conjoined noun phrases (*the horse and the dog are VERBing (together)*) and synchronous events with 2-year-old participants, and Noble et al. (2011) for a discussion on unclear results and further data for the conjoined agent intransitives.

advantage over language-specific (morphological) cues which are only (overtly) specified in some languages (e.g. Lidz et al. 2003, Göksun et al. 2008, Matsuo et al. 2012, Trueswell et al. 2012 and Leischner et al. 2016).

Lidz et al. (2003), based on ideas of universal syntax-semantics mapping and experimental data from speakers of Kannada as a comparison to learners of English, argued that children initially rely on argument number and ignore morphological form to bootstrap verb meaning, even when the morphology provides stronger cues. In the case of Kannada, this was tested by comparing argument number cues to a causative verbal affix. Lidz et al.'s (2003) conclusion has been challenged from various perspectives, one of them being typological evidence against the universality of argument structure cues (Brown & Bowerman 2008). Another critique appears in Göksun et al.'s (2008) paper on morphosyntactic bootstrapping in Turkish, where the findings of Lidz et al. (2003) are partly attributed to the presence of accusative case markers, which were not taken into account in their analysis, but can also be interpreted as markers for causativity. Still, even work on argument-drop languages such as Japanese, Turkish and Tagalog reveals that children use syntactic frames as cues – in addition to e.g. case morphology (Göksun et al. 2008, Matsuo et al. 2012 and Trueswell et al. 2012). In Trueswell et al. (2012), a comparison between Kannada and Tagalog shows that the use of case and causative morphology could rely on parsing conditions. In Tagalog, which is verb-initial, the causative morphology is argued to guide the interpretation, while the same cue forces speakers of Kannada, which is verb-final, to revise parsing commitments (a more taxing task). Research on German (Leischner et al. 2016) additionally shows that children rely less on the number of arguments and more on case when word order is highly flexible, showing that other cues than argument structure come into play when it is less reliable. 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 has these rare properties (Kiparsky 1997, Fedzechkina et al. 2016) and should therefore prove to be a valuable test case in the context of (morpho-)syntactic bootstrapping. Additionally, Icelandic has semantically driven dative productivity in subject and object case, as well as well-documented links between case and lexical

semantics, as discussed in Chapter 1, underlining the robustness of the case-semantics relation in the language.

Whether children learning Icelandic use case as a cue for meaning, with morphosyntactic bootstrapping, therefore is a particularly important question in two contexts. One is the context of the semantic predictability of case (see subsection 1.1.3) and the role it plays in (productive) rule formation. The second is the active debate regarding the advantage of (possibly) universal cues to verb meaning, e.g. the number of arguments, over language specific cues (Lidz et al. 2003, Göksun et al. 2008, Matsuo et al. 2012, Trueswell et al. 2012 and Leischner et al. 2016) such as morphological case marking.

In the context of the morphosyntactic bootstrapping debate, it furthermore is unclear to what extent an advantage of universal over language-specific cues could prove age-dependent. Naigles et al. (1992) noted that children aged 2–12 gradually went from being ‘frame compliant’ in act-out tasks modifying the argument structure of familiar verbs to being ‘verb compliant’ (like adults) as they got older, repairing the frame to fit the more common meaning and usage of the verb. This means that younger children would act out sentences such as *the zebra goes the lion* with the interpretation of the zebra causing the lion to go or by simply removing the lion, choosing the causative interpretation provided by the frame, while older children and adults would show the zebra going to the lion, choosing the usual semantics of the verb root. Building on these findings and their own comparisons between 3-year-old participants and adults speaking Kannada, Lidz et al. (2003) hypothesized that “early in life certain universal cues at the syntax–semantics interface are prepotent for acquisition of the verb lexicon” while adults have “adjusted such procedures to take into account various syntax-semantics relations that he or she has derived from the observation of specifics of the exposure language” (p. 168). This is somewhat in line with Göksun et al.’s (2008) and Leischner et al.’s (2016) results for Turkish and German morphosyntactic bootstrapping, where younger children rely more on the number of arguments than older children and adults. Interestingly, the results from the same studies show that all age groups rely less on argument number than their English speaking peers, compensating with case marking cues which in turn seem to be used more with age (however this age trend for the use of case cues appears only in part

of Göksun et al.'s (2008) results). To summarize, results from English, Kannada, Turkish and German show that younger children rely more on word order/number of arguments when interpreting sentences with familiar verbs in act-out tasks. The results from Turkish (Göksun et al. 2008) and German (Leischner et al. 2016) however show that children also use case as a cue, with some (but not all) results pointing towards an age trend where older children are better able to make use of such language-specific morphological cues.

It is important to note that the morphosyntactic bootstrapping research reviewed so far exclusively targets the cross-linguistically robust link between transitivity and causativity. This does not come as a surprise considering that the transitive-causative context arguably is the most investigated area of the bootstrapping literature and, perhaps as a consequence, has the broadest typological coverage as a research topic (see Fisher et al. (2020) and Lidz (2020) for recent in-depth reviews). The representation of the transitive-causative pattern in the bootstrapping literature can also be considered within the context of an early bias for causative events, with research showing that 27-month-old children show a preference for causative interpretations even in the absence of syntactic or visual information (Naigles & Kako 1993, Arunachalam & Waxman 2010). This perhaps points to the conceptual saliency of event participant role categories such as agent and patient, a notion which appears in various (decompositional) approaches to thematic roles and linking rules, e.g. Dowty's (1991) proto-roles and Baker's (1997) macroroles (see Pearl & Sprouse (2021) for an overview of bootstrapping results in the context of linking rules). The notion of proto- (or macro) roles can also be placed in the context of the idea that agents and patients are more prominent/salient event roles in thematic hierarchies, both from a linguistic and more general cognitive perspective (Jackendoff 1990, Baker 1997, Becker 2014, Ünal et al. 2021), as well as the idea of these roles' possible status as universal (pre-linguistic) components of core knowledge (Rissman & Majid 2019, Rissman & Lupyan 2022).

Indeed, Ünal et al. (2021) experimentally test "the connection between language and perception/cognition cross-linguistically, focusing on the relative salience of causative event components in language and cognition" (p. 1) in a picture-description task with 3–5-year-old learners of Turkish and English, and find evidence for an agent > patient/goal > instrument salience hierarchy both linguistically and cognitively, with a

patient > goal pattern still only emerging linguistically. Somewhat parallel results are found in Rissman and Majid's (2019) work targeting the status of thematic roles as scholarly fictions or core knowledge, where they review typological research, studies from emerging sign languages and data from psycholinguistic and event cognition experiments in search for evidence of category abstraction and universal biases. Rissman & Majid (2019) find strong evidence for abstraction into agent and patient categories and universal biases to differentiate between the two roles, but this could in fact be a bias to differentiate between agent and non-agents more broadly, and that there are indications that the patient role (in English at least) is more heterogeneous and diffuse than the agent role. For recipients and goals, they find evidence for category abstraction but mixed results regarding a universal bias to differentiate between these roles. Finally, they find no evidence for an abstract instrument category nor a universal bias to shape such a category. In sum, it could be inferred that the robustness of the transitive-causative pattern found in the bootstrapping literature is linked to the linguistic and cognitive salience of the causative event itself and the agent-patient contrast it contains. This is further supported by results such as the ones in Noble et al.'s (2011) study on conjoined agentive intransitives, showing that children acquire their mapping to non-causative events much later than the transitive-causative pattern.

In the current context of Icelandic morphosyntactic bootstrapping and the meaning of the dative in particular, it is interesting that the dative is in fact not typically associated with generic causative events and therefore neither agents nor patients, with the distributionally dominant nominative-accusative rather pointing towards such lexical semantics (but recall that Icelandic still has "truly direct object datives" (Wood 2015, p. 129)). This could perhaps mean that the presence of the dative points first towards non-agents and non-patients before being refined into more precise categories, also via bootstrapping. Despite the robustness of the transitive-causative results, it is important to acknowledge that the syntactic bootstrapping literature does contain a body of research, albeit not as extensive and typologically rich, on various contexts beside the transitive-causative mapping where children use (syntactic) form to infer meaning (see Lidz 2020), e.g. in contrasts between novel nouns (objects) and adjectives (object properties) or prepositions (spatial configuration) (Waxman & Booth 2001, Fisher et al. 2006). Within

different approaches to verb meaning, syntactic bootstrapping approaches have also illuminated puzzles on subject versus object experiencers (Hartshorne et al. 2015, 2016), attitude verbs (White et al. 2017, Hacquard & Lidz 2018 and Harrigan et al. 2019) and emotion words (Shablack et al. 2020).

Indeed, Hartshorne et al. (2015, 2016), take on the challenge that psych verbs pose for approaches to the linking problem which posit broad systematic mappings, with the experiencer of *fear*, for example, appearing both in subject position (*I fear dogs*) and object position (*Dogs frighten me*). Using evidence from English, Mandarin, Korean, Japanese, and Russian, Hartshorne et al. (2016, p. 280) reach the conclusion that “psych verbs describing habitual attitudes about some entity (*fear*-type verbs) map their experiencer onto the subject, whereas psych verbs describing a specific episode in which someone is caused to feel some emotion (*frighten*-type verbs) map their experiencer onto the object”. In Hartshorne et al. (2015), the causal component in the *frighten*-type verbs is furthermore interpreted as an explanation for their result that they are comprehended earlier in acquisition than *fear*-type verbs, despite the latter’s higher frequency. In this context, having an experiencer in subject position (versus a stimulus causing an emotion) creates more of a rupture with the ‘privileged’ link between agents and subjects. Interestingly, *fear*-type (but not *frighten*-type) verbs also fall under the category of attitude verbs, taking clausal complements (*I fear it may be too late*). The nature of the syntactic distributions of attitude verbs plays an instrumental role in recent work by White et al. (2017), Hacquard & Lidz (2018) and Harrigan et al. (2019), where it is shown that these distributions can be bootstrapped into various semantic subclasses (e.g. beliefs vs. preferences). In the case of Icelandic, the possibility of a clausal complement and the syntactic properties of said complement can of course also point towards a non-agentive experiencer subject. That is one type of cue to verb meaning, which under the current approach does not exclude the availability of other cues such as case marking, the focus of the current work.

### 2.1.1 Main questions and result summary

The main objective of this chapter is to test whether case marking and the dative in particular can be interpreted as a cue for verb meaning in Icelandic language acquisition.

While investigating comprehension, I focus on the semantic contexts where dative productivity has been observed (as is outlined in Chapter 1): experiencers in subject case and goals/recipients as well as themes of motion verbs in object case (see subsection 4.2.1 for more information about the thematic roles assumed within the project). The specific questions asked are the following:

- (15) a. Do children use case as a cue for verb meaning in Icelandic?
- b. Is the use of case as a cue age-dependent?
- c. Does the use of case as a cue depend on syntactic and/or semantic context and does frequency play a role?
- d. What is the relative importance of case in comparison to other cues such as transitivity/number of arguments?

Using experimental acquisition data from Icelandic targeting comprehension, I provide qualified support for a morphosyntactic bootstrapping account that does not exclusively rely on universal cues. I argue that in specific contexts, morphology can in fact be as salient as the number of arguments. I therefore propose, following e.g. Göksun et al. (2008), Matsuo et al. (2012) and Leischner et al. (2016), to build on the insights from the syntactic bootstrapping literature without limiting the learning possibilities to argument order. The basic theory is the same: Form drives the acquisition of meaning and verbs can be grouped into semantic categories based on their structural behavior (see also Pinker 1989, Levin 1993 and Yang 2016). But behavior other than word order can be relevant, and in languages like Icelandic, morphological case is bootstrapped. This becomes particularly clear when case is the only differentiating factor in minimal pairs such as the ones in examples (6)–(8) (see also (17)–(19) below), warranting the theory that the relationship between case and lexical semantics is informative for learners in other contexts as well. Since a learning model detects the available systematic mappings of form and meaning in a dynamic process (Yang 2016), universality is not the crucial component. Instead, the quality of the information provided by the cue is what matters, and the specific contributions of word order and case are therefore context-dependent and possibly variable over development. In the context of formal case theories, the results

show that learners do associate case with specific semantics despite the non-exact nature of the mappings.

## 2.2 Methods

As has been mentioned, the robust semantically-conditioned case marking patterns, in addition to a relatively rigid word order and very limited argument-drop, make Icelandic a convenient test case in the context of morphosyntactic bootstrapping in language acquisition. In this section, I report on two picture-selection comprehension tasks, with the choice of the relevant thematic role as the outcome variable in a forced-choice pointing paradigm (Noble et al. 2011). The experiments were designed to broadly test whether Icelandic speaking children use case as a cue for verb meaning as well as to answer the specific questions outlined in section 2.1.1. The first experiment is a minimal pair task with existing Icelandic verbs/predicates, while the second one is a novel (nonce) verb task. Both studies target the dative in objects and subjects.

### 2.2.1 Minimal pair task

Minimal pair examples such as the ones discussed in section 1.1.3 (examples (6)-(8)) 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, I conducted a pre-registered experiment with 48 preschoolers<sup>24</sup> aged 2;4–6;4.<sup>25</sup> Eight frequent Icelandic dative/non-dative sentence pairs were used, three subject case pairs and five object case pairs. These pairs are shown in (16) and the stimuli as a whole can be found in Appendix A:

- (16) a. Subject case: *Vera kalt/kaldur* ‘be cold’, *vera heitt/heitur* ‘be hot’, *leiðast* ‘be bored/hold hands’.

---

<sup>24</sup> The original plan was to also test an adult control group for each experiment, but this was disrupted by COVID-19.

<sup>25</sup> Pre-registration information: <https://doi.org/10.17605/OSF.IO/JKPMQ>

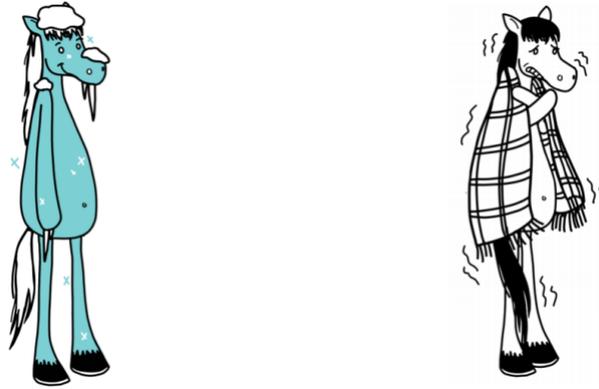
- b. Object case: *Skjóta* ‘shoot’, *sópa* ‘sweep’, *purrka* ‘dry’, and *klóra* ‘scratch’ in two contexts: one with two sentient arguments and one with a sentient agent but inanimate patient.

All participants were presented with the eight sentence pairs, yielding a total of 16 stimuli presented with two possible interpretations in the forms of pictures (also available in Appendix A) presented on paper. The object case sentences were presented in the accusative and dative (read aloud by the investigator) and the subject case pairs in the nominative and dative. The item order was counterbalanced and answers recorded in a spreadsheet in real time.

Figure 2.1 shows an example for the predicate *vera kalt/kaldur* ‘be cold’, where a dative subject indicates an experiencer reading (on the right, see sentence (17b)) where sentience is required. In this example, the horse feels cold without necessarily being cold to the touch. Meanwhile, the nominative indicates a theme (on the left, see sentence (17a)) where sentience is not required: the horse is cold to the touch without necessarily feeling cold.<sup>26</sup> In Figure 2.2, the stimuli for the object case of *klóra* ‘scratch’ are shown, with the test sentences in (18). The benefactive interpretation (left), associated with the dative (18b), was contrasted with a sentient patient (middle) on the one hand and an inanimate one (right) on the other hand (both of which are associated with the accusative in (18a), where the masculine pronoun refers to either the horse or the sofa, both masculine nouns), illustrating that the dative requires sentience.

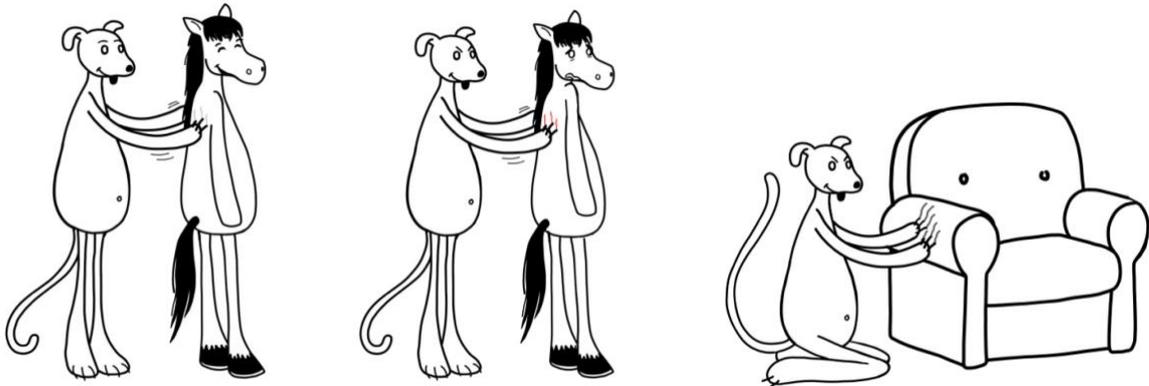
- (17) a. **Hesturinn** er kaldur.  
horse.the.NOM is cold.MASC  
‘The horse is cold.’ (to the touch)
- b. **Hestinum** er kalt.  
horse.the.DAT is cold.NEUT  
‘The horse is cold.’ (feels cold)

<sup>26</sup> Note that the test sentences with *vera kalt/kaldur* ‘be cold’ and *vera heitt/heitur* ‘be hot’, which are both adjectival predicates, also contain an agreement cue. This is not the case with *leiðast*, where the contrast lies in the agentive ‘hold hands’ interpretation with a nominative subject versus the experiencer ‘be bored’ interpretation with a dative subject.



**Figure 2.1.** Minimal pair task stimuli, subject condition: *Vera kaldur/kalt* 'be cold'.<sup>27</sup>

- (18) a. Hundurinn klórar hann.  
 dog.the.NOM scratches him.NOM-ACC  
 'The dog scratches him/it.' (damaging scratch)
- b. Hundurinn klórar honum.  
 dog.the.NOM scratches him.DAT  
 'The dog scratches him/it.' (benefactive scratch)



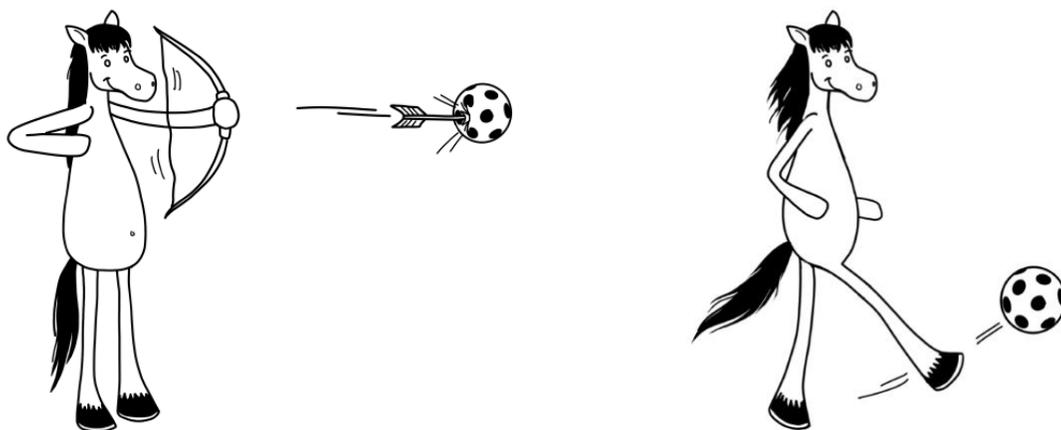
**Figure 2.2.** Minimal pair task stimuli, object condition: *Klóra* 'scratch'.

The dative with *vera kalt* 'be cold' and *klóra* 'scratch' therefore arguably represent the applicative semantics of the dative as experiencers and benefactives

<sup>27</sup> All illustrations are by Auður Ýr Elísabetardóttir (for the MoLiCoDiLaCo-project).

(goals/recipients/positively or negatively affected patients).<sup>28</sup> The association of the dative with object themes of caused motion verbs was also tested and contrasted with an affected patient. An example of this is found in Figure 2.3, with the test sentences in (19):

- (19) a. Hesturinn skýtur **boltann.**  
horse.the.NOM shoots ball.the.ACC  
‘The horse shoots the ball.’ (damaging shot)
- b. Hesturinn skýtur **boltanum.**  
horse.the.NOM shoots ball.the.DAT  
‘The horse shoots the ball.’ (movement shot)



**Figure 2.3.** Minimal pair task stimuli, object condition: *Skjóta* ‘shoot’.

Figure 2.3 shows the stimuli for *skjóta* ‘shoot’, where the dative is associated with the moving ball (right, see (19b)) while the accusative is associated with an affected patient which undergoes a change of state (left, see (19a)).

## 2.2.2 Novel verb task

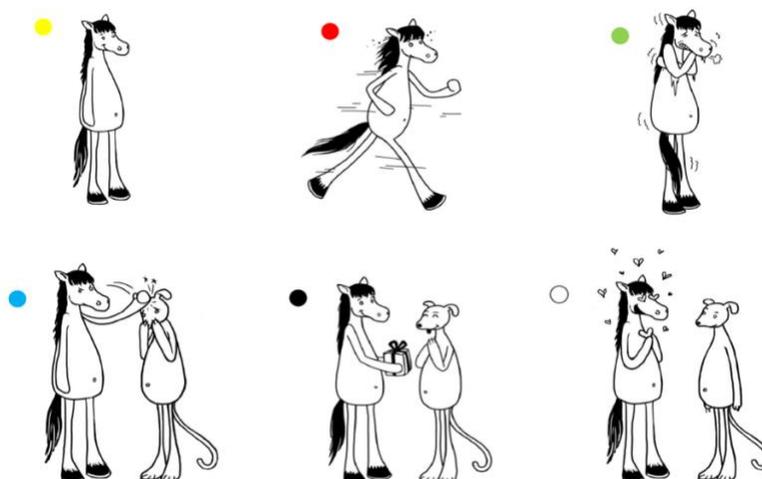
In addition to the minimal pair task, 146 children aged 2–13 (including the participants in the minimal pair task) participated in a novel verb task divided into two parts, one targeting

<sup>28</sup> The terminology and definitions for thematic roles varies tremendously, as is discussed in subsection 4.2.1. See Bosse et al. (2012) for ‘non-selected’ arguments in particular and the distinction between benefactives and affected experiencers.

subject case and the other object case. The first part is 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. All participants were presented with two novel verb stimuli for each condition, eight sentences in total (found in Appendix A). Examples for subject case are shown in (20), with the transitive condition in (20a) and the intransitive one in (20b). Note that the object *hann* 'he/him' is syncretic and could therefore be either in the nominative or accusative case:

- (20) a. **Hesturinn/hestinum** gorparhann.  
horse.the.NOM/DAT gorps he/him.NOMACC  
'The horse gorps him'
- b. **Hesturinn/hestinum** gorpar.  
the.horse.NOM/DAT gorps  
'The horse gorps'

Participants were asked to choose the appropriate picture out of six options (see Figure 2.4), three with a single argument and three with two arguments. The outcome measure was the choice of an experiencer (the first option from the right in both rows).

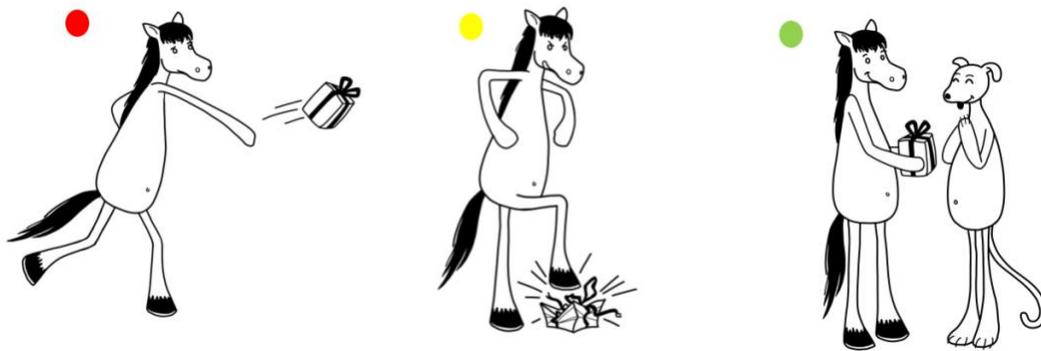


**Figure 2.4.** Pictures used in the novel verb task, subject condition (response to stimuli).

The second part of the novel verb task was an object condition where only case was manipulated, as in (21):

- (21) Hesturinn                    gorpaði                    **gjöfina/gjöfinni.**  
horse.the.NOM            gorp(ed)                    gift.the.ACC/DAT  
'The horse gorp(ed) the gift.'

This yields four items (two accusative and two dative) for which the motion theme interpretation (the first on the left) in Figure 2.5 was coded.



**Figure 2.5.** Pictures used in the novel verb task, object condition (response to stimuli).

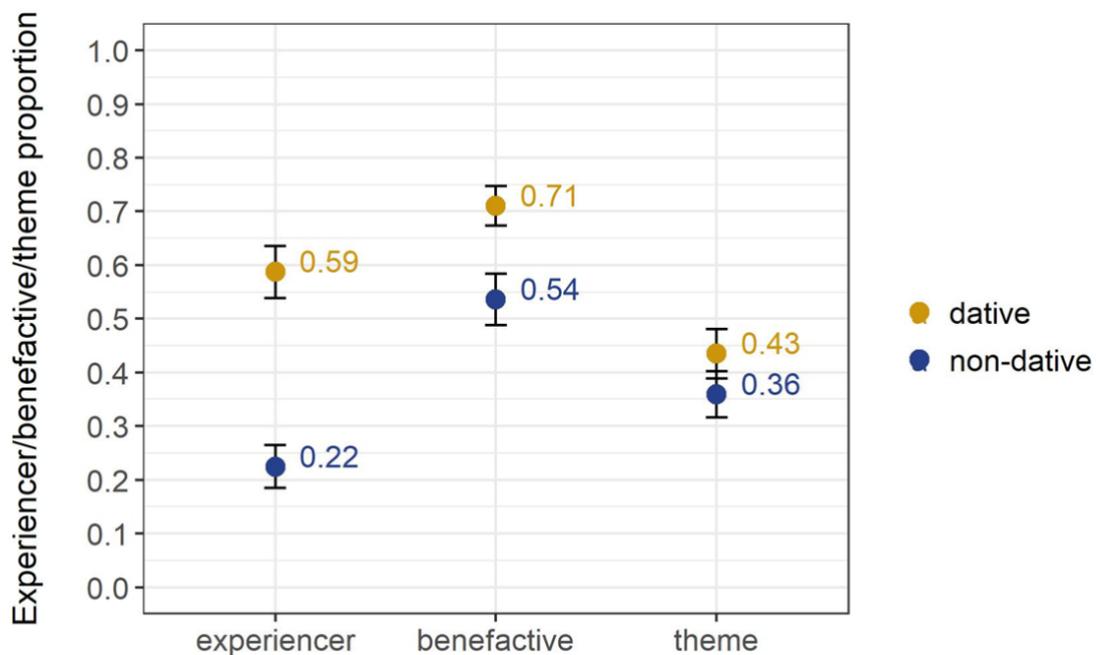
In both parts of the experiment, the order of the pictures was randomized and the order of the items counterbalanced. The pictures were presented on paper and the test sentences read aloud by the investigator. Answers were recorded in real time using a Qualtrics form.

## 2.3 Results

I begin by describing the results from the minimal pair task (2.2.1) before moving on to the novel verb task (2.2.2). The results are then summarized and discussed in subsection 2.4.

### 2.3.1 Minimal pair task

Figures 2.6 and 2.7 illustrate the results from the minimal pair task: Figure 2.6 shows the general results, while Figure 2.7 shows the results by younger (2;4–4;11 years old, N = 27) and older (5;0–6;4 years old, N = 21) participants. Note that the figures show results for both the subject and object condition.

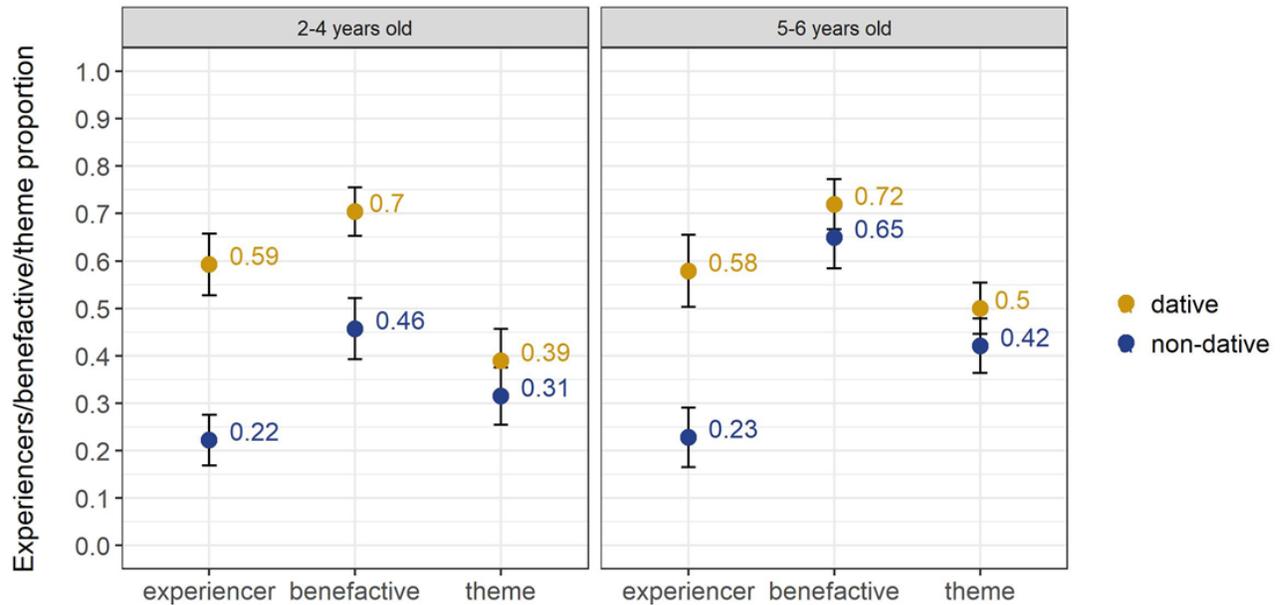


**Figure 2.6.** Results from the minimal pair task. Proportional choice of (dative) thematic role interpretation by case.

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 benefactives or movement themes in object position. Starting with Figure 2.6, it can be seen that preschoolers learning Icelandic are 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<sup>29</sup> as compared to objects (benefactives and motion

<sup>29</sup> Across items, the presence of an agreement cue does not significantly increase the choice of an experiencer interpretation.

themes). Children associate dative subject case with an experiencer reading. With objects, case is more conclusive for benefactives than motion themes.



**Figure 2.7.** Results from the minimal pair task, by age group. Proportional choice of (dative) thematic role interpretation by case.

Turning to Figure 2.7, it can be seen that the subject case effect (for experiencers) is stable across the two age groups. However, younger children seem to associate the dative with benefactive objects more strongly than older children. For motion theme objects, case does not seem to be interpreted as a differentiating factor across age. This distinction between benefactive and motion theme pairs is interesting in light of previous research and the examples discussed in section 1.2, where dative is possibly overgeneralized earlier with benefactives and indirect object goals are acquired before datives in other positions. As a contrast, also discussed in section 1.2, the productivity of the dative for motion themes seems to be better preserved in adult Icelandic, as exemplified in the case marking of many recently coined Icelandic verbs of motion, such as *dánlóða* ‘download’, which take a dative object. For new verbs with recipients/benefactives, on the other hand, dative does not seem to be as productive (see

discussion about the case marking of new Icelandic verbs in Barðdal (2008) and Jónsson & Thórarinsdóttir (2020)). We could therefore speculate that the specific association between dative and benefactives/recipients is particularly strong for a brief time in acquisition, before dative productivity with a wider scope becomes productive. This is further explored in Chapter 3.

To analyze the results from the minimal pair task further, a nested comparison (Likelihood Ratio Test) of mixed effects logistic regression models (lme4 in R, Bates et al. 2015) was conducted. Full model results can be found in Appendix D. Age (as a continuous variable, in months) was included in every model as well as random intercepts for participants and items. A significant improvement to the model fit was found by adding case ( $\chi^2(5) = 31, p < 0.001$ ), syntactic role ( $\chi^2(1) = 9, p < 0.01$ ) and a case-syntactic role interaction ( $\chi^2(4) = 11.1, p < 0.05$ ). The case effects thus appear much more prominently with subjects than objects, as can be deduced from Figures 2.6 and 2.7.

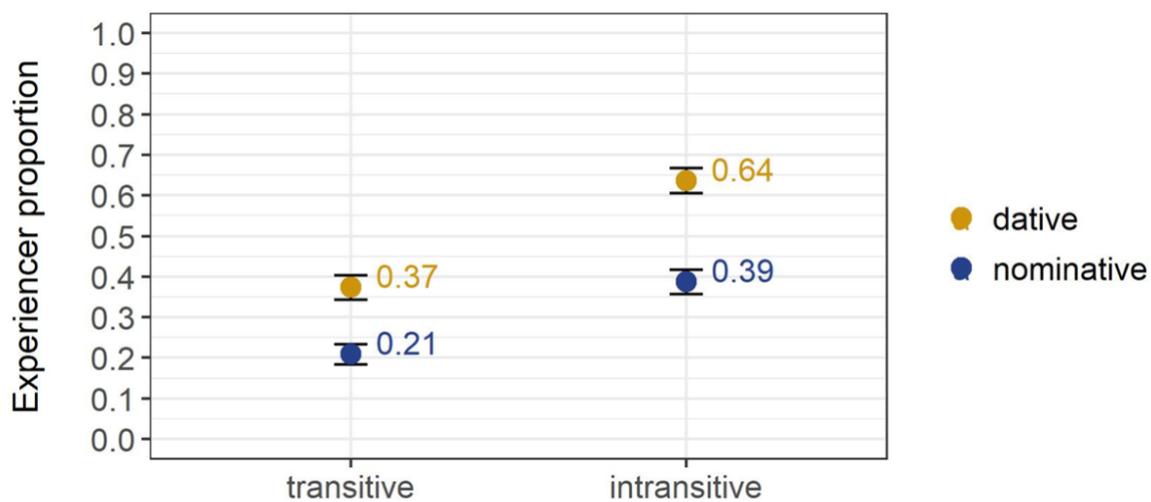
The minimal pair task therefore showed that preschoolers (aged 2;4–6;4) are able to take advantage of some case cues when interpreting sentences where word order is uninformative.<sup>30</sup> The effect is stronger for (experiencer) subjects, although the dative-context selection rate is not higher than for benefactive objects. Within objects, case is more conclusive for benefactives than motion themes, and younger children seem more sensitive to it in that context than older children. This means that within a language, cues can be more or less valuable depending on context and age. The dative on subjects seems to push children towards an experiencer interpretation (as opposed to themes or agents) early on, but the contrasts in object case do not appear as clearly, with the associations to motion themes being the least established. But does this association emerge later? And is the strong link between dative subjects and experiencers preserved with novel verbs, despite the low frequency and lack of productivity with borrowings in Icelandic? The results from the novel verb task in subsection 2.3.2 help illuminate these questions.

---

<sup>30</sup> See also Hreggviðsdóttir (2018) for a study where case is a useful cue for comprehension in contexts where word order is informative.

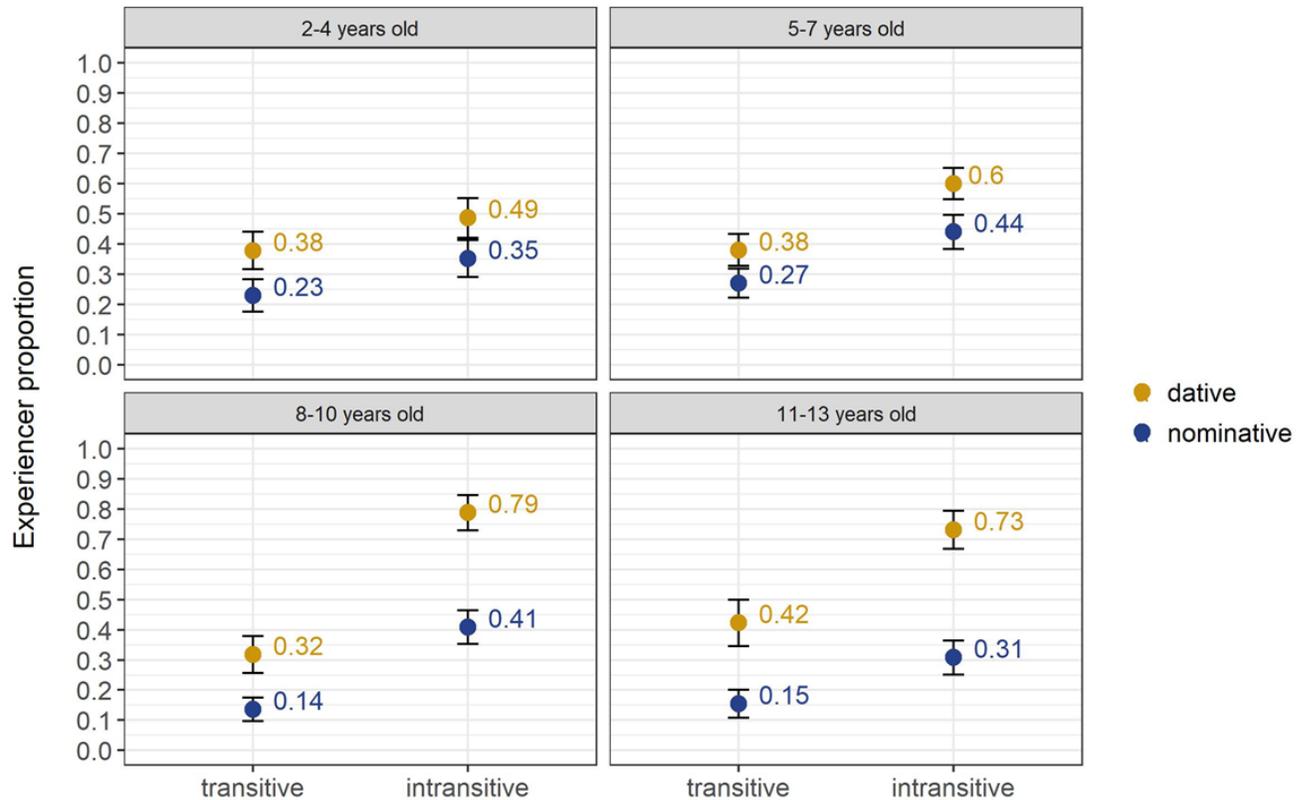
### 2.3.2 Novel verb task

Beginning with the subject condition, Figures 2.8 and 2.9 illustrate the main results from the novel verb task: Figure 2.8 shows the general results, while Figure 2.9 shows the results across four different age groups (2;4–4;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 2.8, the tendency to interpret dative subjects as experiencers in the minimal pair task is confirmed in the novel verb task.



**Figure 2.8.** Results from the novel verb task. Choice of experiencer subject by case and transitivity.

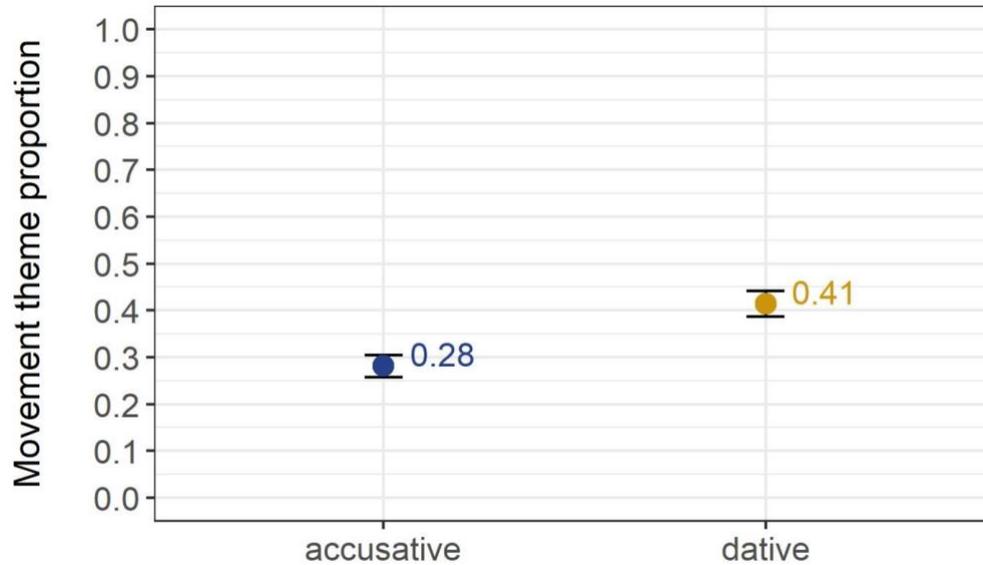
Using nested model comparison again, significant improvements to the fit of the model can be seen when adding transitivity ( $\chi^2(5) = 60.3, p < 0.001$ ) and case ( $\chi^2(6) = 56.3, p < 0.001$ ) but not a transitivity-case interaction to the model ( $p = 0.386$ ) – so the results do not indicate an augmented case effect in the intransitive condition, for example. Both effects (case and transitivity) interact with age, as is illustrated in Figure 2.9.



**Figure 2.9.** Results from the subject condition of the novel verb task, by age group.

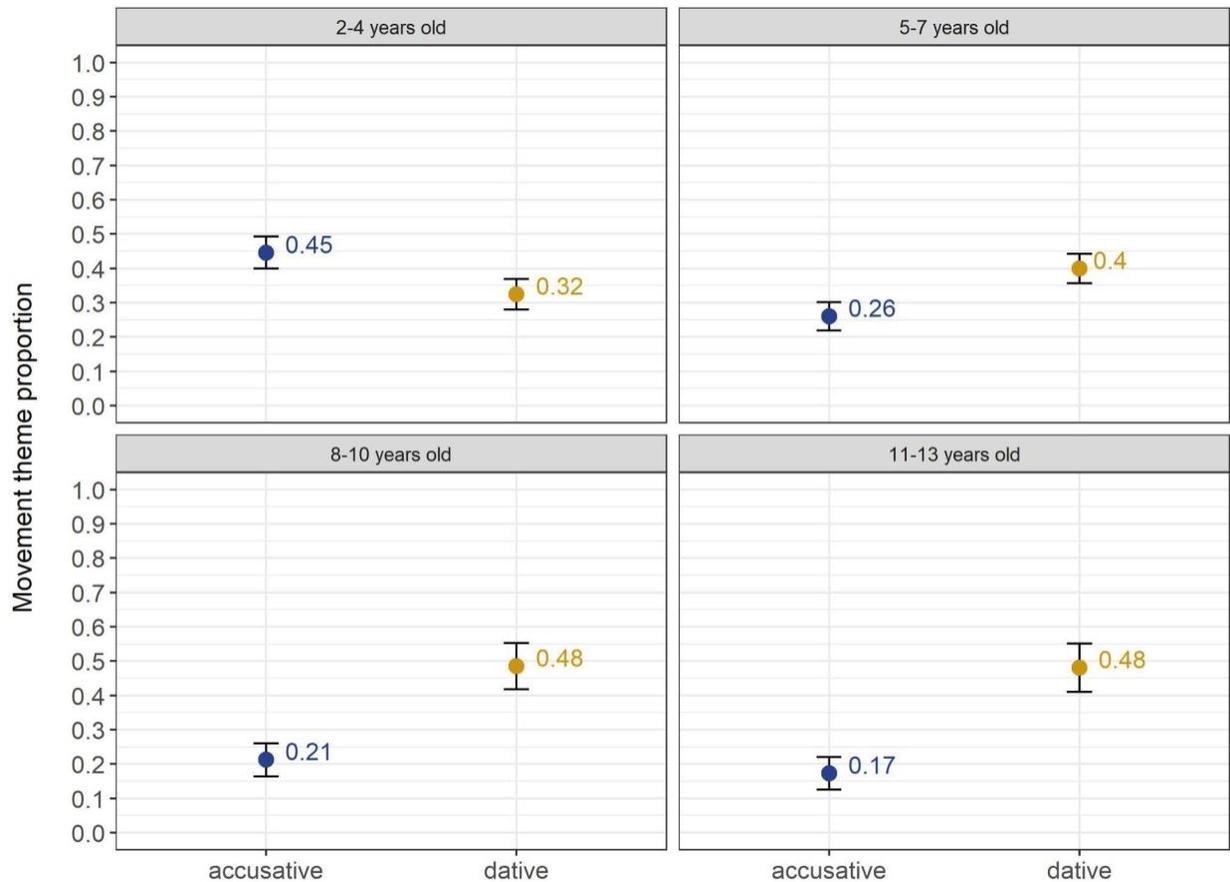
The effects of both variables are augmented as the children grow older. When correcting for age, case (estimate: -1, SE: 0.13) and argument number (estimate: -0.94, SE: 0.13) show comparable effect sizes. So crucially, this does not entail that younger children use the number of arguments while the older participants use case morphology. Instead, the children use both types of cues from age 2 but get better at using them as they get older. This increased cue availability is particularly clear for case with intransitives.

As for the object condition, it can be seen in Figures 2.10 (general results) and 2.11 (results by age groups) that the motion theme interpretation is chosen significantly more often when the object is in the dative, so the distinction which could not be found in the minimal pair task is present in the novel verb task, which additionally has participants aged 6–13.



**Figure 2.10.** Results from the novel verb task. Choice of motion theme by case.

Again using nested model comparison, the fit of the model improves with a case-age interaction ( $\chi^2(6) = 17.2, p < 0.001$ ), as can be deduced from Figure 2.11 on the next page. Indeed, the case effect is not present for the youngest children, which is compatible with the results from the minimal pair task and a scenario where the association between the dative and 'typically' dative semantic contexts is stronger in subject position, with experiencers, than in object position, with themes of motion verbs.



**Figure 2.11.** Results from the object condition of the novel verb task, by age group.

To summarize, the novel verb task confirms the tendency to interpret dative subjects as experiencers. When correcting for age, case and argument number show comparable effect sizes. The effects of both variables are augmented as children grow older, with case in an intransitive context being particularly sensitive to age. For the objects, we saw that the motion themes are selected more frequently with datives. This effect is not present for the youngest age group and the contrasts are smaller (and the dative-context selection rate lower) than in the subject condition. Unfortunately, the association between datives and goals in object position was not tested in the novel verb task, which would have proven important to further understand the unusual developmental path which appears in the results of motion themes from the minimal pair task. The associations between datives and goals are however investigated further in Chapter 3, in the context of production.

## 2.4 Discussion

Putting the experimental comprehension results presented in section 2.3 in the context of the syntactic bootstrapping literature, I 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, Trueswell et al. 2012 and Leischner et al. 2016). The results show that children acquiring Icelandic can, early on, use case to determine verb meaning when word order is uninformative. I furthermore show, by testing transitivity and case morphology within 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, but case particularly. 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. In certain contexts, such as psych predicates with experiencer subjects, the relative importance of case and word order is roughly the same, with no evidence that younger children are ignoring language-specific morphological cues, even though they get better at using them with age. Note however that this does not exclude the possibility of case cues being irrelevant, not present, or completely dependent on age (particularly if younger children would be tested) in other semantic contexts, where word order plays a more prominent role. This should be the case when contrasting causative and synchronous agentive events for example (e.g. Arunachalam & Waxman 2010), where the subject case would not differ between conditions. Additionally, the syntactic properties of Icelandic allow us to show that the use of morphological cues does not depend on argument-drop or word order being less rigid (as is hypothesized in e.g. Leischner et al. 2016).

It is important to bear in mind that the use of case cues varies across contexts in these experimental results. Dative on subjects, as opposed to objects, seems to be 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 the language. 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 (perhaps of the 'privileged' link between agents and subjects mentioned by and Hartshorne et al. (2015), see also Becker 2014). In general, the results indicate that the properties of Icelandic dative subjects could prove important in research on the acquisition of psych verbs (e.g. Hartshorne et al. 2015 and 2016, Harrigan et al. 2019 and Shablack et al. 2020) but also in other wider contexts. For example, one interpretation of the results could be that non-agents in subject position are somehow more salient than non-patients in object position, a finding which is somewhat in line with the work of Rissman & Majid (2019) and Ünal et al. (2021) where agentivity and causativity are more salient cognitively and linguistically.

In the context of language acquisition models, 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). In Chapter 4, I investigate whether the same learning model, through the Tolerance Principle, can provide an account for case marking rule formation and productivity in Icelandic language acquisition. This shows how generalizations which guide verb learning are also relevant in production: Icelandic dative productivity provides a scenario where exceptions, or non-default patterns, can be structured into productive rules and generalized systematically based on distinctions which arise in comprehension. Although the same form-meaning patterns found in the linguistic input are underlying, I assume that the processes of linking together form and meaning in comprehension (word learning) on one hand, and production (with acquired verb roots) on the other hand, are fundamentally distinct in nature. When acquiring the meaning of verbs, the directionality of the hypothesis formation goes from form to meaning. When producing language with acquired verbs, the rules are formed in the meaning-form direction, or even form-form in some cases. This has implications for the association patterns we expect to find in comprehension versus production experiments, as is further explored in Chapter 3, where the focus is on finding the semantic contexts in which children acquiring Icelandic form productive dative case marking rules.

### 3. Producing non-default case

In the preceding chapter, it has been established that child learners of Icelandic can use case as a cue for verb meaning. Dative case marking on subjects provides a more salient cue than dative case marking on objects in comprehension experiments, and the mapping of the dative to applied arguments (experiencers and goals) is seemingly available earlier than the link between datives and themes of motion verbs. We now turn to production and investigate whether the case-semantics patterns found in comprehension can form the basis for (productive) case marking rules in acquisition. How do children form rules/abstractions which can be extended to new items? Can non-default patterns, which are not distributionally dominant (minority rules), like the dative in Icelandic, be productive? How could this type of productivity be determined by children and predicted in models of language acquisition? Studying the productivity of Icelandic datives entails addressing fundamental questions about rule formation and productivity.

#### 3.1 Rule formation and productivity

Rule formation and productivity are at the heart of current but long-standing debates within language acquisition research, e.g. in discussions about stored exemplars versus abstractions (e.g. Ambridge 2020 and responses to his article), rule-based accounts of acquisition or single-/dual-route models (e.g. Yang 2016, Blything et al. 2018) and different approaches to (over)generalization more broadly (e.g. O'Donnell 2015, Yang 2016, Schuler et al. 2016, Yang 2018 and responses to the article, Ambridge et al. 2018 and Goldberg 2019). Before describing experimental data targeting the production and productivity of Icelandic datives in language acquisition and discussing their implications within this line of research (subsection 3.4), I briefly summarize recent developments in research on rule formation and productivity in child language.

So far in this thesis, I have used category labels such as VERB, ARGUMENT, SUBJECT, OBJECT, NOMINATIVE, DATIVE and even AGENT or EXPERIENCER (mostly) following convention, but without defining the categories' properties in detail nor questioning their existence as cognitive entities or pre-linguistic concepts (but see subsections 2.1 and 4.2.1 for thematic roles). Despite the fact that these abstract

categories are for the most part assumed in linguistics and language acquisition, I agree with the call in Ambridge's (2020) target article to question these assumptions, although his alternative proposal where language only consists in *stored exemplars and the ability to analogize* probably is too radical (as he concedes in Ambridge 2020b). What is clear is that arguments against abstractions and rules can be made and should be made in order to finesse our understanding of why abstractions need to be assumed and under what form. For example, in a reply to Ambridge's (2020) target article, Schuler et al. (2020) point towards neural and experimental (as well as computational) evidence of abstraction (or "the higher-level structures that only emerge across experiences") in humans and monkeys, and Zettersten et al. (2020) describe various phenomena which are difficult to account for without abstraction, such as category-based abstraction and the different learnability of categories with the same amount of exemplars (Zettersten & Lupyan 2020). The conclusion, reached by most commentators and Ambridge in his response (2020b), is that the evidence does point towards abstractions being real. Some of the questions which need to be addressed consequently are if and how they are derived, at what level(s) they exist and how they are organized.

These questions have in part been addressed in the enduring debate on the acquisition of past tense in English (e.g. Bybee & Moder 1983, Rumelhart & McClelland 1986, Pinker & Prasada 1993, Pinker 1999, Pinker & Ullman 2002, Yang 2002, Albright & Hayes 2003, Ambridge 2010, O'Donnell 2015, Yang 2016, Blything et al. 2018), where single- and dual-route models contrast with rule-based accounts. In single-route (or connectionist) models (Bybee & Moder 1983, Rumelhart & McClelland 1986, Ambridge 2010 and Blything et al. 2018) of the past tense, no abstract rules are assumed, neither for regulars or irregulars. Instead, the past tense is "generated by analogy across stored exemplars in associative memory" (Blything et al. 2018, p. 61). In the dual-route alternative (Pinker 1999), irregulars are stored in associative memory just as in the single-route models, but regulars are generated with an abstract, default "add -d" rule which can apply to all verbs. Finally, both approaches contrast with the more traditional way in which the past tense has been portrayed in generative linguistics, where both irregulars and regulars are accounted for with abstract rules (Halle & Mohanan 1985). The more unpredictable items still need to be associated with particular patterns or individual forms

in memory, and modern versions of multiple-rule accounts include computational inductive approaches which allow multiple stochastic rules (Albright & Hayes 2003) and take into account rule frequency (Yang 2002). What kind of model fits the Icelandic dative case marking best? A model which assumes neither rules nor abstraction but only stored items and analogy, a model which only assumes one regular default rule, or an approach based on multiple rules? As Yang (2002, 2016) points out, the debate lacks cross-linguistic evidence and the testing of more complex systems than the English past tense, and I argue that the non-default productivity of Icelandic datives can be valuable in this context.

Productivity has in fact been, although perhaps indirectly, an important component in the past tense debate, since a fundamental part of the modeling enterprise has been to explain why regularization occurs much more than irregularization. Is it because regulars always form a default rule (dual-route)? Or simply because they are more frequent (single-route)? One promising line of thinking within rule-based accounts (Yang 2002, 2016 and to some extent O'Donnell 2015) points to two essential components in understanding past tense marking and overgeneralizations: 1) The contexts for rules can be more or less restricted, for example applying to all possible verbs (as in the "add -d" rule) or only verbs having certain phonological properties (e.g. /-ɪŋ/ for *sing-sang*) and 2) Rules (or patterns) can be productive or unproductive. Crucially, productivity is not restricted to regular or statistically dominant forms, and individual speakers or groups (e.g. children and adults) can vary in the ways rules are restricted as well as in their productivity (Yang 2016).

This brings us to another question about the nature of rules and productivity which has shaped different approaches to overgeneralization and the retreat from it in child language: is productivity binary or gradient? Interestingly, single-route and constructionist models tend to assume gradient productivity (Ambridge et al. 2018, Blything et al. 2018, Goldberg 2019) while (at least some) rule-based accounts assume that productivity is binary (Yang 2002, Albright & Hayes 2003, O'Donnell 2015, Yang 2016) and that the different contexts (some very restricted, others broad) over which rules operate can create the appearance of gradience. Note that binary productivity does not exclude stochastic rules. In approaches where productivity is considered gradient, children's

overgeneralizations, or at least the retreat from them, have been accounted for with the notions of entrenchment (raw item frequency, Braine & Brooks 1995), statistical preemption (item frequency in competing contexts, Goldberg 1995 and 2006) or both (Ambridge et al. 2018). But as has been noted (in Yang 2016 and Ambridge et al. 2018), these approaches to language acquisition rarely make quantitative predictions which can be tested experimentally. Fitting data statistically, however successfully, does not automatically become a theory of learning. In binary approaches to productivity, a step is taken towards going beyond the finding that majority patterns are often productive and overgeneralized, an aspect which is promising when it comes to explaining e.g. the productivity of statistically non-dominant rules (as the -s suffix for German plurals, see Yang 2016). In Chapter 4, I explore one such approach to productivity, the Tolerance Principle (Yang 2016), in greater detail and test it in the context of Icelandic case marking. This choice of model is justified with the data in the current chapter, which has the objective of documenting the productivity patterns emerging in the acquisition of the Icelandic case marking system. In addition to discussing the data in the context of rule formation and productivity in language acquisition, I address the relationship between form-meaning mappings in comprehension and production.

### 3.1.1 Main questions and result summary

The main questions of the chapter are found in (22):

- (22)
- a. What overgeneralization patterns are present in Icelandic speaking children's case marking of existing verbs across age?
  - b. What are the optimal semantic and structural contexts for dative productivity with novel verbs?
  - c. What theories of rule formation and productivity are compatible with the patterns of productivity found in the Icelandic case-marking system?
  - d. Do the same patterns of association between form and meaning appear in comprehension and production?

The results replicate previous findings with existing verbs (e.g. Sigurðardóttir 2002, Jónsson & Eythórsson 2003) to some extent and expand them as well. For Dative Substitution (DS) verbs, children overgeneralize both the default nominative and non-default (but 'inherent') dative, with the nominative overgeneralizations decreasing with age while the dative ones increase. Interestingly, accusative is sometimes (but marginally) chosen with the historically dative *finnast* 'find' as well. For object case, the default accusative is unsurprisingly extended to dative and genitive objects, but interestingly, dative is also overgeneralized and nominative overgeneralizations appear as well, mostly in the youngest age group. Dative on indirect objects seems to be acquired sooner than dative on direct objects. Additionally, although frequency effects were not the main focus of this chapter, no clear contrasts between high and low frequency items appear. Although overgeneralizations are a clear sign of productivity, the contexts for the non-default productivity of the dative were tested further with a novel verb experiment. In the subject case, dative is selected more in the expected experiencer contexts, with additional structural components such as the *-st* suffix prompting more dative selection. Dative selection, or productivity, still proves to be significantly higher with objects than subjects, both in typical and atypical semantic contexts, pointing towards productivity that goes beyond the scope of translational equivalents and caused motion (Jónsson & Thórarinsdóttir 2020) and even Barðdal's (2008) more diverse schemas associated with the dative.

The data therefore call for a model of rule formation which accounts for nested non-default productivity and does not hinge on frequency to determine gradient overgeneralization, since patterns which are not statistically dominant can become productive and multiple rules can apply to a context simultaneously. I argue this is most compatible with rule-based accounts (Yang 2002, Albright & Hayes 2003, O'Donnell 2015, Yang 2016), which also provide the best possible framework for the different patterns found in comprehension and productivity. Indeed, rule directionality can be used to account for the fact that dative subjects are produced less than dative objects but provide better cues in comprehension. Finally, I explore the possibility of predicting productivity at an individual level based on vocabulary size and find that the size of productive, but not receptive, vocabulary shows a significant relationship to the

acquisition of dative productivity. This might indicate that productivity relies on a child acquiring a sufficient amount of evidence to derive dative case marking rules, an intuition which is compatible with the Tolerance Principle (Yang 2016) and is further investigated in Chapter 4.

## 3.2 Methods

The method used to map Icelandic dative productivity and its learnability is a forced-choice task with both existing and novel verbs. 101 children aged 3–13 chose between cases for 17 existing Icelandic verbs and 16 novel and semi-novel verbs. The main objective was to document the contexts in which dative is produced the most and provide some contrasting items as well.

### 3.2.1 Existing verbs

For the existing verbs, the test sentences and response options were presented orally by the researcher administrating the task and in written form on a laptop screen, with the case options randomized and a counterbalanced list of items. Based on the fact that children never overgeneralize the genitive in corpus data (Sigurðardóttir 2002) and to lighten the load of options, the participants chose between the nominative, accusative and dative for all the items except the genitive object taking verbs, where the genitive was of course also available as an option. Subject case marking was tested with the following sentences (the list of stimuli is also provided in Appendix B):

- (23) a. **Stelpan/stelpuna/stelpunni**      **hlakkar**      svo      til.  
 girl.the.NOM/ACC/DAT      anticipates      so      to  
 ‘The girl is really looking forward to [this].’
- b. **Strákurinn/strákinn/stráknum**      **langar**      í      nammi.  
 boy.the.NOM/ACC/DAT      wants      in      candy  
 ‘The boy wants candy.’
- c. **Stelpan/stelpuna/stelpunni**      **dreymdi**      um      þetta.  
 girl.the.NOM/ACC/DAT      dreams      about      this  
 ‘The girl dreamt about this.’

- d.     **Strákurinn/strákinn/stráknum**   **klæjar**       alltaf  
           boy.the.NOM/ACC/DAT            itches         always  
           ‘The boy is always itching.’
- e.     **Stelpan/stelpuna/stelpunni**       **finnst** gaman.  
           girl.the.NOM/ACC/DAT           finds fun  
           ‘The girl finds [this] fun.’

Four of the five verbs in (23) are known to appear with variable case marking or Dative Substitution (DS) and were chosen to cover as much of the diversity of DS verbs as possible. *Hlakka* ‘look forward to’ has a historically nominative subject and is known to appear with accusative and dative subjects as well. Despite the historical nominative, it is safe to assume that children acquiring Icelandic regularly encounter this verb with non-nominative subjects. On the other hand, *langa* ‘want’, *dreyma* ‘dream’ and *klæja* ‘itch’ all have historically accusative subjects and should appear much more rarely with a nominative subject, with the variation lying within non-nominative subjects. *Hlakka* ‘look forward to’ and *klæja* ‘itch’ are not common enough to be included in the 410 verb IceCASE corpus but both appear in the child language data used to compile the corpus (rank 541 for *hlakka* and 852 for *klæja*). *Langa* ‘want, long for’ is the most common DS verb in the IceCASE corpus (rank 37) and typically appears with PPs or clausal complements. *Dreyma* ‘dream’ is similar but is the least common DS verb (rank 408) in the IceCASE corpus and has been known to appear with nominative subjects as well (Halldórsson 1982), possibly particularly when used in the context of future aspirations. *Klæja* ‘itch’ is the only one of these five verbs which is not propositional and does not appear with clausal arguments. It also denotes physical discomfort, a semantic component which has been hypothesized to rather appear with accusative subjects (Ingason 2010). Finally, *finnst* ‘find’ is the most common dative subject verb (rank 14), traditionally not displaying any variation (but see Nowenstein & Ingason (2021) for unusual examples of accusative substitution).

For the object condition of the existing verbs, all participants were tested on eight verbs, two of them which historically appear with accusative objects. One of these two accusative verbs, *taka* ‘take’ (example (24a), rank 13 in the IceCASE corpus), only

appears with an accusative object despite having a motion theme object and the possibility of an optional reflexive dative benefactive.

- (24) a. Kennarinn            **tók**    **bókin/bókina/bókinni.**  
 teacher.the.NOM    took    book.the.NOM/ACC/DAT  
 'The teacher took the book'
- b. Mamman            **keyrði**            **stelpun/stelpuna/stelpunni.**  
 mom.the.NOM        drove        girl.the.NOM/ACC/DAT  
 'The mom drove the girl.'

The other one, *keyra* 'drive' (example (24b), rank 44 in the IceCASE corpus), is known to appear with dative objects as well as accusative ones.

Four verbs which appear most frequently with dative objects were then tested, in an attempt to cover the broad range of dative objects present in the language as well as known instances of variation, as can be seen in the examples in (25). *Kasta* 'throw' (rank 181) which is only attested in the dative, *rústa* 'wreck' (does not appear in the IceCASE data despite being very colloquial), a historically accusative verb which now almost exclusively appears in the dative, *gefa* 'give' (rank 28), one of the most common ditransitives with a dative indirect object and finally *þurrka* 'dry' (rank 144), which alternates between the accusative and dative.

- (25) a. Maðurinn            **kastaði**            **boltinn/boltann/boltanum.**  
 man.the.NOM        threw        ball.the.NOM/ACC/DAT  
 'The man threw the ball.'
- b. Stelpan            **rústaði**            **myndin/myndina/myndinni.**  
 girl.the.NOM        wrecked        picture.the.NOM/ACC/DAT  
 'The girl wrecked the picture.'
- c. Konan            **gaf**    **stelpun/stelpuna/stelpunni**        nammi.  
 woman.the.NOM    gave    girl.the.NOM/ACC/DAT        candy.ACC  
 'The woman gave the girl some candy.'

- d. Maðurinn **þurrkaði** strákurinn/strákinn/stráknum.  
 man.the.NOM dried boy.the.NOM/ACC/DAT  
 ‘The man dried the boy.’

I also tested two verbs with genitive objects which provide a non-default/non-structural contrast with the dative (see the examples in (26)). This was done with the verbs *sakna* ‘miss’ (does not appear in the IceCASE data despite being the most common consistently genitive object verb), and *spyrja* ‘ask’ (rank 102), which is known to also appear with accusative objects and most frequently has a PP instead of a direct genitive object. The test sentences are shown in (26).

- (26) a. Strákurinn **saknar** hún/hana/henni/hennar.  
 boy.the.NOM misses she.NOM/ACC/DAT/GEN.  
 ‘The boy misses her.’
- b. Strákurinn **spurði** spurning/spurningu/spurningar.  
 boy.the.NOM asked question.NOM/ACCDAT/GEN.  
 ‘The boy asked a question.’

Finally, participants aged 6 and older chose the object case for four recent technological borrowings which are all attested with dative objects (and do not appear in the IceCASE corpus data, but see the stimuli in the examples in (27)). These are *gúggla* ‘google’, *dánlóða* ‘download’, *kópípeista* ‘copypaste’ and *stríma* ‘stream’ which were presented in the following form:

- (27) a. Strákurinn **gúgglaði** söngvarinn/söngvarann/söngvaranum.  
 boy.the.NOM googled singer.the.NOM/ACC/DAT  
 ‘The boy googled the singer.’
- b. Stelpan **dánlóðaði** myndin/myndina/myndinni.  
 girl.the.NOM downloaded movie.the.NOM/ACC/DAT  
 ‘The girl downloaded the movie’
- c. Strákurinn **kópípeistaði** textinn/textann/textanum.  
 boy.the.NOM copypasted text.the.NOM/ACC/DAT  
 ‘The boy copypasted the text’

- d. Stelpan **strímaði myndin/myndina/myndinni.**  
 girl.the.NOM streamed movie.the.NOM/ACC/DAT  
 'The girl streamed the movie.'

The much less frequent borrowing *kreiva* 'crave' (the only recent verb which has been attested with non-nominative subjects), was tested in the novel verb experiment described in the next section, with a pictorial depiction.

### 3.2.2 Novel verbs

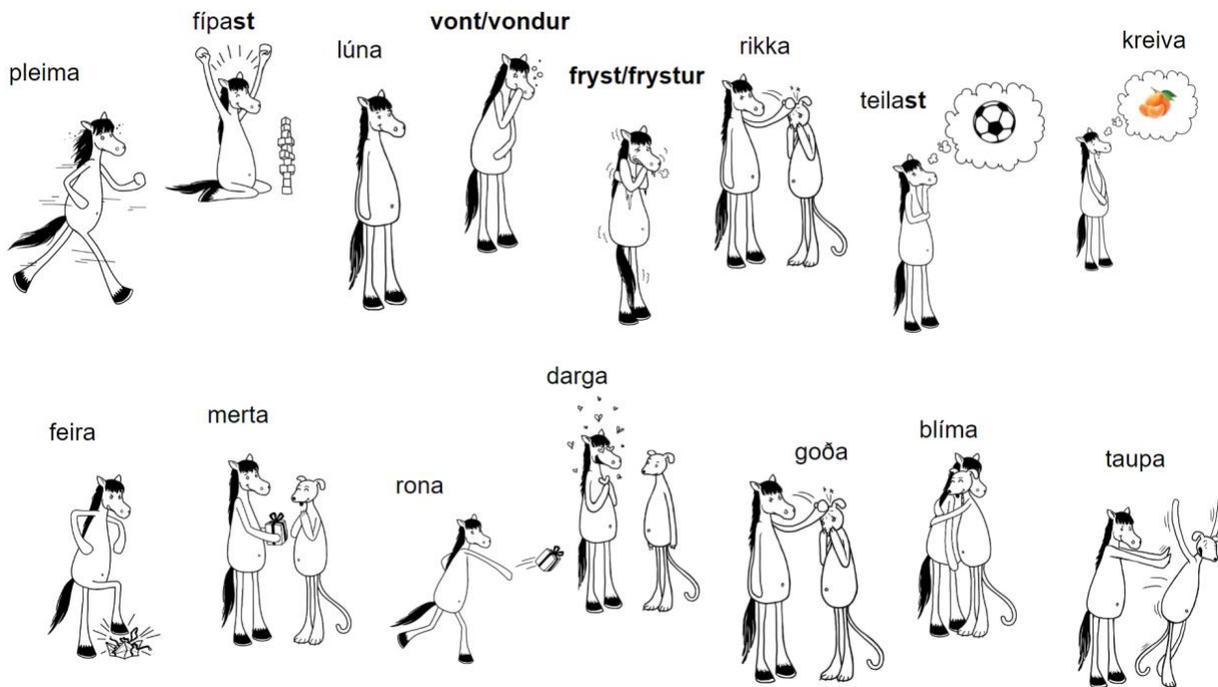
For the novel verbs in the production task, the participants chose between cases in eight subject contexts and seven object contexts. The novel and semi-novel predicates were presented orally by the examiner with (printed) pictures denoting the events/states. These images can be found in Figure 3.1 along with the novel and semi-novel predicates used (test sentences can be found in Appendix B). The items for the subject condition are in the upper segment and the items for the object conditions in the lower one. Although all the novel verbs have Icelandic equivalents, these were not presented along with the pictorial depiction.

In the subject condition, two agentive contexts were tested, one transitive and one intransitive, as a contrast with the semantic contexts where dative case marking would rather be found based on previous research. These are the running (*pleima*, the verb *hlaupa*) and hitting (*rikka*) actions which can be found in Figure 3.1. The existing equivalent verbs *hlaupa* 'run' and *lemja* or *slá* 'hit' appear at ranks 46, 199 and 168 respectively in the IceCASE corpus. For the contexts more prone to non-nominative subjects which contrast with running and hitting, I tested one item with a single argument standing still (and looking bored), *lúna*, which could be considered equivalent to the nominative subject verb *standa* 'stand' (rank 52) or the dative subject verb *leiðast* 'be bored' (rank 326). I also tested the recent borrowing *kreiva* from 'crave', which has to be considered a novel verb instead of a borrowing for some of the participants.<sup>31</sup> Finally, there are four semi-novel predicates. Two of them have *-st* morphology, *teilst* with a

---

<sup>31</sup> This verb was not included in the borrowing part of the existing verbs forced-choice task because of its low frequency and therefore it was necessary to present it with a pictorial depiction.

mental verb representation, which could be considered equivalent to the dative subject *finnast* ‘find’ (rank 14), and *fípast* with happenstance semantics, which could be considered equivalent to the alternating verb *takast* ‘succeed’ (rank 136). The last two are adjectival predicates based on existing roots which typically appear with a nominative subject, *vondur/vont* ‘bad, hurt’ and *fryst/frystur* ‘frozen’ but have synonymous equivalents which appear with the dative, *vera kalt* ‘be cold’ and *vera illt* ‘be hurt’. The IceCASE corpus does not contain adjectival predicates but the respective frequencies of these predicates in the data used for the IceCASE corpus is equivalent to the top 100 ranks for *vera kalt* ‘be cold’ and top 150 rank for *vera illt* ‘be hurt’.



**Figure 3.1.** Stimuli for the novel and semi-novel predicates in the forced-choice production task.

In the object condition, all the verbs are completely novel. The items provide contrasts between atypical and typical dative semantic contexts. There is ruining and throwing (*feira* and *rona*), with an affected patient and a ballistic motion theme. In both cases the object is inanimate as can be seen in Figure 3.1. Many verbs could qualify as equivalent to the novel verb *feira* ‘ruin’, for example the accusative object verbs *skemma*

'damage' (rank 155), *eyðileggja* 'ruin, destroy' (rank 176) and *kremja* 'squish' (not present in the IceCASE corpus data) but also *rústa* 'destroy' which was discussed in subsection 3.2.1 and appears both in the accusative and the dative (not present in the IceCASE corpus data). Equivalent to the novel verb *rona* 'throw' could be the dative object verbs *henda* 'throw away' (rank 87) or *kasta* 'throw' (rank 181). I also contrasted novel verbs of hitting and pushing events (*goða* and *rona*), differentiated by movement as well, with animate objects. Possible equivalents for *goða* 'hit' are the accusative object verbs *meiða* 'hurt' (rank 55), *slá* 'slap' (rank 168) and *lemja* 'hit' (rank 199) while equivalents for the novel verb *rona* 'throw' could be the dative object verbs *ýta* 'push' (rank 83) and *hrinda* 'push (a person violently)' (rank 282). Finally, I tested benefactive goals as direct (*darga* and *blíma*) and indirect (*merta*) objects. Equivalents for *darga* 'love' could be for example the accusative object verbs *elska* 'love' (rank 182) and *dýrka* 'worship' (not in the data used for the IceCASE corpus), and for *blíma* 'hug', the accusative object verbs *knúsa* 'hug' (rank 291) and *faðma* 'hug' (not in the data used for the IceCASE corpus). For the novel verb *merta* 'give', the most direct equivalent would be *gefa* 'give' (rank 28), the most common double-object verb in the IceCASE corpus (with a dative indirect object and accusative direct object). Testing all these different contexts for novel verbs, in addition to the existing verbs in subsection 3.2.1, should therefore provide important insights into the nature of non-default dative productivity in the acquisition of Icelandic.

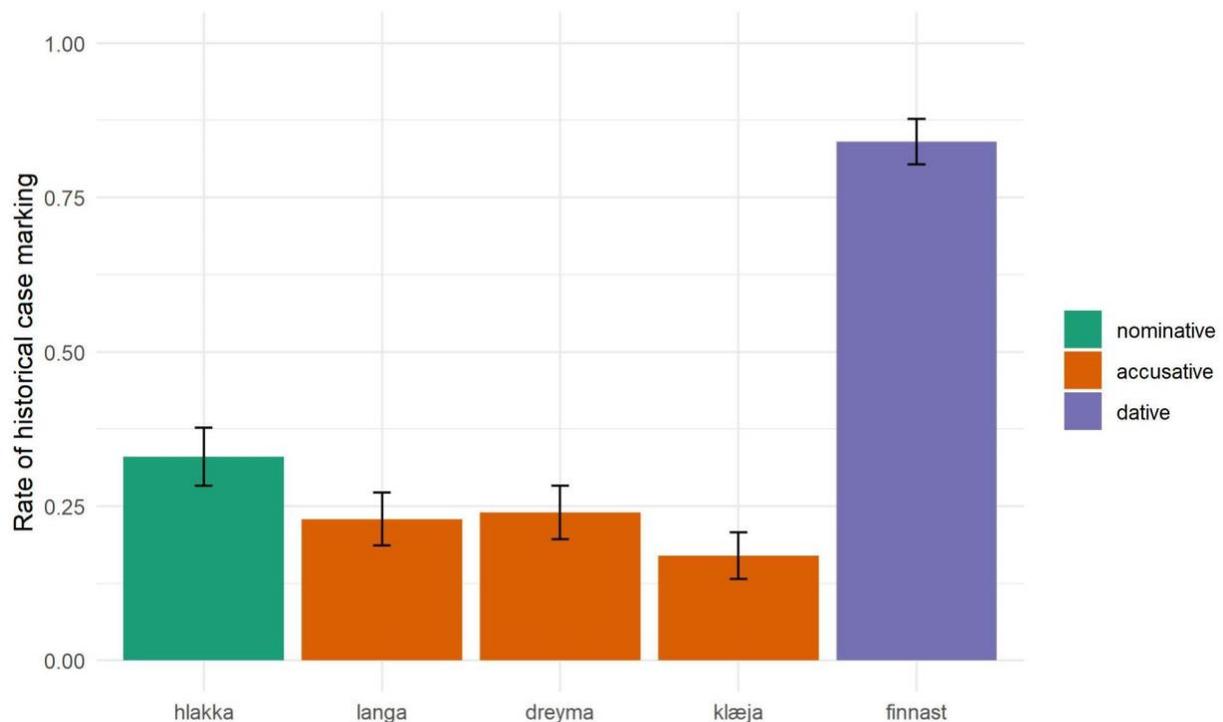
### 3.3 Results

As described in subsection 3.2.1, the forced-choice production task consisted of 1) existing verbs presented orally and in written form and 2) novel verbs presented with descriptive pictures. I begin with a descriptive overview of the forced-choice results with existing verbs before moving on to the novel verb task. The results are both presented across age and within age groups, but note that the older age groups have more participants (N = 17 for 3–5;11 years old, N = 32 for 6–8;11 years old, N = 33 for 9–11;11 years old and N = 21 for 12–13;11 years old) and therefore usually smaller 95% confidence intervals.

### 3.3.1 Existing verbs

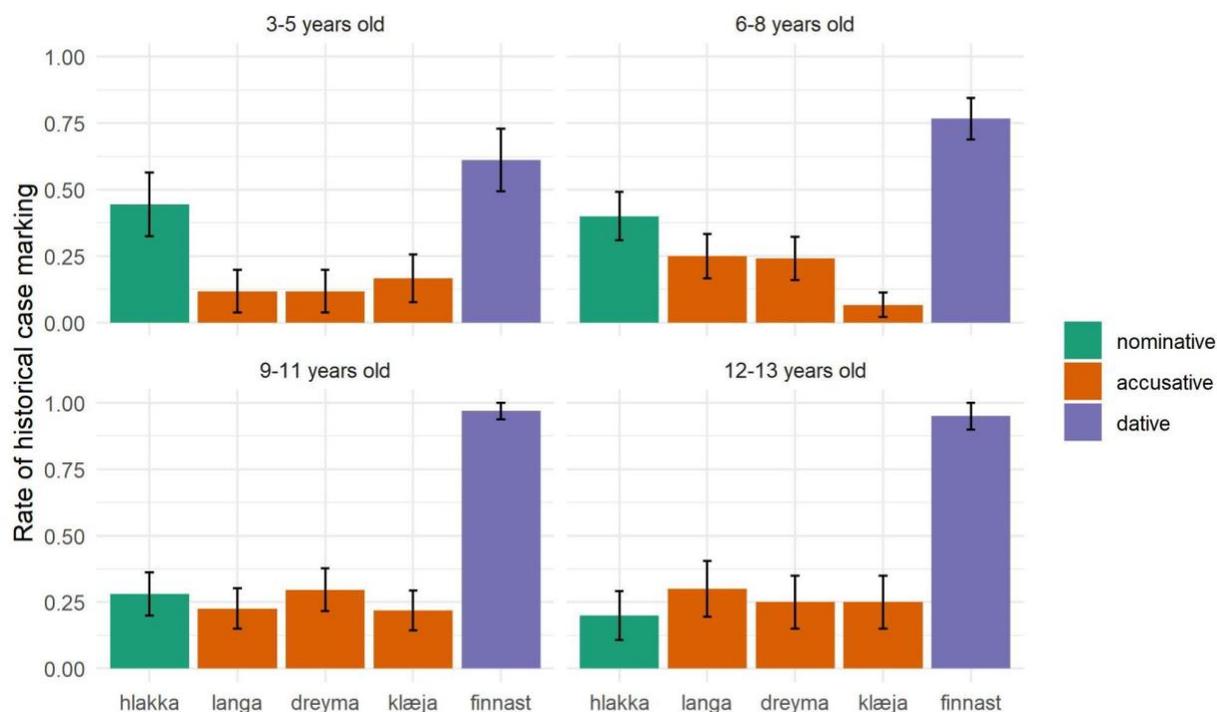
#### 3.3.1.1 Subject case

For the existing verbs, I targeted both subject and object case as described in subsection 3.2.1. I begin by describing the results for the subject case selection, where five experiencer subject verbs were tested, four Dative Substitution verbs and the most common historically dative subject verb *finnast* 'find'. One of the four DS verbs, *hlakka* 'look forward to', historically takes a nominative subject but also appears with both accusative and dative subjects in Modern Icelandic. Figure 3.2, where the rate of historical (and prescriptively standard) case marking is coded, shows that this happens to be the DS verb which appears the most with its historical case marking, in contrast with the three historically accusative DS verbs, ordered in Figure 3.2 based on their frequency/appearance in the IceCASE corpus.



**Figure 3.2.** Forced-choice production task with existing verbs, historical subject case. As can be clearly seen, the low rate of historical subject case in the DS verbs is contrasted with the much higher rate of historical dative with *finnast* 'find'. This can be found across

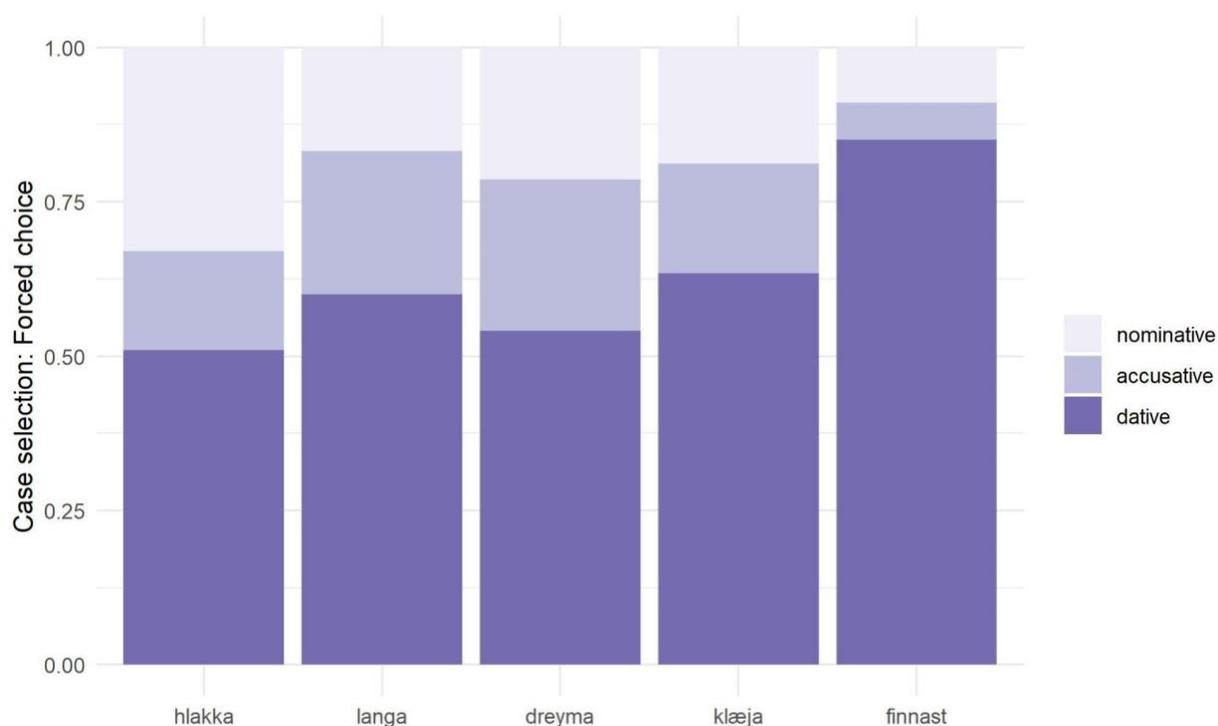
age groups, as Figure 3.3 shows, but notice that the contrast is much smaller for the younger age groups, which have lower rates of dative with *finnast* ‘find’.



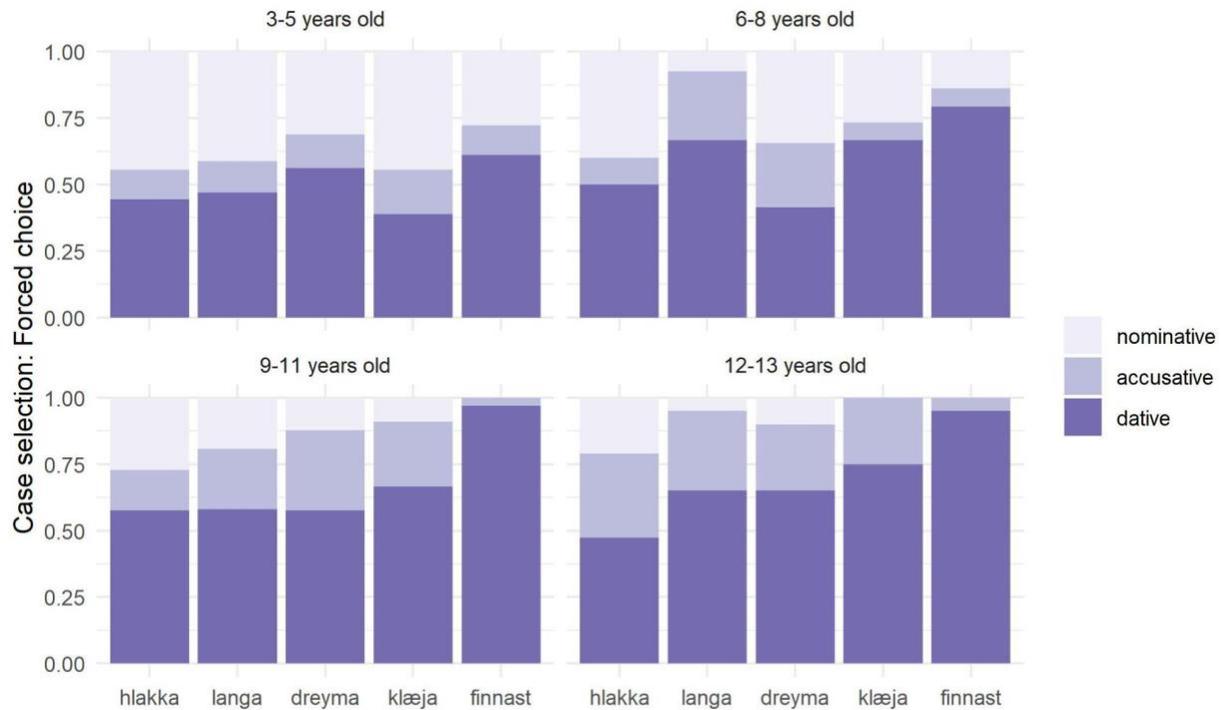
**Figure 3.3.** Forced-choice production task with existing verbs by age group, historical subject case.

As can be seen in Figure 3.4 below, this is explained by a higher rate of nominative overgeneralizations in younger children. These nominative overgeneralizations in younger children also seem to drive the higher (nominative) historical case marking of *hlakka* ‘look forward to’ as compared to the historically accusative DS verbs. This difference disappears with age, with the historical nominative case for *hlakka* ‘look forward to’ being selected less as the children grow older, which points towards increased productivity of non-nominative for experiencer subjects with age. Despite this, the historically accusative DS verbs do not show a clear age trend which might have been expected under an age-grading approach (or simply one where rarer irregular forms are acquired with age), namely more of the historical accusative as the children grow older. Although it has been established that children acquire a paradigm of intra-speaker

variation with these verbs (Nowenstein 2017), instead of e.g. fully acquiring the accusative, the apparent lack of a clear age trend is also explainable when the nominative overgeneralizations are taken into account. Thus, Figure 3.5 shows that they become less frequent with age, being replaced by the dative. Figure 3.5 also shows more clearly that although a clear age trend is missing for the accusative case marking, the rate of the original accusative is for example higher in the oldest age group than the youngest one.



**Figure 3.4.** Forced-choice production task with existing verbs, subject case.



**Figure 3.5.** Forced-choice production task with existing verbs by age group, subject case.

Indeed, breaking down the case selection as is done in Figures 3.4 and 3.5 helps provide a fuller picture of the results. The more frequent nominative overgeneralizations in the younger age groups do not come as a surprise, but various other results in Figures 3.4 and 3.5 might be considered unexpected.

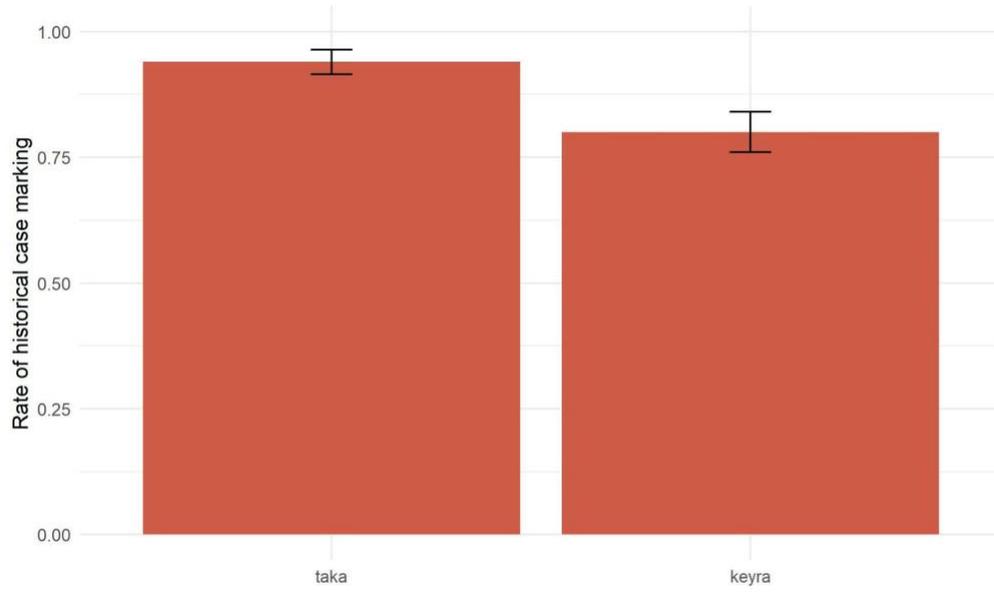
First, it is interesting to see that the nominative overgeneralizations in historically accusative verbs (except for *klæja* ‘itch’) do not disappear completely in the oldest age group, as they do with the historically dative *finnast* ‘find’. This is in line with results from Guðmundsdóttir et al. (2018), where it was found that contrary to previous assumptions (e.g. Jónsson & Eythórsson 2005), Nominative Substitution is also found with experiencer verbs. Second, the non-dative chosen with the relatively stable historically dative *finnast* ‘find’ does not only consist of nominative overgeneralizations. Instead, some children seem to be able to use the accusative productively with experiencers. This is not in line with various approaches to DS in the literature (e.g. Jónsson & Eythórsson 2005 and Barðdal 2008) but fits more recent narratives (Nowenstein 2017, Guðmundsdóttir et al.

2018 and Nowenstein & Ingason 2021) where grammatical conditioning factors result in diffused directionality for well-known instances of language change. This is further discussed in Chapter 5. Third, the results here do not clearly replicate Jónsson & Eythórsson's (2003) result that DS rates are higher with lower frequency verbs, even though *klæja* 'itch' has the highest rate of dative across age groups.

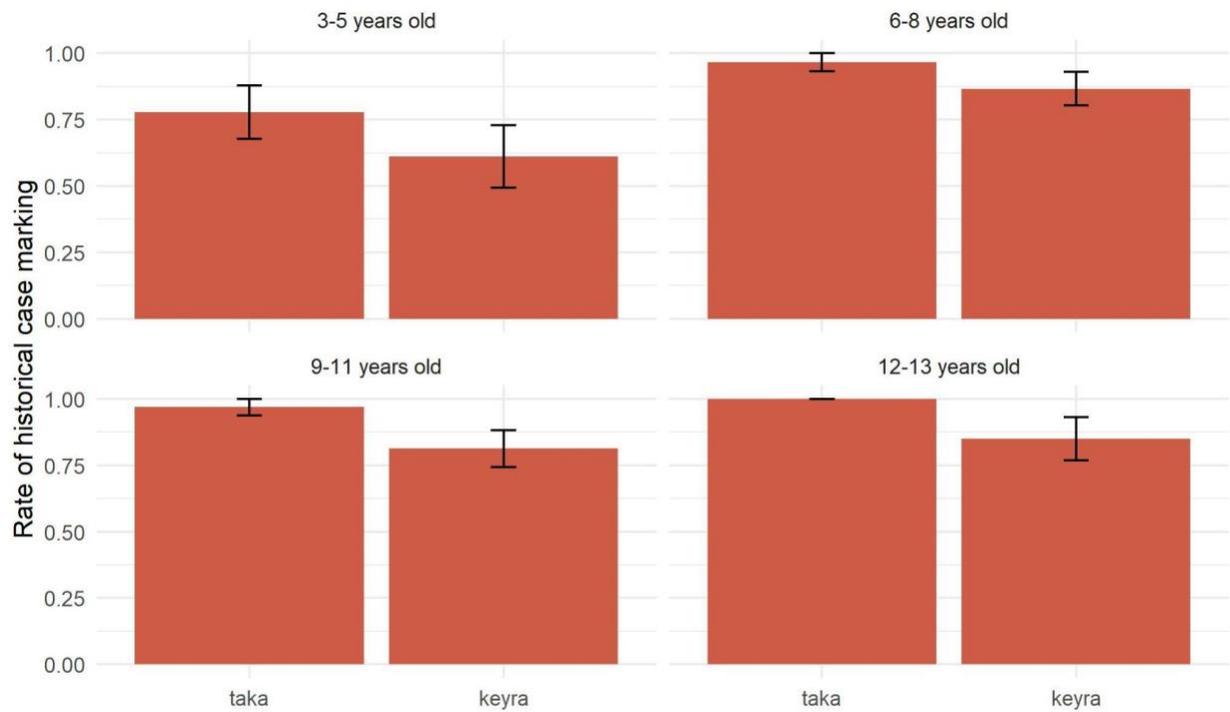
Regardless of these more unexpected results, the main result to retain for now is that children do use the dative for experiencer subjects with existing verbs, both verbs which historically take a dative subject and verbs that do not. In general, and mostly in contrast with nominative overgeneralizations, dative with subjects increases with age. But what about dative case marking with objects?

### 3.3.1.2 Object case: Accusative

I begin by describing the results for the two verbs which traditionally appear with the structural/default (most frequent) accusative on their objects, *taka* 'take' and *keyra* 'drive' (general results in Figure 3.6). There are no known reports of variation in the object case of *taka* 'take'. Still, the argument in object position is not a prototypical change-of-state patient but a theme which undergoes a kind of change in location/possession. *Taka* 'take' also is one of the verbs which can appear with an optional reflexive dative benefactive (see subsection 1.1.2). Additionally, *taka* 'take' can appear with a dative in the sense of taking something (such as news or criticism) well or badly, but this use does not appear in the child language corpus reported on in subsection 3.3.1. On the other hand, it has been reported that a number of speakers can use *keyra* 'drive' in the dative, particularly if the argument is sentient (Barðdal 1993, Maling 2002, Jónsson 2009), which is the case for the object in the test sentence used in the production task as seen in example (24b), *keyra stelpuna/stelpunni* 'drive the girl (ACC./DAT.)'.

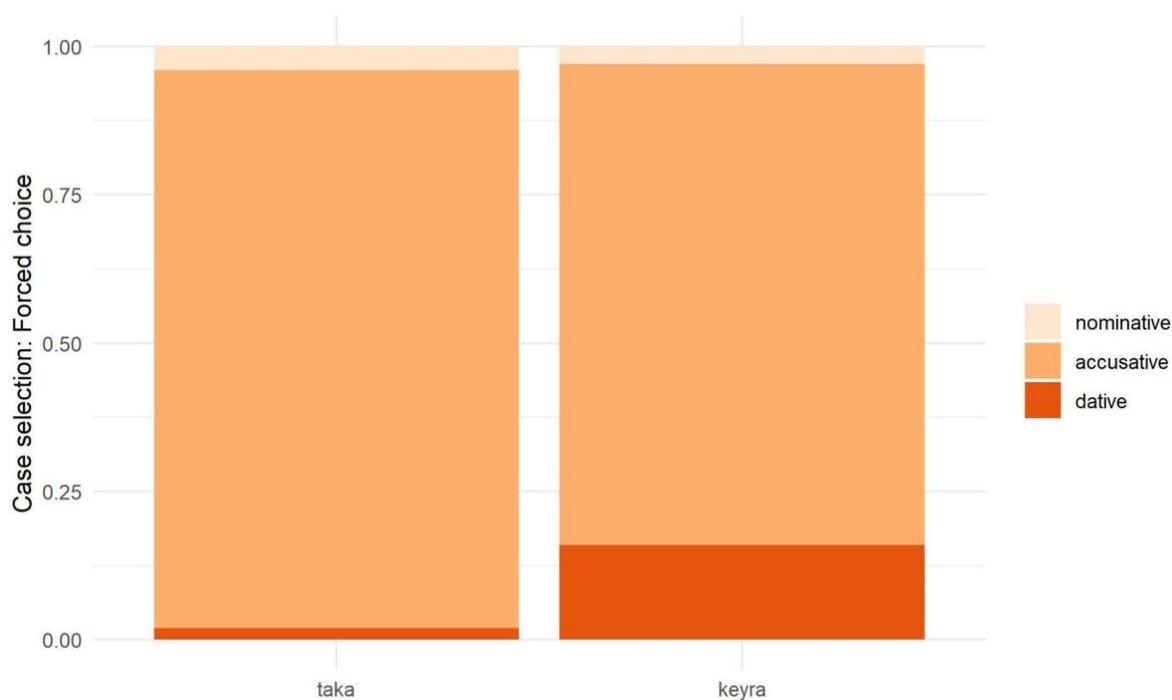


**Figure 3.6.** Forced-choice production task with existing verbs, historical accusative object case.



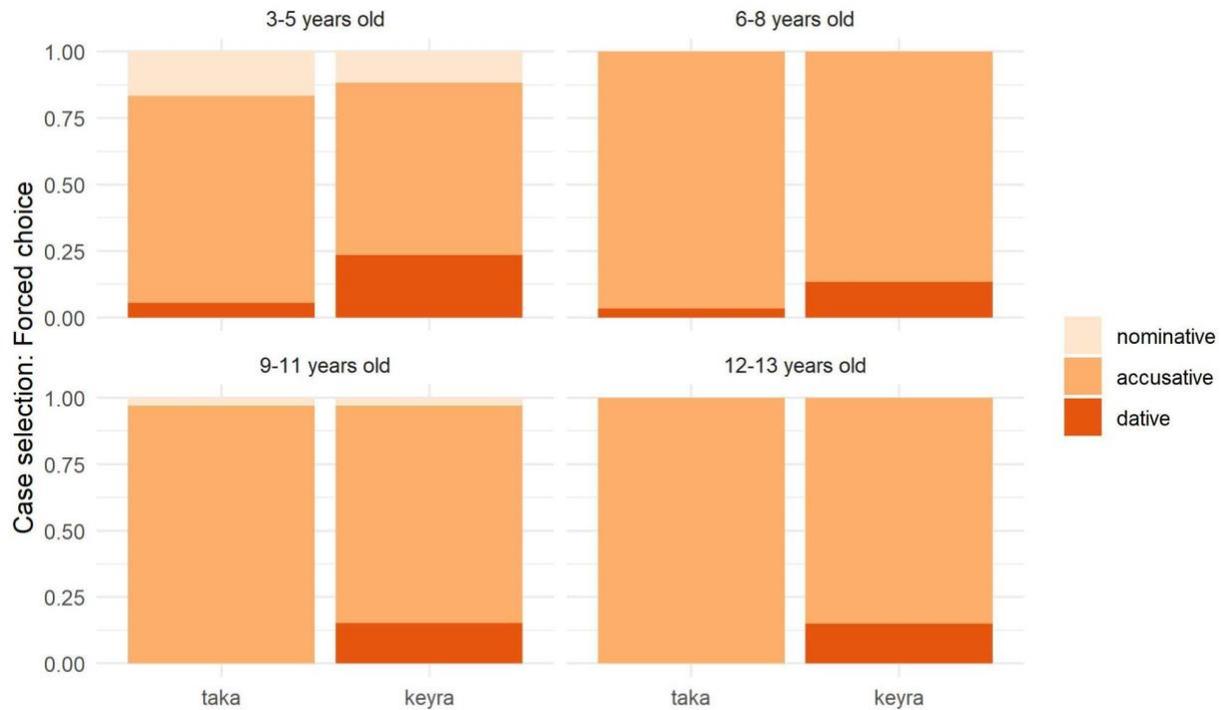
**Figure 3.7.** Forced-choice production task with existing verbs by age group, historical accusative object case.

As can be seen in Figure 3.6, there is a difference in the rate of the historical accusative with the two verbs, with accusative being selected more with *taka* ‘take’ than *keyra* ‘drive’. It is interesting that this contrast is not as clear in the youngest age group, as Figure 3.7 shows. When breaking down the results further though (Figure 3.8), we see that the nominative overgeneralizations play a role again, making the contrast look smaller. In general, the rate of nominative overgeneralizations is similar between the two verbs. The contrast between these verbs is therefore mostly found in the production of the dative.



**Figure 3.8.** Forced-choice production task with existing verbs, object case with historically accusative verbs.

Surprisingly, the dative still is marginally chosen with *taka* ‘take’, but as Figure 3.9 shows, this is only in the youngest two age groups. This could be interpreted in various ways, one of them being a task effect where the results would be expected to be noisier with younger children. The second interpretation would be that younger children show a more general dative productivity early on.



**Figure 3.9.** Forced-choice production task with existing verbs by age group, object case with historically accusative verbs.

In any case, the main information which should be retained from the results with the historically accusative objects is that the accusative is chosen predominantly early on, with nominative overgeneralizations disappearing quickly (in contrast to the non-nominative subject case). Notice though that there is a marginal selection of nominative in the 9–11 years group which I cannot explain.<sup>32</sup> Despite the default/structural/distributionally dominant position of the accusative, variation with the dative can appear in semantically relevant contexts such as animacy/sentience. Still, the present results cannot establish whether dative with a verb like *keyra* ‘drive’ is the result of overgeneralizations based on the productivity of the dative for animate (non-patient) arguments or conditioned variation acquired from the input as was discussed in regard to example (24b). Next, I investigate the acquisition of dative objects.

<sup>32</sup> The nominative overgeneralizations in object position throughout the experimental data might be in part a task effect but they are also attested in corpora (Sigurðardóttir 2002) and might be a remnant of a stage in acquisition where all arguments are marked with the nominative.

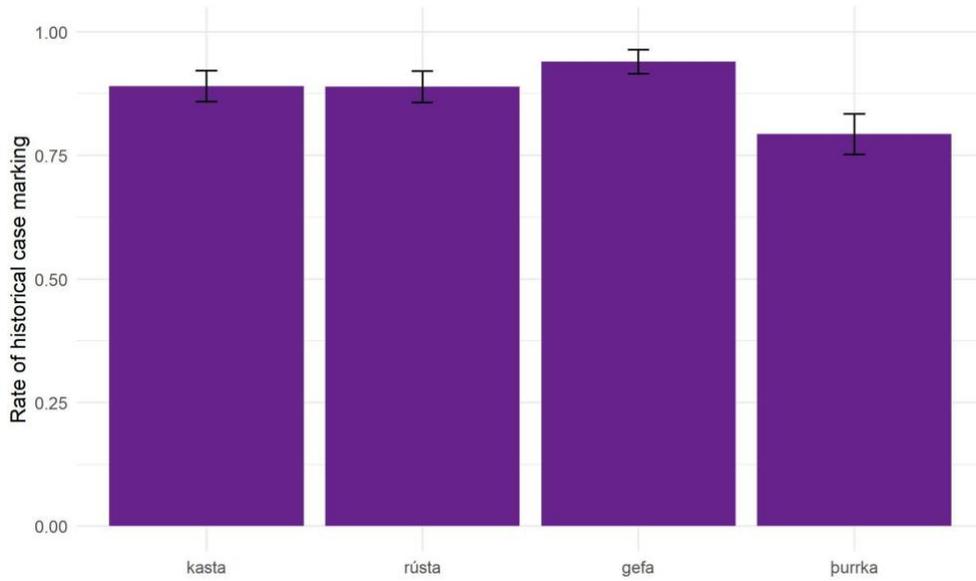
### 3.3.1.3 Object case: Dative

As discussed in subsection 3.2.1, the dative object verbs tested in the forced-choice task consist of one ballistic motion verb, *kasta* 'throw', one verb of splintering/shattering, where the dative has almost completely replaced the accusative,<sup>33</sup> *rústa* 'demolish' (Maling 2002), one of the most common ditransitive verbs, *gefa* 'give', which takes an indirect dative object, and finally one alternating verb, *þurrka* 'dry', where the dative has an animacy requirement while the accusative is used with affected patients and e.g. body parts. Figures 3.10 and 3.11 show the rate of chosen dative for the objects of these verbs in the forced choice task.

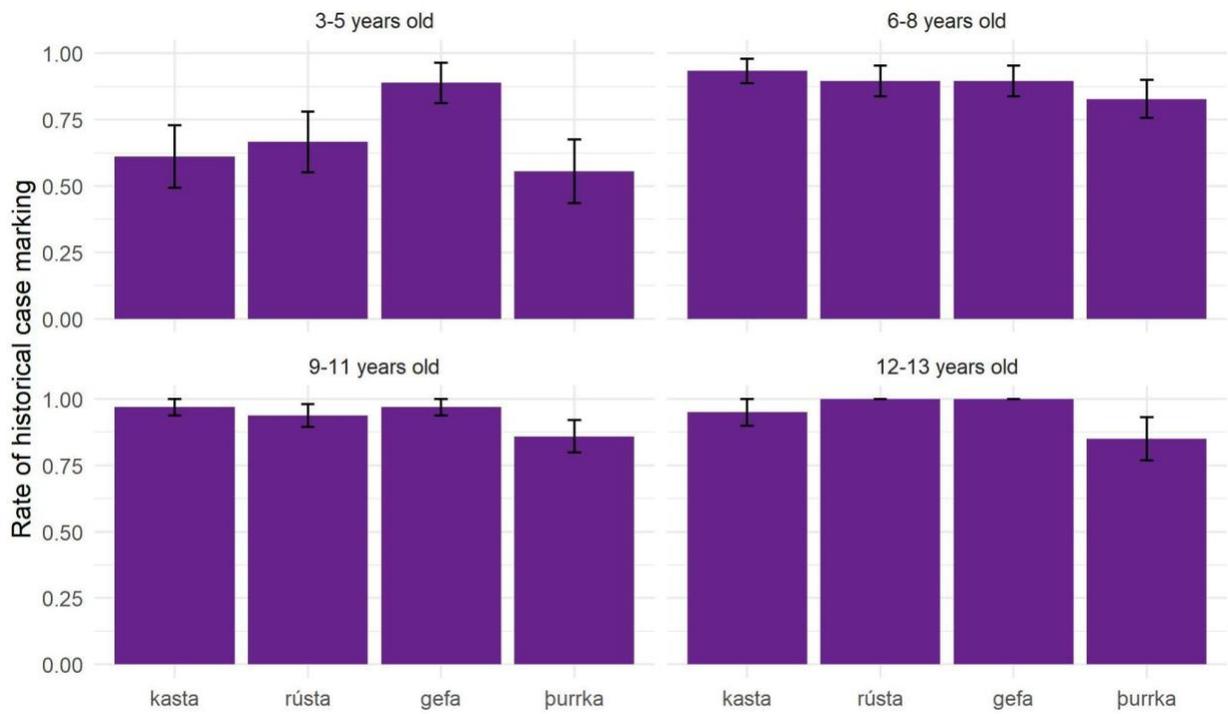
The general picture emerging in Figure 3.10 is that the dative rate is slightly higher in the indirect object of *gefa* 'give' and lower with the alternating verb, *þurrka* 'dry', as compared to *kasta* 'throw' and *rústa* 'demolish'. For *gefa* 'give', the rate of dative is comparable to the rate of accusative with *taka* 'take' (see Figure 3.6 in subsection 3.3.1.2), despite the fact that it is non-structural/non-default. Looking at the results by age group, as in Figure 3.11, it is clear that the contrast between direct and indirect objects can mostly be traced to the distribution in the youngest age group, where it is more pronounced than in the other age groups. This is in line with previous results (Sigurðardóttir 2002), which have shown that children learning Icelandic seem to produce datives on indirect objects consistently sooner than on direct objects. The lower dative rate with *þurrka* 'dry', on the other hand, can be found across age groups, in line with the fact that there is variation in the input. As before, the results are broken down more accurately in Figures 3.12 and 3.13.

---

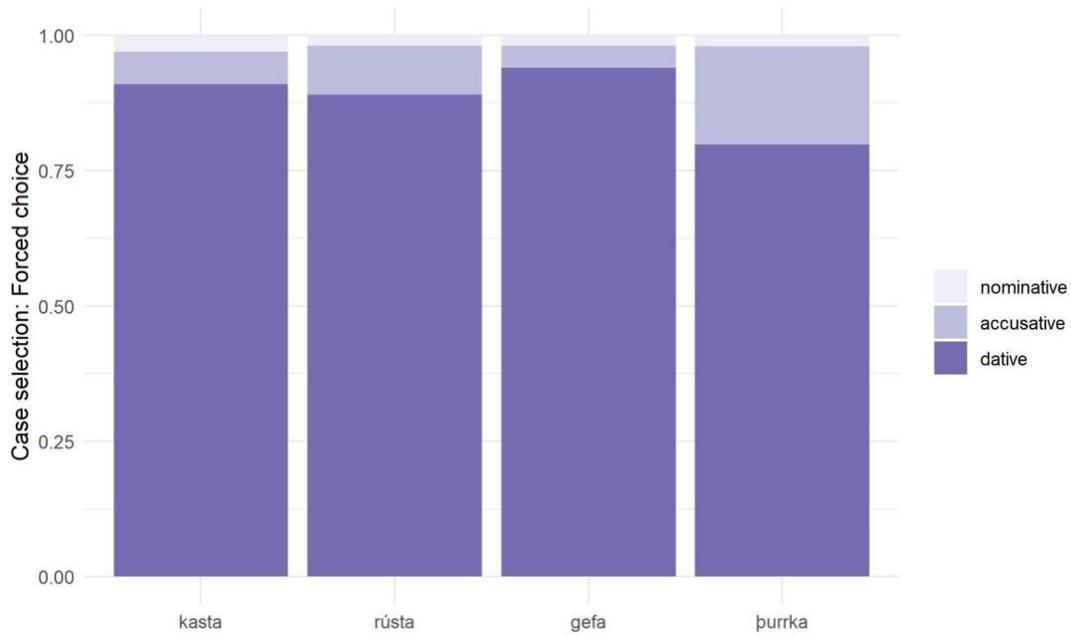
<sup>33</sup> The term 'historical' case marking therefore does not apply as well here but is kept for the sake of consistency.



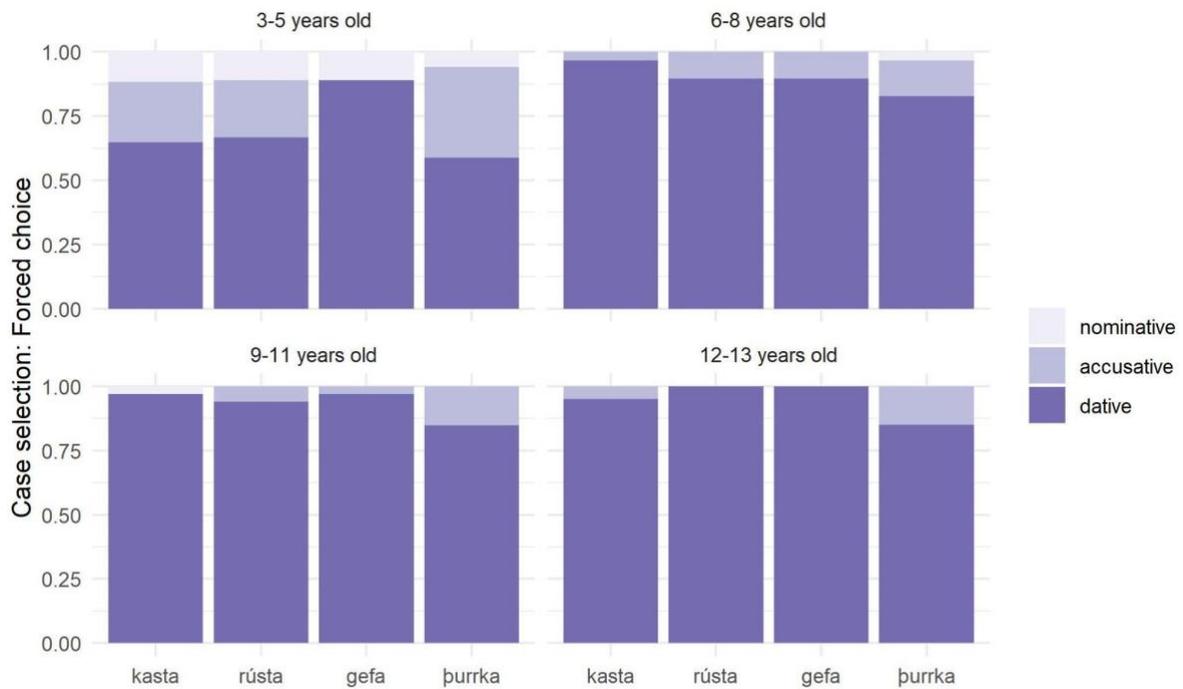
**Figure 3.10.** Forced-choice production task with existing verbs, historical dative object case.



**Figure 3.11.** Forced-choice production task with existing verbs by age group, historical dative object case.



**Figure 3.12.** Forced-choice production task with existing verbs, object case with historically dative verbs.

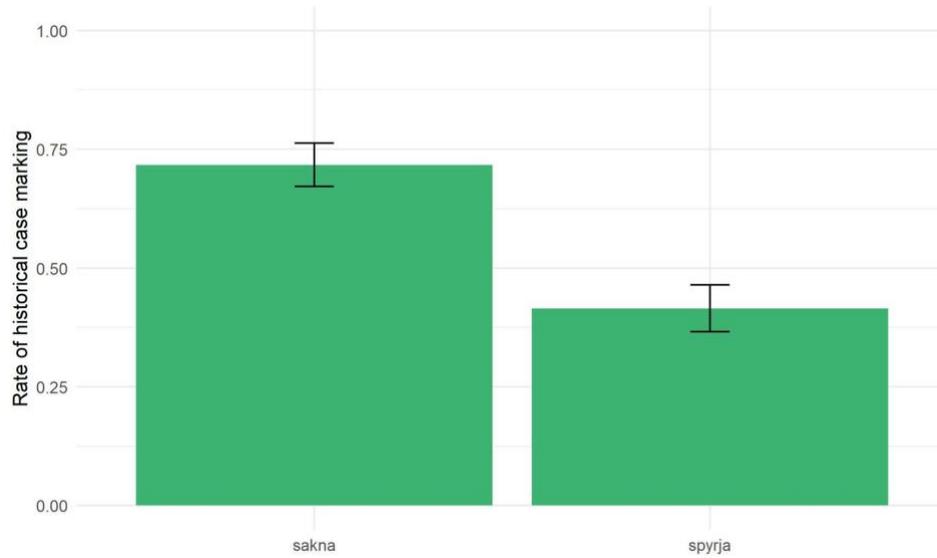


**Figure 3.13.** Forced-choice production task with existing verbs by age group, object case with historically dative verbs.

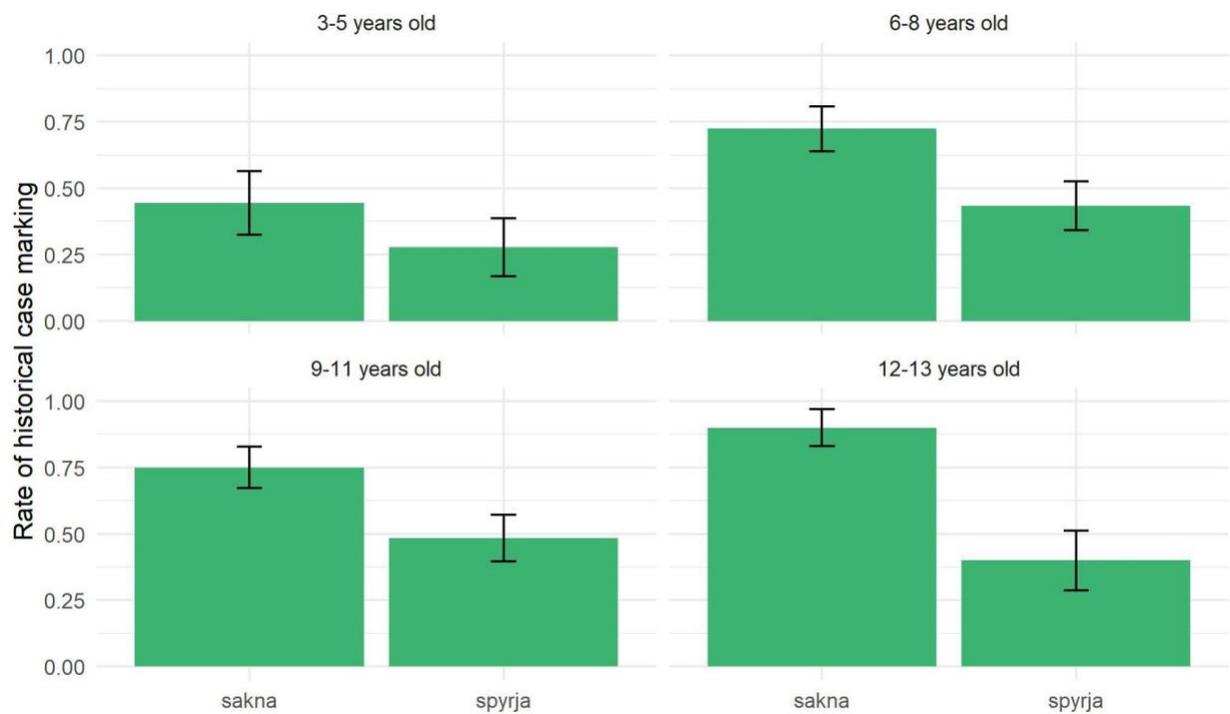
Taking into account the different distribution of nominative and accusative choice does not change the picture just outlined, as the nominative overgeneralizations are quite stable across verbs as well as being almost non-existent beyond the youngest age group. Therefore, the main finding to keep in mind here is the early contrast between datives in direct and indirect objects, as well as the fact that the dative can in certain contexts be as well established as the accusative for direct objects early on. This is somewhat in contrast with the results for dative with subjects, see subsection 3.3.1.1. Additionally, there is an interesting lack of difference between the high frequency verb *kasta* ‘throw’ and the low frequency verb *rústa* ‘demolish’ across age groups. Before looking at the dative with recent borrowings, we turn our attention to genitive objects.

#### 3.3.1.4 Object case: Genitive

The two verbs used to test genitive objects are *sakna* ‘miss’ and *spyrja* ‘ask’ (as was described in 3.2.1). *Sakna* ‘miss’ is the most frequent verb which consistently appears in the genitive (although it still does not appear in the data used in the IceCASE corpus) as *spyrja* ‘ask’ most frequently appears with a PP or in the accusative. The most common nominal object for *spyrja* is the noun *spurning* ‘question’ which therefore was used to test this verb, but it presents accusative-dative syncretism. Figure 3.14 shows the general results for the use of the historical genitive, which appears much more frequently with *sakna* ‘miss’ (although still only reaching 72% of choices). Figure 3.15 shows that the use of the genitive is the lowest in the youngest age group and increases with age, with this trend appearing more clearly in the case of *sakna* ‘miss’ than *spyrja* ‘ask’. The contrast mostly appears between the youngest age group and the rest, instead of a steady increase in genitive as the children get older.

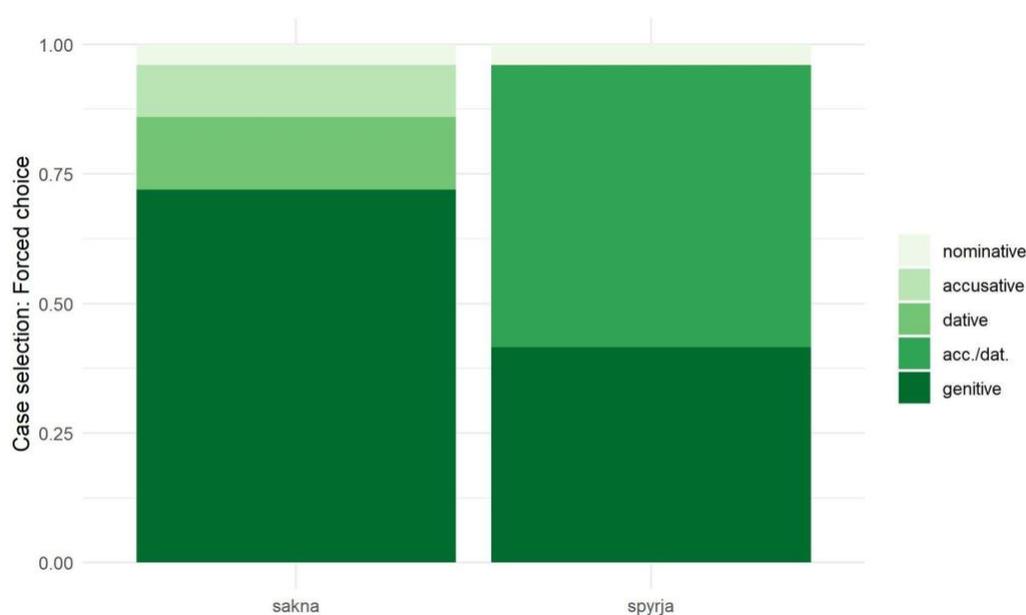


**Figure 3.14.** Forced-choice production task with existing verbs, historical genitive object case.



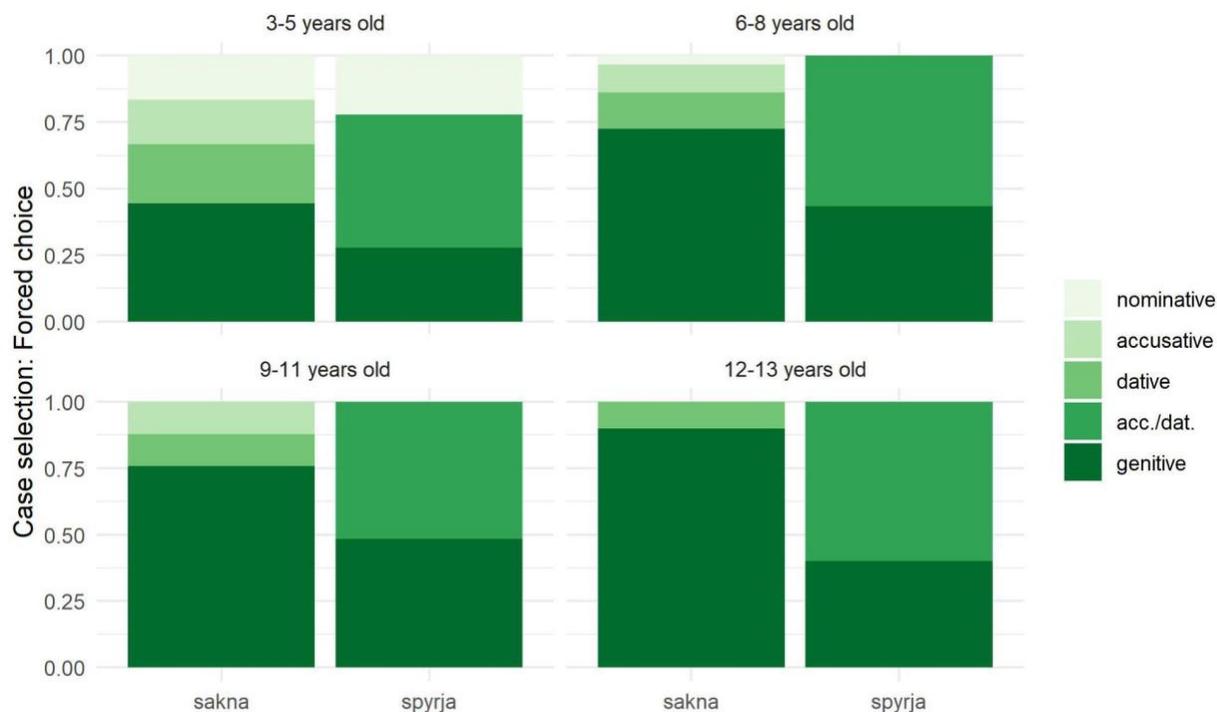
**Figure 3.15.** Forced-choice production task with existing verbs by age group, historical genitive object case.

When breaking the results down by case choice (Figure 3.16), a few interesting additional points emerge. First, nominative overgeneralizations seem to be stable across the two verbs. Second, an unexpected amount of dative is chosen with *sakna* ‘miss’ (however remember that the object of *spyrja* ‘ask’ has accusative-dative syncretism). That is, children do not only overgeneralize the structural/default/distributionally dominant accusative, they also choose the dative for the historically genitive object.



**Figure 3.16.** Forced-choice production task with existing verbs by age group, historical genitive object case.

Figure 3.17 shows that the dative is most prominent in the answers of the youngest age group but still the dative is present throughout. Unexpectedly, in the oldest age group, there is only evidence of dative overgeneralizations and not accusative ones (which are present in the younger three age groups). That is, nominative and accusative overgeneralizations disappear with age but the dative overgeneralizations persist. The object in question was the pronoun *hún* ‘she’, indicating an animate argument, and an interesting question is whether the dative overgeneralizations would not have been as frequent if the object would have been unambiguously inanimate.



**Figure 3.17.** Forced-choice production task with existing verbs by age group, object case with historically genitive verbs.

For the current purposes, the main findings in the historically genitive object results are the low rate of genitive (which does not come as a surprise considering type and token frequency) and the high rate of dative overgeneralizations for *sakna* ‘miss’ in the youngest age group. However, the rate of the genitive increases in the older age groups and the dative overgeneralizations decrease but still are maintained even in the oldest age group. For *spyrja* ‘ask’, the difference in genitive rate is not as steep and might also to a higher degree be explained by regressing nominative overgeneralizations. In general, these results with *sakna* ‘miss’ are yet another demonstration of the robust dative productivity found in Icelandic.<sup>34</sup>

<sup>34</sup> For more on variation with genitive arguments in Icelandic, see e.g. Jónsson (2017).

### 3.3.1.5 A note on borrowings

The productivity of the dative has perhaps, aside from Dative Substitution, mostly been discussed in the context of recent borrowings, most prominently by Barðdal (2008) and Jónsson & Thórarinsdóttir (2020). The main results from those previous studies were discussed briefly in section 1.2. They show that the default/structural/distributionally dominant nominative subjects and accusative objects are the most prominent pattern in the adult language and that no recent borrowing appears with a genitive object or a non-nominative subject (but see Guðmundsdóttir et al. 2018). When it comes to dative objects, it is clear that the dominant view is that datives appear when there is a translational substitute that takes a dative object (even though this option is not considered to be available for the genitive) or when a semantic cluster associated with the dative can attract the verb, e.g. caused motion. In Jónsson & Thórarinsdóttir (2020), the caused motion semantics are the only available abstract category for dative case marking, but Barðdal (2008) assumes that other schemas associated with the dative can attract recent borrowings as well, although the caused motion cluster is considered to be the “most entrenched level of the Nom-Dat construction” (p. 85).

To my knowledge, no studies on the case marking of borrowings or newly coined verbs in Icelandic have been conducted in the context of language acquisition. One reason might be that some of these verbs are acquired quite late. In the production task described here, children 6 years old and older chose case with four technological borrowings: *gúggla* ‘google’, *dánlóða* ‘download’, *kópípeista* ‘copypaste’ and *stríma* ‘stream’. 78 children participated in this part of the study, only choosing the object case for the verbs they claimed to know. The object of the verb *gúggla* ‘google’ (N = 78) received the accusative in 88% of cases (dative 8% and nominative 4%), *dánlóða* ‘download’ (N = 76) received the dative in 79% of cases (accusative 20% and nominative 1%), *kópípeista* ‘copy paste’ (N = 63) received accusative in 76% of cases (dative 24%) and *stríma* ‘stream’ (N = 67), which does not have a clear translational synonym nor caused motion semantics, received accusative in 51% of case (dative 45% and nominative 4%). In Jónsson & Thórarinsdóttir’s (2020) survey, *gúggla* ‘google’ was accusative in 93.6% of cases and *dánlóða* ‘download’ was dative in 93.1% of cases. *Peista* ‘paste,’ and not *kópípeista* ‘copy paste’, was tested, receiving accusative in 47.5%

of cases, slightly more than the dative. *Stríma* ‘stream’ was not tested. Although the case marking of loan verbs in both child and adult language needs to be investigated further, this very brief comparison indicates that the case marking of certain verbs is not as well established for children as for adults. The comparison also shows that the dative is very much available as an option for objects of new verbs, not necessarily only when a translational synonym is available or when caused motion semantics are at play. Finally, the children participating in the forced-choice task also chose a case with the loan verb *kreiva* ‘crave’, which was tested in the novel verb part of the experiment, as no participants claimed they knew it (see more in subsection 3.3.2.1). This is the only loan verb which has been reported to appear (albeit very rarely) with non-nominative subjects. The participants (N = 94) did select the nominative most frequently for the subject of this verb with a rate of 46% but non-nominatives reached 54% (38% dative and 16% accusative), showing that children are prepared to select non-nominative case for the subjects of novel verbs in a production task.

I end this brief note on borrowings with the results from an informal poll conducted on Twitter, where I asked (adult) speakers to select the accusative or dative with the object of *cancella* ‘cancel’ in the ‘cancel culture’ sense. 267 speakers participated, 52.4% choosing the dative. I then added a question with the same verb in the more conventional canceling sense which 203 speakers answered, 82.3% of them choosing the dative.<sup>35</sup> We therefore see semantically motivated conditioning, with the affected patient reading being more biased towards the accusative than the (motion) theme reading. This is despite the fact that the cancel culture interpretation usually involves an animate object and that the two translational synonym candidates for *cancel*, *aflýsa* and *slaufa*, appear with dative objects. I mention this example here because it is an example of the broad scope of dative productivity in Icelandic, with multiple features probably playing a role simultaneously, resulting in extensive intra- and inter-speaker variation, even within predicates. To further establish the contexts in which the non-default dative can be used, I implemented the forced-choice task using novel verbs. The results are described in the next subsection.

---

<sup>35</sup> This was done by having people select the object *einhvorn* (someone.ACC) or *einhverjum* (someone.DAT) following the infinitive *að cancella* ‘to cancel’ for the cancel culture interpretation, and *eitthvað* (something.NOMACC) or *einhverju* (something.DAT) for the conventional canceling interpretation.

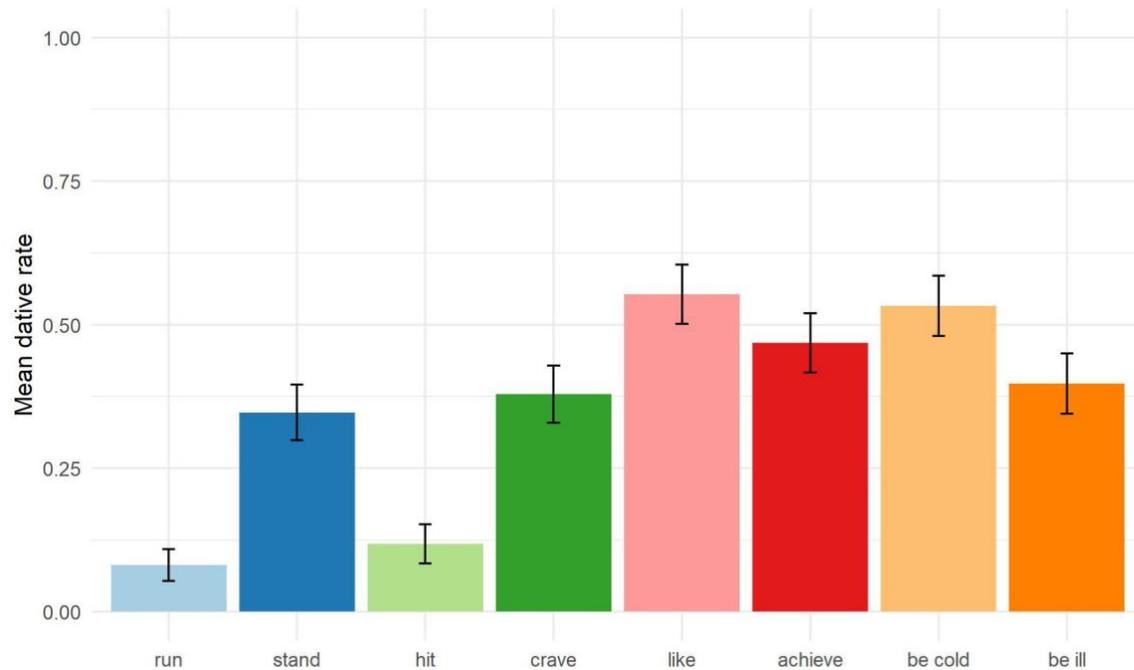
### 3.3.2 Novel verbs

As with the existing verbs, the novel verbs were part of the forced-choice production task which targeted both subject and object case. I begin by describing the results from the subject case items.

#### 3.3.2.1 Subject case

Figure 3.18 shows the results from the eight novel and semi-novel predicates used in the subject case condition of the novel verb forced-choice task (shown in the higher row of Figure 3.1 in subsection 3.2.2). The aim was to discern the forms and semantic contexts most likely to trigger dative subject productivity for Icelandic and to contrast the more typical contexts for dative subjects with atypical ones. The first two pairs of bars show such contrasts (intransitives in blue and transitives in green). The novel verbs presented with pictures of running and hitting agents clearly result in a lower rate of dative subjects as compared to novel verbs presented with pictures depicting arguments standing still (being bored) and desiring (craving) something.

The four remaining verbs all have higher rates of dative selected for their subjects. The *like* verb (*teilst*) which was presented with an *-st* suffix and modeled after *finnast* ‘find’, the most common dative subject verb, triggers the highest rate of dative subject selection. The *achieve* verb (*fípast*), which was also presented with *-st* but is modeled after the happenstance verb *takast* ‘succeed’, did not receive a dative subject as frequently. Finally, the two adjectival predicates presented with existing adjectives that typically do not appear with dative subjects, also received high rates of dative choice. The *be ill* (*vera vondur/vont*) predicate received less dative subjects than *be cold* (*vera frystur/fryst*) but has the highest rate of accusative subject selection (16.7%).

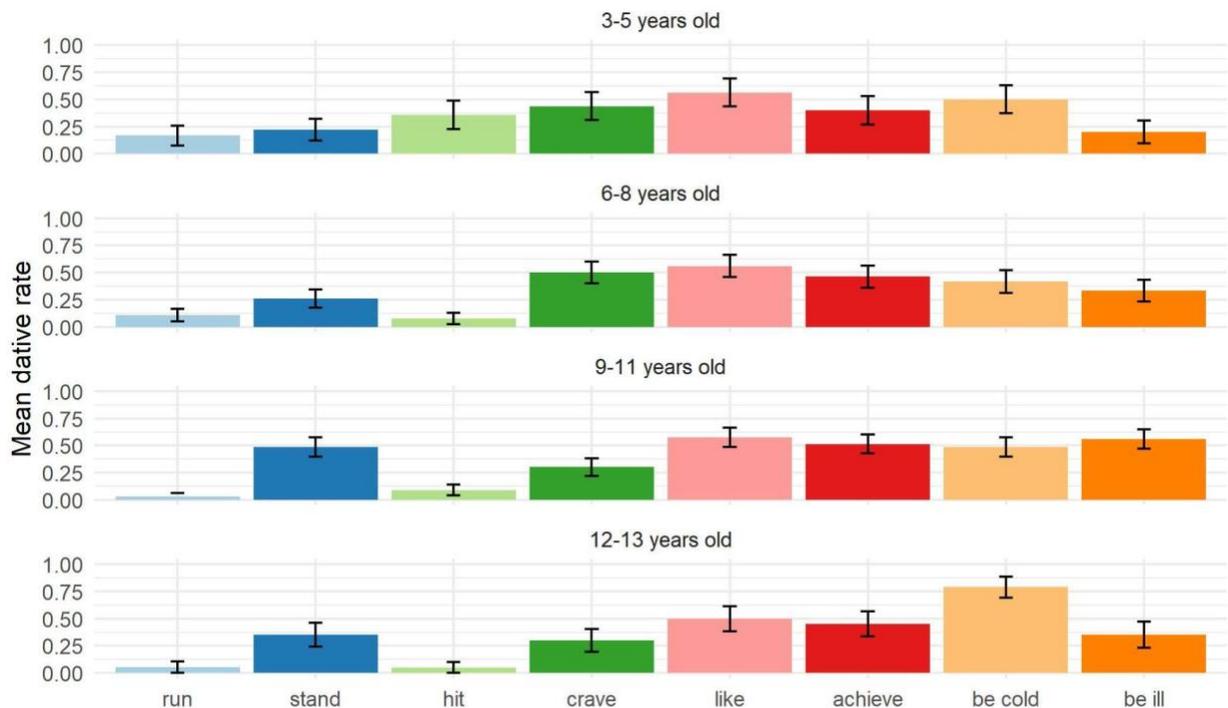


**Figure 3.18.** Forced-choice production task with novel verbs, subject condition.

The accusative subject selection was not included in the coding for Figure 3.18 for the sake of clarity and does not change the distribution depicted. Still, it is interesting to note that the dative is not the only non-nominative subject case produced, and that the choice of accusative seems to follow the same conditioning pattern: Higher with experiencers than agents. This pattern, where the dative is selected more frequently with experiencers rather than agentive subjects, is the relevant finding here, along with the fact that factors such as *-st* morphology also seem to prompt dative subject case in the context of experiencers. It should be noted that the *-st* verbs still are very diverse, and can for example appear with goal and theme subjects (see e.g. Anderson 1990 and Wood 2015).

Finally, it is important to acknowledge that the nature of the task, where all options are presented orally and the child chooses between them, might yield higher rates of unexpected case marking than in a naturalistic setting. I do not expect the rate of dative subjects with the novel verbs in the experiment, from 8% to 55%, to be representative of children's overgeneralizations or usage with newly learned verbs. The results should instead rather be interpreted in terms of contrasts between items, giving information about available case marking patterns and the semantic conditioning of productivity, rather than

being considered to be a quantification or measurement of productivity. In this context, it is interesting to note again the lack of contrasts between comparable higher and lower frequency items if the translational equivalents of the *-st* semi-novel verbs and adjectival predicates are taken into consideration. These possible task effects are reinforced by the age trend in Figure 3.19, where the results are broken down by age group.

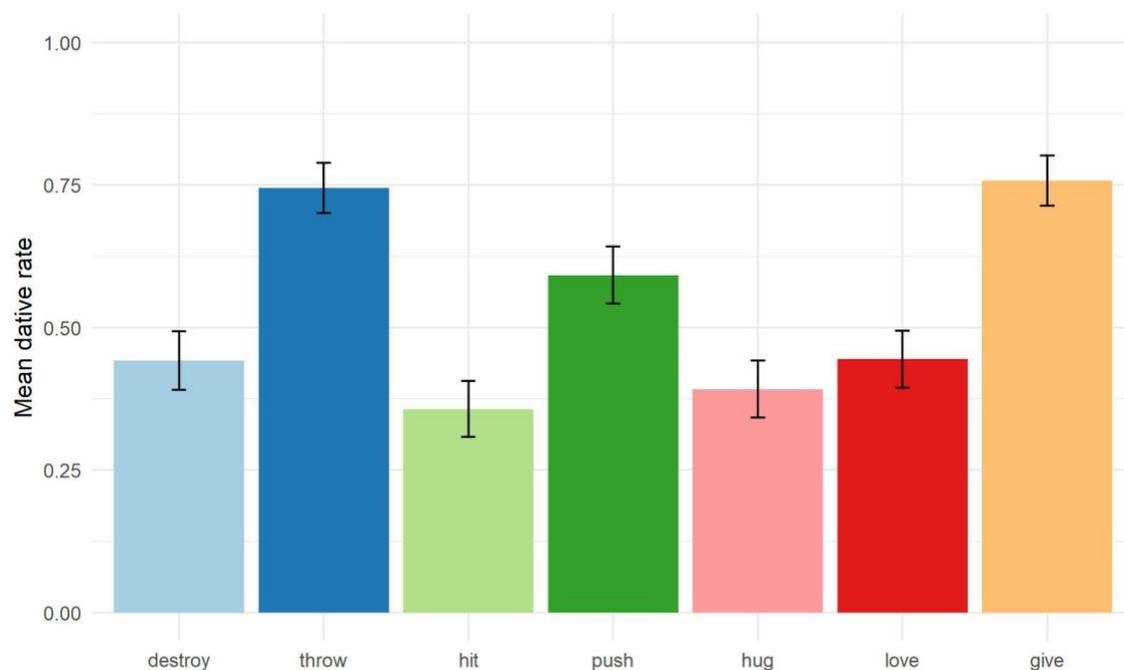


**Figure 3.19.** Forced-choice production task with novel verbs by age group, subject condition.

Relevant here is the lack of contrasts in the youngest age group, with the atypical dative contexts not necessarily appearing with fewer dative subjects. This might indicate that the children have not yet acquired the dative in the appropriate contexts and/or vary their choice of case in the task almost randomly. The contrasts are much clearer in the older age groups. Additionally, it is interesting to observe the higher rate of the dative for the *be cold* predicate in the oldest age group, since one of the only attested examples of dative overgeneralization in subject case outside DS is found with an adjectival predicate describing physical discomfort (see discussion in section 1.2).

### 3.3.2.2 Object case

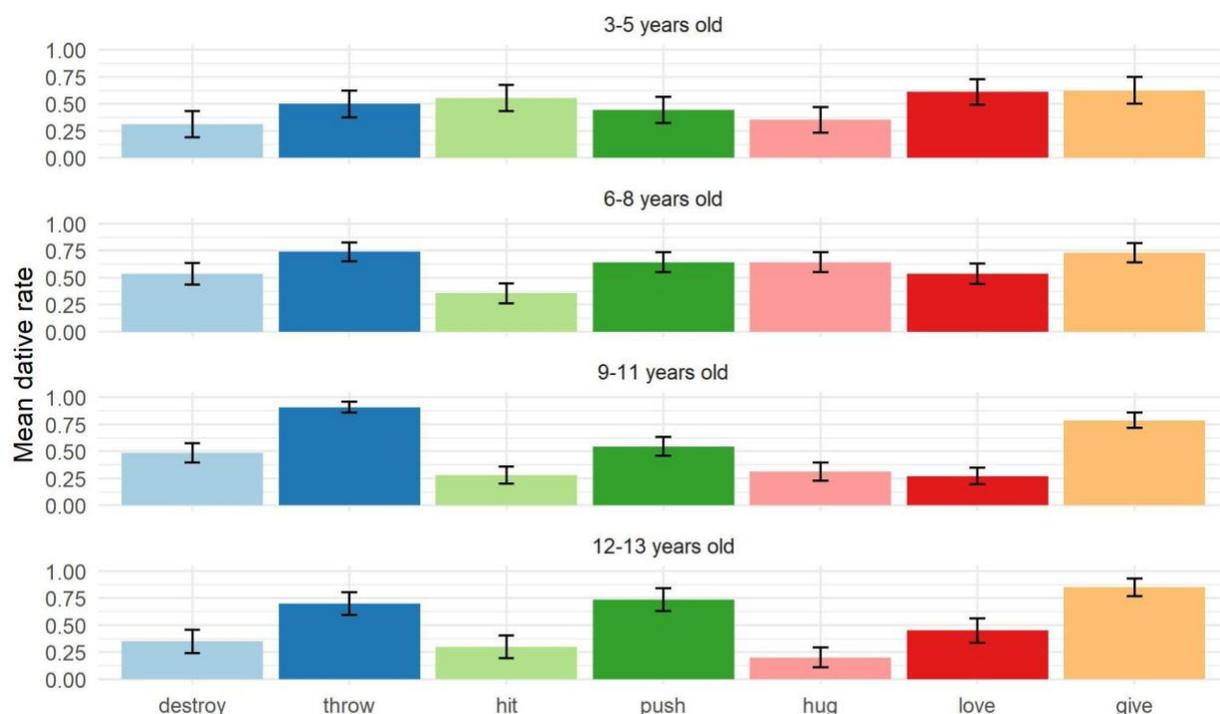
In the context of comparison between conditions, it is interesting to note that the dative is selected to a higher extent with objects than with subjects, as illustrated in a comparison of Figure 3.18 (in subsection 3.3.2.1) and Figure 3.20 below.



**Figure 3.20.** Forced-choice production task with novel verbs, object condition.

Indeed, the lowest rate of dative we see on objects still reaches 37.5%, with the animate patient of a hitting event in Figure 3.18. The highest rate of the dative unsurprisingly appears with the indirect object of a giving event (75.8%) but is also high with the theme of a novel throwing verb. Note that despite children's dative overgeneralizations with benefactives/recipients, the objects of the novel *hug* and *love* verbs (again in Figure 3.20) are not produced in the dative more frequently than the verbs for destroying and hitting. This is in line with the case marking on the objects of the existing verbs that the novel verbs were based on. The object condition furthermore shows similar contrasts between typical and atypical dative contexts as the subject condition, which is explored further in subsection 3.3.2.3.

When breaking down the results in Figure 3.20 by age as is done in Figure 3.21, it can be seen that the youngest age group does not show the expected contrasts for the objects either.



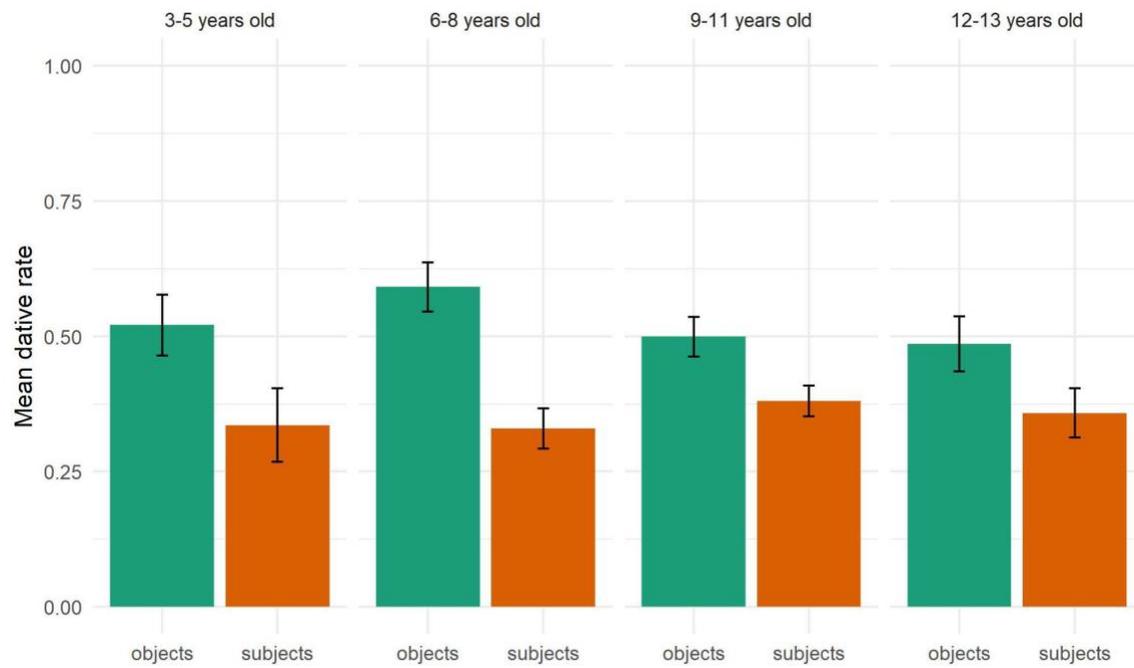
**Figure 3.21.** Forced-choice production task with novel verb by age group, object condition.

Still, there is a higher rate of dative found with the objects as compared to the subjects. The verbs which receive the most frequent choice of a dative object are the verbs corresponding to *love* and *give*, in line with the idea that early dative productivity is somehow more tied to benefactives than e.g. motion themes. The 6–8-year-olds additionally show relatively high rates of dative with the novel *hug* and *love* verbs, but the two oldest age groups show clear contrasts between them and the typically dative goal argument of the novel *give* verb.

In general, we find less of the expected contrasts in the youngest age groups, which still produce the dative in the forced-choice task, and more so with objects than subjects. This is further analyzed in the next subsection, 3.3.2.3, where the results are also compared to the comprehension task results in Chapter 2.

### 3.3.2.3 Typical and atypical contexts for the dative

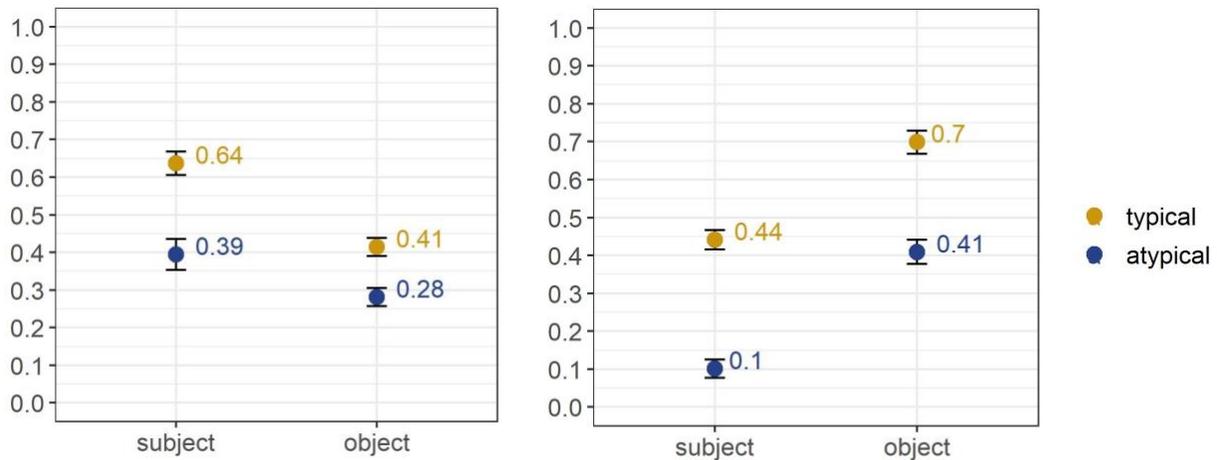
Figure 3.22 shows that the participants chose dative more with objects than subjects.



**Figure 3.22.** Forced-choice production task with novel verbs by age group, subject and object condition.

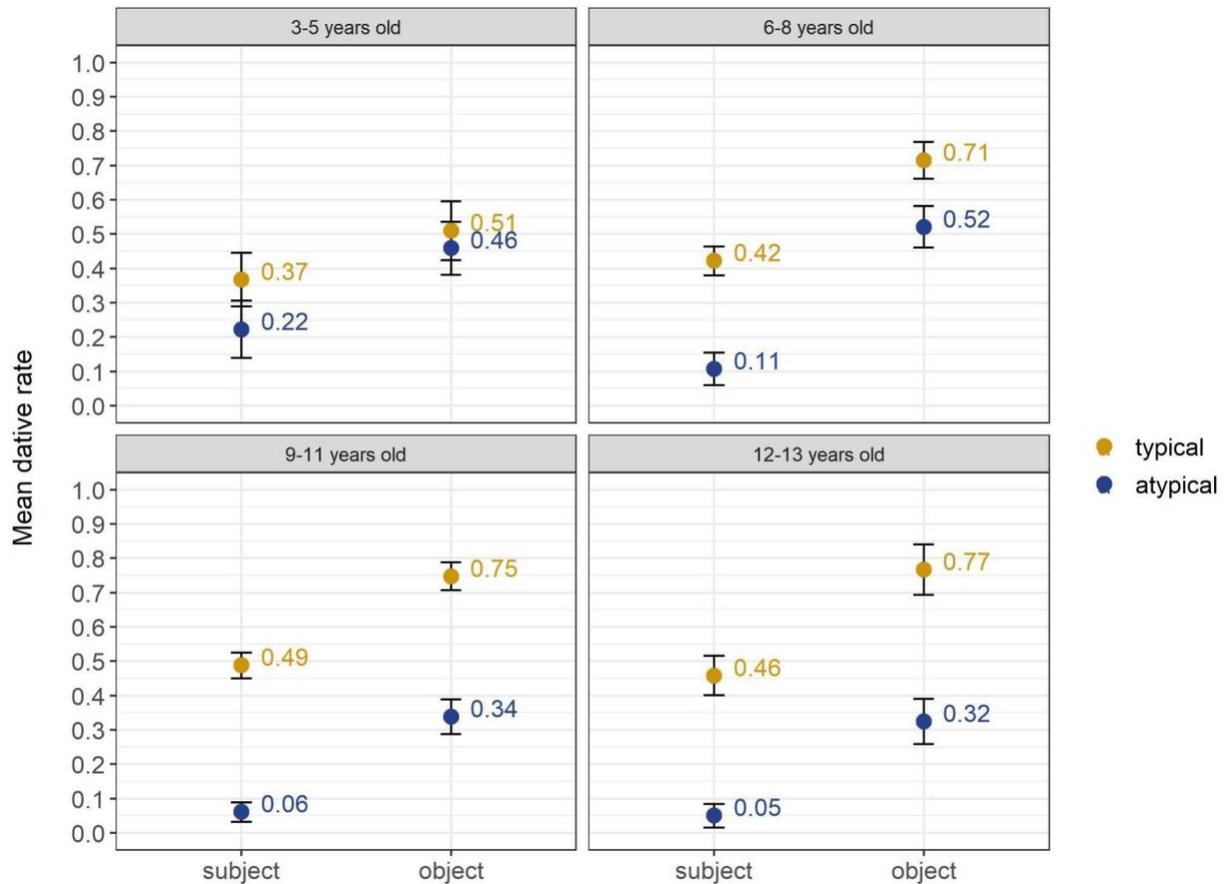
This is in line with the distribution of datives in (adult) Modern Icelandic (Barðdal 2008 for example), where the dative appears more frequently with objects than subjects. As Figure 3.22 also shows, this contrast is more pronounced in younger speakers.

Still, as the results in Figures 3.19, 3.21 and 3.24 below show, the younger participants do not show the expected contrasts between typical and atypical semantic contexts for the dative, despite having shown these associations for subjects (but not as consistently objects) in the comprehension experiments described in Chapter 2. The contrasts do become clearer with age and are reflected in the general results, as Figure 3.23 shows.



**Figure 3.23.** Associations between the dative and typical and atypical semantic contexts, by syntactic role. Comprehension results on the left (mean experimenter/theme rate) and production results (mean dative rate) on the right.

Interestingly, although the youngest children have not fully acquired the appropriate semantic contexts for dative productivity, they seem to have acquired datives based on structural role, with the rate of datives being higher in objects than subjects, as seen in Figure 3.24. To analyze these results further, a nested comparison (Likelihood Ratio Test) of mixed effects logistic regression models (lme4 in R, Bates et al. 2015) was conducted. Age was included in every model as well as random intercepts for participants and items. The analysis shows a significant improvement to the fit of the model when adding syntactic role ( $\chi^2(5) = 17.9, p < 0.001$ ) and semantic context ( $\chi^2(6) = 24.7, p < 0.001$ ) but not when including a syntactic role-semantic context interaction. Context interacts with age, as is clear from the distribution in Figure 3.24 (full modeling results are available in Appendix D).



**Figure 3.24.** Associations between the dative and typical and atypical semantic contexts, by syntactic role. Production results by age group.

In Figure 3.23 above, the general results from the novel verb forced-choice production task are contrasted with the results from the novel verb picture-selection comprehension task discussed in Chapter 2. The comparison shows inverse association patterns. In production, as just reviewed, dative is selected more often with objects showing the semantics typically associated with it, while dative is chosen less with subjects which show the typical non-nominative semantics. In comprehension, on the other hand, the dative yields more salient associations in subject case as compared to the object case, where we also get a smaller effect between typical and atypical contexts. This might be considered a natural consequence of the restrictiveness and low frequency of dative subjects, making them perhaps more salient in comprehension but also less chosen in production (despite additional morphological cues). The fact that the dative

objects are also frequently chosen in atypical semantic contexts might additionally point towards productivity with a wider semantic scope, possibly competing with the structural/default/distributionally dominant accusative. The combined results and age trends point towards a scenario where the semantic context for dative objects remains elusive despite its productivity. Perhaps the tracking down of the appropriate semantic context for datives simply is not possible until children have acquired a larger number of verbs, in particular since we can distinguish two rather distinct semantic contexts within objects (but only one for subjects), animate goals/benefactives and experiencers on the one hand and themes of motion verbs on the other hand.

#### 3.3.2.4 Predicting productivity at the individual level

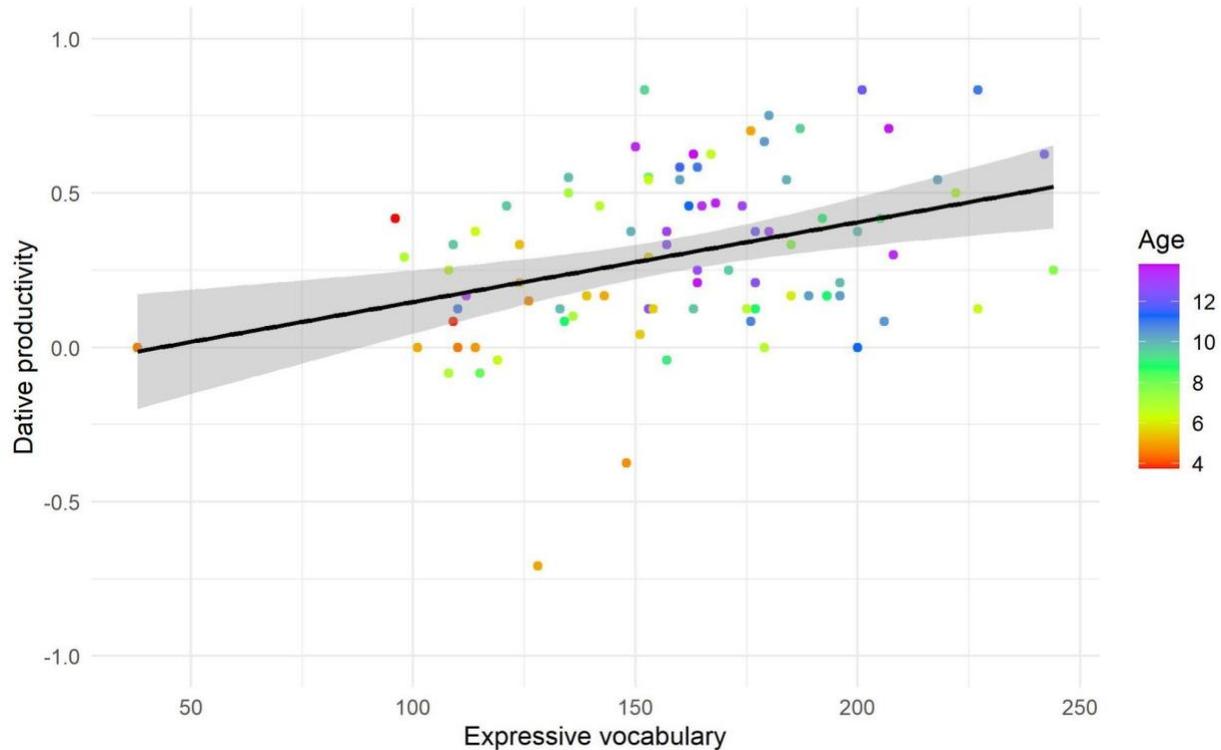
If the productivity of the dative depends on acquiring enough verbs to uncover the associations with meaning, it becomes interesting to investigate whether it is possible to predict dative productivity at the individual level. The vocabularies of the individuals who participated in the forced-choice task were in fact tested, both using an Icelandic translation (Ragnarsdóttir 2018) of the receptive picture-selection assessment PPVT-4, Peabody Picture Vocabulary Test 4<sup>th</sup> edition (Dunn & Dunn 2007), and with language samples in which the number of different words were computed for 50 Icelandic utterances. Of course, the relationship between individual vocabulary size and rule formation (or simply grammatical abilities) is a topic with an extensive research tradition (notably in the context of language development, multilingualism and developmental language disorders) but a lot of uncertainty exists about the nature of this relationship (see e.g. Hoff et al. 2018). Usually, this relationship is investigated using standardized vocabulary and language tests, taking a holistic approach and not targeting specific variables within the grammar. It is therefore interesting to explore the relationship of individual vocabulary size and dative productivity, which I did by computing an individual productivity measure (rate of dative selection in typical contexts – rate of dative selection in atypical contexts) and running a linear regression with age, the number of different words in the language sample (NDW) and the PPVT raw score (standardized scores not available). The results can be found in Table 6.

**Table 6.** Linear regression predicting dative productivity.

	<b>Estimate</b>	<b>Std. error</b>	<b>t</b>	<b>p</b>
<b>(intercept)</b>	-0.2035301	0.1230382	-1.654	0.10182
<b>age</b>	0.0291623	0.0100349	2.906	0.00468**
<b>NDW</b>	0.0017423	0.0008087	2.154	0.03407*
<b>PPVT</b>	-0.0003428	0.0004387	-0.781	0.43683

Adjusted R<sup>2</sup>: 0.1915, F-statistic: 7.87 on 3 and 84 DF, p < 0.001

Not surprisingly, we find an age effect, with older children having a higher rate of dative productivity. More surprising is the significant effect of productive vocabulary (plotted in Figure 3.25) on dative productivity, with no such effect appearing for the receptive vocabulary measure.



**Figure 3.25.** Dative productivity and expressive vocabulary, linear model fitted (95% confidence) and age specified.

The relationship is not straightforward nor the effect large, as Table 6 and Figure 3.25 show, but it is interesting that this analysis at the individual level does fit an approach where there is a relationship between the number of items following a pattern and its abstraction into a productive rule. In this specific case, the expressive vocabulary would be a proxy for the sufficient number of verbs or arguments displaying the productivity patterns acquired by the children. In the corpus analysis described in the following chapter, this intuition is formalized and tested using the Tolerance Principle (Yang 2016).

### 3.4 Discussion

The broad overgeneralization patterns which emerged from the forced-choice production tasks are mostly expected considering previous research on the acquisition of case marking in Icelandic and the productive patterns which emerge within the system (e.g. Svavarsdóttir 1982, Sigurðardóttir 2002, Jónsson & Eythórsson 2003, Barðdal 2008, Nowenstein 2017, Guðmundsdóttir et al. 2019, Jónsson & Thórarinsdóttir 2020 and

Nowenstein & Ingason 2021). Children acquiring Icelandic overgeneralize the default nominative for subjects and accusative for objects, but additionally, there is robust dative productivity and some more marginal phenomena such as nominative on objects and accusative with subjects. The dative productivity, as became clear in the novel verb task experiments, is structurally and semantically conditioned. Datives are much more frequently produced on objects than subjects, and dative on indirect objects emerges earlier than on direct objects. In the subject case, datives are produced more in non-agentive experiencer contexts, and in the object case, datives are more frequent with motion themes and goals than patients. The results furthermore indicate that the association between datives and goals is more prominent earlier in acquisition, but at the same time that children acquire dative productivity based on syntactic role before linking the case marking to the appropriate semantic contexts. This precedence of form over meaning is also present in the corpus results described in Chapter 4 and resonates with previous findings showing differences in the emergence of pattern recognition (e.g. Emond & Shi 2021) and semantically conditioned rule application (e.g. Schuler et al. 2016), where pattern recognition is attested sooner and is possibly a necessary condition for the discovery of form-meaning mappings (Naigles 2002, Yang 2020, Brusini et al. 2021).

In this context, it is important to note that although the dative was used more frequently within the expected semantic contexts, it also was robust in unexpected semantic contexts for objects, pointing towards productivity which goes beyond Jónsson & Thórarinsdóttir's (2020) caused motion context and even Barðdal's (2008) more diverse schemas. This might indicate a scenario where dative even competes with the default accusative for the case-marking of patients. Finally, no clear contrasts between high and low frequency predicates or novel predicate equivalents were found.

The data therefore call for a language acquisition model which can account for nested productivity where statistically non-dominant and non-default patterns can be productive along with, or within, a rule with a broader range. The model also has to be able to account for the productivity of even more marginal patterns (such as accusative for subjects) and variability within and across learners through development, such as learners' early association between datives and goals. Within the approaches reviewed

in section 3.1, and considering the lack of item frequency contrasts as well as the models' potential for testable predictions, a modern rule-based approach such as Yang's Tolerance Principle (2016) seems to be a promising starting point. The abductive search procedure (recursively hypothesizing rules) of recent computational implementations (Belth et al. 2021) furthermore are promising for contexts where rule directionality might be crucial, such as in accounting for the different patterns observed in production and comprehension outlined in the current chapter.

Indeed, the meaning-form or form-form directionality assumed here for production could, depending on the rule formation mechanism assumed, yield different patterns of association than the form-meaning direction assumed for comprehension and the acquisition of root meaning. The restricted semantics and type frequency (range) of non-nominative subjects, for example, provides a very salient cue for non-agentivity in word learning but the same attributes, as well as the competition with nominative experiencers, should yield worse conditions for marking subjects as non-nominative in production. On the other hand, the much broader range of dative objects makes their meaning elusive and reduces the value of dative objects as a cue for meaning. At the same time, it creates conditions where datives are more likely to be used productively with a range of objects.<sup>36</sup>

How the distribution of the data in different contexts shapes comprehension and productivity is one of the questions underlying the application of the Tolerance Principle (Yang 2016) in the next chapter, where I investigate whether it can be used to derive patterns which I have confirmed experimentally using child language data. The vocabulary results in subsection 3.3.2.4, where dative productivity was predicted by expressive vocabulary at an individual level, indicates that an approach which derives rules from the distributional information available in children's language environment is on the right track. In Chapter 5, I put the experimental data described in the current chapter, and the language acquisition model tested in Chapter 4, in the context of language variation and change, showing that the way children derive productive rules from the input informs questions about the progression, stability and specialization of different variants in situations of language change.

---

<sup>36</sup> This is somewhat in line with various of the core components of Barðdal's (2008) account for dative productivity in adults and in diachrony.

## 4. Dative Derivability: A Case for Exceptions

So far, I have shown that despite known exceptions, non-exact correlations and one-to-many and many-to-all relationships, children do discover associations between case and meaning, using case as a cue for meaning and producing non-default datives more with experiencers, goals and motion themes than agents and patients. These associations seem to change through development, with the dative-goal association emerging before the dative-motion theme one for example. Furthermore, the results indicate that children acquire mappings between datives and syntactic role before the semantic correlations have been established, form-form links taking precedence over form-meaning ones. Finally, the form-meaning mappings present differently in comprehension (or word learning) and production. Thus, dative subjects are a more salient cue than dative objects in comprehension, but the pattern is inverted in production, where the dative is produced more on objects than subjects. But how do these associations between form and meaning emerge and how is their productivity determined? Can they be derived from distributional information in the input?

### 4.1 Learning from distributional information

Investigating datives in the acquisition of Icelandic does not only entail addressing questions about (morpho)syntactic bootstrapping (discussed in Chapter 2) and rule formation and productivity (outlined in Chapter 3). It also involves questions about the derivability of linguistic knowledge. How much can children derive from the input? What is not derivable and therefore potentially innate? In a recent paper on the derivability of linking rules (UTAH and rUTAH<sup>37</sup>), Pearl & Sprouse (2021, p. 295) discuss the importance of investigating derivability from realistic child input in general. They argue that in language acquisition theories, “the general direction of investigation is to minimize or eliminate innate scaffolding proposals when (empirically) possible” and that “results that show us which representational components can be derived in principle —and which

---

<sup>37</sup> Uniformity of Theta Assignment Hypothesis (UTAH) and the relativized Uniformity of Theta Assignment Hypothesis (rUTAH).

can't— contribute to that effort". I fully agree with the importance of maximizing derivability approaches, and that innate scaffolding should be considered a proposal of last resort. Of course, this is clearer in the case of overt morphological case marking than the positional licensing of linking rules, where arguments for universality and innateness are more robust, with quite a few approaches to linking rules and syntactic bootstrapping assuming innate mappings between form and meaning (e.g. Lidz et al. 2003 and Fisher et al. 2010). Going back to the results of Chapter 2 which show how case, just as argument number, can be used as a cue for verb meaning, it is interesting to ask what the theoretical implications are for syntactic bootstrapping if morphosyntactic bootstrapping is based on derived mappings between form and meaning.

In this chapter, I test the derivability of case marking rules using the Tolerance Principle (Yang 2016), as Pearl & Sprouse (2021) do in their derivation of UTAH and rUTAH. Although the Tolerance Principle (Yang 2016, 2018, 2021, henceforth the TP) has been mentioned at various points so far in this dissertation, it has only been briefly described. Before using it to test the derivability of case marking rules in Icelandic, I provide a more in-depth description centered around the properties relevant to the current work. Much more extensive accounts of the TP, as well as the mathematical details of its derivation, can be found in Yang's work (e.g. 2016, 2018, 2021). As a reminder, the TP is a parameter-free learning model which provides a threshold for the number of exceptions a productive rule can tolerate. The following definition can be found in Yang's (2016) comprehensive report of the TP, *The Price of Linguistic Productivity*:

- (28) Let a rule  $R$  be defined over a set of  $N$  items.  $R$  is productive if and only if  $e$ , the number of items not supporting  $R$ , does not exceed  $\theta_N$ :

$$e \leq \theta_N = \frac{N}{\ln N}$$

The TP therefore provides a numerical criterion for productivity, entailing the first aspect of the theory I want to highlight: Under the TP, generalizations do not need to be absolute to become productive rules, they just need to be good enough. Speakers can tolerate exceptions in language as in other domains where generalizations occur. This has an

intuitive component which can be materialized into a number of analogies. For example, buses in Reykjavík are traditionally yellow. When waiting for a bus, we expect the next one to arrive to be yellow. In other words, the [bus → yellow] rule is productive. This does not change despite exceptions, for example when a purple bus circulates to raise awareness for dementia, as is currently the case in Reykjavík. Over the past few years, on the other hand, a substantial number of white buses have been taken into service for some of the routes. An interesting question is whether these white buses have reached a number, perhaps only for some contexts/routes, which renders the [bus → yellow] rule unproductive, introducing a stochastic [bus → yellow/white] rule instead or even a [bus → white] rule only for route 6 for example, although yellow is still the default for the system as a whole. Under each of these rules, the purple bus would still be an exception. In the analysis of subsection 4.3, I try to find the yellow, white and purple buses of the Icelandic case marking system.

The second aspect of the TP I want to highlight is the fact that smaller sets of items tolerate relatively more exceptions. This is because  $\theta_N$  is calculated using the natural log of  $N$  and not e.g. a fixed majority proportion such as 0.5 or 0.75. Table 7 (Yang 2021) shows that as the set of  $N$  becomes larger, the proportion of tolerable exceptions becomes lower, requiring a more extreme supermajority for productivity. As Yang (2021, p. 3) discusses, this suggests that “rules defined over a smaller vocabulary can tolerate relatively more exceptions, and are thus easier to learn”. This is interesting in light of the fact that child learners, as opposed to e.g. machines, acquire the rules of language while having relatively small vocabularies, an aspect of acquisition which needs to be considered when building theoretical accounts of language. Some exceptions which have been viewed as counter-arguments to generalizations might not even occur in the input children base their system on (see Yang’s 2016 analysis of dative productivity in English double-object constructions).

**Table 7.** The maximum number of exceptions for a productive rule over N items.

$N$	$\theta_N$	%
10	4	40.0
20	6	30.0
50	12	24.0
100	21	21.0
200	37	18.5
500	80	16.0
1,000	144	14.4
5,000	587	11.7

Additionally, this ‘less is more’ (Newport 1990) component of the TP is important for two key aspects of rule formation. The first one is the room for individual differences which is inherent to the TP as it is hypothesized to operate within individual vocabularies which can be composed differently (Yang 2016). This room for variation in rule formation between individuals arguably becomes bigger in smaller vocabularies which allow for relatively more exceptions. The second one is variation within the developmental trajectory of individuals, since rule formation and productivity might evolve as a child’s vocabulary grows with rules which were productive in a small vocabulary becoming unproductive (and vice versa) as the vocabulary gets bigger or its composition changes.

The third aspect of the TP I want to highlight is the Principle of Maximize Productivity (Yang 2016, p. 75) shown in (29). This yields a “divide-and-conquer” search for productivity: If a productive rule over an entire set cannot be derived, the data is further divided into subsets and the search for productive rules continues (see Belth et al. 2021 for a full implementation).

- (29) Maximize Productivity: Lexicalization is the last resort; when possible, pursue rules that maximize productivity.

Such an approach can yield a scenario of nested productivity, where a restricted, productive subrule also creates the necessary conditions for a general non-restricted rule to become productive by ‘taking care of’ what would previously have been exceptions to the rule applying to the entire set. This means that the order of productivity testing is crucial to the final result. In Belth’s et al. (2021, p. 2) implementation for the German plural, the decision of which feature to split on at each nested level is “selected to maximize consistency: the relative frequency of the most frequent suffix that the instances with that feature take”. This leads me to the last aspect I want to highlight, which really is a specification of what is not part of the TP. The TP does not define the feature space over which the rules are formed, that has to be theoretically and empirically motivated. In the application of the TP for the Icelandic case marking system, I assume that some categories or concepts might be pre-linguistic and/or innate (which does not mean they are completely fixed) while others are constructed (Carey 2009, Becker 2014, Rissman & Majid 2019).

### 4.1.1 Main questions and results

The main questions of the chapter can be found in (30):

- (30) a. Are the links between case, syntactic role and meaning described in Chapters 1, 2 and 3 discoverable and learnable from the input using a Tolerance Principle (Yang 2016, 2018, 2021) approach? Are they present throughout acquisition if verb frequency is used as a proxy for order of acquisition?
- b. Do children use distribution in form, meaning or both as cues to case marking in production? What about cues to verb meaning in comprehension?

Additionally, a major goal of the chapter is to provide an overview of the case frames present in Icelandic child language. This overview reveals various important facts, notably the early prominence of dative goals as opposed to motion themes, in part due to optional reflexives. The results furthermore indicate that many of the broad patterns between case, meaning and syntactic role which have emerged in the experimental data of Chapters 2 and 3 in this dissertation, as well as in previous literature, can in fact be derived from child language data assuming the TP. This refines the role of statistical majority in productivity and redefines the value of exceptions in case theory. In certain contexts, the lack of productivity is also informative for the more unusual results, such as the broad scope of dative productivity with direct objects. In general, I show how non-default productivity can be derived in the form of nested subrules (dative objects) and structured exceptions (dative subjects), and challenge the view that case, both structural and inherent, does not have meaning. This approach also yields novel insights on dative subjects with adjectival predicates and the lack of dative productivity for indirect objects.

## 4.2 Methods

The main objective for the corpus analysis is to incrementally assess the viability of different case marking rules depending on children's acquisition of verbs. This is done using Yang's (2016) Tolerance Principle, a parameter-free learning model which provides

a threshold for the number of exceptions a pattern or rule can tolerate while being productive. The analysis is based on approximately 750.000 words of interactions between children and adults<sup>38</sup> and consists of data from GJEUM, the Einarsdóttir et al. 2019 language sample corpus, the Sigurjónsdóttir (2007) corpora and the Icelandic data available in CHILDES (Strömqvist et al. 1995). Although the corpus has not been prepared for the use of taggers and lemmatizers, 984 verbs were extracted using the ABLtagger (Jónsson & Loftsson 2021) and Nefnir (Ingólfssdóttir et al. 2019). Errors are frequent for hapax legomena<sup>39</sup> but rarely present for verbs appearing more than five times, which yields a dataset of 410 verbs (excluding the copula, auxiliaries and modals). This is the IceCASE corpus (Icelandic Child Argument Structure Evidence), available in Appendix E. A sample of this size (410 verbs) is consistent with the expected verbal repertoires of children who are starting to use the Icelandic dative. Previous research (notably Sigurðardóttir 2002) shows that children learning Icelandic start producing dative objects around the age of 2–3 and dative subjects around the age of 3–4. Accounts of verbal acquisition in other Germanic languages indicate that 100–200 verbs at age 3 and 200–400 at age 4 is an appropriate estimation (e.g. Christensen 2010, Yang 2016, Hadley et al. 2016 and Irani 2019). Based on a literature review of the structural and semantic features which might condition case marking in Icelandic (e.g. Yip, Maling & Jackendoff 1987, Van Valin 1991, Jónsson 1997–1998, 2003, 2009 and 2013, Maling 2002, Svenonius 2002, Barðdal 2008, H.Á. Sigurðsson 2012, Eythórsson et al. 2012, Wood 2015, E.F. Sigurðsson 2017), as well as research on the universality and compositional properties of conceptual event structure categories such as agent and patient (e.g. Dowty 1991, Baker 1997, Levin & Rappaport Hovav 2005, Ramchand 2008, Grimm 2011, Reisinger et al. 2015, Rissman & Majid 2019, Ünal et al. 2021), the following features, in addition to case, were annotated by hand for verbal arguments in the dataset:

- (31) a. **Form:** ± subject, ± direct object, ± indirect object, ± optional reflexive argument  
 b. **Meaning:** ± agent, ± patient, ± theme, ± goal, ± experiencer, ± animacy,  
 ±motion (change of location)

---

<sup>38</sup> Unfortunately, information regarding the proportion of adult utterances in the corpora is not available. This would clarify to what extent the current work measures input (to children) and output (from children).

<sup>39</sup> Items occurring only once in the corpus.

It is important to emphasize that the current annotation is a practical representation of the properties which have been shown to relate to case in research on Icelandic, not a proposal of core generalized thematic roles and their links to structure (e.g. Fillmore 1968, Belletti & Rizzi 1988, Bresnan & Kanerva 1989, Grimshaw 1990, Kiparsky 1997, Baker 1997). As Rissman & Majid (2019) quote, Dowty (1991) states that “there is perhaps no concept in modern syntactic and semantic theory which is so often involved in so wide a range of contexts, but on which there is so little agreement as to its nature and definition, as THEMATIC ROLE” (p. 547) and some twenty years later, Newmeyer (2010) writes that “there is no construct as murky in ANY subdivision of linguistic theory as that of ‘thematic role’.” (p. 689). I do agree with the view that thematic role lists, hierarchies and linking rules have problematic aspects and that decompositional approaches are more viable in general (Dowty 1991, Levin & Rappaport Hovav 2015, Ramchand 2008, Grimm 2011) but also assume that learners will form the structure-meaning associations available from the distribution of the input, and that this mapping can mobilize fine-grained event participant properties such as sentience as well as more complex roles such as Agent or Experiencer. In the current annotation, I build on Jónsson's (1997–1998) and Thráinsson's (2007) thematic role lists and criteria as they have been the basis for analyses on the relationship between Icelandic case (variation) and semantics. I also build on their work (considering additionally Wood 2015 and E.F. Sigurðsson 2017) when it comes to the annotation of the formal features in (31). Additionally, I annotate two semantic argument features which have been shown to be relevant to case marking patterns in Icelandic and binomially annotate them specifically for all verbal arguments, animacy and motion as change of location (Barðdal 2008, Jónsson & Thórarinsdóttir 2020). The thematic role annotation criteria used are outlined below in subsection 4.2.1, with examples from the child language data the IceCASE corpus builds on. This annotation is a first step and attempt and has many limitations, such as a lack of a more precise decomposition of the events, states and roles. In future work, different annotation criteria (as in e.g. Schätzle 2018) should be tested. As might be expected, annotating argument structure in this way still is not a simple task. Predicates are very rarely used in a uniform way, as work by Levin (1993) and Levin and Rappaport-Hovav (2005) made

abundantly clear. In the current work, the annotation, available in Appendix E, is solely based on usage in the child language corpus. Examples are provided below along with the specification on the thematic roles assumed in the annotation.

## 4.2.1 Thematic roles

### 4.2.1.1 Agents

Agents initiate an activity and/or cause a change of state. I do not restrict the category to volitional/sentient entities as Jónsson (1997–1998) does with the agenthood tests he employs. The category therefore includes non-sentient causers and effectors, but the typical agent in the current annotation passes agenthood tests that rely on volition and has all of Dowty's (1991) proto-agent entailments (volition, sentience/perception, causation, movement and independent existence). If sentience/perception, movement or independent existence are the only agent entailments of an argument, the argument will most likely receive the annotation of experiencer or theme as described below. The category is therefore a more restricted version of Dowty's (1991) proto-agent, Van Valin's (1991) actor and Baker's (1997) macrorole Agent but perhaps closer to Ramchand's (2008) initiator and the upper two tiers in Grimm's (2011) proto-agent lattice. Below (32) are examples of a few arguments annotated as agents in the corpus:

- (32) a. *Kaupa* 'buy':      **Amma** [AGENT]      **keypti**      páskaegg  
    grandma.NOM      bought      easter egg.ACC  
    'Grandma bought an easter egg.'  
    (Sigurjónsdóttir longitudinal corpus (2007))
- b. *Kíkja* 'peek':      Má **ég** [AGENT]      **kíkja** bara einu sinni?  
    Can I.NOM      peek just one time  
    'Can I peek just once?'  
    (Einarsdóttir et al. corpus (2019))
- c. *Hringja* 'ring!':      **Síminn** [AGENT]      **hringdi**  
    phone.NOM      rang  
    'The phone rang.'  
    (Sigurjónsdóttir longitudinal corpus (2007))

As can be seen, all the agent arguments have either volition or causation entailments but vary with regards to Dowty's (1991) other proto-agent entailments. An example of the annotation in the corpus is provided below for *kaupa* 'buy' (see Appendix E for all verbs):

verb	english	rank	freq	case	animacy	roles	reflexive	motion
<b>kaupa</b>	<b>buy</b>	<b>30</b>	<b>296</b>	<b>N_A</b>	<b>A_I</b>	<b>A_T</b>	<b>D</b>	<b>0</b>

*Kaup* 'buy' is annotated to have a nominative-accusative case frame (N\_A) despite the fact that some adult speakers accept a dative indirect object for the verb as well. This usage does not appear in the data underlying the corpus. Additionally, in the child language data, all subjects with *kaupa* 'buy' are animate agents and objects appear as inanimate themes (animacy: A\_I and roles: A\_T). This is despite the fact that (adult) speakers can for example in some cases use *kaupa* 'buy' for animate entities. Dative reflexives appear in the child language data (reflexive: D), and the motion (change of location) of the direct object is coded, here the object theme does not move (and is thus 0).

#### 4.2.1.2 Experiencers

Experiencers are sentient and feel/perceive/experience without the proto-agent entailments of causation and/or volition. In the current annotation, experiencers are limited to internally-caused experiences and would therefore fall into Baker's (1997) agent macrorole as well as Dowty's (1991) proto-agent. In Van Valin's (1991) categorization, these experiencers can be considered as actors and undergoers and in Grimm's (2011) approach they occupy the lowest tier of the proto-agent lattice. If the cause is external, the argument receives the annotation of patient (see Belletti & Rizzi 1988, Pesetsky 1995 and Hartshorne et al. 2015 for more about differences between 'subject' and 'object' experiencers). Jónsson (1997–1998) describes how experiencers can be further divided based on lexical semantics and mentions the following categories: (1) feelings and emotions, (2) mental states and perception, (3) bodily state and (4) happenstance. Below in (33) are examples of a few arguments annotated as experiencers in the corpus.

- (33) a. *Sjá* 'see': **Ég** [EXPERIENCER] *geti* **sjáið** það  
 I.NOM can see it.ACC  
 'I can see it.'  
 (Einarsdóttir et al. corpus (2019))
- b. *Vita* 'know': **Ég** [EXPERIENCER] **veit** *alveg hvað leyndarmál* er  
 I.NOM know really what secret is  
 'I definitely know what a secret is.'  
 (Einarsdóttir et al. corpus (2019))
- c. *Langa* 'want': *Út af því að* **mig** [EXPERIENCER] **langaði** það  
 because that me.ACC wanted it.ACC  
 'Because I wanted to.'  
 (Einarsdóttir et al. corpus (2019))

These are all sentient arguments in subject position which systematically fail Jónsson's (1997–1998) agenthood tests. The annotation for *sjá* 'see' is provided below:

verb	english	rank	freq	case	animacy	roles	reflexive	motion
<b>sjá</b>	see	4	1990	N_(A)	A_(AI)	E_(T)	none	0

As can be seen, *sjá* 'see' has a nominative-accusative (N\_A) frame corresponding to an experiencer and optional theme (E\_(T)), but as it can also appear with a clausal complement or an adverbial phrase, the object is placed within parentheses (N\_(A)). The subject is always animate but the theme object (which does not move) can be either animate or inanimate (A\_(AI)). There is no reflexive and no assumed motion of the direct object.

#### 4.2.1.3 Patients

Patients undergo change and/or are affected by external causes, which is in line with Jónsson's (1997–1998) patient but more restrictive than Thráinsson's (2007) theme which includes patients as well as entities which are located somewhere or moving. If considered in light of Dowty's (1991) proto-patient entailments, the current annotation emphasizes (1) change of state (although I do not consider location change as a change of state), (2) being causally affected and (3) existence not being independent of the event.

On the other hand, being stationary relative to another participant is not enough to be considered a patient and some incremental themes (which do not include a change of state) receive the theme annotation described below. The current patient category is therefore more restrictive than Dowty's (1991) proto-patient, as well as Baker's (1997) macrorole patient and Van Valin's (1991) undergoer. It still retains key aspects of those wider categories, namely being causally affected, changing states and coming into existence. Below in (34) are examples of a few arguments annotated as patients in the corpus.

- (34) a. *Borða* 'eat': Þú mátt **borða** allt **pítsuna** [PATIENT]  
you.NOM can eat all.ACC pizza.the.ACC  
'You can eat the whole pizza.'  
(Einarsdóttir et al. corpus (2019))
- b. *Klæða* 'dress': Mamma er að **klæða** **hana** [PATIENT]  
mom.NOM is to dress her.ACC  
'Mom is dressing her.'  
(Sigurjónsdóttir longitudinal corpus (2007))
- c. *Byggja* 'build': Ég get **byggjað** hús [PATIENT]  
I.NOM can built house.ACC  
'I can build a house.'  
(Einarsdóttir et al. corpus (2019))

As can be seen from the above patient examples, patients can both be animate and inanimate entities as long as they fulfill the criteria of being externally affected or changed. The annotation for *borða* 'eat' can be found below:

verb	english	rank	freq	case	animacy	roles	reflexive	motion
borða	eat	19	656	N_(A)	A_(I)	A_(P)	none	0

*Borða* 'eat' usually appears in a nominative-accusative frame although it can also be intransitive, the subject is always an animate agent and the object an inanimate patient (A\_(I) and A\_(P)) (which does not move). There is no reflexive argument.

#### 4.2.1.4 Themes

Themes are entities which are either moving or located somewhere, as in Jónsson's (1997–1998) classification. I also include Dowty's (1991) incremental themes, excluding holistic incremental themes which undergo a change of state. The present theme category therefore includes arguments within the proto-patient (Dowty 1991), macrorole patient (Baker 1997) and undergoer (Van Valin 1991) roles which were excluded in the patient category. This includes what also has been referred to as *Figure* (moving entity) and *Location* or *Locative* (stationary entity), excluding locations which also are goals as described in the next category (see further discussion in e.g. Dowty 1991 and Van Valin 1991). Below (35) are examples of a few arguments annotated as themes in the corpus:

- (35) a. *Skoða* 'look at': **Ég** var að **skoða** **bækur** [THEME]  
 I.NOM was to look books.ACC  
 'I was looking at books.'  
 (Einarsdóttir et al. corpus (2019))
- b. *Keyra* 'drive': **Keyra** **bílinn** [THEME]  
 drive car.the.ACC  
 'Drive (infinitive) the car.'  
 (Sigurjónsdóttir longitudinal corpus (2007))
- c. *Liggja* 'lie down': **Hann** [THEME] **bara** liggur hérna  
 He.NOM just lies here  
 'He just lies here.'  
 (Sigurjónsdóttir longitudinal corpus (2007))

As the examples show, the category of theme has a broad scope in the current annotation, including both animate and inanimate entities which can be moving or stationary. Syntactically, they can appear in subject and object position. The annotation for *skoða* 'look at, examine' can be found below.

verb	english	rank	freq	case	animacy	roles	reflexive	motion
skoða	examine	23	399	N_A	A_AI	A_T	none	0

The verb has a nominative-accusative frame with an animate agent subject and a (non-moving) theme object which can be either animate or inanimate (A\_AI and A\_T). There is no reflexive and the verb does not appear with clausal complements, even though that is possible in adult Icelandic.

#### 4.2.1.5 Goals

I adopt Thráinsson's (2007, p. 200) definition of goals as "the point to or towards which an entity moves or is oriented", including his specifications that "recipient is a related notion, a sentient destination, and so is benefactive/beneficiary", referencing Maling (2002) and Jónsson (2000). Relating the goals in the present annotation to Dowty's (1991) proto-agent and proto-agent entailments, the only proto-patient entailment present in the category is being stationary relative to another participant. There is also a proto-agent entailment present, namely the existence of the argument being independent of the event. In Van Valin's (1991) classification, the current goals fall under the undergoer category but within Baker's (1997) UTAH, they are not part of the Agent and Patient macroroles but categorized as 'Other' and map onto the oblique sentence position in Pearl & Sprouse's (2021) characterization. Note that some of the arguments classified as themes in the current annotation (e.g. locations) would also fall into Baker's (1997) 'Other' category. Examples of a few arguments annotated as goals in the corpus are found in (36).

- (36) a. *Segja* 'tell':           Mamma    **segir mér** [GOAL] hvað   ég    get    kaupt  
 Mom.NOM   tells   me.DAT   what.ACC I.NOM can   bought  
 'Mom tells me what I can buy.'  
 (Einarsdóttir et al. corpus (2019))
- b. *Fá* 'get':             Ég    **fékk mér** [GOAL] ekki  
 I.NOM got   me.DAT   not  
 'I didn't get myself any.'  
 (Einarsdóttir et al. corpus (2019))

c. *Gefa* 'give':      Ég   ætla  að   **gefa** **apa** [GOAL]   banana  
 I.NOM will   to   give   monkey.DAT banana  
 'I'm going to give monkey a banana'  
 (Einarsdóttir et al. corpus (2019))

As is shown in the examples, goals frequently appear as indirect objects or optional reflexives. Interestingly, in the case of the verb *fá* 'get', the presence of the reflexive argument entails that the subject is an agent, but without it the same argument is interpreted as a goal itself (AG\_(T)). The annotation representing this is shown below:

verb	english	rank	freq	case	animacy	roles	reflexive	motion
fá	get	7	1211	N_(A)	AI_AI	AG_(T)	D	0

An overview of the case frames found in the corpus is found in subsection 4.3.1. Before describing it, I turn to the methods used to test case marking rules with the TP.

#### 4.2.2 Patterns and hypothesis formation

It is important to mention again that the broad case-semantics patterns which have been described in formal approaches to case crucially also are present in the semantically motivated productivity of the dative in Modern Icelandic which has been documented in previous research. This is apparent in subject case variation and change (Jónsson & Eythórsson 2003), the coinage of new verbs (Barðdal 2008, Jónsson & Thórarinsdóttir 2020) and finally in children's overgeneralizations (Sigurðardóttir 2002, and Chapters 2 and 3 of this dissertation). Furthermore, following Yang's (2016) Tolerance Principle, those mappings could theoretically take the form of productive rules even in the presence of exceptions, as long as those exceptions do not exceed the threshold which would render the rule non-viable.

I consider the current approach to hypothesis formation as maximal in nature, both in terms of the possible size of the set a rule might apply to and the categories/features I assume. As has been described, when a productive rule is not found, the set is further divided. I apply the possible rules in a discovery procedure at 50 verb intervals in addition to generating random verb samples at each 100 verb interval (with  $N - 50$  verbs). I begin

with distinctions in form before proceeding to associations with meaning and eventually associating formal and semantic features within rules when necessary. The precedence of form over meaning is based on known differences in the emergence of pattern recognition (e.g. Emond & Shi 2021) and semantically conditioned rule application (e.g. Schuler et al. 2016), where pattern recognition is attested much sooner and possibly is a necessary condition for the discovery of form-meaning mappings (Yang 2020). The insights from the literature mentioned above yield the following base patterns, ordered here in function of expected discoverability based on linguistic distribution and complexity (one vs. two features).

(37) a. Form:

[external argument/subject → nominative]

[direct object → accusative]

[indirect object → dative]

[optional reflexive → dative]

b. Meaning:

[agent → nominative]

[patient → accusative]

[goal → dative]

[(motion) theme → dative]

[experiencer → dative]

When necessary for the discovery of productive patterns, I expect narrower rules to be formed. Note that this approach shares various insights with LFG-accounts to case marking and positional licensing (Kiparsky 1997, Galbraith 2018 and Schätzle 2018), as well as the linking theory testing in Pearl & Sprouse (2021). One aspect which is not present in the current approach but should be tested in future work (as in Pearl & Sprouse 2021 for rUTAH) is the notion of relativity/dependency in case marking (Yip et al. 1987, Marantz 1991, Baker 2015), where e.g. accusative ‘assignment’ depends on the presence of a nominative argument. Presumably, the added complexity of testing hierarchical case

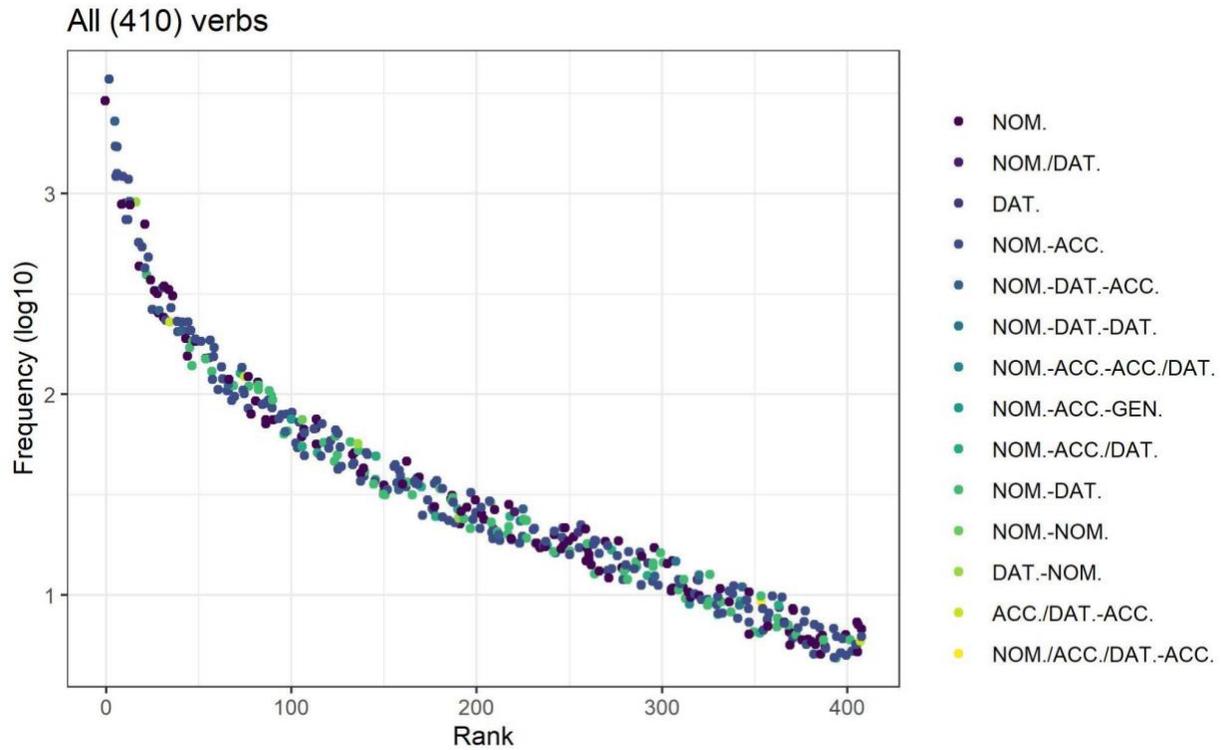
marking, as opposed to direct linking rules, also entails more complexity in language acquisition. The rules or patterns presented here could therefore be thought of as a simplified first case marking step which must be considered theoretically and developmentally but might not be sufficient.

## 4.3 Results

The major goal of the corpus analysis was to get an overview of the case frames present in Icelandic child language and to incrementally assess the productivity of case marking rules using the Tolerance Principle (Yang 2016). I start with an overview of the case frames appearing with the verbs in the dataset before proceeding to testing rules, for production and comprehension, based on form, meaning and finally both form and meaning. I include a note on the productivity of dative subjects with adjectival predicates and different ditransitive case patterns and finally discuss the implications of the results in the context of the data presented in previous chapters as well as dative productivity in Icelandic more broadly.

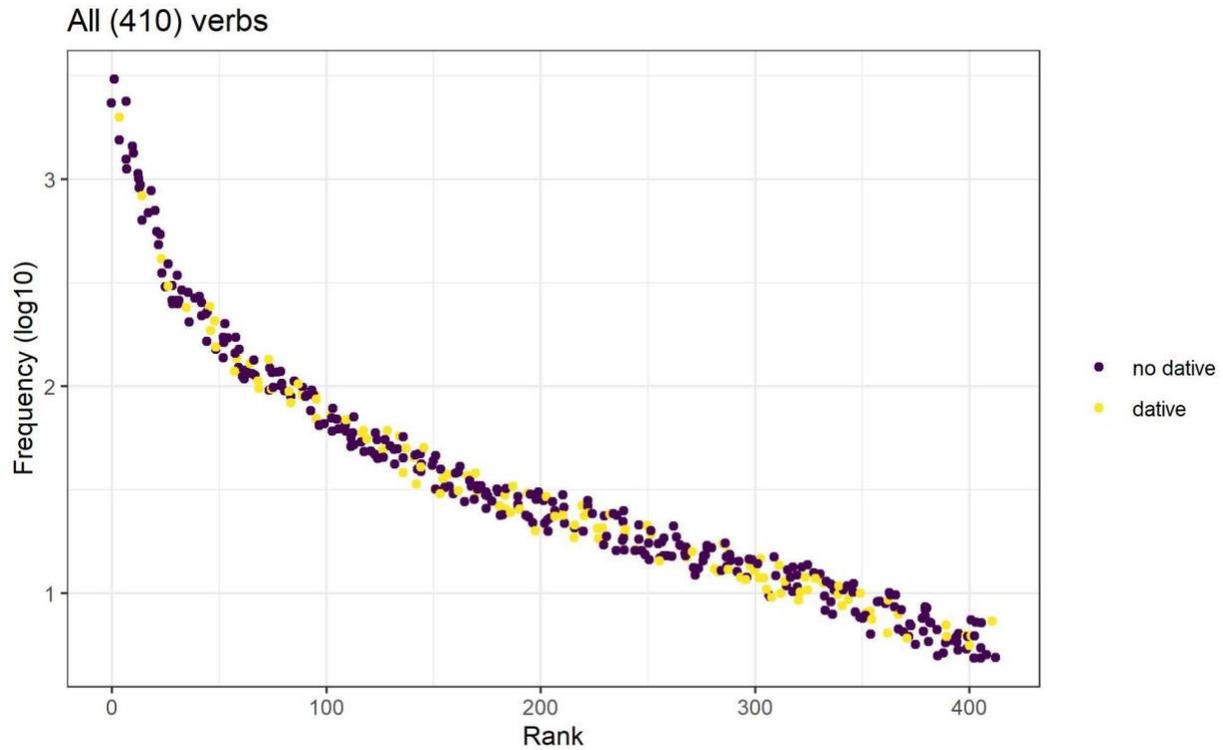
### 4.3.1 Overview

In Figure 4.1, the distribution of the verbs in the dataset is plotted with the corresponding case frames. A point plot with the log-transformed frequency shows the data is somewhat consistent with a Zipfian distribution (short head and long tail), an important assumption in the original derivation of the TP (Yang 2005, 2016).



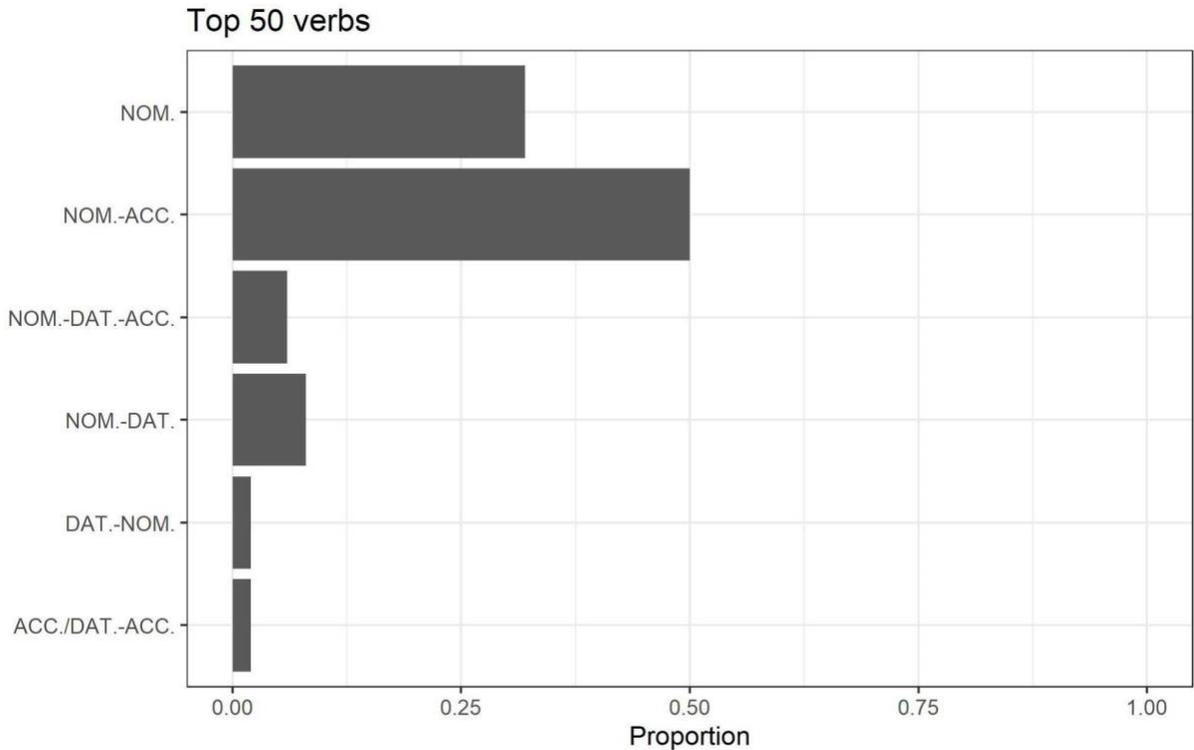
**Figure 4.1.** Distribution of analyzed verbs in the IceCASE corpus by case frame.

The legend illustrates the variety of case frames found within the analyzed verbs of the corpus (14 categories), which makes a visual overview complicated. Still, it should be apparent that intransitives with a nominative subject (the darkest points, for the category at the top of the legend) make up an important proportion of the data and are spread throughout the dataset. Within transitives, the same can be said about verbs with a nominative subject and an accusative object (NOM.-ACC.). To simplify this picture and zoom in on the datives, I plotted the same data in Figure 4.2 but only included information about the presence of a dative subject or (direct/indirect) object.



**Figure 4.2.** Distribution of analyzed verbs in the IceCASE corpus, presence of a dative argument.

Figure 4.2 shows that the dative arguments are also spread throughout the data, with a few notable characteristics in the distribution. First, the distribution of the dative arguments seems to cluster in a few areas, for example between ranks 100 and 150, and then between 200 and 250. The clearest cluster then appears around rank 300. Second, the top 100 verbs contain surprisingly few datives considering the reported early productivity, with the top 50 in particular showing few verbs with dative arguments. This can be viewed more clearly in Figures 4.3–4.6 which illustrate the proportion of different case frames at different intervals in the dataset.

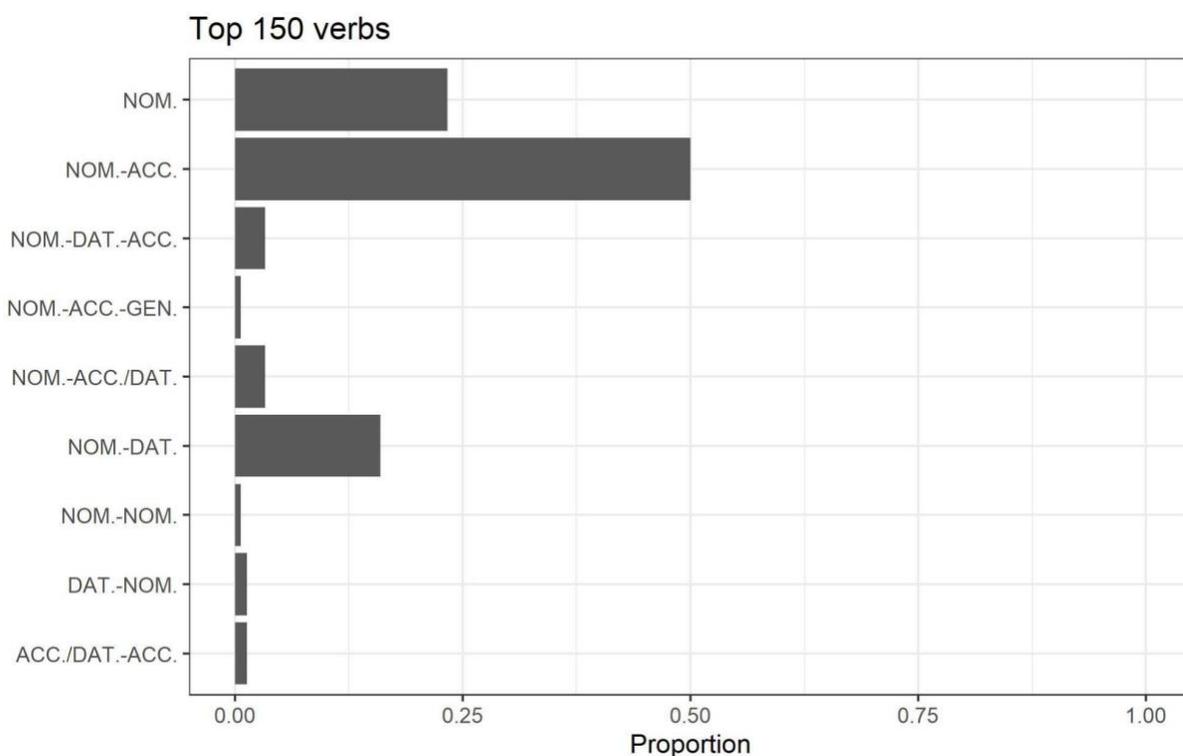


**Figure 4.3.** Proportion of different case frames in the IceCASE corpus, top 50 most frequent verbs.

Beginning with the top 50 verbs, Figure 4.3 shows that the NOM.-ACC. frame makes up for half of the data points, followed by intransitives with a nominative subject (16 items, 32%). Other frames, all transitive including datives, are 18% of the verbs, with the NOM.-DAT. construction reaching 8%.

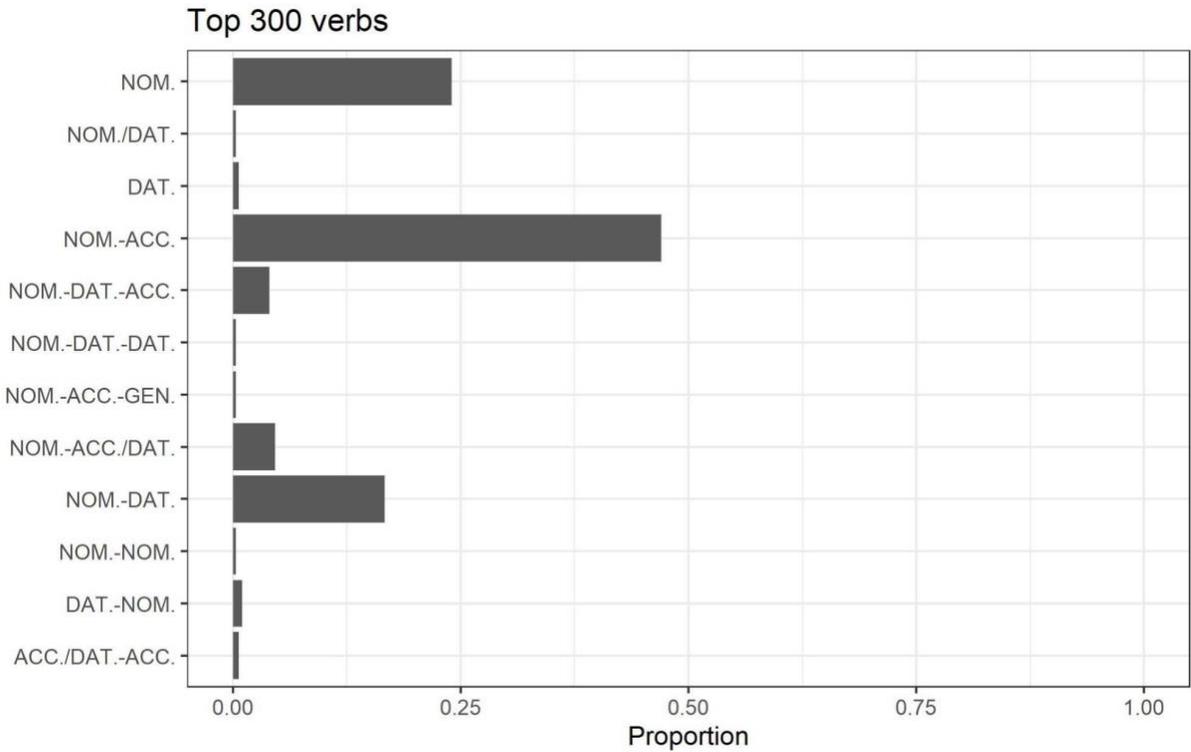
In contrast, the NOM.-DAT. frame reaches 16% in the top 150 verbs, doubling its proportion of the data, as can be seen in Figure 4.4. Verbs that alternate between accusative and dative also make their entrance between ranks 50 and 150, bringing case frames that include the dative up to 25%. The NOM.-ACC. frame still is the most prominent at 50% and intransitives with a nominative subject are at 23.3%. So even though the pattern is roughly the same between the top 50 and top 150 (NOM.-ACC first, intransitive NOM. second and NOM.-DAT. and NOM.-DAT.-ACC. fourth and fifth, there is an important increase in the proportion of the datives, particularly in the context of direct objects. Despite this, the proportion of ditransitives with dative goals goes down, from 6%

in the top 50 to 3.3% in the top 150. This is an interesting fact given the early acquisition of datives with indirect objects discussed previously.



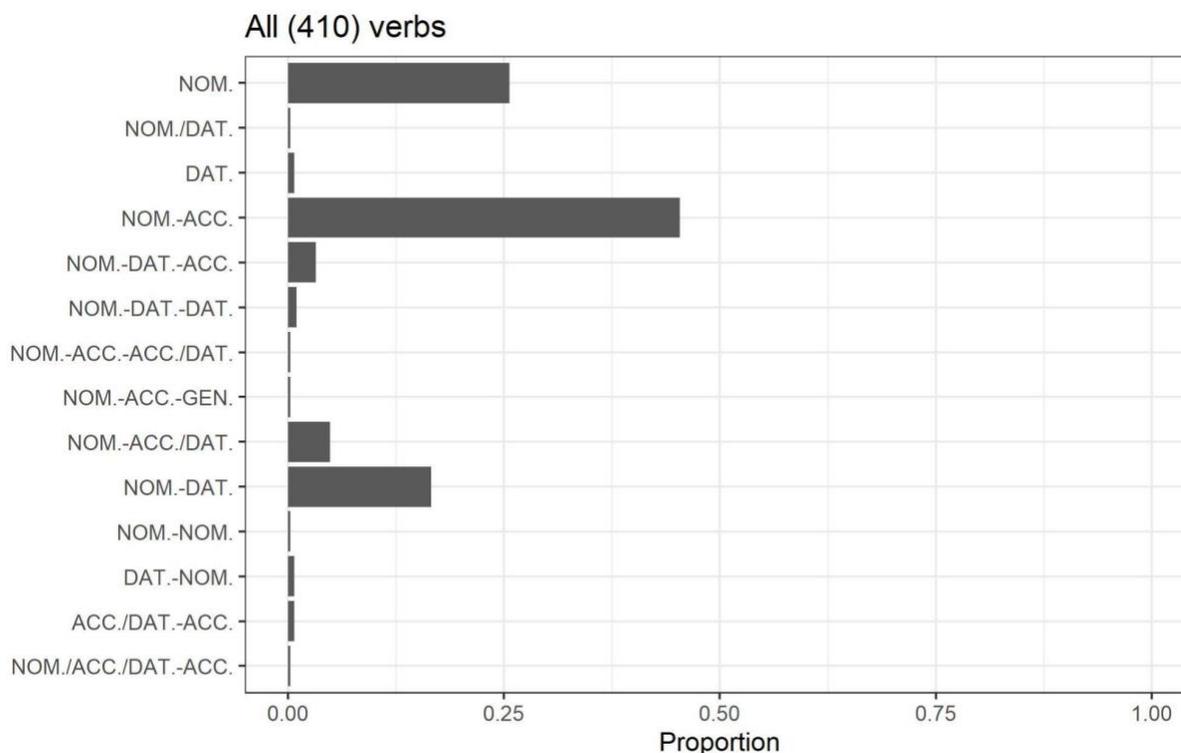
**Figure 4.4.** Proportion of different case frames in the IceCASE corpus, top 150 most frequent verbs.

Figure 4.5 shows the proportions for the top 300 verbs, displaying an almost identical picture as in the top 150, showing that distributional differences occur early on and then stabilize.



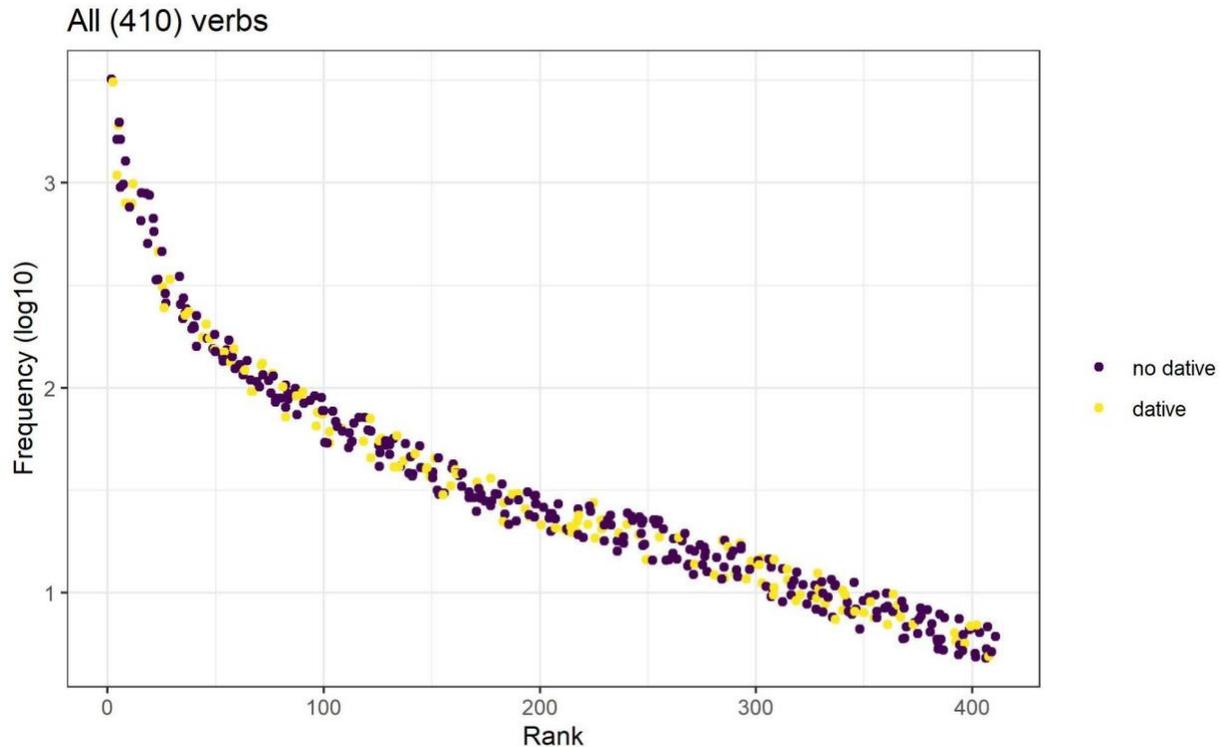
**Figure 4.5.** Proportion of different case frames in the IceCASE corpus, top 300 most frequent verbs.

Finally, Figure 4.6 contains the proportions for the whole dataset. The NOM.-ACC. frame constitutes 45.3% of the dataset, the intransitives with nominative subjects reach 25.6% and frames containing dative arguments are at 25.4%. This is again a picture very similar to the ones in the top 150 and 300 intervals, with frames containing a dative reaching 25.4%, most of them NOM.-DAT. (16.5%). 4.9% of the verbs alternate between accusative and dative and the rate of frames containing dative indirect objects goes up to 4.2%, still not as high as the 6% of the top 50 interval.



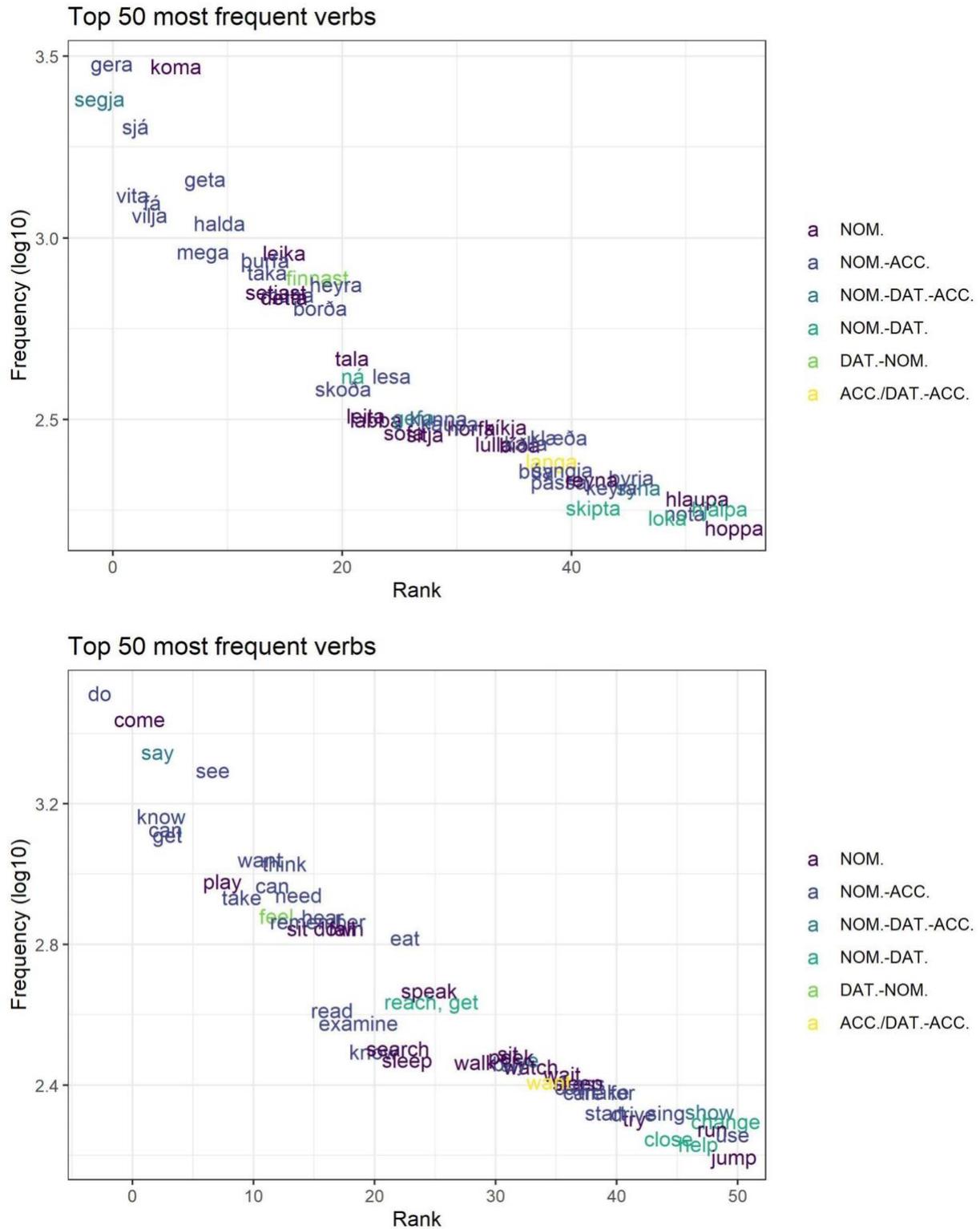
**Figure 4.6.** Proportion of different case frames in the IceCASE corpus, all verbs (410).

These results entail that distributionally, the association between the dative and goals might be stronger than the association with e.g. motion themes early in acquisition (as the experimental results in Chapters 2 and 3 seem to indicate), if frequency in the corpus is used as a proxy for frequency in the input and order of acquisition. This is further reinforced if optional dative reflexives (the *get yourself* construction shown in (5d)) are included in the analysis. Fourteen verbs appear with them in the corpus, six of those in the top 50 and all but one within the top 150 verbs. Figure 4.7 is a modified version of Figure 4.2 as it includes the reflexives, changing the picture considerably in the most frequent verbs, particularly the top 50. As Figure 4.7 shows in comparison to Figure 4.2, including the optional dative reflexives creates a scenario much richer in dative arguments with the most frequent verbs, all of them in the form of sentient goals. This suggests that initially, children build themselves dative case marking rules based on the more cross-linguistically expected association between the dative and applied arguments. The association with (motion) themes should emerge later.

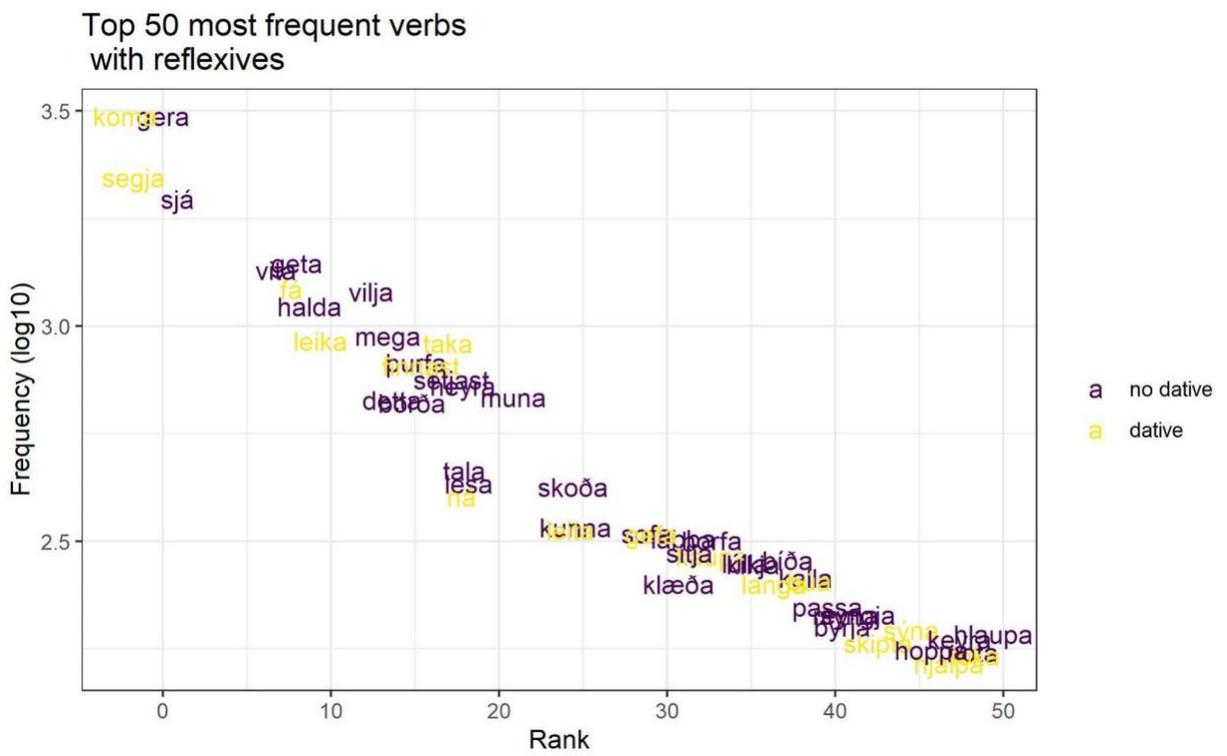
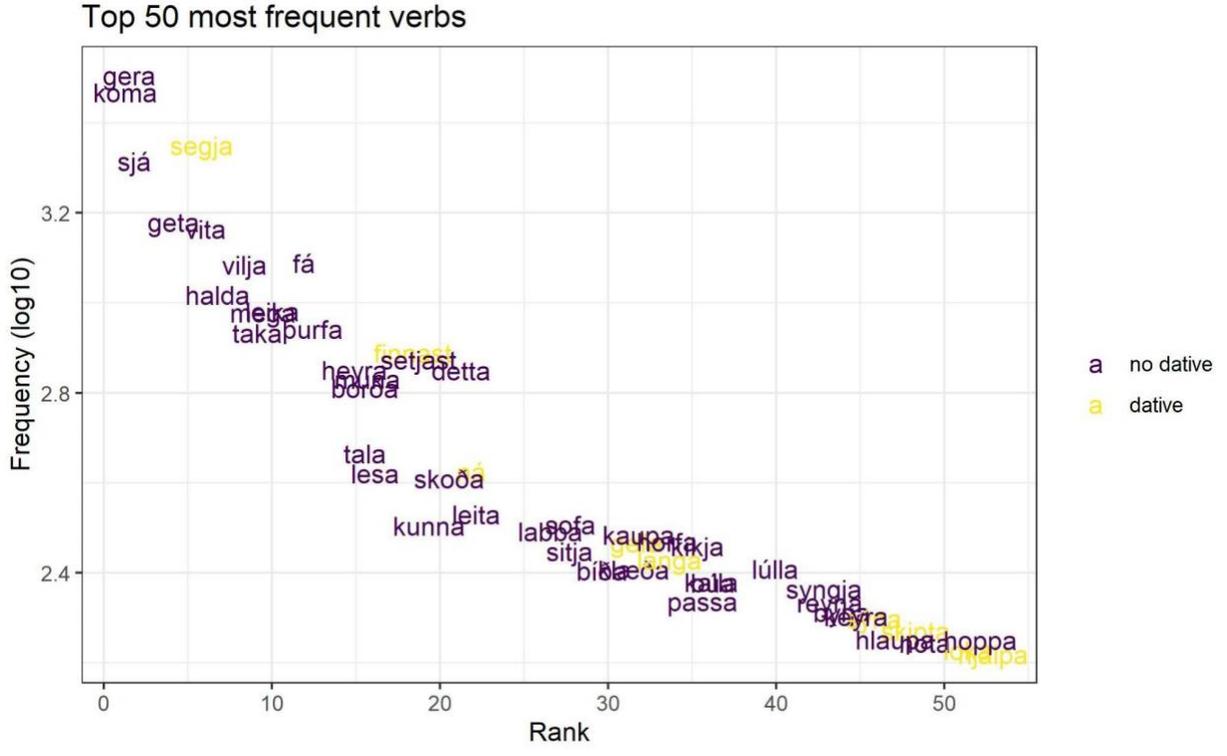


**Figure 4.7.** Distribution of analyzed verbs in the IceCASE corpus, by presence of a dative argument. Reflexives included.

This is further illustrated in Figures 4.8 and 4.9, where the verbs are used to label the data points in the 50 most frequent verbs. In the lower part of Figure 4.8, which includes all the case frames, translations are provided. In this figure we see the high rate and even distribution of the NOM.-ACC. and NOM. frames. It is interesting to see that two verbs with non-nominative subjects are present in the top 50 most frequent verbs, *finnast* ‘find’ and *langa* ‘want’. The verbs which take dative objects are a quite disparate set semantically (although none of them have proto-typical patients), including for example *ná* ‘get’, *skipta* ‘(ex)change, divide’, *gefa* ‘give’, *loka* ‘close’ and *hjálpa* ‘help’. As Figure 4.8 shows, the dative in general is also sparse in the top 50 verbs if the reflexives are not included. When they are included, as in the lower frame of Figure 4.9, the distribution is more even and the bias towards dative goals clearer.



**Figure 4.8.** Distribution of the top 50 most frequent verbs in the IceCASE corpus by case frame. Icelandic verbs above and translation below. Differences in verb position between languages are due to the jitter added to the plot for clarity.



**Figure 4.9.** Distribution of the top 50 most frequent verbs in the IceCASE corpus by presence of a dative argument, reflexives included in the lower frame. Differences in verb position between languages are due to the jitter added to the plot for clarity.

But does the semantic coherence of the dative matter for its productivity? Is it perhaps enough for children to assume that the dative is associated with indirect objects and reflexives (optional arguments), using form alone to discover rules? What about the extensive rate of the dative with direct objects and the non-agentive non-nominative subjects? In the following subsections, I attempt to implement and evaluate a possible developmental path for the Icelandic case marking system based on the distributional information available in children's language environment and the TP, a parameter-free learning model which seeks productivity over a maximum number of items as long as a viable number of exceptions is not exceeded. I begin by looking at the case in production before turning to the type of hypotheses which permit children to use the case as a cue for meaning in comprehension.

## 4.3.2 Datives in production

### 4.3.2.1 Form

I begin with form because previous research gives reason to believe that children begin with form (Yang 2020, Emond & Shi 2021). Table 8 provides an overview of the distribution of case by the annotated syntactic roles in the totality of the dataset, which includes 746 arguments. The emerging patterns are familiar, subjects are mostly nominative, direct objects are mostly accusative and indirect objects and reflexives are mostly dative. All these general patterns, except for the reflexives, furthermore have exceptions. The relevant question here is whether these exceptions are few enough for the rules to be productive, and whether this holds true with smaller as well as bigger vocabularies. As has been outlined, we take a maximal approach to the discovering of possible rules, trying to generalize over the highest number of items with the widest categories possible.

**Table 8.** Argument case by syntactic role, the IceCASE corpus.

	<b>subject</b>	<b>direct object</b>	<b>indirect object</b>	<b>reflexive</b>	<b>total</b>
<b>nominative</b>	399	4	0	0	403
<b>accusative</b>	0	209	2	0	211
<b>dative</b>	6	67	18	14	105
<b>genitive</b>	0	1	0	0	1
<b>nom./acc./dat.</b>	1	0	0	0	1
<b>nom./dat.</b>	1	0	0	0	1
<b>acc./dat.</b>	3	21	0	0	24
<b>total</b>	410	302	20	14	746

Following this, the first rule which is tested is [+subject  $\rightarrow$  +nominative]. The TP (see (28)) calculations are found in Table 9. It can be seen straight away that the rule is productive in every rank tier, with the number of exceptions (e) never exceeding the threshold ( $\theta_N$ ), and the number of items following therefore reaching sufficiency.

**Table 9.** Evaluation of the [+subject  $\rightarrow$  +nominative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b><math>\theta_N</math></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	50	38	12	2	yes
<b>top 100</b>	100	79	21	3	yes
<b>top 200</b>	150	121	29	4	yes
<b>top 200</b>	200	163	37	5	yes
<b>top 250</b>	250	205	45	6	yes
<b>top 300</b>	300	248	52	8	yes
<b>top 350</b>	350	291	59	9	yes
<b>all (410)</b>	410	342	68	11	yes

The randomized sampling results, starting at the top 100 tier (the sample tiers always have 50 less verbs, as described in section 4.2.2) also show productive nominative case marking for subjects in every tier. This means that under a TP approach, children should have a productive nominative case marking rule for subjects early on based on form only. This finding is relevant in the context of Dative Substitution, as will be further elaborated on in sections 4.3.2.2 and 4.4. For now, it is important to note that the exceptions to the nominative of subjects in Icelandic are not numerous enough to challenge the pattern's status as a productive rule.

The next association tested is the inverse pattern, [-subject → -nominative] (in Table 10). Of course, non-nominative should not be a possible category in the long-term, with different forms forcing distinction, but the syncretism patterns discussed in section 1.1.1 suggest that a differentiation between the accusative and the dative is not necessarily straightforwardly acquired, as they often co-occur on the surface. A similar argument can be made for a non-subject category, which I do not assume for the long-term either in the context of case marking.

**Table 10.** Evaluation of the [-subject → -nominative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	44	33	11	1	yes
<b>top 100</b>	92	72	20	1	yes
<b>top 200</b>	135	108	27	3	yes
<b>top 200</b>	176	142	34	4	yes
<b>top 250</b>	217	177	40	4	yes
<b>top 300</b>	253	208	45	4	yes
<b>top 350</b>	293	242	51	4	yes
<b>all (410)</b>	336	279	57	4	yes

Nevertheless, Table 10 shows that if this type of rule emerges (perhaps for some but not all learners), it would (unsurprisingly) be productive for non-subjects at all times, with a maximum of four exceptions.<sup>40</sup> The random sample testing yields the same results, but as I have just outlined, the case marking system itself does not warrant this type of rule, even though it might be justified in a developmentally transient system where oblique cases are not distinguished beyond this category.

I therefore proceed to testing a narrower rule, [-subject → +accusative]. Table 11 shows that this rule is not productive at any point in time. The counts here do not include verbs taking both accusative and dative objects as exceptions, and still the number of exceptions exceeds the threshold for productivity at every level.

**Table 11.** Evaluation of the [-subject → +accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	44	33	11	15	no
<b>top 100</b>	92	72	20	33	no
<b>top 200</b>	135	108	27	44	no
<b>top 200</b>	176	142	34	55	no
<b>top 250</b>	217	177	40	69	no
<b>top 300</b>	253	208	45	78	no
<b>top 350</b>	293	242	51	95	no
<b>all (410)</b>	336	279	57	104	no

The randomized sample analysis yields the same result, so I divide the data further, as the learner is expected to in the learning model. This directs us to the [+direct object → +accusative] rule tested in Table 12.

<sup>40</sup> Nominative objects appear with transitive dative subject verbs in Icelandic (see e.g. Árnadóttir & E.F. Sigurðsson (2012)). Interestingly, DS does not entail that the accusative objects of DS verbs become nominative.

**Table 12.** Evaluation of the [+direct object → +accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	34	25	9	5	yes
<b>top 100</b>	75	58	17	16	yes
<b>top 150</b>	116	92	24	26	no
<b>top 200</b>	155	125	30	35	no
<b>top 250</b>	191	155	36	44	no
<b>top 300</b>	226	185	41	52	no
<b>top 350</b>	261	215	46	65	no
<b>all (410)</b>	302	250	52	72	no

Here, the [+direct object → +accusative] pattern only results in a productive rule in the first two tiers. Counting objects which appear both in the accusative and dative as exceptions does not change the results and the randomized sample testing furthermore yields the same results as in Table 12. If we keep using frequency as a proxy for order of acquisition, this means that associating direct objects with the accusative is only viable early on, with the exceptions becoming too numerous as more verbs are learned. This is despite the fact that the accusative is the dominant case marking for direct objects throughout, and a consequence of the logarithmic nature of the TP, where bigger sets tolerate less exceptions proportionally. This finding entails that the direct object category needs to be divided further at a certain point in acquisition, an attempt at this is found in section 4.3.2.3. Another option is that the category is not divided further, resulting in no productive rule for direct objects in Icelandic. This would predict that case on direct objects would be variably assigned based on the distributions in the input instead. This idea is further discussed in section 4.4. Whatever the conclusion, the initial productivity of accusatives for direct objects presumably pushes the learner to attempt the discovery of a rule for the rest of the data, arguments which are neither subjects nor direct objects.

As these arguments are majoritarily dative, this is the rule tested in Table 13: [-subject, -direct object → +dative]. This rule is productive throughout as the table shows, and the randomized sample testing yields the same results.

**Table 13.** Evaluation of the [-subject, -direct object → +dative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b><math>\theta_N</math></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	10	6	4	0	yes
<b>top 100</b>	17	11	6	0	yes
<b>top 200</b>	19	13	6	1	yes
<b>top 200</b>	20	14	6	1	yes
<b>top 250</b>	25	18	7	1	yes
<b>top 300</b>	26	19	7	1	yes
<b>top 350</b>	30	22	8	2	yes
<b>all (410)</b>	34	25	9	2	yes

Therefore, form should be enough to create a productive dative case marking rule for these arguments, which consist of indirect objects and optional reflexives. If we assume that the [-subject, -direct object] category should be divided further into those two sets, the dative would still be productive for both. Interestingly, the arguments in question are semantically coherent, with all of them falling under the category of goals. This coherence, along with the fact that form seems to be enough to determine case early on, points towards a scenario where distributional regularities could lead to the discovery of form-meaning mappings where the dative is associated with applicative semantics early on.

On the other hand, the lack of a productive case marking rule for direct objects might push the learner towards partitioning the data based on other factors, with the semantic coherence of the dative with indirect objects and reflexives perhaps reinforcing the idea that the dative should be associated with particular semantic classes. In section

4.3.2.3, we test rules combining semantic features and regularities in form, but we begin with a productivity analysis which is purely based on meaning.

#### 4.3.2.2 Meaning

Although there is independent evidence for the precedence of pattern recognition in form as compared to meaning, and various analyses point towards morphological case being *recessive* in Icelandic as compared to syntactic position when it comes to marking argument role (Kiparsky 1997),<sup>41</sup> it is of interest to investigate whether viable productive rules arise based on associations between case and meaning alone. After all, the only absolute case marking generalization discussed so far is that all agents are nominative. This information can be found in Table 14, which summarizes the distribution of case based on the argument role annotation of the IceCASE corpus. As expected from the description of the annotation methodology in section 4.2.1, this summary table contains more layers than the form summary in section 4.3.2.1. The annotation was based on the use of each verb in the corpus, meaning that the arguments of some verbs have variable argument roles, e.g. the object of *sópa* ‘sweep’ (used in the minimal pair task as well) which appears both as a patient and theme in the corpus and therefore receives the annotation ‘PT’ in the table below.

Despite this complexity in the annotation, various familiar patterns can be found in Table 14 as well. The nominative agents have already been mentioned. Additionally, the majority of patients are accusative and the majority of goals dative. For the themes and experiencers, no case has a convincing majority, although most themes are accusative and nominative is the most frequent case for the experiencers. Again, the question here is whether these patterns are viable as productive rules, and whether they are so throughout the dataset.

---

<sup>41</sup> Kiparsky (1997, p. 480) describes how morphosyntactic case features can be recessive in the following way: “The abstract case features of an argument may fail to unify with the morphosyntactic case features assigned to it either by morphology (“quirky case”), or by position (“scrambling”).”

**Table 14.** Argument case by thematic role, the IceCASE corpus.

	<b>nominative</b>	<b>accusative</b>	<b>dative</b>	<b>genitive</b>	<b>NOM./ACC./DAT.</b>	<b>NOM./DAT.</b>	<b>ACC./DAT.</b>	<b>total</b>
<b>agent</b>	324	0	0	0	0	0	0	324
<b>patient</b>	0	97	7	0	0	0	1	105
<b>theme</b>	42	106	51	1	0	0	2	202
<b>goal</b>	0	3	39	0	0	0	0	42
<b>exp.*</b>	14	0	7	0	0	0	3	24
<b>AT</b>	20	0	0	0	0	0	0	20
<b>AG</b>	1	0	0	0	0	0	0	1
<b>AE</b>	0	0	0	0	1	0	0	1
<b>ATE</b>	1	0	0	0	0	0	0	1
<b>PG</b>	0	0	0	0	0	0	6	6
<b>PT</b>	0	5	1	0	0	0	11	17
<b>TG</b>	0	0	0	0	0	0	1	1
<b>TE</b>	1	0	0	0	0	1	0	2
<b>total</b>	403	211	105	1	1	1	24	746

\*A = agent, P = patient, T = theme, G = goal, E/exp = experiencer

The first rule which I would test under the maximal approach taken here is the association of the nominative with agents. But as this rule has no exceptions, it is not necessary to test it. I therefore turn to the [-agent → -nominative] rule, where similar caveats as in the [-subject → -nominative] rule (Table 10 in subsection 4.3.2.1) apply: The categories are not considered sustainable in the long term in the context of case marking.

**Table 15.** Evaluation of the [-agent → -nominative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	62	47	15	13 (17)	yes (no)*
<b>top 100</b>	124	99	25	17 (30)	yes (no)*
<b>top 150</b>	163	131	32	27 (43)	yes (no)*
<b>top 200</b>	227	186	41	32 (50)	yes (no)*
<b>top 250</b>	276	227	49	38 (58)	yes (no)*
<b>top 300</b>	319	264	55	41 (63)	yes (no)*
<b>top 350</b>	364	303	61	45 (67)	yes (no)*
<b>all (410)</b>	419	350	69	58 (80)	yes (no)*

\*including polythematic arguments as exceptions changes the result

Despite this, Table 15 shows that the rule, if it would be formed, would be productive throughout the dataset. This changes if we include polythematic arguments (e.g. a verb argument which appears as an agent and a theme) as exceptions (as shown in parentheses). The random samples yield the same picture, except for samples 1 (rank <100 – 50) and 4 (rank <250 – 50) where the rule is productive regardless of whether polythematic arguments are included or not. Narrowing the rule with the accusative, on the other hand, does not yield any productivity (Table 16) in any conditions tested (samples included), despite the fact that the majority of non-agents are accusative.

**Table 16.** Evaluation of the [-agent → +accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	62	47	15	28	no
<b>top 100</b>	124	99	25	50	no
<b>top 150</b>	163	131	32	70	no
<b>top 200</b>	227	186	41	86	no
<b>top 250</b>	276	227	49	105	no
<b>top 300</b>	319	264	55	122	no
<b>top 350</b>	364	303	61	142	no
<b>all (410)</b>	419	350	69	164	no

On the other hand, if the rule is narrowed down to [+patient → +accusative] (Table 17), productivity is observed in every condition (including the maximum number of exceptions).

**Table 17.** Evaluation of the [+patient → +accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	8	4	4	2	yes
<b>top 100</b>	23	16	7	2	yes
<b>top 150</b>	37	27	10	2	yes
<b>top 200</b>	57	43	14	3	yes
<b>top 250</b>	73	56	17	4	yes
<b>top 300</b>	86	67	19	6	yes
<b>top 350</b>	107	84	23	6	yes
<b>all (410)</b>	128	102	26	8	yes

Having accounted for the agents and patients, I move on to the remaining arguments, testing first the productivity of a dative rule for arguments which are neither agents nor patients (i.e. goals, experiencers and themes) in Table 18, since a [-agent, -patient → -nominative] rule is productive but not sufficiently specified. The dative never reaches the majority of items in such a rule (and accusative never reaches sufficiency either), yielding a clear lack of productivity in every condition.

**Table 18.** Evaluation of the [-agent, -patient → +dative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	55	41	14	36	no
<b>top 100</b>	102	80	22	54	no
<b>top 150</b>	149	119	30	84	no
<b>top 200</b>	182	147	35	101	no
<b>top 250</b>	212	172	40	119	no
<b>top 300</b>	243	199	44	132	no
<b>top 350</b>	273	224	49	142	no
<b>all (410)</b>	316	261	55	167	no

I therefore narrow the rule down to two of the three remaining roles within this category of [-agent, -patient] based on their frequent animacy requirement, and test whether the [+goal, +experiencer → +dative] pattern is productive in Table 19.<sup>42</sup> The results are mixed, with the rule attaining productivity when the whole dataset is considered (but without including ambiguous arguments as exceptions (number in parentheses)).

<sup>42</sup> Since the remaining set of themes is larger, they should be tested before, but I present the rules in this order for the sake of clarity.

**Table 19.** Evaluation of the [+goal, +experiencer → +dative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	21	14	7	8	no
<b>top 100</b>	33	24	9	13	no
<b>top 150</b>	44	32	12	14	no
<b>top 200</b>	49	36	13	16	no
<b>top 250</b>	57	43	14	17	no
<b>top 300</b>	62	47	15	17	no
<b>top 350</b>	69	53	16	17	no
<b>all (410)</b>	75	58	17	17 (21)	yes (no)

Including polythematic arguments and arguments which can appear both in the accusative and dative (variation in subject case) as exceptions does not change the results except in the last tier. The sample testing yields more productivity, with samples 1 (rank <100 - 50), 3 (rank <200 - 50) and 4 (rank <250 - 50) all showing a distribution where the exceptions do not exceed the threshold if ambiguous items are not included (sample 1 still shows productivity in all conditions). Since this is inconclusive but could potentially be a viable rule early on, I break the category down further and test the [+goal → +dative] rule in Table 20, which yields productivity in every condition.

**Table 20.** Evaluation of the [+goal → +dative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	12	8	4	1	yes
<b>top 100</b>	19	13	6	2	yes
<b>top 150</b>	26	18	8	3	yes
<b>top 200</b>	28	20	8	3	yes

<b>top 250</b>	35	26	9	3	yes
<b>top 300</b>	38	28	10	3	yes
<b>top 350</b>	44	33	11	3	yes
<b>all (410)</b>	49	37	12	4	yes

I now move on to the slightly trickier testing of the experiencers in Table 21, before addressing themes. One of the issues which needs to be addressed when determining how to test experiencers as a set in the context of case marking is the adjectival predicates which appear with dative subjects.<sup>43</sup> In the corpora used to build the IceCASE dataset, there are four of them: *vera kalt* ‘be cold’ (frequencies fit with the top 100 predicates), *vera heitt* ‘be hot’ (top 100 as well), *vera illt* ‘be sore, hurt’ (top 150) and *vera hlýtt* ‘be warm’ (top 350).

**Table 21.** Evaluation of the [+experiencer → +dative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	9	5	4	6	nominative
<b>top 100</b>	15 (17)	10 (11)	5 (6)	9	nominative
<b>top 150</b>	18 (21)	12 (15)	6 (6)	11 / 13	<i>nominative</i>
<b>top 200</b>	21 (24)	15 (17)	6 (7)	13 / 15	<i>nominative</i>
<b>top 250</b>	22 (25)	15 (18)	7 (7)	13 / 16	<i>nominative</i>
<b>top 300</b>	24 (27)	17 (19)	7 (8)	13 / 16	<i>nominative</i>
<b>top 350</b>	25 (29)	18 (21)	7 (8)	13	no
<b>all (410)</b>	28 (32)	20 (21)	8 (9)	14	no

It is possible to argue for developmental scenarios where these are included in the experiencer category and scenarios where they are not, so both scenarios are included

<sup>43</sup> The dataset also includes one adjective that licenses dative on an object, *líkur* ‘similar’.

in the testing above, with the numbers in parentheses showing the counts if the adjectival predicates are included. Here, nominative experiencers (not accusative ones since they all display DS) are counted as exceptions. Adding the four adjectives which appear with dative subjects does not change the results for the (non-)productivity of the dative but does in part for the productivity of the nominative which is not maintained in the italicized cases if polythematic arguments are counted as exceptions. The main finding to draw from the evaluation of the [+experiencer → +dative] rule is that the dative is never productive for experiencers. Nominative is in fact a better candidate, but this productivity is fragile and non-existent when the whole dataset is taken into account, pointing towards either subrules or a scenario without a productive rule.

The testing with the random sampling yields similar results (Table 22), with the nominative being a better candidate than the dative in general.

**Table 22.** Evaluation of the [+experiencer → +dative] rule in the IceCASE corpus, randomized samples.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>Sample 1</b>	7 (9)	3 (5)	4 (4)	4 / 5	yes - nominative
<b>Sample 2</b>	12 (15)	8 (10)	4 (5)	7 / 8	<i>nominative</i>
<b>Sample 3</b>	14 (17)	9 (11)	5 (6)	9 / 11	nominative
<b>Sample 4</b>	18 (21)	12 (15)	6 (6)	11 / 13	<i>nominative</i>
<b>Sample 5</b>	20 (23)	14 (16)	6 (7)	11 / 13	no
<b>Sample 6</b>	24 (28)	17 (20)	7 (8)	12 / 14	no
<b>Sample 7</b>	24 (28)	17 (20)	7 (8)	13 / 14	no

This is only true under certain testing conditions, but notice that the dative is productive under one scenario in the first tier.

Since the productivity of the dative is uncontested in the context of Dative Substitution, I decided to dig deeper for the experiencer rule, restricting the set to non-

nominative (oblique) experiencers and therefore counting historically nominative and accusative experiencers as exceptions.<sup>44</sup> The calculations can be found in Table 23, which shows that the dative is productive throughout for oblique experiencers.

**Table 23.** Evaluation of the [+obl. experiencer → +dative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	2	?	?	1	?
<b>top 100</b>	4 (6)	2 (3)	2 (3)	2	yes*
<b>top 150</b>	5 (8)	2 (5)	3 (3)	2	yes
<b>top 200</b>	6 (9)	3 (5)	3 (4)	2	yes
<b>top 250</b>	7 (10)	3 (6)	4 (4)	2	yes
<b>top 300</b>	9 (12)	5 (8)	4 (4)	2	yes
<b>top 350</b>	10 (14)	6 (9)	4 (5)	2	yes
<b>all (410)</b>	12 (16)	8 (11)	4 (5)	4	yes

\*accusative and dative are productive

These results hold in all testing conditions. This means that within the subclass (exceptions) of subjects which appear as non-nominative, the dative is productive. But why should such a subrule arise? This is discussed in subsection 4.4.

To summarize quickly the status of the rule discovery procedure based on meaning alone, for now I have reached the following results: [+agent → +nominative], [+patient → +accusative] and [+goal → +dative] are all productive rules throughout the dataset and viable in the long-term. On the other hand, a productive and sustainable case marking rule for experiencers was not found. Nominative is a better candidate than the dative in this context, but the results entail a need for more subrules or a lack of productive rule. In the subset of non-nominative experiencers though, dative is found to be productive.

<sup>44</sup> This is in fact a form and meaning case marking (see section 4.3.2.3) rule, but I describe it here for the sake of continuation.

I now turn to the remaining role category, themes, and begin by testing the productivity of the majority pattern, [+theme → +accusative]. As Table 24 shows, this is only viable in the first tier of the data and only if ambiguous arguments are not included. The randomized sample testing yields no productivity at all.

**Table 24.** Evaluation of the [+theme → +accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	35	26	9	7 / 12	yes
<b>top 100</b>	72	56	16	19	no
<b>top 150</b>	105	78	22	33	no
<b>top 200</b>	134	107	27	43	no
<b>top 250</b>	161	130	31	55	no
<b>top 300</b>	182	148	34	63	no
<b>top 350</b>	206	168	38	89	no
<b>all (410)</b>	242	198	44	93	no

In line with previous research and the results from the novel verb production task (in section 3.3.2.2), I restrict the rule to themes which do not undergo change of location, as motion themes are more frequently dative. As Table 25 shows, this still does not yield a productive rule under any conditions. The sample testing yields the same results except for the accusative being productive in the first sample.

**Table 25.** Evaluation of the [+theme, -motion → +accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	27	19	8	9	no
<b>top 100</b>	52	39	13	20	no
<b>top 150</b>	73	56	17	32	no
<b>top 200</b>	82	64	18	35	no
<b>top 250</b>	97	76	21	44	no
<b>top 300</b>	107	85	22	47	no
<b>top 350</b>	116	92	24	53	no
<b>all (410)</b>	139	111	28	65	no

A similar picture emerges when the [+theme, +motion → +dative] rule is tested in Table 26.

**Table 26.** Evaluation of the [+theme, +motion → +dative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	8	4	4	5	accusative
<b>top 100</b>	20	13	7	11	no
<b>top 150</b>	34	24	10	18	no
<b>top 200</b>	53	40	13	28	no
<b>top 250</b>	66	50	16	34	no
<b>top 300</b>	76	58	18	41	no
<b>top 350</b>	91	71	20	43	no
<b>all (410)</b>	103	81	22	53	no

The only productive rule is found in the first tier, where accusative is in fact productive. In the randomized sample testing, a productive rule can also only be found in the first tier, where the dative turns out to be productive. This serves to show how split the theme category is. Additionally, this lack of productive rules in the theme category can in part be traced to the fact that it includes many nominative theme subjects under an approach that does not distinguish between subjects and objects. This might point towards the necessity of taking both form and meaning into account.

#### 4.3.2.3 Form and meaning

Having assessed case marking rule productivity based on form and meaning (mostly separately without successfully accounting for the system as a whole perhaps suggests that the two approaches need to be merged. If this is the case, it remains to be determined which initial path should be assumed. Based on the results presented in the previous two subsections, as well as previous research which (1) points to the early acquisition of distributive patterns without association to meaning and (2) shows that regularities in form are used to acquire meaning, an initial discovery procedure which relies on form alone seems more viable. In fact, for the top 100 verbs, productive rules based on form can be discovered for the entire dataset. After the top 100 verbs, it is only in the category of direct objects that no productive rules are found. Based on this, the last series of evaluations conducted here are within the direct object category.

I started by testing whether simpler features such as +animate and +motion would yield productive rules with direct objects, but this did not prove to be the case (although the [+direct object, +motion → dative] pattern is productive in the first randomized sample). I therefore proceeded to the largest thematic role category within the direct objects, testing [+direct object, +patient → accusative] in Table 27. This yields productivity in all the conditions tested, leaving only direct objects which are goals and themes unaccounted for. Table 28 shows that for those arguments (direct objects which are goals and themes), the accusative is not productive except for the first tier, and this holds true in all conditions tested.

**Table 27.** Evaluation of the [+direct object, +patient → accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b><math>\theta_N</math></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	8	5	3	2	yes
<b>top 100</b>	23	16	7	2	yes
<b>top 150</b>	37	27	10	2	yes
<b>top 200</b>	57	43	14	3	yes
<b>top 250</b>	73	56	17	4	yes
<b>top 300</b>	91	71	20	6	yes
<b>top 350</b>	107	85	22	6	yes
<b>all (410)</b>	128	102	26	8	yes

**Table 28.** Evaluation of the [+direct object, -patient → accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b><math>\theta_N</math></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	27	21	8	4	yes
<b>top 100</b>	55	42	13	15	no
<b>top 150</b>	83	65	18	25	no
<b>top 200</b>	101	80	21	35	no
<b>top 250</b>	132	105	27	41	no
<b>top 300</b>	150	121	29	47	no
<b>top 350</b>	171	138	33	60	no
<b>all (410)</b>	197	160	37	65	no

Discarding the direct objects which also are goals is a logical next step (although perhaps taken later in acquisition) as they are all (N = 9) dative except for one, yielding clear productivity once sufficiency is acquired. I then only have direct object themes to account for, which are majoritarily accusative as Table 14 shows. Still, the exceptions become too numerous after the first tier, with the randomized sampling giving the same results.

**Table 29.** Evaluation of the [+direct object, +theme → accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	26	19	7	3	yes
<b>top 100</b>	53	40	13	14	no
<b>top 150</b>	79	61	18	22	no
<b>top 200</b>	103	81	22	30	no
<b>top 250</b>	123	98	25	37	no
<b>top 300</b>	139	111	28	42	no
<b>top 350</b>	159	128	31	54	no
<b>all (410)</b>	182	148	34	57	no

This results in an additional partitioning of the data, excluding the direct object themes which undergo a change of location. Such a split yields a productive accusative case marking rule under all conditions, as illustrated in Table 30. This leaves the learner with only 90 arguments left to account for, direct object themes which undergo a change of location in Table 31. Although these are majoritarily dative in the dataset as a whole, the first tiers does yield the accusative as productive. After this, no productive case marking rule is viable. Randomized sample testing yields the same pattern, but adds dative productivity in sample 2.

**Table 30.** Evaluation of the [+direct object, +theme, -motion → accusative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	20	14	6	2	yes
<b>top 100</b>	37	22	10	5	yes
<b>top 150</b>	51	39	12	10	yes
<b>top 200</b>	58	44	14	11	yes
<b>top 250</b>	66	51	15	13	yes
<b>top 300</b>	74	57	17	14	yes
<b>top 350</b>	79	61	18	16	yes
<b>all (410)</b>	91	71	20	17	yes

**Table 31.** Evaluation of the [+direct object, +theme, +motion → dative] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	6	3	3	5	accusative
<b>top 100</b>	16	11	5	6	no
<b>top 150</b>	28	20	8	14	no
<b>top 200</b>	45	32	11	20	no
<b>top 250</b>	57	43	14	26	no
<b>top 300</b>	64	49	15	29	no
<b>top 350</b>	79	61	18	31	no
<b>all (410)</b>	90	70	20	37	no

For the case marking of these arguments, learners are therefore left with either no productive rule (which should result in variation between the accusative and the dative)

or the task to divide the data further, perhaps into more fine-grained verb classes as the ones discussed by Maling (2002) and Jónsson (2013) (although those might not be distributionally or conceptually available to young children).

#### 4.3.2.4 A note on adjectival predicates and ditransitives

The last two counts I conduct in the production direction are within adjectival predicates on the one hand and ditransitives on the other. As I described earlier (in section 4.3.2.2), four adjectives in the corpus appear with dative subjects, all denoting some type of physical (dis)comfort (although in adult Icelandic, the semantic scope of adjectival predicates with dative subjects is much wider). This is in line with the overgeneralization example discussed in section 1.2 (in the text above the examples in (9)), where a dative subject was used with *vera sveitt* ‘be sweaty’ (which usually appears with a nominative subject). In fact, an interesting distribution comes to light if all adjectives denoting physical (dis)comfort in the child language corpus are counted. There are 19 of them, a set which should tolerate 6 exceptions. This means that the dative subject option should not be productive. But interestingly, if we look at the top four most frequent adjectives, three of them can appear with a dative subject. Early on, it might therefore be possible for children to form a productive dative subject rule for adjectives denoting physical (dis)comfort.

The second question I want to address using a TP approach is the idea that dative indirect objects perhaps are not productive in Icelandic because of the direct object often being dative and speakers of Icelandic preferring a NOM.-DAT.-ACC. ditransitive frame to the NOM.-DAT.-DAT. This would account for the Icelandic twist on Baker’s paradox mentioned in Chapter 1. The productivity of the dative for direct objects would therefore in a sense block the possibility of a dative indirect objects. One way to approach this is to evaluate the productivity of NOM.-DAT.-ACC. for ditransitives in Icelandic. The child language corpus in fact contains 19 ditransitives and 13 of them are NOM.-DAT.-ACC.. A set of 19 allows for 6 exceptions, which fits the number of NOM.-DAT.-DAT. items perfectly and should therefore result in productivity of the NOM.-DAT.-ACC. frame with NOM.-DAT.-DAT. as exceptions. Another possible explanation for the lack of indirect objects with novel verbs in Icelandic is that the double object construction itself is not productive (see Salome Lilja Sigurðardóttir’s forthcoming MA-thesis) and that indirect

objects in the form of PPs are the only available productive option for novel verbs denoting events with a recipient/goal in addition to a direct object theme.

### 4.3.3 Datives in comprehension

So far, I have only addressed possible hypotheses in the direction assumed for production, where I evaluate whether children can derive case marking rules based on syntactic or thematic roles. On the other hand, as I showed in Chapter 2, children learning Icelandic can also to some extent use case as a cue for verb meaning in comprehension. This entails the formation of rules with case as a starting point, which is the type of patterns I now turn to.

#### 4.3.3.1 Form

As in subsection 4.3.2.1, I begin with associations between case and syntactic role. On the other hand, it could be argued that a step where children use case as a cue for syntactic role is unnecessary, and that it should be assumed that children use case as a cue for meaning within categories such as subject and object. This is the approach that I take in subsection 4.3.3.3 and the one I ultimately adopt, but I begin with the case-position mappings for the sake of consistency. Table 8, as described before, provides an overview of the distribution of case by the annotated syntactic roles in the totality of the dataset, which includes 746 arguments. The inverted patterns provide a slightly different, but still expected, picture. Most nominative arguments are subjects and most accusative arguments are direct objects. Dative arguments are more split, with the majority of them still occurring as direct objects, but indirect objects and reflexives also have a sizable share. Like before, the question is whether the evaluation of those patterns with the TP yields productivity and what that tells us about the interpretation of the experimental data in Chapter 2.

I start with the evaluation of the [+nominative → +subject] rule in Table 32, which yields productivity in every condition. As before, it can be seen straight away that the rule is productive in every rank tier of Table 32, with the number of exceptions ( $e$ ) never exceeding the threshold ( $\theta_N$ ), and the items following the rule thus reaching sufficiency.

Children can therefore assume that a nominative argument is the subject of a verb, despite the few nominative objects which appear in the IceCASE corpus.

**Table 32.** Evaluation of the [+nominative → +subject] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	49	38	12	1	yes
<b>top 100</b>	98	77	21	1	yes
<b>top 200</b>	149	120	29	3	yes
<b>top 200</b>	199	162	37	4	yes
<b>top 250</b>	249	204	45	4	yes
<b>top 300</b>	297	245	52	4	yes
<b>top 350</b>	346	287	59	4	yes
<b>all (410)</b>	405	338	67	4	yes

The reverse rule, [-nominative → -subject] is also productive in every condition, see Table 33, meaning that children could reliably form the assumption that non-nominative arguments are non-subjects, despite the exceptions of oblique subject verbs. Including the adjectival predicate oblique subjects as exceptions does not change the picture and specifying the non-nominative case into accusative, dative or genitive also yields the same results in all conditions (randomized samples as well). I therefore move to more specific categories, testing the [+accusative → +direct object] rule in Table 34.

**Table 33.** Evaluation of the [-nominative → -subject] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	45	34	11	2 (4)	yes
<b>top 100</b>	94	74	20	3 (5)	yes
<b>top 150</b>	136	109	27	4 (7)	yes
<b>top 200</b>	177	143	34	5 (8)	yes
<b>top 250</b>	219	279	40	5 (8)	yes
<b>top 300</b>	257	211	46	7 (10)	yes
<b>top 350</b>	298	246	52	8 (12)	yes
<b>all (410)</b>	343	285	58	9 (13)	yes

**Table 34.** Evaluation of the [+accusative → +direct object] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	30	22	8	0	yes
<b>top 100</b>	61	14	47	0	yes
<b>top 150</b>	93	73	20	1	yes
<b>top 200</b>	123	98	25	1	yes
<b>top 250</b>	150	121	29	1	yes
<b>top 300</b>	177	143	34	1	yes
<b>top 350</b>	200	163	37	2	yes
<b>all (410)</b>	236	193	43	2	yes

It is productive in every condition, suggesting that children can assume accusative arguments are direct objects. But what could the dative on an argument suggest in this context? I start by testing the broadest possible option, [+dative → -subject, -direct object] in Table 35.

**Table 35.** Evaluation of the [+dative → -subject, -direct object] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	16 (18)	10 (12)	5 (6)	6 (8)	yes (no)
<b>top 100</b>	36	26	10	16	no
<b>top 150</b>	49	37	12	24	no
<b>top 200</b>	64	39	15	33	<i>no</i>
<b>top 250</b>	80	62	18	42	no
<b>top 300</b>	94	74	20	52	no
<b>top 350</b>	116	92	24	66	no
<b>all (410)</b>	131	105	26	73	no

This does not yield productivity in any condition except for the first tier and only if adjectival predicates with non-nominative subjects are excluded. Additionally, the rule is productive in the top 200 tier if alternating ACC./DAT. arguments are excluded as well as the adjectival predicates. In the randomized samples, the rule never reaches productivity. This is not surprising considering the fact that the most frequent type of dative argument actually is direct objects, as mentioned before and seen by the fact that the maximum exceptions reach majority by the top 200 tier. Despite this, the number of direct objects never reaches sufficiency so a [+dative → +direct object] rule does not reach productivity either.

To summarize, the results show that on the current approach, nominative could be used as a cue for subjecthood and accusative as a cue for (direct) objecthood, but the dative does not yield consistent information about argument type. But what about case as a cue for meaning regardless of syntactic role?

#### 4.3.3.2 Meaning

When looking again at Table 14 in subsection 4.3.2.2, we can see that most (but far from all) nominative arguments are agents, accusative arguments are quite evenly split

between themes and patients (with themes being slightly more frequent) and datives are majoritarilly themes and goals with a rather even split as well. Of course, these are only the broad patterns, but could they be viable cues when encountering novel verbs?

I start by testing the [+nominative → +agent] rule which only is productive if arguments which also can appear as non-agents with the same verb are excluded.

**Table 36.** Evaluation of the [+nominative → +agent] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	49	37	12	13 (17)	yes (no)
<b>top 100</b>	98	77	21	17 (30)	yes (no)
<b>top 150</b>	149	120	29	27 (43)	yes (no)
<b>top 200</b>	199	162	37	32 (50)	yes (no)
<b>top 250</b>	249	204	45	38 (58)	yes (no)
<b>top 300</b>	296	244	52	40 (62)	yes (no)
<b>top 350</b>	346	287	59	45 (67)	yes (no)
<b>all (410)</b>	405	338	67	58 (81)	yes (no)

The randomized samples yield the same results. There is no need to test the reverse rule, [-nominative → -agent] because, as has been mentioned at various occasions, all agents are nominative.

I therefore turn to a [+accusative → +patient] rule in Table 37. Bearing in mind that accusative arguments are almost as often patients and themes in the IceCASE corpus, it is not unexpected that no productivity is found except for the first tier, where an [+accusative → +theme] is productive (and no productivity at all emerges in the randomized samples).

**Table 37.** Evaluation of the [+accusative → +patient] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	30	22	8	24	theme
<b>top 100</b>	61	47	14	38	no
<b>top 150</b>	93	73	20	56	no
<b>top 200</b>	123	98	25	67	no
<b>top 250</b>	150	121	29	79	no
<b>top 300</b>	177	143	34	89	no
<b>top 350</b>	200	163	37	95	no
<b>all (410)</b>	237	194	43	114	no

To illustrate this, I show in Table 38 that a [+accusative → +patient/+theme] would be productive in all conditions (including randomized samples), but exactly the same holds true for [+dative → +goal/+theme] as is shown in Table 39.

**Table 38.** Evaluation of the [+accusative → +patient OR +theme] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	30	22	8	0	yes
<b>top 100</b>	61	47	14	1	yes
<b>top 150</b>	93	73	20	2	yes
<b>top 200</b>	123	98	25	2	yes
<b>top 250</b>	150	121	29	2	yes
<b>top 300</b>	177	143	34	2	yes
<b>top 350</b>	200	163	37	3	yes
<b>all (410)</b>	237	194	43	3	yes

**Table 39.** Evaluation of the [+dative → +goal OR +theme] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b><math>\theta_N</math></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	16	11	5	3	yes
<b>top 100</b>	36	26	910	5	yes
<b>top 150</b>	47	35	12	6	yes
<b>top 200</b>	64	49	15	8	yes
<b>top 250</b>	80	62	18	9	yes
<b>top 300</b>	94	74	20	14	yes
<b>top 350</b>	110	89	21	15	yes
<b>all (410)</b>	128	102	26	18	yes

On the other hand, a [+dative → +goal] rule is only productive in the top 50 tier in Table 40, with the randomized samples showing no such initial productivity. This is in line with the overview in 4.3.1, where the comparison between Figures 4.3 and 4.4. shows there is a proportional richness of dative goals in the top 50 verbs which is very much enhanced by reflexive arguments (Figure 4.9).

**Table 40.** Evaluation of the [+dative → +goal] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b><math>\theta_N</math></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	16	11	5	5	yes
<b>top 100</b>	36	26	10	19	no
<b>top 150</b>	47	35	12	25	no
<b>top 200</b>	64	49	15	39	no
<b>top 250</b>	80	62	18	48	no
<b>top 300</b>	94	74	20	59	no

<b>top 350</b>	110	89	21	70	no
<b>all (410)</b>	128	102	26	82	no

Judging by the results in the current section, case-meaning cues are relatively sparse in the absence of a more precise context. Nominative points to agents in limited conditions and accusative and dative both point to a theme or another role (patients for the accusative and goals for the dative). In addition, the dative indicates a goal in the top 50 tier which does not hold in the first tier of the randomized samples. In the following section, we narrow down the context for case-role rules based on syntactic role (form+ approach).

#### 4.3.3.3 Form+ and meaning

Finishing with case as a cue for thematic roles within syntactic roles, I start with the evaluation of the [+subject, +nominative → +agent] rule in Table 41. Unsurprisingly, the results are the same as for the [+nominative → +agent] rule; it is productive in all conditions only if polythematic arguments are not counted as exceptions.

**Table 41.** Evaluation of the [+subject, +nominative → +agent] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	48	36	12	12 (16)	yes (no)
<b>top 100</b>	97	76	21	16 (29)	yes (no)
<b>top 150</b>	146	117	29	24 (40)	yes (no)
<b>top 200</b>	195	159	36	28 (46)	yes (no)
<b>top 250</b>	244	200	44	33 (53)	yes (no)
<b>top 300</b>	292	241	51	36 (58)	yes (no)
<b>top 350</b>	341	283	58	40 (62)	yes (no)
<b>all (410)</b>	401	335	66	54 (77)	yes (no)

The lack of clear productivity for this rule in general comes from the fact that, unsurprisingly for a default, nominative subjects have the broadest semantic scope. Although they are mostly agents, they can also be themes and experiencers. On the other hand, the non-nominative subjects in the IceCASE corpus are all experiencers. This means that in the context of subjects, a rule like [+subject, -nominative → -agent] is automatically productive and can be finessed into [+subject, -nominative → +experiencer] rule in the IceCASE corpus. If a subject is non-nominative, the child learner can assume it is an experiencer.<sup>45</sup> This has no exceptions whatsoever and the productivity of the rule is therefore never challenged, a fact which sheds light on the finding (in Chapter 2) that the dative is particularly useful as a cue for verb meaning in the context of subjects. An interesting question is whether the non-nominative category is further divided into +dative and +accusative (which would both yield a productive experiencer rule of course), as all the accusative subjects are DS verbs which also appear in the dative.

Turning to the context of objects, I omit various tables to avoid repeating the numbers from the previous sections, which are almost identical. The [+object, +accusative → +patient] rule yields identical productivity patterns as the ones in Table 37 with initial [+accusative → +theme] in the first tier and nothing else. As in Tables 38 and 39, the [+object, +accusative → +patient/+theme] and [+object, +dative → +goal/+theme] rules are productive but uninformative.<sup>46</sup> If objects are further divided with ±direct object, the dative yields clearer cues while the accusative still points to either themes or patients. But what about the dative and themes (Table 42)?

---

<sup>45</sup> Recall that the current approach focuses on canonical word order based on evidence that children ‘filter’ out movement to some extent in the acquisition of argument structure (Perkins 2019).

<sup>46</sup> The [+object, +dative → +goal] yields the same results as [+dative → +goal] with productivity only in the top 50 tier.

**Table 42.** Evaluation of the [+direct object, +dative → +theme] rule in the IceCASE corpus.

	<b>N</b>	<b>sufficient</b>	<b>θ<sub>N</sub></b>	<b>e</b>	<b>productive</b>
<b>top 50</b>	4	2	2	2	yes
<b>top 100</b>	16	11	5	2	yes
<b>top 150</b>	25	18	7	7	yes
<b>top 200</b>	39	29	10	8	yes
<b>top 250</b>	50	38	12	11	yes
<b>top 300</b>	61	47	14	16	no
<b>top 350</b>	77	60	17	21	no
<b>all (410)</b>	86	67	19	22	no

A [+direct object, +dative → +theme] rule is productive initially, but not beyond the top 250. Finally, as could be deduced from e.g. Table 14 in subsection 4.3.2.1, a [-direct object, +dative → +goal] rule is productive in all conditions.

To recapitulate, the form+ approach in the current subsection yields a scenario in which nominative subjects are indicative of agents in some testing conditions and non-nominative subjects consistently indicate experiencers. On the other hand, accusative and dative cases on objects do not give any clear cues unless objects are divided further with [±direct object], which yields a clear indication for goals in the case of non-direct dative objects and partial indications of direct dative objects being themes. This means that the case marking yields the most salient cues for experiencers, followed by agents, goals and themes. No case cue emerges in the case of patients beyond a [+patient OR +theme] rule. The dative cue for themes is not only worse because it hinges on a [±direct object] division but also because the accusative also in part indicates a theme in the same conditions, arguably blurring the signal further. Interestingly, the cue hierarchy which emerges in these results is somewhat in line with the experimental results from the morphosyntactic bootstrapping experiments discussed in Chapter 2.

## 4.4 Discussion

I now summarize and try to simplify the general findings of the chapter, answering the main questions and discussing the results in the context of derivability. An overview of the IceCASE corpus shows that the NOM.-ACC. frame is the most frequent frame, followed by intransitives with nominative subjects, the NOM.-DAT. frame, alternating NOM.-ACC./DAT. predicates, NOM.-DAT-ACC. ditransitives and finally less frequent frames, with non-nominative subjects for example. Looking at the distribution of the verbs incrementally, a few important facts emerge. First, the NOM.-DAT. frame is relatively infrequent at first (in the top 100 verbs) compared to the sample as a whole, with ditransitives and therefore dative goals initially making up a higher proportion of dative verbs. This early prominence of dative goals is greatly enhanced when optional reflexive arguments are included in the overview, as their distribution is concentrated in the top 100 most frequent verbs. This might explain the results in Chapters 2 and 3, which suggest that the more cross-linguistically expected association between the dative and applied arguments emerges earlier than the association with datives and themes. When a closer look is taken at the arguments themselves, it also becomes clear that the association between datives and goals is supported by other regularities in form: indirect objects and optional reflexive arguments are almost exclusively dative and a TP analysis of a [-subject, -direct object → +dative] rule, as well as a [+goal → +dative] rule, yields productivity at all levels of testing, despite the presence of exceptions. In the comprehension form-meaning direction, a [+dative → +goal] rule reaches productivity with the top 50 most frequent verbs but exceptions become too numerous after that point.

I have already mentioned that most indirect objects and optional reflexives in the IceCASE corpus are dative, as well as goals in general. Other expected patterns emerged: Most subjects, and all agents, are nominative, yielding productivity at all levels for [+subject → +nominative] and [+agent → +nominative]. In the word learning/comprehension direction, [+nominative → +subject] is productive at all levels but [+nominative → +agent] only is productive if polythematic arguments are excluded. Most direct objects and patients were accusative, but the majority for these categories is less pronounced. As a consequence, [+direct object → +accusative] is only productive for the top 100 most frequent verbs, while a [+patient → +accusative] rule is productive at all

levels. In the comprehension direction, [+accusative → direct object] is productive at every level but [+accusative → +patient] never is, since the majority of accusative arguments in the IceCASE corpus turn out to be themes (with [+accusative → +theme] reaching productivity in the top 50 verbs). Themes are in fact the most disparate category, with no case reaching productivity for them in the directionality assumed for production. All theme subjects are nominative, and adding features such as [±motion] or [±animacy] did not yield a system where all items are accounted for. This disparity appears in the comprehension direction as well, where the possible [+accusative → +patient OR +theme] and [+dative → +goal OR +theme] are both productive, potentially creating conditions where the accusative and dative compete in the overlapping context of [+theme]. As a consequence of these results, and considering work suggesting the precedence of form over meaning (Naigles 2002 and the comparison between Schuler et al. 2016 and Emond & Shi 2016) as well as the possible recessive nature of case in Icelandic (Kiparsky 1997, Schätzle 2018 and Galbraith 2018),<sup>47</sup> the paths assumed in case marking rule formation are based on syntactic and thematic roles, form and meaning. Table 43 contains a summary of the final rules assumed in the tedious discovery for productivity conducted in section 4.3.

---

<sup>47</sup> With mismatches between morphosyntactic case and syntactic position being allowed in e.g. the form of non-nominative subjects.

**Table 43.** Final rules assumed in the incremental corpus analyses with the TP. Rule productivity in yellow, lack of productivity in blue.

<b>Production rules</b>	N	50	100	150	200	250	300	350	all
+subject → NOM.	410	Yellow							
+direct object → ACC.	302	Yellow	Yellow	Blue	Blue	Blue	Blue	Blue	Blue
+subject, -direct object → DAT.	20	Yellow							
+direct object, +patient → ACC.	128	Yellow							
+direct object, -patient → ACC.	197	Yellow	Blue						
+direct object, +theme, -motion → ACC.	91	Yellow							
+direct object, +theme, +motion → DAT	90	Blue	Yellow	Blue	Blue	Blue	Blue	Blue	Blue
+direct object, +goal → DAT.	4	Yellow							
+oblique, +experiencer → DAT.	11	Yellow							

<b>Comprehension rules</b>	N	50	100	150	200	250	300	350	all
+subject, +NOM. → agent	401	Yellow							
+subject, -NOM. → experiencer	11	Yellow							
+direct object, +ACC. → patient	230	Blue							
+direct object, +DAT. → theme	88	Yellow	Yellow	Yellow	Yellow	Yellow	Blue	Blue	Blue
-subject, -direct object, +DAT. → goal	18	Yellow							

For production, where the end result is case, this yields a system where syntactic role is enough to assign case productively for subjects (nominative), indirect objects (dative) and reflexives (dative). For direct objects, which initially have a productive accusative rule, adding a [+patient] feature permits assigning accusative consistently. When it comes to the remaining direct object themes, adding [-motion] furthermore yields a consistently productive accusative rule. This means that the system, through nested productivity, has accounted for all arguments except for direct object themes which change location ([+motion]). For this category, no productive rule (neither accusative or dative) is found, pointing to either a scenario without productivity where both accusative and dative are possible, or the necessity to discover more precise subrules (as in Maling

2002 and Jónsson 2013). In any case, this lack of productivity is in line with children's difficulty to use case as a cue for motion (see Chapter 2), and the complexity of reaching productive rules for direct objects is furthermore in line with the broad semantic scope of dative productivity in Chapter 3. In this context, it's important to reiterate that the developmental paths derived here are assumed to vary between individuals which of course have different compositions of vocabularies but could also vary in the way they divide their vocabulary in rule formation. As Rissman and Majid (2019) review and discuss for example, thematic roles may vary in universality, innateness and status as pre-linguistic core knowledge. This suggests that there is also a developmental path with room for individual variation in the construction of some of the features/concepts encoded in language (see Carey 2009).

Before describing the final rule formation assumed for comprehension, I want to briefly discuss the status of non-nominative subjects in the context of a productive [+subject → +nominative] rule. This means that non-nominative subjects are listed as exceptions, but most rule-based accounts assume that exceptions can be structured into productive rules (as in the possible transient irregularization which appears in the *bring-brang* examples some learners of English show, see Yang 2016, p. 24). In the case of non-nominative subjects, I argue that a productive [+dative] rule can be found within the set of non-nominative subjects. This is the driving force of Dative Substitution, the topic of Chapter 5. Such a productive rule, emerging within exceptions, could benefit from being already salient as a cue in comprehension, which is the case for non-nominative subjects.

Indeed, the final rule formation path assumed for comprehension entails that case is a cue for meaning within syntactic roles. This yields a system where nominative subjects are (under some conditions) agents and non-nominative subjects indicate an experiencer (without exceptions), but objects have to be partitioned further with [ $\pm$ direct object]. This yields a clear indication for goals in the case of non-direct dative objects and partial indications of direct dative objects being themes, but direct accusative objects do not indicate a thematic role conclusively. Again, this complements the experimental data from Chapter 2 rather well, with a deeper explanation for the particular salience of non-nominative subjects as experiencers, the association between datives and goals and the later emergence of a distinction between patients and themes based on case. Broadly,

this analysis indicates that form-meaning mappings in the context of morphosyntactic bootstrapping can be derived from distributional information in the input, a finding which pairs well with the findings of Pearl & Sprouse (2021), where linking rules (rUTAH) are derivable with the TP.

## 5. Acquiring and shaping case marking variation

Non-nominative experiencer subjects are the topic of this final chapter, where I finally address Dative Substitution (DS) in detail, the best known instance of case variation in Icelandic. As has been discussed (see subsections 1.3 and 3.3.1.1), DS is a stigmatized but widespread instance of grammatically conditioned morphosyntactic variation which consists in dative being extended to historically accusative experiencer subjects, as well as to two verbs with historically nominative experiencer subjects. For children currently acquiring Icelandic, it is fair to assume that all DS verbs appear with variable case marking in their language environment.<sup>48</sup> Intra-speaker variation within items is widespread (Svavarsdóttir 2013, Thráinsson 2013, Nowenstein 2017) and more frequent verbs seem to show less DS (Jónsson & Eythórsson 2003). Additionally, this intra-speaker variation is conditioned by various factors, both sociostylistic and grammatical, and can be put in the context of Tamminga et al.'s (2016) recent framework on the dynamics of variation in individuals, which recognizes three types of factors conditioning variation: sociostylistic (s-), internal linguistic (i-) and psychophysiological (p-). P-conditioning can only entail variable alternations as it is extragrammatical, universal and automatic.<sup>49</sup> It applies to effects which are constant or at least predictably-distributed across the population. I-conditioning, on the other hand, can entail both variable and categorical alternations, as it is grammar internal and language- or variety- specific, arbitrary and therefore learned. In this chapter I focus on i-conditioning with person and number effects (the Person-Specific Retention, PSR), which has already been documented quite extensively in DS (e.g. Nowenstein & Ingason 2021), as well as on syncretism effects (see more in subsection 1.3).

---

<sup>48</sup> With *hlakka* and *kvíða* (which historically take nominative subjects) appearing with accusative and dative subjects.

<sup>49</sup> Tamminga et al. (2016) describe p-conditioning, supposedly the least transparent of these three categories, in the following way: "What we call p-conditioning factors arise from cognitive and physiological systems that are shared by all humans, like working memory capacity, articulatory pressures arising from the physiology of the speech apparatus, resting activation levels for words (of the type that are implicated in priming), and so on." (p. 303).

This type of distinction is important when making predictions on the directionality of language change based on the type of conditioning factor. In other words, if the possible variants are specializing in one way or another through conditioning (subrules under the TP), the nature of the dimension this specialization occurs on appears to be relevant in determining whether replacement, complete specialization or stable variation will follow. Fruehwald and Wallenberg (2013) provide evidence for this through a few diachronic case studies which form the basis for their Minimalist Theory of Variation. This approach assumes that if categorical variants specialize along a categorical dimension, complete specialization should eventually result. On the other hand, if categorical variants specialize along a continuous or ordinal dimension, then complete specialization can never result. Instead, stable variation/optionality is established. Furthermore, Wallenberg (2019) notes that specialization can entail a coordination problem in the speech community, slowing the speed of eventual specialization as speakers need to converge on a dimension of specialization as well as converging on which variant specializes in which direction along the dimension of specialization. The coordination problem is amplified by each potential dimension of specialization, adding to the possible diachronic outcomes. DS seems to be an ideal opportunity to observe complex specialization mechanisms in real time, individually and across the speech community. But how do children acquire variation and how does it affect language change?<sup>50</sup>

## 5.1 Inconsistent input

In fact, one of the fundamental questions within developmental sociolinguistics, and language acquisition research more broadly, has to do with children's reaction to variability in their input, or primary linguistic data (e.g. Labov 1989, Yang 2002, Hudson Kam & Newport 2005, Smith et al. 2009, Cournane & Pérez-Leroux 2020). Repetti-Ludlow and MacKenzie (2022) describe that at a glance, the literature points to a paradox: Children are both expected to diverge from and match their caregivers' speech. As has been extensively documented, children overgeneralize and regularize both consistent (Marcus et al. 1992) and inconsistent (Hudson Kam & Newport 2005) input. These

---

<sup>50</sup> Part of the findings in this chapter were reported in NWAV proceedings: Nowenstein et al. (2021).

generalizations can be interpreted as part of productive rule formation regardless of language variation and change (Schuler et al. 2016) but have also been described as a source for innovation and incrementation in the language community (e.g. Lightfoot 1979, Labov 2001, Cournane 2019, Cournane & Pérez-Leroux 2020 and Hall & Maddeaux 2020). Despite this tendency to go beyond the input, we do expect children to learn their caregivers' dialect, and they have in fact been known to match the rates of variation found in their environment (Labov 1989, Johnson & White 2020). When a closer look is taken at the growing body of literature targeting children's acquisition of variation, it becomes clear that the paradox represents reality. Both regularization and matching occur, but under different circumstances. Indeed, the nature of the developmental path can depend on a number of factors such as the learner's age and the amount and consistency of exposure to different dialects, but also the variable type (language domain), the complexity (or existence) of the conditioning factors and the social salience of the variable (e.g. Smith et al. 2009, Hendricks et al. 2018).

As has already been mentioned, previous work shows both matching and regularization when variation is present in the learners' input, often within the same study. The pioneering work of Smith and colleagues (Smith et al. 2007 and 2009) shows that correlations between variant rate in child-caregiver dyads may depend on the age of the child, the variable type, conditioning complexity and social context (i.e. style-shifting) as well as social salience. Research in cross-linguistic and multilingual contexts shows that amount and consistency of data might matter as well, which might also be related to conditioning complexity (Hendricks et al. 2018). Lessons can furthermore be drawn from the artificial language experiment literature, where it has been shown that adults probability match inconsistent input while children regularize it, but conditioning changes this picture with (at least older) children acquiring contrasts but still not matching the input (Hudson Kam & Newport 2009, Hudson Kam 2015). This is the third option I want to emphasize in the current chapter, where children neither regularize nor match the rate exactly. Instead, there is a replication of the contrasts present in the input.

This contrast replication is present in previous work with child-caregiver dyads and in studies comparing patterns and rates in child and adult language corpora. For example, Smith et al. (2009) found a significant relationship between children and their caregivers

for the lexically conditioned *house-hoose*<sup>51</sup> variable in Scots (11 children aged 2;10–3;6), with the parents' rate predicting the children's rate, but not for third-person-plural -s and *t/d*-deletion. Crucially, group contrasts were still replicated for those variables. A similar pattern can be found in Hall & Maddeaux's (2020) work on /u/-fronting and /æ/-raising in 19 Toronto families (children aged 4–12) where children replicate the broad contrasts in their parents' input (and do so more closely as they get older) while still advancing change. Finally, the work of Repetti-Ludlow and MacKenzie (2022), comparing stem-final fricative plurals in child and adult corpora, shows that children acquire the hierarchy of phonemes found in adult production but simultaneously diverge from adults in their rate of irregular voicing of /f/-final stems. To summarize, a growing body of work points toward what could be called regularization within matched conditions. In the next section, I describe the methods used to investigate whether this is the case for children acquiring Dative Substitution.

But how does variation and change fit into models of language acquisition in general? This has been addressed from very different perspectives (e.g. Yang 2002 and 2016, Christiansen & Chater 2016 and Kapatsinski 2018), but I have in previous work attempted to account for DS in Icelandic by combining the TP with Yang's (2002) Variational Model of Language Acquisition, henceforth VM (Nowenstein 2017, Nowenstein & Ingason 2021, see also Woods et al. 2021 for auxiliary-initial declaratives). Yang's (2002) VM assumes weighted grammatical rules that can apply with a probability <100% and a mechanism of Elsewhere Condition Serial Search (ECSS), where more specific rules are applied first. Additionally, as was reviewed in section 1.3, Yang's (2016) *Price of Productivity* book contains an analysis of DS in synchrony which also addresses the rise and distribution of the change diachronically, showing how the dative is productive in the context of non-nominative experiencer subjects but not non-nominative theme subjects (which are known to appear with Nominative Substitution instead). As non-nominative theme subjects do not appear in the IceCASE corpus, the current focus is on DS with experiencers and its grammatical conditioning instead.

---

<sup>51</sup> Alternation between a diphthong (*house*) and the monophthong (*hoose*).

### 5.1.1 Main questions and results

The main questions/objectives of the chapter can be found below:

- (38) a. How do children acquire grammatically conditioned case marking variation in function of caregiver language within and across conditions and how can that be accounted for within models of language acquisition?
- b. Can judgments and forced choice be used as a proxy for usage in developmental sociolinguistic studies targeting low-frequency variables?

Additionally, the main empirical objective of this chapter is to document variation in Icelandic child-caretaker dyads for the first time and confirm the presence of grammatical conditioning factors for children and adults showing DS.

The results show that children neither regularize nor match their caregivers' usage. Instead, they replicate the systematic contrasts they encounter. Indeed, by investigating DS in 101 children aged 3–13 and their caregivers (using forced-choice tasks and grammaticality judgments across multiple items as a proxy for case use), I find that the caregivers' general variant rate does not predict the rate at which their children select DS, regardless of age. On the other hand, when analyzing the data within conditioning factors, I find that children replicate the contrasts present in their caregivers' speech, both at the group and individual level, and that this is in part dependent on age. This is to some extent in line with previous studies (Smith et al. 2009, Hendricks et al. 2018, Cournane 2019, Hall & Maddeaux 2020, Repetti-Ludlow & MacKenzie 2022), and has implications for studies on the role of specialization (Wallenberg 2019) and the dynamics of variation in individuals (Tamminga et al. 2016). Additionally, I argue that a combination of the TP and Yang's (2002) VM can account for the patterns found in the Icelandic data, and that a case should be made for stochastic exceptions in addition to stochastic rules.

## 5.2 Methods

As has been mentioned, most of the data used in the present project were collected within the MoLiCoDiLaCo-project at the University of Iceland (PIs: Sigríður Sigurjónsdóttir and

Eiríkur Rögnvaldsson), where the aim of the project was not to investigate the main questions of the present chapter but to analyze the presence of English in the Icelandic language community and evaluate its possible effects on the development of Icelandic in a broad sense (Sigurjónsdóttir & Nowenstein 2021, Nowenstein & Sigurjónsdóttir 2021). This contributes to the study in the current chapter being possibly underpowered within conditions, despite a high number of participants, i.e. 101 children aged 3–13 which participated in the case production forced-choice task described in Chapter 3. 95 caregivers additionally participated in a task containing grammaticality judgments as well as forced-choice items (with a higher number of total items). Using these data, I was able to connect children and caregivers into 84 valid dyads (with the child and caregivers having at least answered two items each so a rate could be computed).

As a reminder, in the children’s forced choice task (described in section 3.2.1), stimuli were presented orally and in written form (only orally for children below 6 years old). Seven DS items were presented, two of them testing the Person-Specific Retention (PSR, described in section 1.3) with first and third person non-syncretic pronouns with the same verb (*langa* ‘want’, for which only one item was presented in subsection 3.3.1.1), and two of them testing syncretism effects (also described in section 1.3) with third person plural full NPs (also one verb). The three other items were in the third person singular as can be seen in Appendix C. Caregivers were presented with 13 DS items in written form. Within these items, four were used to test the PSR (two verbs, *langa* ‘want’ and *klæja* ‘itch’) and four to test syncretism effects (two verbs also). Both surveys were counterbalanced and the options in the forced-choice items randomized. As mentioned before, we used forced-choice production data and grammaticality judgments as a proxy for usage, computing the variant rate of each individual based on their answers. As the judgments were made on a 5-point Likert-scale (example of a test sentence in (39)), they were first transformed into a binomial accept/reject variable.

- (39) **Strákana**                      **langar**                      í                      aðra                      ferð.  
 boys.the.ACC                      wants                      in                      another                      ride.  
 ‘The boys want to go for another ride’

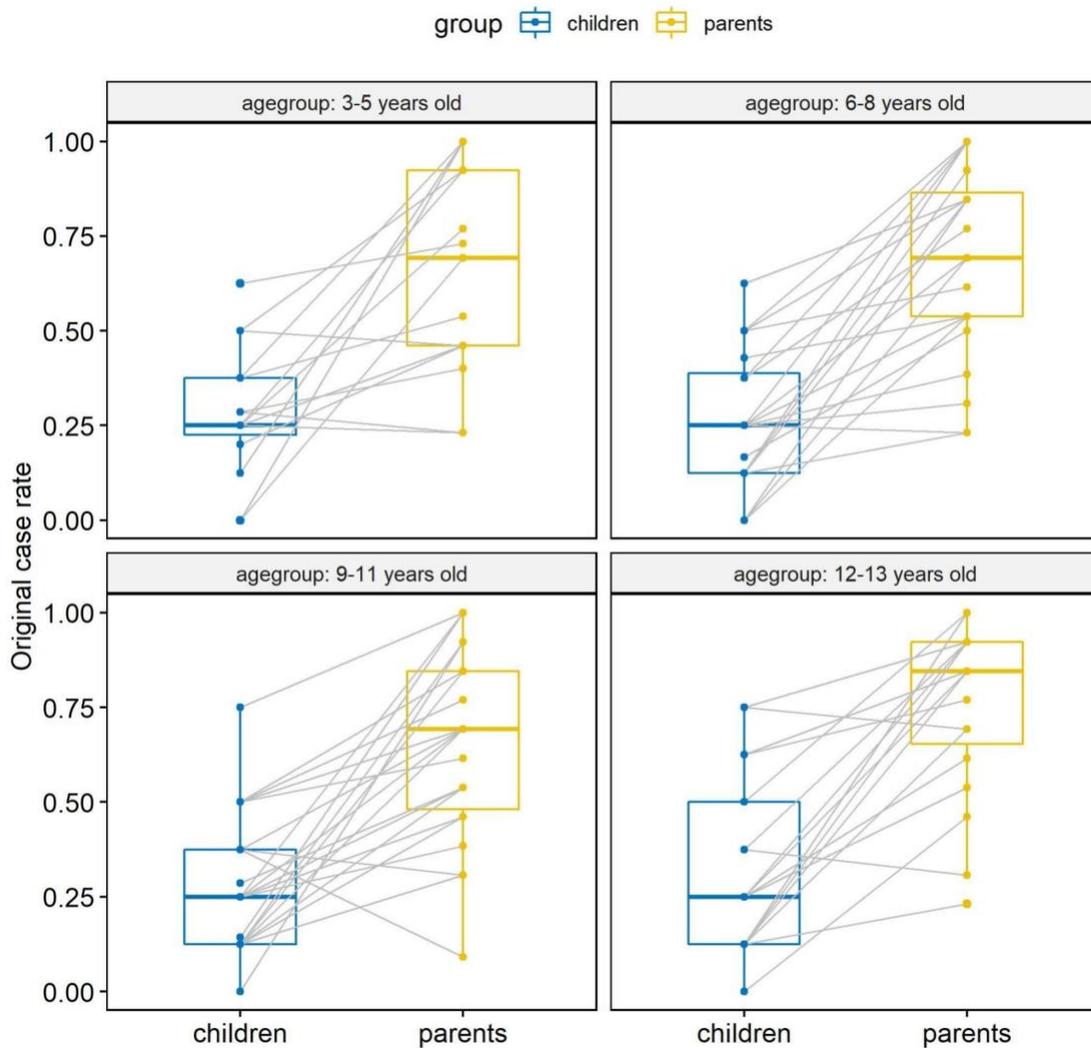
Although this methodology is not commonly found in developmental sociolinguistics, DS is a low-frequency variable and therefore not the ideal candidate for corpus-based research either, particularly when the conditioning factors discussed in section 1.3 are also under investigation. Additionally, I believe there is methodological value in exploring experimental data in the context of developmental sociolinguistics, and future work on DS with corpora and increased statistical power in experiments will hopefully confirm the validity of such methods.

## 5.3 Results

Before going over the results within conditioning factors at the group and individual level, I begin with a description of the relationship between the rate at which the original case was selected in the answers of the children and their caregivers.

### 5.3.1 Parental variant rate

I start with the rate of historical/original case (nominative or accusative) – as opposed to the innovative dative – and the correspondence between the children and their caregivers. In Figure 5.1, we can for example see whether the caregivers with the highest rate of original case use are associated with the children who select the original case the most. However, this is not the case. Instead, it is visually clear, from the directionality of the lines linking children and their caregivers together, that there is not a strong relationship between children and caregivers' variant rates. Modeling results are in line with the lack of a relationship shown in the figure. A linear regression analysis with an age-caregiver rate interaction predicting the children's rate shows no significant effects ( $F(3, 80) = 1.005, p = 0.3951$ ). The children do not match their caregivers' usage rate as measured by my proxy, which was computed across conditions, and this relationship is not affected by age as can also be visualized in Figure 5.1.



**Figure 5.1.** Rate of original case retention in child-caregiver dyads, by age group.

Interestingly, the rate of original case does not increase with the children's age. As can be seen, the pattern between children and adults is stable across age groups, with no correspondence in the dyads, but the caregivers always using the original case more than their children, as would be expected with a change in progress. Here it is important to stress that the computed rate is the original case use (nominative/accusative), not the rate of overgeneralizations which can be either dative, as in DS, or nominative, as is mainly present in language acquisition. If we compute the rate of nominative use only, as was done in subsection 3.3.1.1, an age trend does appear. This still does not change the fact that the present results do not show the trend observed in the literature in which

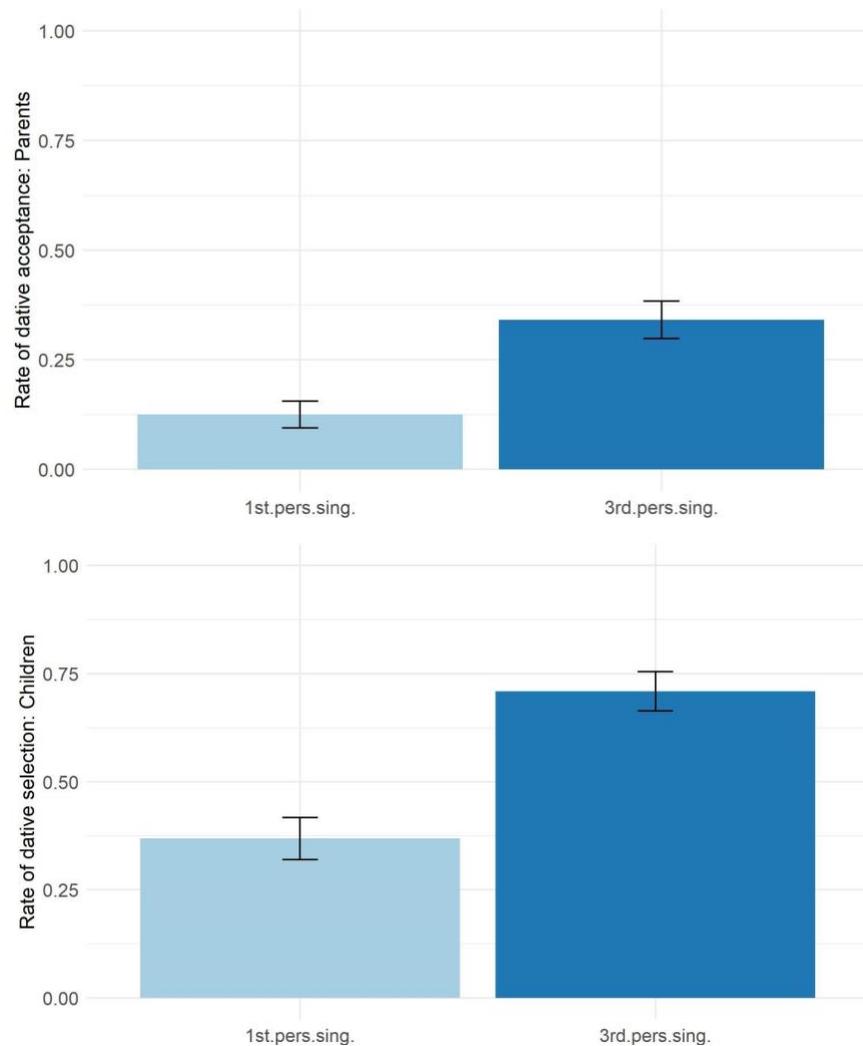
younger children are categorical but older children acquire variation. Instead, the acquisition of variation is also present in our younger participants. Seven children show categorical results, with no original case, but they are spread across age groups. If we look at the categorical results of caregivers on the other hand, 15 of them show the original case throughout, but the answers of the children associated with them do not differ from the rest of the group, neither in age or original case use.

To summarize, when looking at the rate of original case use with DS verbs across conditions in child-caregiver dyads, there is no evidence of matching even though the majority of the children acquire variation instead of regularizing their input. I additionally find no age pattern in the rate of original case use, which came as a surprise considering previous research on the acquisition of variation. One reason for this might be that the overall usage rate of a grammatically conditioned variable is not the right place to look in search for correspondence between children and caregivers or age patterns. Instead, it might be necessary to look within the conditioning factors.

## 5.3.2 Conditioning factors

### 5.3.2.1 Person-Specific Retention

Starting with the PSR, Figure 5.2 shows a similar pattern emerging in both the caregivers and the children's results.

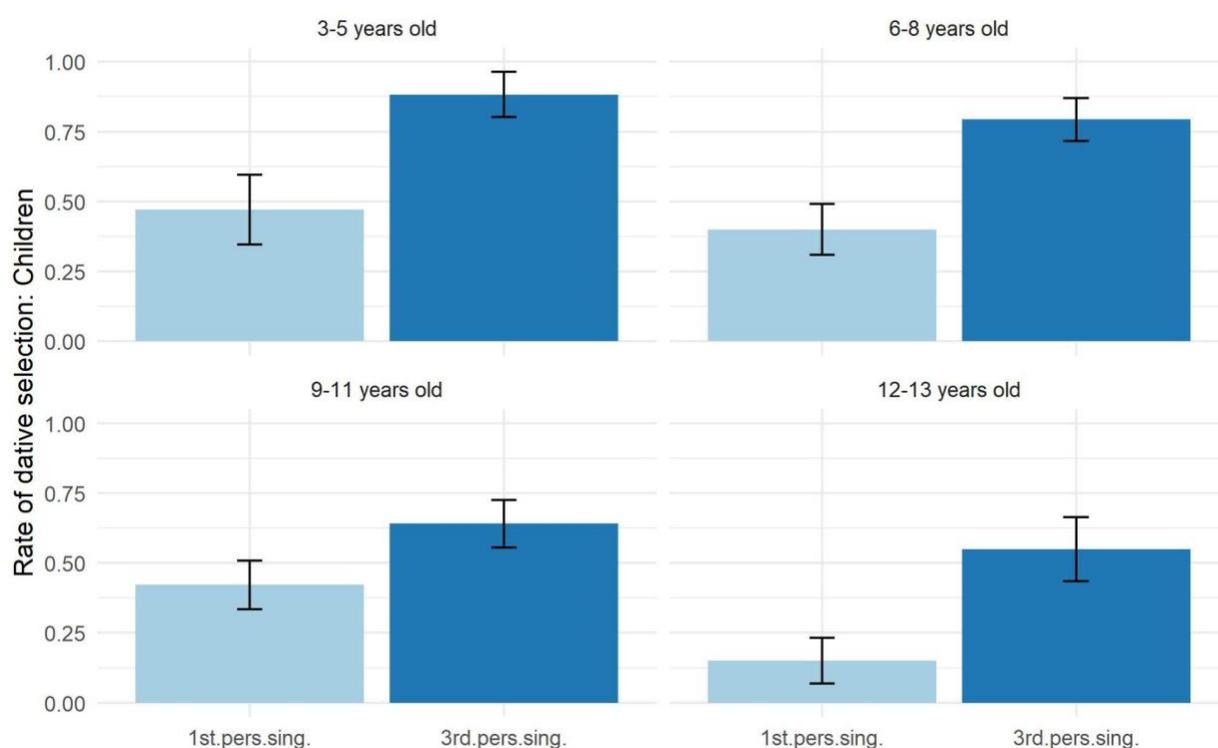


**Figure 5.2.** Dative selection/acceptance in the first and third person singular, by group (children and caregivers).

As expected, the rate of Dative Substitution is higher with a third person singular pronoun as compared to the first person. The conditioning effect is slightly bigger for the caregivers and the rate of DS overall lower, as would be expected when documenting ongoing

linguistic change. Here it still is important to keep in mind that the methods differed between groups and that the caregivers had more items per condition, resulting in more opportunities for contrasts.

To explore the results further, I conducted a nested comparison (Likelihood Ratio Test) using mixed effects logistic regression models (lme4 in R, Bates et al. 2015).<sup>52</sup> Random intercepts for participants and items were included in every model, as well as age and an age-pronoun type interaction for the children. As can be seen in Figure 5.3, the PSR is clearly present from a young age.



**Figure 5.3.** Dative selection in the first and third person singular, by age group (children).

The comparison shows a significant improvement to the fit of the model when adding pronoun type as a variable, both for caregivers ( $\chi^2(1) = 37.6, p < 0.001$ ) and children ( $\chi^2$

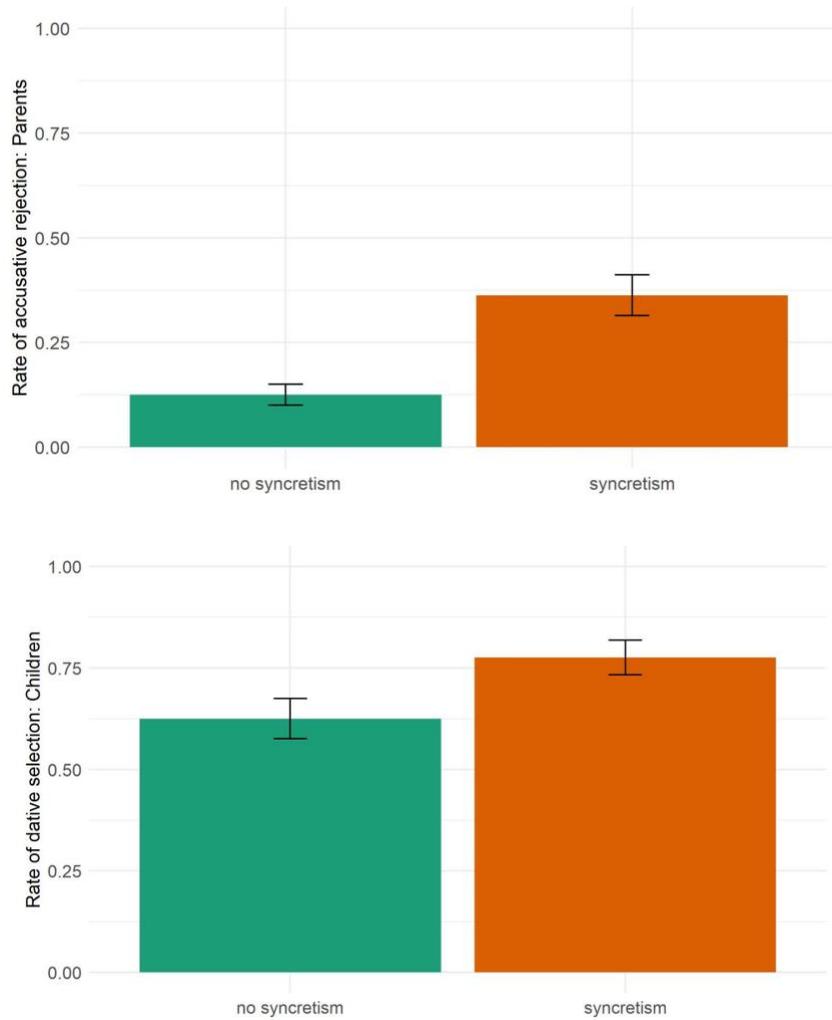
<sup>52</sup> Full modeling in Appendix D as before.

(1) = 23.9,  $p < 0.001$ ). Additionally, age improves the children's model significantly ( $\chi^2 (1) = 7.1$ ,  $p < 0.01$ ) but the age-pronoun type interaction does not ( $p = 0.74$ ).

The difference between the youngest and oldest children in the experiment is therefore found in the overall rate of DS within PSR condition, not the presence of the conditioning itself: Younger children have more DS while older children match the patterns found in the caregiver group more closely. Finally, I looked at the PSR correspondence within dyads, where 48/99 children (48.4%) showed intra-speaker variation within the two items used to test the condition and 85% of them had the expected PSR. Out of those 48 children, 19 had caregivers who also displayed intra-speaker variation (four items) and 13 out of the 19 dyads (68.5%) had matching PSR patterns. For the six remaining dyads, the PSR was present in the parents' answers but not the child's in four cases. Even though the overall trend therefore shows matching child-caregiver PSR patterns, this needs to be tested carefully in a study with more statistical power.

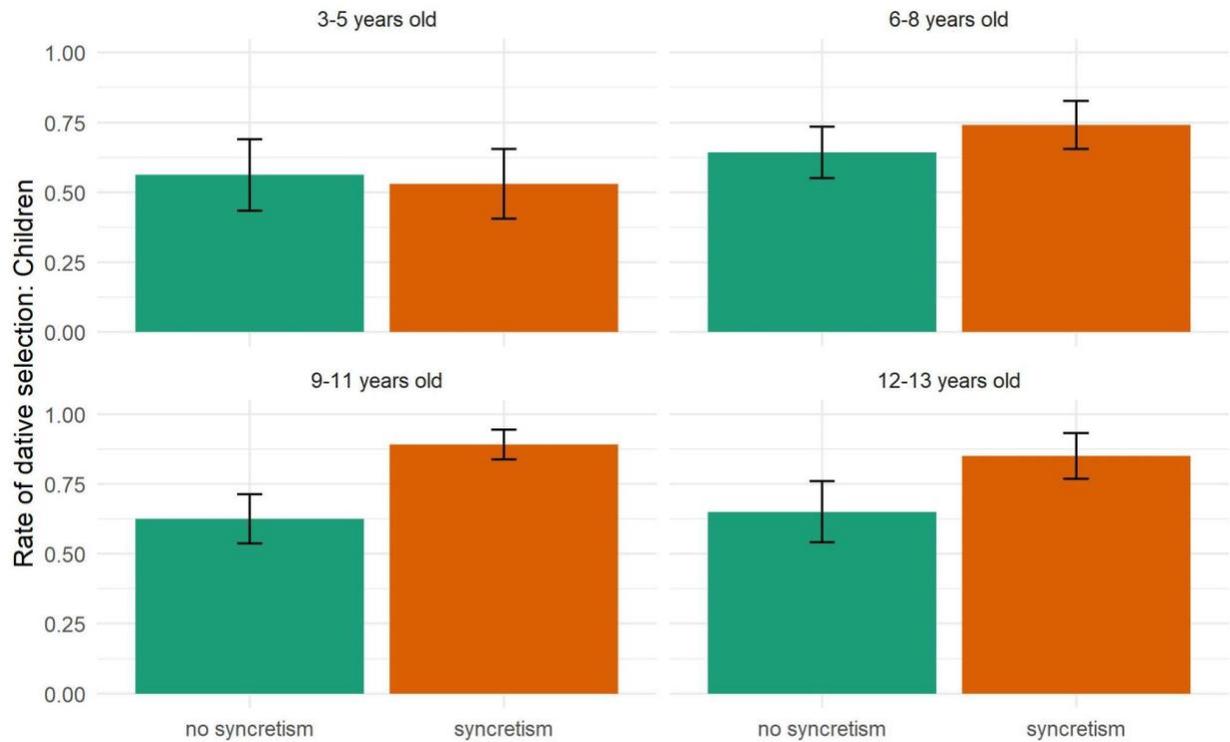
#### 5.3.2.2 Syncretism effects

I now move on to the syncretism effects results shown in Figures 5.4 and 5.5. As is the case for the PSR, Figure 5.4 shows that the syncretism effect is present in both groups (children and caregivers). The dative substitution rate, or accusative rejection for the caregivers, rises when the subjects show nominative-accusative syncretism, and the effect is larger in the adult group. I also conducted a nested comparison of mixed effects logistic models for the syncretism effects. For the adults, a significant improvement to the fit of the model is observed when adding syncretism ( $\chi^2 (1) = 37.4$ ,  $p < 0.001$ ) and this is present in the children's data as well ( $\chi^2 (1) = 6.1$ ,  $p < 0.05$ ). On the other hand, neither age ( $p = 0.07$ ) nor a syncretism-age interaction ( $p = 0.07$ ) were significant.



**Figure 5.4.** Dative selection/accusative rejection with/without syncretism on the subject, by group (children and caregivers).

The visualization in Figure 5.5 indicates that the lack of a significant age-syncretism interaction might be due to a power issue, as the syncretism effects seem to emerge late in acquisition.



**Figure 5.5.** Dative selection with/without syncretism on the subject, by age group (children).

At the individual level, 33/99 children (33.3%) show intra-speaker variation within the two items used to test syncretism effects, 73% showing the expected pattern (DS with syncretic subjects). Out of those 33 participants, 16 had caregivers also displaying intra-speaker variation. 11/16 dyads (69%) had matching patterns and 9 of them with the expected syncretism effects. In 4/5 remaining cases, the caregiver displayed the effect but the child did not.

## 5.4 Discussion

The main contributions of this chapter can be put into three different categories. Theoretically, I aimed to investigate how children acquire grammatically conditioned case marking variation as a function of caregiver language within and across conditions and how this might be dependent on age. The results show that when acquiring the grammatically conditioned case marking variation of Icelandic Dative Substitution,

children neither regularize nor match their parents' variant rate. Instead, systematic contrast replication at the group and individual level is observed, or regularization within matched conditions. This is in line with some of the findings from previous work (e.g. Smith et al. 2007 and 2009, Hall & Maddeaux 2020 and Repetti-Ludlow & MacKenzie 2022), but it is worth mentioning that most previous studies within developmental sociolinguistics target phonetic or phonological variables, which might be hypothesized to prompt more matching due to increased social salience as compared to more structural and abstract variables (Sankoff 2019). Future work should address how different domains of language might entail different scenarios of caregiver-child convergence or divergence.

Empirically, this is the first attempt at documenting a variable in caregiver-child dyads in Icelandic. Although children do not match the parental variant rate, they do acquire the grammatical conditioning present in their caregivers' language, including the syncretism effects which have previously not been investigated in Icelandic. DS therefore shows robust i-conditioning along a categorical dimension with s-conditioning also present and is supposedly acquired in specific contexts (Tamminga et al. 2016). This might point towards stabilizing variation within specialized variants or subrules (Wallenberg 2019), although we also observe results which are in line with the "two steps forward, one step back" approach (Hall & Maddeaux 2020).

Methodologically, I wanted to evaluate the use of grammaticality judgments and forced-choice tasks as a proxy for usage in developmental sociolinguistic studies targeting low-frequency variables. The methods in the study contrast with the more frequently applied corpus analysis used when investigating variation in child-caregiver dyads (but see Hall & Maddeaux 2020 and Hall 2020 for elicitation tasks), but the results' correspondence with previous work suggests that judgments and forced choice tasks are viable methodological options, as long as the limitations of the present work are addressed, in particular the statistical power within conditions.

To conclude the chapter, I address the second theoretical objective, which was to put the results in the context of language acquisition models. Such models need to be able to account for the non-default productivity which drives the grammatically conditioned intra-speaker variation which DS consists in, and I consider a combination of Yang's VM (2002) and the TP (2016) to be a promising approach. This combination is

something I have attempted to implement in previous work on DS (Nowenstein 2017 and Nowenstein & Ingason 2021) but a new account is presented here in light of the evidence from previous chapters, with the novel addition of stochastic exceptions under the TP.

Following the results in Chapter 4, I assume that the drive behind DS is an [+oblique, +subject → +dative] rule which emerges within the exceptions to the [+subject → +nominative] rule. The [+oblique] (non-nominative) feature can be justified with various different arguments, e.g. the commonly observed syncretism between accusative and dative and the fact that all subjects of DS verbs appear both in the accusative and dative.<sup>53</sup> Based on the strong mapping between non-nominative subjects and experiencers, we could imagine that this rule would include the thematic role as a feature as well. Here it is also important to bear in mind that, as is outlined in Chapter 1, the rich productivity of the dative within oblique experiencer subjects in Icelandic is very rarely extended to nominative experiencers or new experiencer verbs. So within the [+oblique, +subject, +experiencer → +dative] rule, accusative experiencer subjects are the exception, not nominative ones, *unless* they have also been observed in non-nominative forms. But how does this rule account for DS as variation and not only transient productivity in acquisition? This could be schematized in the following way:

- (40) **IF** [+oblique, +subject, +experiencer] **THEN** [+dative]<sub>productive</sub>  
**e** = [*langa, vanta...*]<sub>ACC</sub>  
**ELSE** [+subject, +experiencer] **THEN** [+nominative]<sub>default</sub>  
**e** = [*finnast, leiðast...*]<sub>DAT</sub> and [*langa, vanta...*]<sub>ACC</sub>

Once a frequent predicate like *langa* ‘want’ is listed as an accusative exception, it should not consistently appear with the dative, as we know it does in adult language with DS. The dative exceptions to the default rule are covered by the narrower productive rule.

In this situation, we would expect children acquiring Icelandic to 1) overgeneralize the nominative (which they do) and 2) occasionally overgeneralize the dative on subjects which they have encountered in non-nominative forms (which they also do), but they

---

<sup>53</sup> As well as the nominative for *hlakka* ‘look forward to’ and *kvíða* ‘be anxious about’ subjects.

would eventually acquire the list of exceptions since *langa* ‘want’, *vanta* ‘need’ and other comparable verbs would always appear with subjects in the accusative. To put this in other words, the predicates’ status as exceptions reaches the probability of 1. But as is clear from the results in the current chapter, there is variation in the adult language, so children do not consistently get this information. Instead, the exceptions also appear in the dative in the children’s language environment, providing contradictory evidence – accusative indicates an exception and dative indicates that the predicate follows the rule. In situations like these (but maybe more generally), there might be a need to assume that learning whether an item is an exception or not is a variational learning process. Sometimes the child hears *langa* ‘want’ with an accusative subject (exception), sometimes with a dative (productive rule). With such variation, the status of the item as an exception does not reach 1. This could be schematized in the following way:

- (41) **IF** [+oblique, +subject, +experiencer] **THEN** [+dative]<sub>productive</sub>  
**e** = [*langa* 0.7, *vanta* 0.4...]<sub>ACC</sub>  
**ELSE** [+subject, +experiencer] **THEN** [+nominative]<sub>default</sub>  
**e** = [*finnast* 1.0, *leidast* 1.0...]<sub>DAT</sub> and [*langa* 1.0, *vanta* 1.0...]<sub>ACC</sub>

Here *langa* ‘want’ has the exception value of 0.7, meaning that based on distributional information from the input, it is considered an exception in 7 out of 10 times the rule is applied. This approach, with stochastic exceptions instead of stochastic subrules (which might be appropriate in another context, see below) could additionally explain differences in DS rate between verbs, as the application of the rule depends on the strength of the evidence for the verb to be listed as an exception. If an exception is less frequent, it might mean less information about its status as an exception and that the rule is more likely to be applied for that item. As for historically nominative verbs like *hlakka* ‘look forward to’ and *kvíða* ‘be anxious about’, it would be possible to imagine a scenario where the stochastic component to exceptions would allow verbs to be listed as exceptions at various levels of the nested productivity.

But what about the conditioning described in this chapter? For now, I do not assume that the syncretism effects materialize into a subrule like the PSR, following

previous work (Nowenstein & Ingason 2021) which argues that the feature inventory of natural language constrains change and forces a specialization based on e.g. differential aspects of phi-features (Harley & Ritter 2002 and Wiltschko 2008). In other words, the division of DS rules based on person and number, for example using Noyer's (1992) [ $\pm$ participant] feature, is theoretically motivated. This is not as clear for syncretism effects, but a subrule based on the [ $\pm$ participant] feature might either look like (42a) or (42b), with (42a) indicating a stochastic subrule and (42b) showing the difference through stronger evidence for the (now dative) exceptions:

- (42) a. **IF** [+oblique, +subject, +participant] **THEN** 0.7 [+accusative]  
           e = [*finnast* 1.0, *leiðast* 1.0...]DAT  
**ELSE IF** [+oblique, +subject, +experiencer] **THEN** [+dative]<sub>productive</sub>  
           e = [*langa* 0.7, *vanta* 0.4...]ACC  
**ELSE** [+subject, +experiencer] **THEN** [+nominative]<sub>default</sub>  
           e = [*finnast* 1.0, *leiðast* 1.0...]DAT and [*langa* 1.0, *vanta* 1.0...]ACC
- b. **IF** [+oblique, +subject, +participant] **THEN** [+accusative]  
           e = [*finnast* 1.0, *leiðast* 1.0, *langa* 0.1, *vanta* 0.2...]DAT  
**ELSE** +oblique, +subject, +experiencer] **THEN** [+dative]<sub>productive</sub>  
           e = [*langa* 0.7, *vanta* 0.4...]ACC  
**ELSE IF** [+subject, +experiencer] **THEN** [+nominative]<sub>default</sub>  
           e = [*finnast* 1.0, *leiðast* 1.0...]DAT and [*langa* 1.0, *vanta* 1.0...]ACC

Both of these approaches could additionally account for unusual phenomena like the productivity of accusative with historically dative subjects (see arguments in Nowenstein & Ingason 2021), where a productive accusative rule emerges in limited contexts. Perhaps most importantly, this would be an example about how the way children acquire language, in a constant search for productive rules while learning exceptions, shapes the directionality and dynamics of language variation and change, possibly yielding sub rules

which formalize grammatical i-conditioning (Tamminga et al. 2016) into specialized variation (Wallenberg 2019).

## 6. Synthesis

The main objective of this thesis was to shed light on how learners of Icelandic acquire the case marking system of the language, which is characterized by rich morphological marking despite a relatively rigid word order, robust non-default productive patterns and well known instances of grammatically conditioned variability. More generally speaking, this work is an investigation into how children map form, link it to meaning and generate productive rules, and how the mechanisms of language acquisition can shape language variation and change. Although this is done through the lens of Icelandic case marking acquisition, the contribution extends to language learning, language variation and change and case theory more broadly. In the following synthesis of the results, I summarize the main findings and discuss them in the context of previous research, showing how the different components of the dissertation create a concrete account of the acquisition of Icelandic case marking and its non-default productive dative. I finish with directions for future work.

### 6.1 Case has meaning and children know it

In Chapter 2, *The Meaning of Case*, I used two comprehension experiments, with a total of 146 participants aged 2–13, to investigate whether children use case as a cue for verb meaning. This is an important question in two active research areas, (1) work on (morpho)syntactic bootstrapping in language acquisition and (2) work on the nature of case and its relationship to semantics.

In the context of (morpho)syntactic bootstrapping, a prominent question has been whether children favor possibly universal cues such as word order over language-specific ones such as morphological case (Lidz et al. 2003, Göksun et al. 2008, Matsuo et al. 2012, Trueswell et al. 2012, Leischner et al. 2016). Despite earlier work arguing for an advantage of universal cues (Lidz et al. 2003), later contributions clearly demonstrated that children learning languages with overt case marking can use it as a cue early on, with some evidence that the ability to use case as a cue grows with age (Göksun et al. 2008, Matsuo et al. 2012, Leischner et al. 2016). Crucially, it has been maintained that case cues are used in languages where word order is less reliable (Leischner et al. 2016).

In the context of (morpho)syntactic bootstrapping, the rare feature of marking argument function with case as well as word order (Kiparsky 1997, Fedzechkina et al. 2016), creating redundant cues (Tal & Arnon 2022), make Icelandic valuable testing grounds. Indeed, the results from the two comprehension experiments show that preschoolers, as well as older children, are able to use case as a cue for verb meaning, with some evidence for this ability growing with age. Interestingly, the cue value is context dependent and case proves to be particularly useful in subject case, where a dative subject (as opposed to a nominative one) is interpreted as an experiencer. In this context, when correcting for age, case is as salient as transitivity, showing that cue universality is not necessarily an indispensable component when it comes to mapping form to meaning. To push this further, the results in subsection 4.3.3, where links between agents and nominative can be derived, might in some respects be interpreted as a challenge to the existence of cue universality altogether (see Pearl & Sprouse 2021 for linking rule derivability).

Using the context of Icelandic dative experiencer subjects as a contrast to nominative agents, these findings furthermore push research on verb meaning and (morpho)syntactic bootstrapping in languages other than English beyond the cognitively and linguistically prominent causative-transitive association<sup>54</sup> (but see Hacquard & Lidz 2018). They might also be evidence for a bootstrapping account grounded in parsing (Trueswell et al. 2012), where an early case cue, on a subject instead of an object in Icelandic, forces an earlier reinterpretation of the privileged, but perhaps also derivable (see Chapter 4), link between agents and subjects mentioned by e.g. Hartshorne et al. (2015) and Becker (2014).

In general, the results clearly show that child learners take advantage of associations between case and meaning. This reveals the semantic nature of case in Icelandic and argues against formal approaches to case marking where case is considered independent of meaning (most prominently H.Á. Sigurðsson 2012 for Icelandic) and for approaches in which case marking is (1) at least in part, determined by semantics, and (2) the broad generalizations which have been observed are not discarded on the grounds of correlations not being exact or one-to-one (e.g. Thráinsson 2007 and Wood 2015). Putting these criteria in the context of recent formal approaches

---

<sup>54</sup> See Rissman & Majid (2019) and Ünal et al. (2021).

to Icelandic case, this means that the very diverse work by Svenonius (2002), Jónsson (2003), Barðdal (2008), Schätzle (2018) and Galbraith (2018) is on the right track when it comes to acknowledging this semantic component in the nature of case.<sup>55</sup> But as discussed in section 1.1, even very ‘structural’ approaches to case marking, as in Wood 2015 and E.F. Sigurðsson 2017, conserve semantic criteria to case marking through e.g. applicative heads. This usually only applies to non-structural case, and structural case is in contrast traditionally (see overview in Butt 2006) considered completely independent of meaning. One question which arises from the results in this thesis, and the work on the transitive-causative association in language acquisition more broadly, is whether the default nominative-accusative pattern in Icelandic can in fact be considered completely devoid of meaning if it does indeed provide learners with cues to a causative event. Similarly, the generalization that all agents are nominative in canonical contexts, and that non-nominative subjects cannot be agents, seem to be powerful distributional cues in language acquisition which point towards an association between structural case and meaning. There are many ways in which these insights from language acquisition could be incorporated into formal accounts of case marking, and while it is beyond the scope of this thesis to present such an account, it is in line with the approach taken here to reach the conclusion that formalizations of case marking should draw on findings from language acquisition. Language acquisition does, after all, form the basis of adult grammars.

When it comes to formal models of language acquisition, on the other hand, the findings in Chapter 2 are not problematic if we assume a learning model that simply detects the available systematic mappings of form and meaning and does not require them to be absolute to become productive (Yang 2016). Under such an approach, the various cues are exploited by the learner in what seems to be greedy fashion. No cue is privileged: the learner simply picks out those that work best in the sense of local computation (e.g. Belth et al. 2021). In this sense, it overcomes the difficulty of approaches that assert the priority of some cues over others. Such accounts do not easily generalize cross-linguistically (as the growing literature on morphosyntactic literature

---

<sup>55</sup> See McFadden (2021) for an analysis of how different approaches to the derivation of case as a morphological (e.g. Marantz 1991) or syntactic (e.g. Caha 2009) phenomenon make different predictions about the relationship between case and semantics, and his attempt to merge the two lines of formalization into one comprehensive account where structural case is syntactic but inherent case is morphological.

shows), and are furthermore theoretically incomplete as they do not account for how the higher priority cue is “demoted”.

Simpler discovery procedures like the Tolerance Principle (Yang 2016) additionally do not rule out the productivity of non-default minority patterns, the main subject of Chapter 3. Although the same form-meaning patterns found in the linguistic input underlie comprehension and production, I do assume that the processes of linking together form and meaning in comprehension (word learning, Chapter 2), on the one hand, and production (with acquired verb roots, Chapter 3), on the other hand, are fundamentally distinct in nature. When acquiring the meaning of verbs, the directionality of the hypothesis formation goes from form to meaning. When producing language with acquired verbs, the rules are formed in the meaning to form direction, or even form to form in some cases. This has implications for the association patterns I expected to find in the comprehension versus the production experiments, as was further explored in Chapter 3 (as well as derived in Chapter 4), where I focused on finding the semantic contexts in which children acquiring Icelandic form productive dative case marking rules.

## 6.2 Non-defaults can be productive

In Chapter 3, Producing non-default case, I used data from forced-choice tasks 101 children aged 3–13 participated in to find the semantic and structural contexts where the dative is most likely to be produced and (over)generalized. The results are important in the context of dative productivity in general (e.g. Maling 2002, Jónsson 2003, Barðdal 2008, Ingason 2011 and Jónsson & Thórarinsdóttir 2020) as well as work on rule formation and productivity more broadly. These are subjects which are at the heart of current but long-standing debates within language acquisition research, where rule-based accounts of acquisition which assume abstraction (e.g. Albright & Hayes 2003, O’Donnell 2015, Yang 2016) contrast with single- (and dual-, e.g. Pinker 1999) route models which often assume gradient productivity which largely depends on frequency (e.g. Blything et al. 2018, Ambridge et al. 2018, Goldberg 2019).

The broad case productivity patterns which emerged from the forced-choice tasks are mostly in line with previous research but still provide novel insights into the productivity of the dative. Children acquiring Icelandic overgeneralize the default nominative for

subjects and accusative for objects, both direct and indirect (as in Sigurðardóttir 2002), but additionally, there is robust dative productivity (in line with e.g. Maling 2002, Jónsson 2003, Barðdal 2008) and some more marginal phenomena such as nominative on objects and accusative on subjects (also found in Nowenstein 2017, Guðmundsdóttir et al. 2019 and Nowenstein & Ingason 2021). The dative productivity, as became clear in the novel verb task in subsection 3.3.2, is structurally and semantically conditioned. Datives are much more frequently produced on objects than subjects, and dative on indirect objects emerges earlier than dative on direct objects. For subjects, datives are produced more in non-agentive experiencer contexts and for objects, datives are more frequent with motion themes and goals than patients. The results furthermore indicate that the widely cross-linguistically acknowledged association between datives and goals is more prominent earlier in acquisition (in comparison to datives with motion themes). But at the same time, they indicate that children acquire dative productivity based on syntactic role before linking the case marking to the appropriate semantic contexts. This precedence of form over meaning fits with accounts of the early emergence of pattern recognition (e.g. Emond & Shi 2021) and later acquisition of semantically conditioned rule application (e.g. Schuler et al. 2016), suggesting that in language acquisition, form is easy but meaning is hard (Naigles 2002).

Although the dative was in fact used more frequently within the semantic contexts typically associated with it (experiencers, goals, motion themes), it also was robust in unexpected semantic contexts for objects, pointing towards productivity which goes beyond Jónsson & Thórarinsdóttir's (2020) caused motion context as well as Barðdal's (2008) more diverse schemas. This indicates that the dative might even compete with the default accusative for the case-marking of patients.

These findings of semantically restricted dative productivity with subjects but more general/elusive productivity with objects, paired with the fact that no clear contrasts between high and low frequency predicates or novel predicate equivalents were found, calls for a language acquisition model which can account for nested productivity. The model should allow statistically non-dominant and non-default patterns to be productive along with, or within, a rule with a broader range. The model also has to be able to account for the productivity of even more marginal patterns (such as accusative for subjects) and

variability within and across learners through development, such as learners' early association between datives and goals. I argue that a modern rule-based approach such as Yang's Tolerance Principle (2016) is a promising starting point. The abductive search procedure (recursively hypothesizing rules) of recent computational implementations (Belth et al. 2021) is also promising for contexts where rule directionality might be crucial, such as in accounting for the different patterns observed in production and comprehension. Although comprehension-production asymmetries are a somewhat understudied domain within work on form-meaning mappings, it is fairly clear that different directions in hypothesis formation have to be assumed. Although all agents are nominative, not all nominative arguments are agents.

How the distribution of the data in different contexts shapes comprehension and productivity is one of the questions underlying the application of the Tolerance Principle (Yang 2016) in Chapter 4, where I investigated whether this parameter-free discovery procedure can be used to derive the patterns which I confirmed experimentally in Chapters 2 and 3.

### 6.3 Case marking rules can be derived from the input

The main objectives of Chapter 4, *Dative Derivability: A Case for Exceptions*, was to provide an overview of the case frames present in Icelandic child language and investigate whether the links between case, syntactic role and meaning described in Chapters 1, 2 and 3 are discoverable and learnable from the input. This was done with an incremental analysis of the arguments of verbs appearing with a frequency  $>5$  ( $N = 410$ ) in an approximately 750,000 word child language corpus (Einarsdóttir et al. 2019, Sigurjónsdóttir 2007, Strömqvist et al. 1995). The arguments were syntactically and thematically annotated, building on previously proposed generalizations on the associations between case, structural role and meaning. This yields the IceCASE corpus, in which I then implemented a maximal approach to hypothesis formation, both in terms of the size of the set a rule applies to, rule specificity and the conceptual categories assumed. Using frequency as a proxy for order of acquisition, I tested possible rules at 50 verb intervals as well as in randomly generated verb samples at each 100 verb interval

(with  $N - 50$ ). When a productive rule was not discovered, the set was further divided (Yang 2016, Belth et al. 2021).

Rule productivity was evaluated using the Tolerance Principle (TP, Yang 2016), a parameter-free learning model which provides a threshold for the number of exceptions a productive rule can tolerate. I argue that the TP, with its quantification of what is (1) sufficient evidence of a rule's productivity and (2) a tolerable number of exceptions, while maintaining said productivity, provides a cognitively plausible account of how children can learn to (over)generalize mappings which are not exact. Such an acquisition standpoint has been missing in formal approaches to case in Icelandic, and it could be argued that any formalization of case needs to acknowledge that the case marking system is acquired by children with relatively small vocabularies. Additionally, the words in those vocabularies have to be linked to concepts which have been developmentally accounted for. As is discussed in subsection 1.1.3, there is independent evidence for children's early sensitivity to agentivity (Rissman & Majid 2019), animacy (Opfer & Gelman 2011, see Becker 2014 for an overview) and (caused) motion (Kellman et al. 1987), features which were annotated in the IceCASE corpus.

Although acquisition has not been a central component of accounts on Icelandic case, the approach to case marking in Chapter 4 can be argued to have a (very) long theoretical tradition, as broad generalizations between case, form and meaning are e.g. found both in Panini's work and the Greco-Roman approach to case (see Blake 2001 and Butt 2006). The route I take is therefore not new at all, although its distributional derivation from child-language corpora is a novel contribution. In general, the results from the chapter show distributions which are in line with previous work: Subjects are mostly nominative, indirect objects and optional reflexives are almost entirely dative and the majority of direct objects are accusative, but a robust number of them are dative as well. This distribution correlates with agents<sup>56</sup> being nominative and (most) goals being dative, but the case marking of patients and themes is more split, with patients still leaning towards the accusative and motion themes towards the dative. Additionally, a case for exceptions is made – I show that various generalizations can, despite exceptions, be derived from the input and formed into productive rules, subrules and structured

---

<sup>56</sup> And most experiencers, although a considerable number of them are dative.

exceptions. This shows that children do extract case-meaning mappings from their input (as Pearl & Sprouse 2021 show they can for rUTAH), but the results also show that in various contexts, an association between case and form (syntactic role) takes precedence and is enough to account for the data. Although case can clearly index meaning as is shown in Chapter 2, it is also determined structurally.

The overview of the IceCASE corpus shows that the NOM.-ACC. frame is the most frequent frame, followed by intransitives with nominative subjects, the NOM.-DAT. frame, alternating NOM.-ACC./DAT. predicates, NOM.-DAT-ACC. ditransitives and finally less frequent frames (e.g. those with non-nominative subjects). The incremental approach taken in the analysis, where a proxy for smaller vocabularies is explored first, reveals various novel insights. First, an early prominence of dative goals is detected, as the NOM.-DAT. frame is relatively infrequent at first (in the top 100 verbs) compared to the sample as a whole, with ditransitives and optional reflexive arguments initially making up a higher proportion of dative verbs. Children therefore seem to start by building themselves a more cross-linguistically expected association between the dative and applied arguments, with the association between datives and motion themes appearing later. This is in line with the experimental results and could be supported by other regularities in form: as indirect objects and optional reflexive arguments are almost exclusively dative and a TP analysis of a [-subject, -direct object → +dative] rules, as well as a [+goal → +dative] rule, yields productivity at all levels of testing, despite the presence of exceptions. Other expected patterns also emerge and are more stable across the sample: Most subjects, and all agents, are nominative, yielding productivity at all levels for [+subject → +nominative] and [+agent → +nominative]. In the word learning/comprehension direction, [+nominative → +subject] is productive at all levels but [+nominative → +agent] only is productive if polythematic arguments are excluded, showing that the different directionality of hypothesis formation in production and comprehension yields different results.

For production, where the end result is case, this yields a system where syntactic role is initially enough to assign case productively for subjects (nominative), indirect objects (dative) and optional reflexives (dative). For direct objects, which only initially have a productive accusative rule, the dative exceptions are so numerous that they compete

with the default accusative and more partitioning is needed. Adding a [+patient] feature permits assigning accusative consistently. When it comes to the remaining direct object themes, adding [-motion] yields a consistently productive accusative rule. But for direct object themes, which change location ([+motion]), no productive rule (neither accusative or dative) is found, pointing to either a scenario without productivity, where both accusative and dative are possible, or the necessity to discover more precise subrules (as in Maling 2002 and Jónsson 2013). This lack of productivity is in line with children's difficulty to use case as a cue for motion (Chapter 2) and the complexity of reaching productive rules for direct objects is furthermore in line with the broad semantic scope of dative productivity in Chapter 3. Within the rule-based models reviewed in Chapter 3, Yang's (2016) Tolerance Principle is unique in not needing to identify productive rules if the data does not support any. Other models (e.g. Albright & Hayes 2003 and O'Donnell 2015) always have a highest valued option, something which does not seem to be supported by the evidence when it comes to the diffuse case marking patterns of Icelandic direct objects. Additionally, these results show that the status of the accusative as a default case for direct objects in Icelandic is more fragile than the status of the dative case for indirect objects and optional reflexives. The status of the dative for subjects, on the other hand, is unequivocally that of exceptions, despite its productivity within Dative Substitution. Still, most rule-based accounts assume that exceptions can be structured into productive minority rules (see Yang 2016 for the English past tense), and a TP analysis shows that the dative is in fact productive for non-nominative subjects, which are all experiencers in the IceCASE corpus.

This creates a very salient [+oblique subject → experiencer] cue (without any exceptions) in the comprehension hypothesis formation. The rule is even more reliable than the association between nominative subjects and agents, which still is productive under certain conditions. When it comes to objects, dative arguments which are not direct objects again clearly indicate goals. But in the context of direct objects, a parallel with the production results appears, with no clear patterns emerging for neither accusative or dative. Most accusative direct objects are either patients or themes and most dative direct objects are either themes or goals, with no consistent cues beyond this division. These results also complement the experimental data from Chapter 2 rather well, providing a

deeper explanation for the association between datives and applied arguments and the later emergence of a distinction between patients and themes based on case, along with the particular salience of non-nominative subjects as a cue. A closer look at the acquisition of non-nominative subjects, their variation and change was taken in Chapter 5, where I investigated how children learn to use these exceptions productively as is predicted by the TP. In general, the results from Chapter 4 show that children do derive generalizations, but what happens when the input is not consistent and variation is grammatically conditioned?

## 6.4 Acquisition specializes variation

Finally, in Chapter 5: Acquiring and Shaping Case Marking Variation, I put the findings regarding the meaning of case, its non-default productivity and derivability in the context of language variation and change. This is done with a case study on Dative Substitution (DS) in 101 children aged 3–13 and 95 parents, which were additionally matched into 84 child-caretaker dyads.

DS was described in subsection 1.3 and various findings in Chapters 2, 3 and 4 provide insights into the nature of this change in the case marking of subjects. In Chapter 2, the results clearly show that children associate non-nominative subjects with experiencers, the exact context where the dative is extended to historically accusative (and to a limited extent nominative) subjects (Jónsson 2003 and Jónsson & Eythórsson 2005 i.a.). In Chapter 3, I confirmed that children do produce non-nominative subjects with novel verbs depicting experiencers and also replicated previous findings (Nowenstein 2017, Nowenstein & Ingason 2021) indicating that children acquire variational case marking with DS verbs. Finally, the Tolerance Principle analysis (TP, Yang 2016) in Chapter 4 showed that the dative should be productive within a cluster of non-nominative subjects which all denote experiencers. Although the objective of Chapter 5 was to investigate how children acquire grammatically conditioned variation already present in their input, it is clear that the findings just described point to a scenario where the driving force behind DS is found in language acquisition, supporting a child innovator approach to language variation and change (see e.g. Lightfoot 1979 and Cournane 2015,

2017 and 2019, as well as Yang 2016 for Icelandic subject case variation in particular). But how do children acquire variation once it is the norm in their language environment?

As was described in Chapter 5, one of the fundamental questions within developmental sociolinguistics, and language acquisition research more broadly, has to do with children's reaction to variability in their input, or primary linguistic data (e.g. Labov 1989, Yang 2002, Hudson Kam & Newport 2005, Smith et al. 2009, Cournane & Pérez-Leroux 2020). Interestingly, children are both known to converge with and diverge from the input. They usually acquire their caregivers' dialect and can in some circumstances even match their variant usage rate (Labov 1989, Smith et al. 2007, Johnson & White 2019). But they are also master 'generalizers' (Sneller & Newport 2019) who can innovate, push incrementation forward and determine the directionality of language change (e.g. Cournane 2017, Cournane & Pérez-Leroux 2020 and Hall & Maddeaux 2020). Although this mixture of divergence and convergence might seem paradoxical (Repetti-Ludlow & MacKenzie 2022), it is supported by the findings in Chapter 5. Indeed, the results show that when acquiring the grammatically conditioned case marking variation of Icelandic Dative Substitution, children neither regularize nor match their parents' variant rate. Instead, systematic contrast replication at the group and individual level is observed, or regularization within matched conditions. In other words, children acquire a system which contains the same grammatical conditioning that is observed in their caregivers' results, while also advancing the change.

In the case of DS, the grammatical conditioning which these findings apply to is the Person-Specific Retention (PSR, Svavarsdóttir 2013, Nowenstein 2017, Nowenstein & Ingason 2021) and syncretism effects, confirmed for the first time in the present work, which also is the first attempt at documenting a variable in caregiver-child dyads in Icelandic. I furthermore make an attempt at formalizing DS and its conditioning with a combination of the TP and Yang's (2002) Variational Model of Language Acquisition (VM), with the novel addition of stochastic exceptions in addition to the VM's stochastic rules. I argue that in cases of variation such as DS, there might be a need to assume that learning whether an item is an exception or not is a variational learning process. Sometimes the child hears *langa* 'want' with an accusative subject (exception), sometimes with a dative (productive rule). This would account for the varying rates of DS

between verbs and could additionally be applied within subrules. One of the advantages of using a combination of the TP and the VM is the possibility for nested rules, where more specific rules are applied first, which is the mechanism I assume for the PSR (where accusative is applied at a higher rate with first and second person pronoun subjects than with third person subjects). For now, I do not assume that the syncretism effects, which are not present in the youngest participants, materialize into a subrule like the PSR. Instead, I assume that they could be interpreted as a preference for contrast saliency when the input is variable, because the nominative-accusative syncretism entails a less clear signal for an oblique subject construction.

In general, the results of Chapter 5, building on the work in previous chapters, show how the way children acquire language, in a constant search for productive rules while learning exceptions, shapes the directionality and dynamics of language variation and change, possibly yielding subrules which formalize grammatical i-conditioning (Tamminga et al. 2016) into specialized variation (Wallenberg 2019). Paired with sociostylistic conditioning, this potentially creates scenarios where learners contribute to incrementation in certain contexts while stabilizing variation in others.

## 6.5 Future work

Various limitations to the current project have been mentioned at different points in this thesis and will hopefully be addressed in future work. When it comes to the experiments, one major limitation is the very broad age range of the participants. While it does provide a useful overview of the development of case marking in toddlers to teenagers, it sacrifices statistical power when it comes to providing a precise developmental timeline for the acquisition of Icelandic case marking and its variation. In the same vein, all experiments should have included an adult control group. In the context of the corpus analysis, various other routes could have been taken in the annotation, including a more decompositional and/or event structure based approach (see e.g. Schätzle 2018). Additionally, exploratory clustering analyses (e.g. White et al. 2017, Becker & Guzmán Naranjo 2020) could have been conducted to potentially reveal novel mappings and links between case, semantics and syntactic role. It is also crucial to contrast the Tolerance Principle (Yang 2016) with other accounts of language acquisition which depend on e.g.

entrenchment and pre-emption (Ambridge et al. 2018, Goldberg 2019). This type of comparison requires a different annotation of the child language data (with usage frequencies in addition to type frequency), which was unfortunately outside the scope of the current work. Finally, I would have liked to address individual participant profiles to a much greater extent.<sup>57</sup> While I hope to get the opportunity to address these limitations, and many others, in future work, there are also new directions which could prove fruitful. One of them is briefly outlined below, the Tolerance Principle as an adaptation in an evolutionary context.

### 6.5.1 The Tolerance Principle as an adaptation

In future work, I furthermore aim to explore the Tolerance Principle as an adaptation in an evolutionary context and I briefly outline the direction this work might take here.<sup>58</sup> The data presented in this thesis are relevant in such an endeavor because an essential trait in the inherited suite of adaptations that constitute the human language faculty is the child language learner's ability to induce rules from noisy, stochastic linguistic observations. Derived generalizations are not consistently exact but productive rules do still underlie the combinatorial nature of human language, making *infinite use of finite means* as observed by von Humboldt, and allowing language learners to generalize beyond the input they hear from the previous generation. This is a fact which has been demonstrated experimentally and replicated numerous times, in this thesis for example, since the seminal work of Berko-Gleason (1958). Learners are able to induce productive rules in the face of considerable noise (e.g. speech errors and perception errors) and while acquiring massive amounts of complex sociolinguistic variation. It is therefore a central part of linguistic cognition that children are able to abstract rules from the mess, while also realizing that not everything is a rule: true exceptions must be recognized, learned, and listed. In an evolutionary context, we propose that productive rules exist by virtue of the Tolerance Principle (Yang 2016) as a specific adaptation. But while Yang intended

---

<sup>57</sup> For example, the Icelandic dative productivity could be particularly interesting in the context of language learning and memory (e.g. Ullman 2001, Lum et al. 2012 and 2014, Soh & Yang 2021), since it is an example of lexically dependent rule formation which presumably taps into both declarative and procedural memory.

<sup>58</sup> This approach was developed with Joel Wallenberg.

the TP as a model of language learning, we suggest this calculation is itself the tool that the evolution of language learning has selected for.

The TP was proposed from a non-evolutionary perspective and the equation derived under the likely faulty assumptions (Yang 2021) that children roughly have the processing constraints of a digital computer, and that any vocabulary follows a Zipfian distribution (Zipf, 1949). The TP is intended to only allow the positing of rules that will save computational time for a learner in production, assuming the learner always searches sequentially through a list of exceptions ordered by frequency rank. This will only save computational time for Zipfian vocabularies. Nevertheless, the Tolerance Principle makes highly accurate empirical predictions in both corpus and experimental studies, as is described in section 4.1. Most surprisingly, it does so even when children are experimentally forced to consider rules over a vocabulary with a non-Zipfian distribution. Koulaguina and Shi (2019) conducted an experiment in which artificial verbs undergoing a certain rule had uniform distributions, and therefore violate the Zipfian distribution assumption of the Tolerance Principle. Even so, the TP correctly predicted when their child participants posited rules and when they did not. Additionally, the equation above uses an approximation of the Nth Harmonic number in a Zipfian distribution,  $\ln N$ . Bizarrely, Schuler et al. (2016)'s artificial language study showed that the equation in (1) and repeated below, with its approximation, matched experimental results better than an exact calculation of the Nth Harmonic number based on the experimental conditions (Yang 2021).

- (43) Let a rule  $R$  be defined over a set of  $N$  items.  $R$  is productive if and only if  $e$ , the number of items not supporting  $R$ , does not exceed  $\theta_N$ :

$$e \leq \theta_N = \frac{N}{\ln N}$$

The question is: Why should the Tolerance Principle work for vocabularies where it does not save computational time at all, and why should an approximation work more accurately than a precise calculation?

We propose to solve this seeming paradox by taking an evolutionary perspective: even if the children do not have access to very large vocabularies that are always Zipfian and amenable to an InN approximation, natural selection does. Given that a threshold for productivity is necessary for language learners/users to have the considerable benefits of not having to memorize everything as an exception, humans must have evolved a threshold of some kind. If multiple phenotypes of the threshold trait were innovated and competed during the evolution of rule-based language, and the Tolerance Principle was one such phenotype, then it would have a selectional advantage over others by virtue of the fact that vocabularies are often Zipfian, and an InN approximation is correct as much of the time as can be expected from an easily calculable heuristic. Baixeries et al. (2013) explored Zipfian distributions in child language and child-directed speech, and found that while no individual speakers represent the Zipfian distributions found in nature very well, children change in the direction of the adults' representation over time. Thus, child language learners do not have access to a true representation of the distribution of large vocabularies, and no individual speaker does, but the experimental results show that they use a heuristic which assumes exactly the large vocabulary distributions that they are failing to actually estimate.

If such a heuristic were innovated in evolutionary time, it would be more useful even than an adult's exact calculations based on their estimation of a large vocabulary distribution, because the heuristic would more closely match true distributions in communities of speakers. As in the famous "outfielder problem" (McBeath et al. 1995), we propose a model in which natural selection gave us a better heuristic than anything we could calculate.

# References

- Alexiadou, A., Anagnostopoulou, E. & Sevdali, C. 2014. Opaque and transparent datives, and how they behave in passives. *Journal of Comparative Germanic Linguistics*, 17, 1–34. doi:10.1007/s10828-014-9064-8
- Albright, A., & Hayes, B. 2003. Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition*, 90(2), 119–161. doi:10.1016/S0010-0277(03)00146-X
- Ambridge, B. 2010. Children's judgments of regular and irregular novel past-tense forms: New data on the English past-tense debate. *Developmental Psychology*, 46(6), 1497–1504. doi:10.1037/a0020668
- Ambridge, B. 2020. Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, 40(5–6), 509–559. doi:10.1177/0142723719869731
- Ambridge, B. 2020b. Abstractions made of exemplars or 'You're all right, and I've changed my mind': Response to commentators. *First Language*, 40(5–6), 640–659. doi:0.1177/0142723720949723
- Ambridge, B., Barak, L., Wonnacott, E., Bannard, C., Sala, G. 2018. Effects of Both Preemption and Entrenchment in the Retreat from Verb Overgeneralization Errors: Four Reanalyses, an Extended Replication, and a Meta-Analytic Synthesis. *Collabra: Psychology*, 4(1), 23. doi:10.1525/collabra.133
- Anagnostopoulou, E., & Sevdali, C. 2015. Case alternations in Ancient Greek passives and the typology of Case. *Language*, 91, 442–481.
- Anderson, S.R. 1990. The Grammar of Icelandic Verbs in -st. In J. Maling & A. Zaenen, *Modern Icelandic Syntax*, pp. 235–273. Brill.
- Andrews, A. 1976. The VP Complement Analysis in Modern Icelandic. *Proceedings of the North East Linguistic Society (NELS 6)*, pp. 1–21.
- Arunachalam, S., & Waxman, S.R. 2010. Meaning from syntax: Evidence from 2-year-olds. *Cognition*, 114(3), 442–446. doi:10.1016/j.cognition.2009.10.015
- Árnadóttir, H. & Sigurðsson, E.F. 2012. Case in disguise. In B. Fernandez & R. Etxepare (eds.), *Variation in datives: A microcomparative perspective*, pp. 96–143. Oxford University Press.

- Baixeries, J., Elvevåg, B. & Ferrer-i-Cancho, R. 2013. The evolution of the exponent of Zipf's law in language ontogeny. *PloS one*, 8(3), e53227.  
doi:0.1371/journal.pone.0053227
- Baker, C. L. 1979. Syntactic theory and the projection problem. *Linguistic Inquiry*, 10, 533–581.
- Baker, M.C. 1997. Thematic Roles and Syntactic Structure. In L. Haegeman (ed.), *Elements of Grammar*, pp. 73–137. Springer. doi:10.1007/978-94-011-5420-8\_2
- Baker, M.C. 2015. *Case*. Cambridge University Press.
- Barðdal, J. 1993. Accusative and dative case of objects of some transitive verbs in Icelandic and the semantic distinction between them. In *Flyktförsök: Kalasbok till Christer Platzack*, 1–13.
- Barðdal, J. 2001. Case in Icelandic: A Synchronic, Diachronic and Comparative Approach. Doctoral dissertation, Lund University.
- Barðdal, J. 2008. Productivity: *Evidence from Case and Argument Structure in Icelandic*. *Constructional Approaches to Language*, 8. John Benjamins.  
doi:10.1075/cal.8
- Barðdal, J. 2011. Lexical vs. structural case: a false dichotomy. *Morphology*, 21, 619–654. doi:10.1007/s11525-010-9174-1
- Barðdal, J. & Eythórsson, Th. 2003. The change that never happened: The story of oblique subjects. *Journal of Linguistics*, 39(3), 439–472.  
doi:10.1017/S002222670300207X
- Barðdal, J., Eythórsson, Th. & Dewey, T.K. 2014. Alternating Predicates in Icelandic and German: A Sign-Based Construction Grammar Account. *Working Papers in Scandinavian Syntax*, 93, 50–101.
- Barðdal, J., Smitherman, T., Bjarnadóttir, V., Danesi, S., Jensen, G.B. & McGillivray, B. 2012. Reconstructing Constructional Semantics: The Dative Subject Construction in Old Norse-Icelandic, Latin, Ancient Greek, Old Russian and Old Lithuanian. *Studies in Language*, 36(3), 511–547. doi:10.1075/sl.36.3.03bar
- Bates, D., Mächler, M., Bolker, B., & Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.  
doi:10.18637/jss.v067.i01

- Becker, L. & Guzmán Naranjo, M. 2020. Psych predicates in European languages: A parallel corpus study. *STUF – Language Typology and Universals*, 73(4), 483–523. doi:10.1515/stuf-2020-1017
- Becker, M. 2014. *The Acquisition of Syntactic Structure: Animacy and Thematic Alignment*. Cambridge University Press. doi:10.1017/CBO9781139022033
- Belletti, A., Rizzi, L. 1988. Psych-verbs and  $\theta$ -theory. *Natural Language & Linguistic Theory*, 6, 291–352. doi:10.1007/BF00133902
- Belth, C., Payne, S., Beser, D., Kodner, J., & Yang, C. 2021. The greedy and recursive search for morphological productivity. arXiv preprint arXiv:2105.05790
- Berko Gleason, J. 1958. The child's learning of English morphology. *Word*, 14(2–3), 150–177.
- Blake, B. 2001. *Case*. Cambridge University Press. doi:10.1017/CBO9781139164894
- Blything, R.P., Ambridge, B., & Lieven, E.V.M. 2018. Children's acquisition of the English past-tense: Evidence for a single-route account from novel verb production data. *Cognitive Science*, 42(Suppl 2), 621–639. doi:10.1111/cogs.12581
- Bobaljik, J.D. 2008. Where's phi? Agreement as a postsyntactic operation. In D. Harbour, D. Adger & S. Béjar (eds.): *Phi Theory: Phi-features across modules and interfaces*, 295–328. Oxford University Press.
- Bornkessel-Schlesewsky, I., Roehm, D., Mailhammer, R. & Schlewsky, M. 2020. Language Processing as a Precursor to Language Change: Evidence From Icelandic. *Frontiers in Psychology*, 10, 3013. doi:10.3389/fpsyg.2019.03013
- Bosse, S., Bruening, B. & Yamada, M. 2012. Affected Experiencers. *Natural Language and Linguistic Theory*, 30, 1185–1230. doi:10.1007/s11049-012-9177-1
- Bowerman, M. & Brown, P. 2008. Introduction. In M. Bowerman & P. Brown (eds): *Crosslinguistic Perspectives on Argument Structure*, pp. 1–29. Taylor & Francis.
- Braine, M. & Brooks, P. 1995. Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello & W.E. Merriman (eds.): *Beyond names for things: Young children's acquisition of verbs*, 352–376. Erlbaum.
- Bresnan, J., & Kanerva, J.M. 1989. Locative Inversion in Chicheŵa: A Case Study of Factorization in Grammar. *Linguistic Inquiry*, 20(1), 1–50.

- Brusini, P., Seminck, O., Amsili, P., & Christophe, A. 2021. The Acquisition of Noun and Verb Categories by Bootstrapping From a Few Known Words: A Computational Model. *Frontiers in Psychology*, 12, Article 661479. doi:10.3389/fpsyg.2021.661479
- Butt, M. 2006. *Theories of Case*. Cambridge University Press.
- Butt, M. & King, T.H. 2004. The Status of Case. In V. Dayal and A. Mahajan (eds.) *Clause Structure in South Asian Languages*, pp. 153–198. Kluwer Academic Publishers.
- Bybee, J.L., & Moder, C.L. 1983. Morphological Classes as Natural Categories. *Language*, 59 (2), 251–270.
- Caha, P. 2009. The nanosyntax of case. Doctoral dissertation, University of Tromsø.
- Carey, S. 2009. *The Origin of Concepts*. Oxford University Press.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Foris.
- Christensen, L. 2010. Early verbs in child Swedish – a diary study on two boys Part I: Verb spurts and the grammar burst. *Nordlund*, 30.
- Christiansen, M.H. & Chater, N. 2016. *Creating language: Integrating evolution, acquisition, and processing*. MIT Press.
- Clahsen, H., Eisenbeiss, S., & Penke, M. 1996. Lexical learning in early syntactic development. *Language Acquisition and Language Disorders*, 14, 129–160.
- Cournane, A. 2015. Modal Development: Input-Divergent L1 Acquisition in the Direction of Diachronic Reanalysis. Doctoral dissertation, University of Toronto.
- Cournane, A. 2017. In defence of the child innovator. In E. Mathieu & R. Truswell (eds), *Micro-change and Macro-change in Diachronic Syntax*, Oxford Studies in Diachronic and Historical Linguistics, pp. 10–24. doi:10.1093/oso/9780198747840.003.0002
- Cournane, A. 2019. A developmental view on incrementation in language change. *Theoretical Linguistics*, 45(3-4), 127–150. doi:10.1515/tl-2019-0010
- Cournane, A. & Pérez-Leroux, A. T. 2020. Leaving Obligations Behind: Epistemic Incrementation in Preschool English. *Language Learning and Development*, 16(3), 270–291. doi:10.1080/15475441.2020.1738233
- Dowty, D. 1991. Thematic Proto-Roles and Argument Selection. *Language*, 67(3),

- 547–619. doi:10.2307/415037
- Dunn, L.M., & Dunn, D.M. 2007. *Peabody Picture Vocabulary Test: Fourth Edition* (PPVT-4). APA PsycTests. doi:10.1037/t15144-000
- Einarsdóttir, J.T., Pétursdóttir, A.L. & Rúnarsdóttir, Í.D. 2019. *Tíðni orða í tali barna*. Háskólaútgáfan.
- Eisenbeiss, S., Bartke, S., & Clahsen, H. 2006. Structural and lexical case in child German: Evidence from language-impaired and typically developing children. *Language Acquisition*, 13(1), 3–32. doi:10.1207/s15327817la1301\_2
- Emond, E. & Shi, R. 2021. Infants' Rule Generalization is Governed by the Tolerance Principle. In D. Dionne & L.-A. Vidal Covas (eds.): *Proceedings of the 45th annual Boston University Conference on Language Development*. Cascadilla Press.
- Eythórsson, Th. & Barðdal, J. 2005. Oblique Subjects: A Common Germanic Inheritance. *Language*, 81, 824–881. doi:10.1353/lan.2005.0173
- Eythórsson, Th., Johannessen, J.B., Laake, S. & Áfarli, T.A. 2012. Dative case in Norwegian, Icelandic and Faroese: Preservation and nonpreservation. *Nordic Journal of Linguistics*, 35, 219–249 doi:10.1017/S0332586513000036
- Eythórsson, Th. & Thráinsson, H. 2017. *Variation in oblique subject constructions in Insular Scandinavian*. In H. Thráinsson, C. Heycock, H. Petersen & Z. Hansen (eds.), *Syntactic Variation in Insular Scandinavia*, pp. 53–90. John Benjamins.
- Fedzechkina, M., Newport, E.L., & Jaeger, T.F. 2017. Balancing Effort and Information Transmission During Language Acquisition: Evidence From Word Order and Case Marking. *Cognitive Science*, 41(2), 416–446. doi:10.1111/cogs.12346
- Fillmore, C.J. 1968. The case for case. In E. Bach, & R.T. Harms (eds.), *Universals in Linguistic Theory*, pp. 1–88. Holt, Rinehart, and Winston.
- Fisher, C., Gertner, Y., Scott, R.M., & Yuan, S. 2010. Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(2), 143–149. doi:10.1002/wcs.17
- Fisher, C., Jin, K.-s. & Scott, R.M. 2020. The Developmental Origins of Syntactic Bootstrapping. *Topics in Cognitive Science*, 12, 48–77. doi:10.1111/tops.12447
- Fisher, C., Klingler, S.L. & Song, H.J. 2006. What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions. *Cognition*, 101(1),

- 19–29. doi:10.1016/j.cognition.2005.10.002
- Fruehwald, J. & Wallenberg, J.C. 2013. Optionality is Stable Variation is Competing Grammars. Talk presented at 25th Scandinavian Conference of Linguistics, Formal Ways of Analyzing Variation (FWAV) Workshop.
- Galbraith, D. 2018. The Predictable Case of Faroese. Doctoral dissertation, Stanford University.
- Gleitman, L. 1990. The structural sources of verb meanings. *Language Acquisition*, 1(1), 3–55. doi:10.1207/s15327817la0101\_2
- Goldberg, A. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.
- Goldberg, A. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press.
- Goldberg, A. 2019. *Explain me this*. Princeton University Press.
- Göksun, T., Küntay, A.C. & Naigles, L. 2008. Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. *Journal of Child Language*, 35(2), 291–323. doi:10.1017/S0305000907008471.
- Grimm, S. 2011. Semantics of case. *Morphology*, 21, 515–544. doi:10.1007/s11525-010-9176-z
- Grimshaw, J. 1990. *Argument structure*. MIT Press.
- Guðmundsdóttir, D., Nowenstein, I. & Sigurjónsdóttir, S. 2019. Changing Variation: Diffuse Directionality in Icelandic Subject Case Substitution. *Penn Working Papers in Linguistics*, 25.1.
- Gunnarsdóttir, H.G. 1996. Aukafallsfrumlög: Athugun á máli barna. BA-thesis, University of Iceland.
- Hacquard, V. & Lidz, J. 2018. Children's attitude problems: Bootstrapping verb meaning from syntax and pragmatics. *Mind & Language*, 34, 73–96. doi:10.1111/mila.12192
- Hadley, P.A., Rispoli, M., & Hsu, N. 2016. Toddlers' Verb Lexicon Diversity and Grammatical Outcomes. *Language, speech, and hearing services in schools*, 47(1), 44–58. doi:10.1044/2015\_LSHSS-15-0018
- Hall, E. 2020. Child participation in linguistic changes in progress in Ontario English.

- Doctoral dissertation, University of Toronto.
- Hall, E. & Maddeaux, R. 2020. /u/-fronting and /æ/-raising in Toronto families. *University of Pennsylvania Working Papers in Linguistics*, 25(2), 7.
- Halldórsson, H. 1982. Um méranir: Drög að samtímalegri og sögulegri athugun. *Íslenskt mál og almenn málfræði*, 4, 159–189.
- Halle, M. & Mohanan, K.P. 1985. Segmental Phonology of Modern English. *Linguistic Inquiry*, 16(1), 57–116.
- Harðarson, G.R. 2016. A case for a Weak Case Contiguity hypothesis – a reply to Caha. *Natural Language and Linguistic Theory*, 34(4), 1329–1343.  
doi:10.1007/s11049-016-9328-x
- Harley, H. 1995. Subjects, events, and licensing. Doctoral Dissertation, Massachusetts Institute of Technology.
- Harley, H. & Ritter, E. 2002. Person and number in pronouns: A feature-geometric analysis. *Language*, 78(3), 482–526.
- Harrigan, K., Hacquard, V. & Lidz, J. 2019. Hope for syntactic bootstrapping. *Language*, 95, 642–682. doi:10.1353/Lan.2019.0069
- Hartshorne, J.K., O'Donnell, T.J., Sudo, Y., Uruwashii, M., Lee, M., & Snedeker, J. 2016. Psych verbs, the linking problem, and the acquisition of language. *Cognition*, 157, 268–288. doi:10.1016/j.cognition.2016.08.008
- Hartshorne, J.K., Pogue, A., & Snedeker, J. 2015. Love is hard to understand: the relationship between transitivity and caused events in the acquisition of emotion verbs. *Journal of Child Language*, 42(3), 467–504.  
doi:10.1017/S0305000914000178
- Hendricks, A.E., Miller, K. & Jackson, C.N. 2018. Regularizing Unpredictable Variation: Evidence from a Natural Language Setting. *Language Learning and Development*, 14(1), 42–60, doi:10.1080/15475441.2017.1340842
- Hoff, E., Quinn, J.M., & Giguere, D. 2018. What explains the correlation between growth in vocabulary and grammar? New evidence from latent change score analyses of simultaneous bilingual development. *Developmental Science*, 21(2), 10.1111/desc.12536. doi:10.1111/desc.12536
- Hreggviðsdóttir, H. 2018. Skilningur íslenskra barna á mismunandi setningagerðum,

- með og án sýnilegrar fallmörkunar. MS-thesis, University of Iceland.
- Hudson Kam, C.L. 2015. The impact of conditioning variables on the acquisition of variation in adult and child learners. *Language*, 91(4), 906–937.  
doi:10.1353/lan.2015.0051.
- Hudson Kam, C.L. & Newport, E. 2005. Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Language Learning and Development*, 1(2), 151–195.  
doi:10.1080/15475441.2005.9684215
- Hudson Kam, C.L. & Newport, E.L. 2009. Getting it right by getting it wrong: when learners change languages. *Cognitive psychology*, 59(1), 30–66.  
doi:10.1016/j.cogpsych.2009.01.001
- Ingason, A.K. 2010. Productivity of Non-Default Case. *Working Papers in Scandinavian Syntax*, 85, 65–117.
- Ingason, A.K. 2015. Rhythmic preferences in morphosyntactic variation and the theory of loser candidates. In R. Vogel & R. van de Vijver (eds.), *Rhythm in Cognition and Grammar: A Germanic Perspective*, pp. 235–254. De Gruyter Mouton.
- Ingason, A.K. 2016. Realizing Morphemes in the Icelandic Noun Phrase. Doctoral dissertation, University of Pennsylvania.
- Ingason, A.K. & Sigurðsson, E.F. 2017. The interaction of adjectival structure, concord and affixation. In A. Lamont & K. Tetzloff (eds.), *Proceedings of the 47th Meeting of the North East Linguistic Society (NELS 47)*, vol. 2, pp. 89–98.
- Ingólfssdóttir, S.L., Loftsson, H., Daðason, J.F. & Bjarnadóttir, K. 2019. Nefnir: A high accuracy lemmatizer for Icelandic. In *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, pp. 310–315.
- Irani, A. 2019. Learning From Positive Evidence: The Acquisition Of Verb Argument Structure. Doctoral dissertation, University of Pennsylvania.
- Jackendoff, R. 1990. *Semantic structures*. The MIT Press.
- Jin, K.S. 2015. The Role of Syntactic and Discourse Information in Verb Learning. Doctoral dissertation, University of Illinois at Urbana-Champaign.
- Johnson, E.K. & White, K.S. 2020. Developmental sociolinguistics: Children's acquisition of language variation. *WIREs Cognitive Science*, 11, 1515.

doi:10.1002/wcs.1515

Jónsson, H.P. and Loftsson, H. 2021. *ABLTagger* (Lemmatizer) – 3.1.0. CLARIN-IS, <http://hdl.handle.net/20.500.12537/134>.

Jónsson, J.G. 1996. Clausal architecture and case in Icelandic. Doctoral dissertation, University of Massachusetts, Amherst.

Jónsson, J.G. 1997–98. Sagnir með aukafallsfrumlagi. *Íslenskt mál og almenn málfræði*, 19–20, 11–43.

Jónsson, J.G. 2000. Case Assignment and Double Objects in Icelandic. In D. Nelson & P. Foulkes (eds.), *Leeds Working Papers in Linguistics*, 8, pp. 71–94.

Jónsson, J.G. 2003. Not so quirky: On subject case in Icelandic. In E. Brandner & H. Zinsmeister (eds.), *New perspectives on Case Theory*, 127–163. CSLI.

Jónsson, J.G. 2005. Merkingarhlutverk, rökliðir og fallmörkun. In H. Thráinsson (ed.): *Setningar, Íslensk tunga III*, pp. 265–349. Almenna bókafélagið.

Jónsson, J.G. 2009. Verb classes and dative objects in Insular Scandinavian. In J. Barðdal & S.L. Shelliah (eds.), *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, pp. 203–224. John Benjamins.

doi:10.1075/slcs.108.12jon

Jónsson, J.G. 2013. Dative versus Accusative and the Nature of Inherent Case. In B. Fernandez & R. Etxepare (eds.), *Variation in Datives: A Microcomparative Perspective*, Oxford Studies in Comparative Syntax. Oxford University Press.

doi:10.1093/acprof:oso/9780199937363.003.0005

Jónsson, J.G. 2017. Avoiding Genitive in Icelandic. In H. Thráinsson, C. Heycock, Z. Svabo & H.P. Petersen (eds.), *Syntactic Variation in Insular Scandinavian*, pp. 141–164. John Benjamins.

Jónsson, J.G. & Eythórsson, Th. 2003. Breytingar á frumlagsfalli í íslensku. *Íslenskt mál og almenn málfræði*, 25, 7–40.

Jónsson, J.G., & Eythórsson, T. 2005. Variation in subject case marking in Insular Scandinavian. *Nordic Journal of Linguistics*, 28(2), 223–245.

doi:10.1017/S0332586505001435

Jónsson, J.G. & Thórarinsdóttir, R. 2020. Dative objects with novel verbs in Icelandic. In

- A. Pineda & J. Mateu (eds.), *Dative constructions in Romance and beyond*, pp. 297–315. Language Science Press.
- Kapatsinski, V. 2018. *Changing Minds Changing Tools: From Learning Theory to Language Acquisition to Language Change*. MIT Press.
- Kellman, P.J., Gleitman, H. & Spelke, E.S. 1987. Object and observer motion in the perception of objects by infants. *Journal of Experimental Psychology*, 13(4), 586–593. doi:10.1037//0096-1523.13.4.586
- Kibort, A. 2014. Mapping out a construction inventory with (Lexical) Mapping Theory. In M. Butt and T.H. King (eds.), *Proceedings of the LFG14 Conference*, pp. 262–282. CSLI.
- Kiparsky, P. 1997. The rise of positional licensing. In A. van Kemenade & N. Vincent (eds.), *Parameters of Morphosyntactic Change*, pp. 460–494. Cambridge University Press.
- Kiparsky, P. 2001. Structural case in Finnish. *Lingua*, 111(4-7), 315–376
- Koulaguina, E. & Shi, R. 2019. Rule Generalization from Inconsistent Input in Early Infancy. *Language Acquisition*, 26(4), 416–435.  
doi:10.1080/10489223.2019.1572148
- Kratzer, A. 2005. Building resultatives. In C. Maienborn & Angelika Wöllstein-Leisten (eds.), *Events in syntax, semantics, and discourse*, pp. 177–212. Niemeyer.
- Labov, W. 1989. The child as linguistic historian. *Language Variation and Change*, 1(1), 85–97. doi:10.1017/S0954394500000120
- Labov, W. 2001. *Principles of linguistic change: Social factors* (vol. 2). Wiley-Blackwell.
- Landau, B., & Gleitman, L. 1985. *Language and experience: Evidence from the blind child*. Harvard University Press.
- Legate, J.A. 2008. Morphological and Abstract Case. *Linguistic Inquiry*, 39(1), 55–101. doi:10.1162/ling.2008.39.1.55
- Leischner, F.N., Weissenborn, J. & Naigles, L. Universal and Language-Specific Patterns in the Acquisition of Verb Argument Structures in German. *Language Learning and Development*, 12(2), 116–137.  
doi:10.1080/15475441.2015.1052450

- Levin, B. 1993. *English verb classes and alternations: A preliminary investigation*. University of Chicago Press.
- Levin, B. & Rappaport-Hovav, M. 2005. *Argument Realization*. Cambridge University Press. doi:10.1017/CBO9780511610479
- Lidz, J. 2020. Learning, Memory, and Syntactic Bootstrapping: A Meditation. *Topics in Cognitive Science*, 12, 78–90. doi:10.1111/tops.12411
- Lidz, J., Gleitman, H., & Gleitman, L. 2003. Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition*, 87(3), 151–178. doi:10.1016/S0010-0277(02)00230-5
- Lightfoot, D. 1979. *Principles of Diachronic Syntax*. Cambridge University Press.
- Lum, J.A., Conti-Ramsden, G., Morgan, A.T., & Ullman, M.T. 2014. Procedural learning deficits in specific language impairment (SLI): a meta-analysis of serial reaction time task performance. *Cortex*, 51(100), 1–10. doi:10.1016/j.cortex.2013.10.011
- Lum, J.A., Conti-Ramsden, G., Page, D. & Ullman, M.T. 2012. Working, declarative and procedural memory in specific language impairment. *Cortex*, 48(9), 1138–1154. doi:10.1016/j.cortex.2011.06.001
- Maling, J. 2001. Dative: The heterogeneity of the mapping among morphological case, grammatical functions, and thematic roles. *Lingua*, 111(4-7), 419–464.
- Maling, J. 2002. Það rignir þágufalli á Íslandi: Verbs with dative objects in Icelandic. *Íslenskt mál og almenn málfræði*, 24, 31–106. doi:10.1016/S0024-3841(00)00039-5
- Marantz, A. 1991. Case and licensing. In G.F. Westphal, B. Ao, & H.-R. Chae (eds.), *Proceedings of ESCOL '91*, pp. 234–253.
- Marcus, G.F. 2001. *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. The MIT Press.
- Marcus, G.F., Pinker, S., Ullman, M., Hollander, M., Rosen, T.J., Xu, F., & Clahsen, H. 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, 57(4), i–182. doi:10.2307/1166115
- Matsuo, A., Kita, S., Shinya, Y., Wood, G.C., & Naigles, L. 2012. Japanese

- two-year-olds use morphosyntax to learn novel verb meanings. *Journal of Child Language*, 39(3), 637–663. doi:10.1017/S0305000911000213
- McBeath, M.K., Shaffer, D.M., & Kaiser, M.K. 1995. How baseball outfielders determine where to run to catch fly balls. *Science*, 268(5210), 569–573. doi:10.1126/science.7725104
- McCormack, T. & Hoerl, C. 2017. The Development of Temporal Concepts: Learning to Locate Events in Time, *Timing & Time Perception*, 5(3–4), 297–327. doi:10.1163/22134468-00002094
- McFadden, T. 2004. The position of morphological case in the derivation: A study on the Syntax-Morphology Interface. Doctoral Dissertation, University of Pennsylvania.
- McFadden, T. 2018. \*ABA in stem-allomorphy and the emptiness of the nominative”, *Glossa*, 3(1), 8. doi:10.5334/gjgl.373
- McFadden, T. 2020. Case in Germanic. In M. Putnam & B. Page (eds.), *The Cambridge Handbook of Germanic Linguistics*, pp. 282–312. Cambridge University Press. doi:10.1017/9781108378291.014.
- McFadden, T. 2021. The Structural/Inherent Case Distinction and the Implementation of Dependent Case. Abralín ao Vivo talk.
- Miller, K. & Schmitt, C. 2012. Variable Input and the Acquisition of Plural Morphology, *Language Acquisition*, 19(3), 223–261. doi:10.1080/10489223.2012.685026
- Naigles, L. 1990. Children use syntax to learn verb meanings. *Journal of Child Language*, 17(2), 357–374. doi:10.1017/S0305000900013817
- Naigles L. R. 2002. Form is easy, meaning is hard: resolving a paradox in early child language. *Cognition*, 86(2), 157–199. doi:10.1016/s0010-0277(02)00177-4
- Naigles, L.G., Fowler, A., & Helm, A. 1992. Developmental shifts in the construction of verb meanings. *Cognitive Development*, 7(4), 403–427. doi:10.1016/0885-2014(92)80001-V
- Naigles, L.G., & Kako, E.T. 1993. First contact in verb acquisition: defining a role for syntax. *Child Development*, 64(6), 1665–1687. doi:10.2307/1131462
- Newmeyer, F.J. 2010. On comparative concepts and descriptive categories: A reply to Haspelmath. *Language*, 86(3), 688–695.
- Newport, E. 1990. Maturational Constraints on Language Learning. *Cognitive Science*,

- 14, 11–28. doi:10.1207/s15516709cog1401\_2
- Noble, C.H., Rowland, C.F. and Pine, J.M. 2011. Comprehension of Argument Structure and Semantic Roles: Evidence from English-Learning Children and the Forced-Choice Pointing Paradigm. *Cognitive Science*, 35, 963–982.  
doi:10.1111/j.1551-6709.2011.01175.x
- Norris, Mark. 2014. A Theory of Nominal Concord. Doctoral Dissertation, University of California, Santa Cruz.
- Nowenstein, I. 2012. Mig langar sjálfri til þess: Rannsókn á innri breytileika í fallmörkun frumlaga. BA-thesis, University of Iceland.
- Nowenstein, I. 2014. Intra-speaker Variation in Subject Case: Icelandic. *University of Pennsylvania Working Papers in Linguistics*, 20(1), 28.
- Nowenstein, I. 2017. Determining the nature of intra-speaker variation. In H. Thráinsson, C. Heycock, Z. Svabo & H.P. Petersen (eds.), *Syntactic Variation in Insular Scandinavian*, pp. 92–112. John Benjamins.
- Nowenstein, I. & Ingason, A.K. 2021. Featural dynamics in morphosyntactic change. In J.G. Jónsson & Th. Eythórsson (eds.), *Syntactic features and the limits of syntactic change*, pp. 301–320. Oxford University Press.  
doi:10.1093/oso/9780198832584.003.0013
- Nowenstein, I., Ingason, A.K. & Wallenberg, J. 2022. Conditioned Variation: Children Replicate Contrasts, not Parental Variant Rate. *University of Pennsylvania Working Papers in Linguistics*, 28(2), 14.
- Nowenstein, I. & Sigurjónsdóttir, S. 2021. Stafrænt málsambýli íslensku og ensku: Áhrif ensks ílags og málnotkunar á málfærni íslenskra barna. *Ritið*, 21(3), 11–56.  
doi:10.33112/ritid.21.3.2
- Nowenstein, I., Sigurjónsdóttir, S., Yang, C., Ingason, A.K. & Wallenberg, J. 2020. The Meaning of Case: Morphosyntactic Bootstrapping and Icelandic Datives. In M.M. Brown & A. Kohut (eds.), *Proceedings of the 44th annual Boston University Conference on Language Development*, pp. 402–415. Cascadilla Press.
- Noyer, R. 1992. Features, Positions, and Affixes in Autonomous Morphological Structure. Doctoral dissertation, MIT.
- O'Donnell, T. 2015. *Productivity and Reuse in Language*. The MIT Press.

- Opfer, J.E. & Gelman, S.A. 2011. Development of the animate-inanimate distinction. In U. Goswami (ed.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, pp. 213–238. Wiley-Blackwell.
- Parrott, J.K. 2009. Danish vestigial case and the acquisition of Vocabulary in Distributed Morphology. *Biolinguistics*, 3(2–3), 270–304. doi:10.5964/bioling.8711
- Pearl, L., & Sprouse, J. 2021. The acquisition of linking theories: A Tolerance and Sufficiency Principle approach to deriving UTAH and rUTAH. *Language Acquisition*, 28(3), 294–325. doi:10.1080/10489223.2021.1888295
- Perkins, L. 2019. Argument Structure and the Acquisition of Non-Basic Syntax. Doctoral dissertation, University of Maryland.
- Pesetsky, D. 1995. *Zero Syntax: Experiencers and Cascades*. MIT Press.
- Pfaff, A. 2017. Adjectival inflection as diagnostic for structural position: inside and outside the Icelandic definiteness domain. *Journal of Comparative Germanic Linguistics*, 20, 283–322. doi:10.1007/s10828-017-9091-3
- Pinker, S. 1984. *Language Learnability and Language Development*. Harvard University Press.
- Pinker, S. 1989. *Learnability and cognition*. The MIT Press.
- Pinker, S. 1999. *Words and rules*. Basic Books.
- Pinker, S., & Ullman, M.T. 2002. The past and future of the past tense. *Trends in Cognitive Sciences*, 6, 456–463.
- Prasada, S., & Pinker, S. 1993. Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8, 1–56.
- Preminger, O. 2014. *Agreement and its Failures*. The MIT Press.
- Preminger, O. 2020. Taxonomies of case and ontologies of case. Pre-print, lingbuzz/005463.
- Ragnarsdóttir, H. 2018. Orðaforði íslenskra barna frá 4 til 8 ára aldurs: Langtímarannsókn á vaxtarhraða og stöðugleika. *Netla*. doi:10.24270/netla.2018.15
- Ramchand, G. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge University Press. doi:10.1017/CBO9780511486319
- Reisinger, D., Rudinger, R., Ferraro, F., Harman, C., Rawlins, K., Van Durme, B. 2015.

- Semantic Proto-Roles. *Transactions of the Association for Computational Linguistics*, 3, 475–488. doi:10.1162/tacl\_a\_00152
- Repetti-Ludlow, C., & MacKenzie, L. 2022. Regularization in the face of variable input: Children's acquisition of stem-final fricative plurals in American English. *Linguistics Vanguard*. doi:10.1515/lingvan-2021-0064
- Rissman, L. & Lupyan, G. 2022. A dissociation between conceptual prominence and explicit category learning: Evidence from agent and patient event roles. *Journal of Experimental Psychology*, 151(7), 1707–1732. doi:10.1037/xge0001146
- Rissman, L., & Majid, A. 2019. Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, 26(6), 1850–1869. doi:10.3758/s13423-019-01634-5
- Rosen, C.G. 1984. The Interface between Semantic Roles and Initial Grammatical Relations. In D.M. Perlmutter and C.G. Rosen (eds.), *Studies in Relational Grammar 2*, 38–77. The University of Chicago Press.
- Rouveret, A, & Vergnaud, J.-R. 1980. Specifying Reference to the Subject: French Causatives and Conditions on Representations. *Linguistic Inquiry*, 11(1), 97–202.
- Rumelhart, D.E., & McClelland, J. 1986. On learning the past tense of English verbs. In D. Rumelhart, J. McClelland & T.P. Group (eds.), *Parallel Distributed Processing*, 216–271. MIT Press.
- Sankoff, G. 2019. *Language change across the lifespan: Three trajectory types*. *Language*, 95(2), 197–229. doi:10.1353/lan.2019.0029.
- Shablack, H., Becker, M., & Lindquist, K.A. 2020. How do children learn novel emotion words? A study of emotion concept acquisition in preschoolers. *Journal of Experimental Psychology*, 149(8), 1537–1553. doi:10.1037/xge0000727
- Schäfer, F. 2008. *The syntax of (anti-)causatives*. John Benjamins.
- Schätzle (Beck), K. 2018. Dative Subjects: Historical Change Visualized. Doctoral dissertation, University of Konstanz.
- Scherger A.L. 2018. German dative case marking in monolingual and simultaneous bilingual children with and without SLI. *Journal of Communication Disorders*, 75, 87–101. doi:10.1016/j.jcomdis.2018.06.004
- Schmitz, K. 2006. Indirect objects and dative case in monolingual German and bilingual

- German/Romance language acquisition. In D. Hole, A. Meinunger & W. Abraham (eds.): *Datives and other cases: Between argument structure and event structure*, pp. 239–268. John Benjamins.
- Schuler, K., Kodner, J., & Caplan, S. 2020. Abstractions are good for brains and machines: A commentary on Ambridge (2020). *First Language*, 40(5–6), 631–635. doi:10.1177/0142723720906233
- Schuler, K., Yang, C., & Newport, E. 2016. Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so. *Cogsci*, 38, 2321–2326.
- Schuler, K., Yang, C. & Newport, E. 2021. Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient. Pre-print. doi:10.31234/osf.io/utgds
- Schütze, C. 1997. INFL in child and adult language: Agreement, case and licensing. Doctoral dissertation, Massachusetts Institute of Technology.
- Schütze, C. 2001. On the Nature of Default Case. *Syntax*, 4, 205–238.
- Sharon, T. & Wynn, K. 1998. Individuation of Actions from Continuous Motion. *Psychological Science*, 9(5), 357–362. doi:10.1111/1467-9280.00068
- Sigurðardóttir, H.Þ. 2002. Fall í íslensku. Hvernig læra íslensk börn að nota föll? MA-thesis, University of Iceland.
- Sigurðardóttir, S.L. Forthcoming MA-thesis, University of Iceland.
- Sigurðsson, E.F. 2017. Deriving case, agreement and Voice phenomena in syntax. Doctoral dissertation, University of Pennsylvania.
- Sigurðsson, H.Á. 1989. Verbal syntax and case in Icelandic. Doctoral Dissertation, University of Lund.
- Sigurðsson, H.Á. 2012. Case variation: Viruses and star wars. *Nordic Journal of Linguistics*, 35(3), 313–342. doi:10.1017/S033258651300005X
- Sigurjónsdóttir, S. 2005. Máltaka og setningafræði. In H. Thráinsson (ed.): *Setningar, Íslensk tunga III*, pp. 636–655. Almenna bókafélagið.
- Sigurjónsdóttir, S. 2007. The Eva & Fía corpora. Manuscript, University of Iceland.
- Sigurjónsdóttir, S. & Nowenstein, I. 2021. Language acquisition in the digital age: L2

- English input effects on children's L1 Icelandic. *Second Language Research*, 37(4), 697–723. doi:10.1177/02676583211005505
- Slabakova, R. 2002. Recent research on the acquisition of aspect: an embarrassment of riches? *Second Language Research*, 18(2), 172–188. doi:10.1191/0267658302sr202ra
- Smith, J., Durham, M., & Fortune, L. 2007. “Mam, my trousers is fa'in doon!": Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change*, 19(1), 63–99. doi:10.1017/S0954394507070044
- Smith, J., Durham, M., & Fortune, L. 2009. Universal and dialect-specific pathways of acquisition: Caregivers, children, and t/d deletion. *Language Variation and Change*, 21(1), 69–95. doi:10.1017/S0954394509000039
- Sneller, B. & Newport, E. 2019. The Acquisition of Phonological Variation: Evidence from artificial language learning. Talk at BUCLD 44.
- Soh, C., & Yang, C. 2021. Memory Constraints on Cross Situational Word Learning. Pre-print. doi:10.31234/osf.io/ard24
- Starke, M. 2017. Resolving (DAT = ACC) ≠ GEN. *Glossa*, 2(1), 104. doi:10.5334/gjgl.408
- Strömqvist, S., Ragnarsdóttir, H., Engstrand, O., Jonsdóttir, H., Lanza, E., Leiwo, M., ... Toivainen, K. 1995. The Inter-Nordic Study of Language Acquisition. *Nordic Journal of Linguistics*, 18(1), 3–29. doi:10.1017/S0332586500003085
- Svavarsdóttir, Á. 1982. „Þágufallssýki“. *Íslenskt mál og almenn málfræði*, 4, 19–62.
- Svavarsdóttir, Á. 1993. *Beygingarkerfi nafnorða í íslensku*. Institute of Linguistics, University of Iceland.
- Svavarsdóttir, Á. 2013. Þágufallshneigð í sjón og raun: Niðurstöður spurningakannana í samanburði við málnotkun. In H. Thráinsson, Á. Angantýsson & E.F. Sigurðsson (eds.), *Tilbrigði í íslenskri setningagerð I*, pp. 83–109. Málvísindastofnun Háskóla Íslands.
- Svenonius, P. 2002. Icelandic Case and the Structure of Events. *The Journal of Comparative Germanic Linguistics*, 5, 197–225. doi:10.1023/A:1021252206904
- Tal, S. & Arnon, I. 2022. Redundancy can benefit learning: Evidence from word order

- and case marking. *Cognition*, 224, Article 105055.  
doi:10.1016/j.cognition.2022.105055
- Tamminga, M., MacKenzie, L., & Embick, D. 2016. The dynamics of variation in individuals. *Linguistic Variation*, 16(2), 300–336. doi:10.1075/lv.16.2.06tam
- Thráinsson, H. 1979. *On complementation in Icelandic*. Garland.
- Thráinsson, H. 1995. Icelandic. In E. König & J. van der Auwera (eds.), *The Germanic Languages*, 142–189. Routledge.
- Thráinsson, H. 2007. *The Syntax of Icelandic*. Cambridge University Press.  
doi:10.1017/CBO9780511619441
- Thráinsson, H. 2013. Ideal Speakers and Other Speakers: The Case of Dative and Some Other Cases. In B. Fernandez & R. Etxepare (eds.), *Variation in Datives: A Microcomparative Perspective*, pp. 161–188. Oxford University Press. doi:10.1093/acprof:oso/9780199937363.003.0006
- Tomasello, M. 2000. First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11(1–2), 61–82. doi:10.1515/cogl.2001.012
- Trueswell, J.C., Kaufman, D., Hafri, A. & Lidz, J. 2012. Development of Parsing Abilities Interacts with Grammar Learning: Evidence from Tagalog and Kannada. In A.K. Biller, E.Y. Chung & A.E. Kimball (eds.), *Proceedings of the 36th Annual Boston University Conference on Language Development*, pp. 620–632. Cascadilla Press.
- Ullman, M.T. 2001. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30(1), 37–69. doi:10.1023/A:1005204207369
- Ünal, E., Richards, C., Trueswell, J.C. & Papafragou, A. 2021. Representing agents, patients, goals and instruments in causative events: A cross-linguistic investigation of early language and cognition. *Developmental Science*, 24(6), e13116. doi:10.1111/desc.13116
- Van Hout, A. 2005. Imperfect Imperfectives. In P. Kempchinsky & R. Slabakova (eds.), *Aspectual Inquiries*, pp. 317–343. Springer. doi:10.1007/1-4020-3033-9\_14
- Van Valin, R.D. 1991. Another look at Icelandic case marking and grammatical relations. *Natural Language and Linguistic Theory*, 9, 145–194.  
doi:10.1007/BF00133328

- Viðarsson, H.F. 2022. From Old to Modern Icelandic: Dative applicatives and NP/DP configurationality. In B. Egedi & V. Hegedús (eds.), *Functional Heads Across Time: Syntactic Reanalysis and Change*, pp. 210–246. Oxford University Press. doi:10.1093/oso/9780198871538.003.0010
- Wagner, L. 2006. Aspectual Bootstrapping in Language Acquisition: Telicity and Transitivity, *Language Learning and Development*, 2(1), 51–76. doi:10.1207/s15473341lld0201\_3
- Wallenberg, J. 2019. A variational theory of specialization in acquisition and diachrony. In A. Breitbarth, M. Bouzouita, L. Danckaert & M. Farasyn (eds.), *The Determinants of Diachronic Stability*, pp. 245–262. John Benjamins. doi:10.1075/la.254.10wal
- Waxman, S.R., & Booth, A. 2001. Seeing pink elephants: fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology*, 43 (3), 217–242.
- White, A.S., Hacquard, V., & Lidz, J. 2017. Semantic information and the syntax of propositional attitude verbs. *Cognitive Science*, 42(2), 416–456. doi:10.1111/cogs.12512.
- Wiltschko, M. 2008. Person-hierarchy effects without a person-hierarchy. In R. D'Alessandro, S. Fischer & G. Hrafnbjargason (ed.), *Agreement Restrictions*, 281–314. De Gruyter Mouton. doi:10.1515/9783110207835.281
- Wood, J. 2015. *Icelandic Morphosyntax and Argument Structure*. Springer. doi:10.1007/978-3-319-09138-9
- Wood, J. 2016. Reflexive Datives and Argument Structure. Talk at the 22nd Germanic Linguistics Annual Conference (GLAC).
- Wood, J., Barros, M. & Sigurðsson, E.F. 2020. Case mismatching in Icelandic clausal ellipsis. *Journal of Linguistics*, 56(2), 399–439. doi:10.1017/S0022226719000161
- Woods, R., Heim, J. & Wallenberg, J. 2021. Input beyond the Threshold: Explaining Auxiliary Initial Assertions in a British English Early Talker. *WCCFL39*. Cascadilla Proceedings Press.
- Woolford, E. 2006. Lexical Case, Inherent Case, and Argument Structure. *Linguistic Inquiry*, 37(1), 111–130. doi:10.1162/002438906775321175

- Yang, C. 2002. *Knowledge and Learning in Natural Language*. Oxford University Press.
- Yang, C. 2005. On productivity. *Linguistic Variation Yearbook*, 5(1), 265–302.  
doi:10.1075/livy.5.09yan
- Yang, C. 2016. *The Price of Productivity*. The MIT Press.
- Yang, C. 2018. A formalist perspective on language acquisition. *Linguistic Approaches to Bilingualism*, 8(6), 665–706.
- Yang, C. 2020. How to Make the Most out of Very Little. *Topics in Cognitive Science*, 12, 136–152. doi:10.1111/tops.12415
- Yang, C. 2021. A User's Guide to the Tolerance Principle. *lingbuzz/004146*
- Yip, M., Maling, J., & Jackendoff, R. 1987. Case in tiers. *Language*, 63, 217–250.
- Yuan, M. 2022. Case as an Anaphor Agreement Effect: Evidence from Inuktitut. *Linguistic Inquiry*. doi:10.1162/ling\_a\_00443
- Zaenen, A. 1993. Unaccusativity in Dutch: Integrating Syntax and Lexical Semantics. In J. Pustejovsky (ed.), *Semantics and the Lexicon*, pp. 129–161. Kluwer
- Zaenen, A., Maling, J. & Thráinsson, H. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language & Linguistic Theory*, 3, 441–483.  
doi:10.1007/BF0013328
- Zettersten M., & Lupyan G. 2020. Finding categories through words: More nameable features improve category learning. *Cognition*, 196, Article 104135.  
doi:10.1016/j.cognition.2019.104135
- Zettersten, M., Schonberg, C., & Lupyan, G. 2020. What does a radical exemplar view not predict? A commentary on Ambridge (2020). *First Language*, 40(5–6), 636–639. doi:10.1177/0142723720903895
- Zipf, G.K. 1949. *Human Behavior and the Principle of Least Effort*. Addison-Wesley Press.

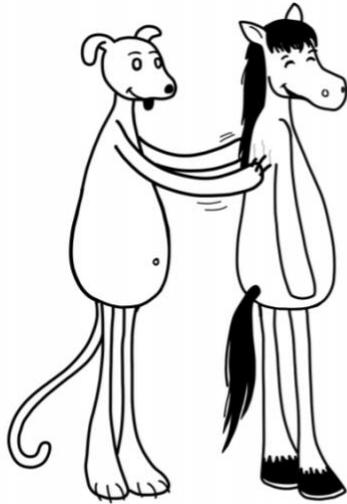
# Appendix A: Comprehension task stimuli

## I. Minimal pair task

(1) 'The dog scratches him':

Hundurinn **klórar** **honum.**  
dog.the.NOM scratches him.DAT

Hundurinn **klórar** **hann.**  
dog.the.NOM scratches him.ACC.

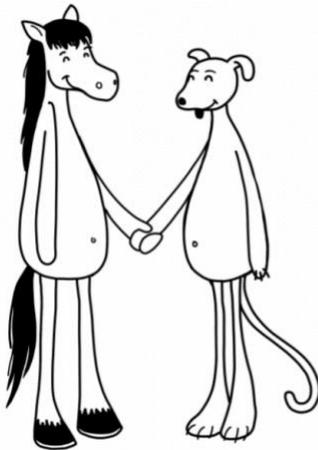


(2) 'I want them to hold hands (nominative subject)/be bored (dative subject)':

Ég vil að...  
I.NOM want that  
þeir leiðist.  
they.NOM bore

[elicits subjunctive and therefore homonymous forms]

Þeim leiðist.  
they.DAT bore



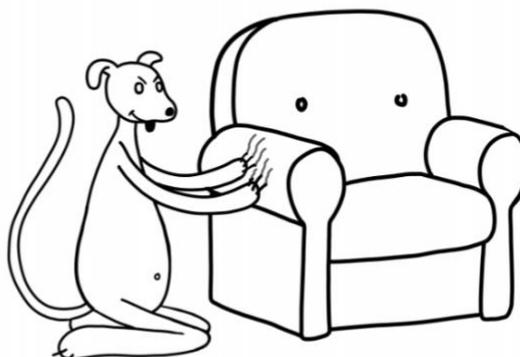
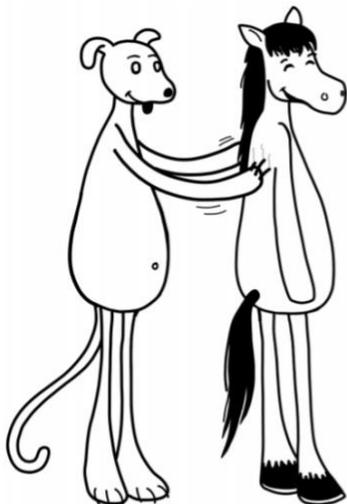
(3) 'The dog scratches him (dative → animate)/it (accusative → inanimate)':

Hundurinn **klórar**  
dog.the.NOM scratches

**honum.**  
him.DAT

Hundurinn **klórar**  
dog.the.NOM scratches

**hann.**  
him.ACC.



(4) 'The dog washes it (accusative → inanimate)/him (dative → animate)':

Hesturinn **þvær**  
dog.the.NOM washes

**hann.**  
him.ACC

Hesturinn **þvær**  
dog.the.NOM washes

**honum.**  
him.DAT.

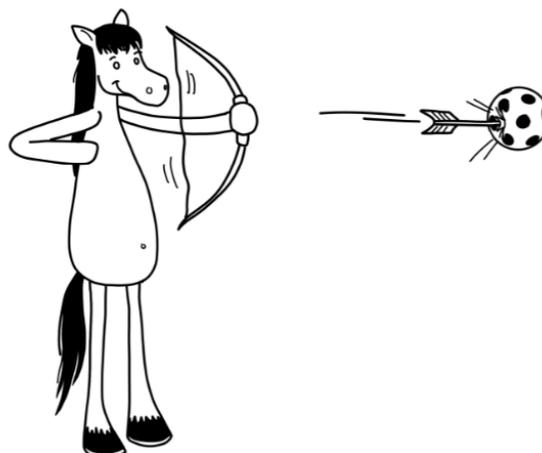


(5) 'The horse shoots the ball':

Hesturinn **skýtur** **boltanum.**  
horse.the.NOM shoots ball.the.DAT.

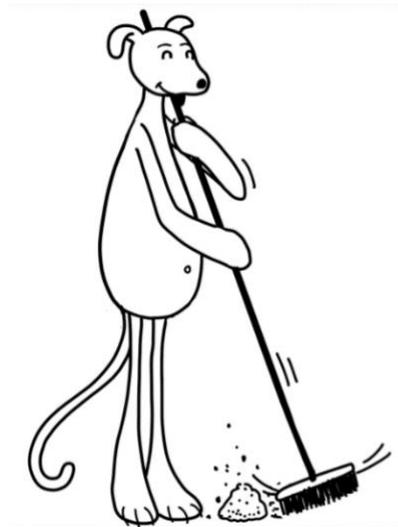


Hesturinn **skýtur** **boltann.**  
horse.the.NOM shoots ball.the.ACC.

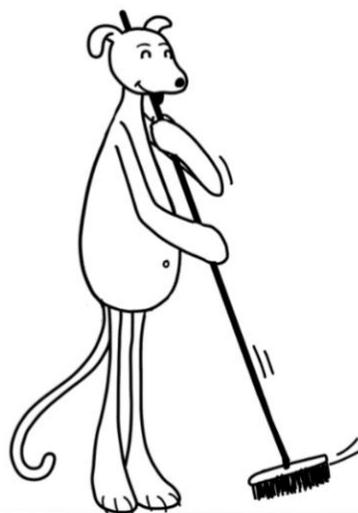


(6) 'The dog sweeps it.':

Hundurinn **sópar** **því.**  
dog.the.NOM sweeps it.DAT



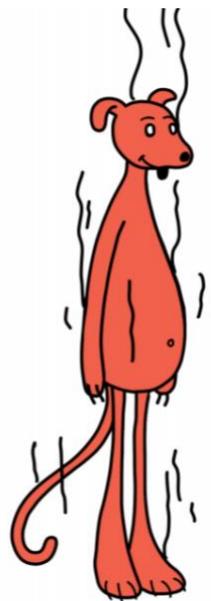
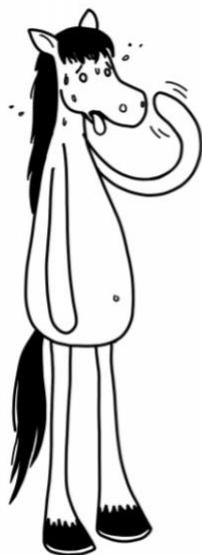
Hundurinn **sópar** **það.**  
dog.the.NOM sweeps it.ACC.



(7) 'He is hot.' (different characters depicted by mistake):

**Honum** er **heitt.**  
he.DAT is hot.NEUT.

**Hann** er **heitur.**  
he.NOM is hot.MASC.



(7) 'The horse is cold.':

**Hann** er **kaldur.**  
he.NOM is cold.MASC.

**Honum** er **kalt.**  
him.DAT is cold.NEUT.



## II. Novel verb task

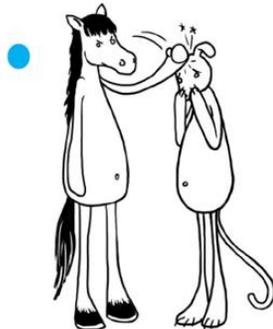
### 1) Subject condition, sentences and picture stimuli

- a. **Hesturinn/hestinum** gorparhann  
horse.the.NOM/DAT gorps he/him.NOM/ACC  
'The horse gorps him'

Novel verbs tested: *merta* and *fípa* (training), *taupa*, *táfa*, *búnsa* and *konda* (test).

- b. **Hesturinn/hestinum** gorpar  
the.horse.NOM/DAT gorps  
'The horse gorps'

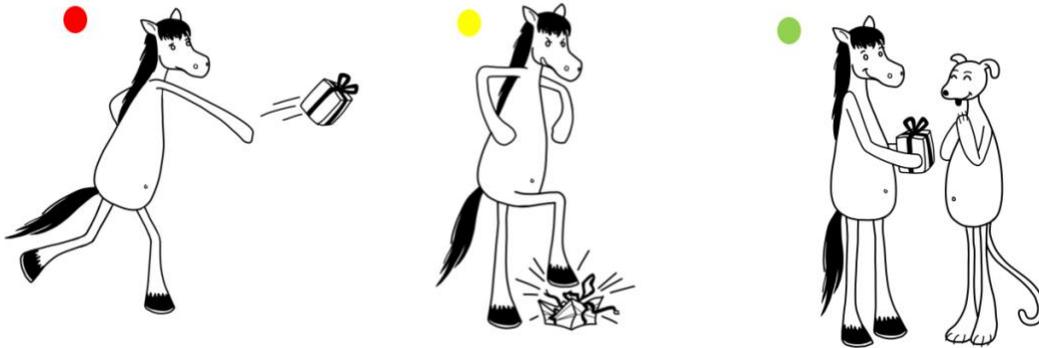
Nonce verbs tested: *glambra* and *lúna* (training), *rikka*, *plosa*, *folsa* and *melba* (test).



2) Object condition, sentences and picture stimuli

Hesturinn	gorpaði	<b>gjöfina/gjöfinni</b>
horse.the.NOM	gorped	gift.the.ACC/DAT
'The horse gorped the gift.'		

Nonce verbs tested: *merta* (training), *goða*, *darga*, *næfa*, *prúna*.



## Appendix B: Production task stimuli

### I. Existing verb task

#### Subject case

- a. **Stelpan/stelpuna/stelpunni** **hlakkar** svo til.  
girl.the.NOM/ACC/DAT anticipates so to  
'The girl is really looking forward to [this].'
- b. **Strákurinn/strákinn/stráknum** **langar** í nammi.  
boy.the.NOM/ACC/DAT wants in candy  
'The boy wants candy.'
- c. **Stelpan/stelpuna/stelpunni** **dreymdi** um þetta.  
girl.the.NOM/ACC/DAT dreams about this  
'The girl dreamt about this.'
- d. **Strákurinn/strákinn/stráknum** **klæjar** alltaf  
boy.the.NOM/ACC/DAT itches always  
'The boy is always itching.'
- e. **Stelpan/stelpuna/stelpunni** **finnst** gaman.  
girl.the.NOM/ACC/DAT finds fun  
'The girl finds [this] fun.'

#### Object case

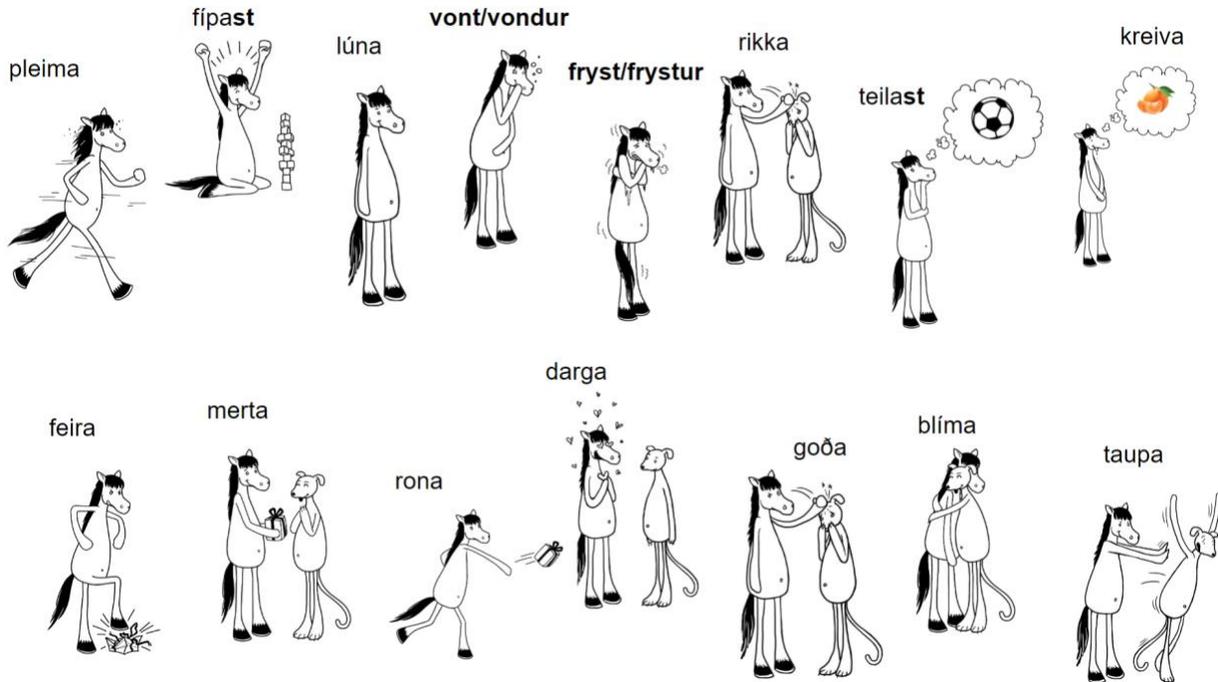
- a. Kennarinn **tók** **bókin/bókina/bókinni.**  
teacher.the.NOM took book.the.NOM/ACC/DAT  
'The teacher took the book'
- b. Mamman **keyrði** **stelpan/stelpuna/stelpunni.**  
mom.the.NOM drove girl.the.NOM/ACC/DAT  
'The mom drove the girl.'
- c. Kennarinn **tók** **bókin/bókina/bókinni.**  
teacher.the.NOM took book.the.NOM/ACC/DAT  
'The teacher took the book'

- d. Mamman **keyrði** **stelpan/stelpuna/stelpunni.**  
mom.the.NOM drove girl.the.NOM/ACC/DAT  
‘The mom drove the girl.’
- e. Maðurinn **kastaði** **boltinn/boltann/boltanum.**  
man.the.NOM threw ball.the.NOM/ACC/DAT  
‘The man threw the ball.’
- f. Stelpan **rústaði** **myndin/myndina/myndinni.**  
girl.the.NOM wrecked picture.the.NOM/ACC/DAT  
‘The girl wrecked the picture.’
- g. Konan **gaf** **stelpan/stelpuna/stelpunni** nammi.  
woman.the.NOM gave girl.the.NOM/ACC/DAT candy.ACC  
‘The woman gave the girl some candy.’
- h. Maðurinn **þurrkaði** **strákurinn/strákinn/stráknum.**  
man.the.NOM dried boy.the.NOM/ACC/DAT  
‘The man dried the boy.’
- i. Strákurinn **saknar** **hún/hana/henni/hennar.**  
boy.the.NOM misses she.NOM/ACC/DAT/GEN.  
‘The boy misses her.’
- j. Strákurinn **spurði** **spurning/spurningu/spurningar.**  
boy.the.NOM asked question.NOM/ACCDAT/GEN.
- k. Strákurinn **gúglaði** **söngvarinn/söngvarann/söngvaranum.**  
boy.the.NOM googled singer.the.NOM/ACC/DAT  
‘The boy googled the singer.’
- l. Stelpan **dánlódaði** **myndin/myndina/myndinni.**  
girl.the.NOM downloaded movie.the.NOM/ACC/DAT  
‘The girl downloaded the movie’
- m. Strákurinn **kópíeistaði** **textinn/textann/textanum.**  
boy.the.NOM copypasted text.the.NOM/ACC/DAT  
‘The boy copypasted the text’

- n. Stelpan **strímaði myndin/myndina/myndinni.**  
 girl.the.NOM streamed movie.the.NOM/ACC/DAT  
 'The girl streamed the movie.'

## II. Novel verb task

Subject case items and pictures in the top row, object case items and pictures in the bottom row.



Subject case

- a. **Hesturinn/hestinn/hestinum** **pleimar.**  
 horse.the.NOM/ACC/DAT NOVEL  
 'The horse Xs.'
- b. **Hesturinn/hestinn/hestinum** **fípaðist** að byggja turninn.  
 horse.the.NOM/ACC/DAT NOVEL to build tower.the  
 'The horse Xed to build the tower.' (novel verb with *-st* morpheme)
- c. **Hesturinn/hestinn/hestinum** **lúnar.**  
 horse.the.NOM/ACC/DAT NOVEL  
 'The horse Xs.'

- d. **Hesturinn/hestinn/hestinum** er **vont/vondur** í maganum.  
horse.the.NOM/ACC/DAT is bad.NEUT./MASC. in stomach.the  
‘The horse is bad in the stomach.’ (semi-novel predicate)
- e. **Hesturinn/hestinn/hestinum** er **fryst/frystur**.  
horse.the.NOM/ACC/DAT is frozen.NEUT./MASC.  
‘The horse is freezed.’ (semi-novel predicate)
- f. **Hesturinn/hestinn/hestinum** **rikkaði** hann.  
horse.the.NOM/ACC/DAT NOVEL him.NOMACC  
‘The horse Xed him.’
- g. **Hesturinn/hestinn/hestinum** **teilst** gaman að spila fótbolta.  
horse.the.NOM/ACC/DAT NOVEL fun to play soccer.  
‘The horse Xs fun to play soccer.’ (novel verb with -st morpheme)
- h. **Hesturinn/hestinn/hestinum** **kreivar** mandarínur.  
horse.the.NOM/ACC/DAT craves mandarins.NOMACC  
‘The horse craves mandarins.’ (recent borrowing)

#### Object case

- a. Hesturinn **feiraði** **gjöfin/gjöfina/gjöfinni**.  
horse.the.NOM NOVEL gift.the.NOM/ACC/DAT  
‘The horse Xed the gift.’
- b. Hesturinn **mertaði** **hundurinn/hundinn/hundinum** **gjöf**.  
horse.the.NOM NOVEL dog.the.NOM/ACC/DAT gift.NOMACC  
‘The horse Xed the dog a gift.’
- c. Hesturinn **ronaði** **gjöfin/gjöfina/gjöfinni**.  
horse.the.NOM NOVEL gift.the.NOM/ACC/DAT  
‘The horse Xed the gift.’
- d. Hesturinn **dargaði** **hundurinn/hundinn/hundinum**  
horse.the.NOM NOVEL dog.the.NOM/ACC/DAT  
‘The horse Xed the dog.’
- e. Hesturinn **goðaði** **hundurinn/hundinn/hundinum**  
horse.the.NOM NOVEL dog.the.NOM/ACC/DAT  
‘The horse Xed the dog.’

- f. Hundurinn blímaði hesturinn/hestinn/hestinum  
horse.the.NOM NOVEL dog.the.NOM/ACC/DAT  
‘The dog Xed the horse.’
- g. Hesturinn taupaði hundurinn/hundinn/hundinum  
horse.the.NOM NOVEL dog.the.NOM/ACC/DAT  
‘The horse Xed the dog.’

## Appendix C: Variation task stimuli

- I. Children
- a. **Strákurinn/strákinn/stráknum** langar í nammi.  
boy.the.NOM/ACC/DAT wants in candy  
‘The boy wants candy.’
- b. **Strákurinn/strákinn/stráknum** klæjar alltaf.  
boy.the.NOM/ACC/DAT itches alltaf.  
‘The boy is always itching.’
- c. **Stelpan/stelpuna/stelpunni** hlakkar svo til.  
girl.the.NOM/ACC/DAT anticipate so to.  
‘The girl looks so forward to (this).’
- d. **Stelpan/stelpuna/stelpunni** dreymdi um þetta.  
girl.the.NOM/ACC/DAT dreamt about this.NOMACC  
‘The girl dreamt about this.’
- e. **Ég/mig/mér** langa(r) í ís.  
I.NOM/ACC/DAT want(s) in icecream  
‘I want ice cream.’
- f. **Hún/hana/henni** langar í ís.  
she.NOM/ACC/DAT wants in icecream  
‘She wants ice cream.’
- g. **Strákarnir/strákana/strákunum** langa(r) út.  
boys.the.NOM/ACC/DAT want(s) out  
‘The boys want to go outside.’

h. **Stelpurnar/stelpunum langa(r) út.**  
 girls.the.NOM/ACC/DAT want(s) out  
 'The girls want to go outside.'

II. Caregivers

a. **Strákurinn/strákinn/stráknum langaði** í súkkulaðiköku.  
 boy.the.NOM/ACC/DAT wanted in chocolate cake.  
 'The boy wanted chocolate cake.'

b. **Strákurinn/strákinn/stráknum er** búið að langa í þetta lengi.  
 boy.the.NOM/ACC/DAT is done to want in this long  
 'The boy has wanted this for a long time.'

c. **Stelpan/stelpuna/stelpunni** klæjaði í hársvörðinn.  
 girl.the.NOM/ACC/DAT itched in scalp.the.  
 'The girl's scalp itched.'

d. **Konan/konuna/konunni** hlakkar mikið til sumarsins.  
 woman.the.NOM/ACC/DAT anticipates much to summer.the.  
 'The woman looks very much forward to the summer.'

e. **Maðurinn/manninn/manninum** dreymir um nýjan bíl.  
 man.the.NOM/ACC/DAT dreams of new car.  
 'The man dreams of a new car.'

f. **Mér** langar að borða pasta.  
 me.DAT wants to eat pasta.  
 'I want to eat pasta.'

g. **Honum** langar að borða pasta.  
 him.DAT wants to eat pasta.  
 'He wants to eat pasta.'

h. **Mér** klæjar þegar fólk talar um lús.  
 me.DAT itches when people talk about lice  
 'I get itchy when people talk about lice.'

i. **Honum** klæjar þegar hann fær ofnæmi.  
 him.DAT itches when he gets allergy

'He itches when he gets an allergy.'

j. **Strákana** langar í þessa bók.  
boys.the.ACC wants in this book.  
'The boys want this book.'

k. **Stelpurnar** langar í þessa bók.  
girls.the.NOMACC wants in this book.  
'The girls want this book.'

l. **Strákana** klæjar greinilega.  
boys.the.ACC itches obviously  
'The boys are clearly itching.'

m. **Stelpurnar** klæjar greinilega.  
girls.the.NOMACC itches obviously  
'The girls are clearly itching.'

# Appendix D: Modeling results

Significance codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1

## I. Comprehension

### Minimal pairs:

```
null_model <- glmer(role ~ (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
977.0	990.8	-485.5	971.0	733

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.08087	0.20069	-0.403	0.687

```
age_model <- glmer(role ~ age + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
973.8	992.2	-482.9	965.8	732

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.9575	0.4244	-2.256	0.0240 *
age	0.1925	0.0823	2.340	0.0193 *

```
case_model <- glmer(role ~ age + case + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
944.8	967.8	-467.4	934.8	731

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.40801	0.45138	-0.904	0.3660
age	0.20137	0.08637	2.332	0.0197 *
non-dative	-1.19938	0.22577	-5.312	1.08e-07 ***

```
position_model <- glmer(role ~ age + case + position + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
937.9	965.5	-462.9	925.9	730

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.19556	0.46064	-0.425	0.67118
age	0.20313	0.08739	2.324	0.02010 *
non-dative	-1.23759	0.23030	-5.374	7.71e-08 ***
subject	-0.56243	0.18919	-2.973	0.00295 **

```
c*p_model <- glmer(role ~ age*case*position + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
934.7	980.8	-457.4	914.7	726

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	0.07972	0.66903	0.119	0.9052
age	0.07222	0.13583	0.532	0.5950
non-dative	-2.10329	0.88308	-2.382	0.0172 *
subject	0.01166	0.98388	0.012	0.9905
age*non-dative	0.32766	0.18005	1.820	0.0688 .
age*subject	0.02375	0.20571	0.115	0.9081
non-dative*subject	-0.42014	1.51413	-0.277	0.7814
age*non-dative*subject	-0.21380	0.31000	-0.690	0.4904

```
anova(null_model, age_model, case_model, position_model, c*p_model)
```

	AIC	BIC	logLik	deviance	Chisq	p value
null	977.02	990.82	-485.51	971.02		
age	973.79	992.20	-482.90	965.79	5.2224	0.022297 *
case	944.83	967.83	-467.41	934.83	30.9647	2.628e-08 ***
pos.	937.88	965.48	-462.94	925.88	8.9519	0.002772 **
c*p	934.74	980.75	-457.37	914.74	11.1368	0.025069 *

### Novel verbs: Subjects

```
null_model <- glmer(experiencer ~ (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1578.6	1593.8	-786.3	1572.6	1164

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.39762	0.05971	-6.66	2.75e-11 ***

```
age_model <- glmer(experiencer ~ age + (1|participant) + (1|item) family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1580.5	1600.7	-786.2	1572.5	1163

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.447904	0.153978	-2.909	0.00363 **
age	0.006626	0.018684	0.355	0.72287

```
t_model <- glmer(experiencer ~ age + transitivity + (1|participant) + (1|sentence), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1522.2	1547.5	-756.1	1512.2	1162

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.000236	0.168279	-0.001	0.999
age2	0.006821	0.019232	0.355	0.723
transitive	-0.942205	0.124056	-7.595	3.08e-14 ***

```
case_model <- glmer(experiencer ~ age + transitivity + case + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1467.8	1498.2	-727.9	1455.8	1161

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	0.472553	0.189526	2.493	0.0127 *
age	0.007051	0.020301	0.347	0.7283
transitive	-0.998880	0.128438	-7.777	7.42e-15 ***
nominative	-0.940903	0.128235	-7.337	2.18e-13 ***

```
t*c_model <- glmer(experiencer ~ age + transitivity*age + case*age + (1|participant) +
(1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1458.1	1498.6	-721.1	1442.1	1159

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.28828	0.28230	-1.021	0.30717
age	0.11025	0.03558	3.099	0.00194 **
transitive	-0.26181	0.32809	-0.798	0.42487
nominative	-0.02024	0.32788	-0.062	0.95078
age*t	-0.10090	0.04137	-2.439	0.01472 *
age*n	-0.12512	0.04139	-3.023	0.00250 **

```
anova(null_model, age_model, transitivity_model, case_model, t*c_model)
```

	AIC	BIC	logLik	deviance	Chisq	p value
null	1578.6	1593.8	-786.29	1572.6		
age	1580.5	1600.7	-786.23	1572.5	0.1257	0.722937
t.	1522.2	1547.5	-756.08	1512.2	60.2999	8.145e-15 ***
case	1467.8	1498.2	-727.92	1455.8	56.3175	6.166e-14 ***
t*case	1458.1	1498.6	-721.07	1442.1	13.6961	0.001062 **

### Novel verbs: Objects

```
null_model <- glmer(motion theme ~ (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
742.7	755.8	-368.3	736.7	578

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.6505	0.2325	-2.797	0.00515 **

```
age_model <- glmer(motion theme ~ age + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
744.3	761.7	-368.1	736.3	577

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.52216	0.31310	-1.668	0.0954 .
age	-0.01697	0.02785	-0.609	0.5423

```
case_model <- glmer(motion theme ~ age + case + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
744.1	765.9	-367.0	734.1	576

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.82489	0.32805	-2.515	0.0119 *
age	-0.01696	0.02786	-0.609	0.5426
dative	0.60519	0.35285	1.715	0.0863 .

```
c*a_model <- glmer(motion theme ~ age + case*age + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
728.9	755.1	-358.4	716.9	575

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	0.15130	0.40738	0.371	0.710341
age	-0.15250	0.04546	-3.355	0.000794 ***
dative	-1.16220	0.56124	-2.071	0.038379 *
dative*age	0.23933	0.05923	4.041	5.33e-05 ***

```
anova(null_model, age_model, case_model, c*a_model)
```

	AIC	BIC	logLik	deviance	Chisq	p value
null	742.66	755.75	-368.33	736.66		
age	744.29	761.75	-368.14	736.29	0.3727	0.5415
case	744.06	765.89	-367.03	734.06	2.2247	0.1358
c*a	728.89	755.08	-358.44	716.89	17.1755	3.408e-05 ***

## II. Production

```
null_model <- glmer(dative ~ (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1807.6	1823.4	-900.8	1801.6	1429

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.2300	0.2403	-0.957	0.339

```
age_model <- glmer(dative ~ age + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1809.6	1830.6	-900.8	1801.6	1428

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.2281225	0.3488826	-0.654	0.513
age	-0.0002023	0.0275040	-0.007	0.994

```
position_model <- glmer(dative ~ age + position + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1793.6	1819.9	-891.8	1783.6	1427

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	0.391645	0.372760	1.051	0.293
age	0.001001	0.027807	0.036	0.971
subject	-1.195710	0.294564	-4.059	4.92e-05 ***

```
context_model <- glmer(response ~ age + position + context + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1770.9	1802.5	-879.4	1758.9	1426

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	1.075088	0.312290	3.443	0.000576 ***
age	0.001731	0.027840	0.062	0.950417
subject	-1.409179	0.188875	-7.461	8.59e-14 ***
atypical	-1.596352	0.211734	-7.539	4.72e-14 ***

```
p*c_modelp <- glmer(dative ~ age*position*context + (1|participant) + (1|item), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
1750.7	1803.4	-865.4	1730.7	1422

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.19354	0.48471	-0.399	0.6897
age	0.12198	0.05029	2.425	0.0153 *
subject	-0.53366	0.55074	-0.969	0.3325
atypical	0.81194	0.58411	1.390	0.1645
age*subject	-0.06882	0.05620	-1.224	0.2208
age*atypical	-0.23561	0.06019	-3.914	9.06e-05 ***
subject*atypical	-0.47680	1.03101	-0.462	0.6437
age*subj.*atyp.	-0.03581	0.11554	-0.310	0.7566

anova(null\_model, age\_model, position\_model, context\_model, p\*c\_model)

	AIC	BIC	logLik	deviance	Chisq	p value
null	1807.5	1823.4	-900.78	1801.5		
age	1809.5	1830.6	-900.78	1801.5	0.0001	0.9942
p.	1793.6	1819.9	-891.80	1783.6	17.9574	2.259e-05 ***
c.	1770.9	1802.5	-879.43	1758.9	24.7293	6.597e-07 ***
p*c	1750.7	1803.4	-865.36	1730.7	28.1408	1.168e-05 ***

### III. Variation

#### Child and caregiver rates

lm(formula = child ~ caregiver \* age)

Coefficients:

	Estimate	Std. Error	t value	p value
(Intercept)	0.39533	0.19053	2.075	0.0412 *
caregiver	-0.20374	0.25494	-0.799	0.4265
age	-0.01885	0.02197	-0.858	0.3936
caregiver:age	0.03329	0.02867	1.161	0.2490

Residual standard error: 0.1863 on 80 degrees of freedom

Multiple R-squared: 0.03631, Adjusted R-squared: 0.0001735

F-statistic: 1.005 on 3 and 80 DF, p-value: 0.3951

#### PSR: Parents

null\_model <- glmer(dative ~ (1| participant), family = binomial)

AIC	BIC	logLik	deviance	df.resid
398.5	406.4	-197.2	394.5	376

psr\_model <- glmer(dative ~ (1| participant) + person, family = binomial)

AIC	BIC	logLik	deviance	df.resid
362.8	374.6	-178.4	356.8	375

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-2.9717	0.4444	-6.687	2.28e-11 ***
3.p.sg.	1.9235	0.3577	5.377	7.56e-08 ***

anova(null\_model, psr\_model)

	AIC	BIC	logLik	deviance	Chisq	p value
null	398.49	406.36	-197.24	394.49		
psr	362.84	374.64	-178.42	356.84	37.647	8.476e-10 ***

### PSR: Children

null\_model <- glmer(dative ~ (1| participant), family = binomial)

AIC	BIC	logLik	deviance	df.resid
278.7	285.3	-137.4	274.7	197

psr\_model <- glmer(dative ~ (1| participant) + pronoun, family = binomial)

AIC	BIC	logLik	deviance	df.resid
256.8	266.7	-125.4	250.8	196

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.5821	0.2354	-2.472	0.0134 *
3.p.sg.	1.5485	0.3657	4.235	2.29e-05 ***

psr+age\_model <- glmer(dative ~ (1| participant) + pronoun + age, family = binomial)

AIC	BIC	logLik	deviance	df.resid
251.7	264.9	-121.9	243.7	195

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	0.75183	0.54982	1.367	0.1715
3.p.sg.	1.57372	0.36467	4.316	1.59e-05 ***
age	-0.15512	0.06129	-2.531	0.0114 *

psr\*age\_model <- glmer(dative ~ (1| participant) + pronoun\*age, family = binomial)

AIC	BIC	logLik	deviance	df.resid
253.6	270.1	-121.8	243.6	194

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	0.60707	0.70257	0.864	0.3876
3.p.sg.	1.90078	1.06613	1.783	0.0746 .
age	-0.13782	0.08027	-1.717	0.0860 .
3.p.sg*age	-0.03719	0.11325	-0.328	0.7426

---

anova(null\_model, psr\_model, psr+age\_model, psr\*age\_model)

	AIC	BIC	logLik	deviance	Chisq	p value
null	278.74	285.33	-137.37	274.74		
psr	256.79	266.67	-125.39	250.79	23.9546	9.864e-07 ***
psr+a	251.72	264.89	-121.86	243.72	7.0702	0.007838 **
psr*a	253.61	270.07	-121.80	243.61	0.1080	0.742390

### Syncretism: Caregivers

null\_model <- glmer(dative ~ (1| participant) + (1|item), family = binomial)

AIC	BIC	logLik	deviance	df.resid
458.7	470.5	-226.3	452.7	375

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.9596	0.1393	-6.888	5.67e-12 ***

syn\_model <- glmer(dative ~ (1| participant) + (1|item) + syn, family = binomial)

AIC	BIC	logLik	deviance	df.resid
404.6	420.3	-198.3	396.6	374

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-2.2296	0.2900	-7.687	1.50e-14 ***
syncretism	2.0028	0.3083	6.496	8.27e-11 ***

anova(null\_model, syn\_model)

	AIC	BIC	logLik	deviance	Chisq	p value
null	458.69	470.49	-226.34	452.69		
syn	404.58	420.32	-198.29	396.58	56.107	6.864e-14 ***

## Syncretism: Children

```
null_model <- glmer(dative ~ (1| participant), family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
241.3	247.8	-118.7	237.3	192

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	0.918753	0.003483	263.8	<2e-16 ***

```
syn_model <- glmer(dative ~ (1| participant) + syn, family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
237.2	247.0	-115.6	231.2	191

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	1.3958	0.3131	4.458	8.26e-06 ***
syncretism	-0.8366	0.3522	-2.376	0.0175 *

```
syn+age_model <- glmer(dative ~ (1| participant) + syn + age, family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
235.9	249.0	-114.0	227.9	190

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	0.38652	0.61045	0.633	0.5266
syncretism	-0.85137	0.35509	-2.398	0.0165 *
age	0.11816	0.06735	1.754	0.0794 .

```
syn*age_modelc <- glmer(dative ~ (1| participant) + syn*age, family = binomial)
```

AIC	BIC	logLik	deviance	df.resid
234.6	251.0	-112.3	224.6	189

Fixed effects:

	Estimate	Std. Error	z value	p value
(Intercept)	-0.6487	0.8437	-0.769	0.4420
syncretism	0.9883	1.0793	0.916	0.3598
age	0.2478	0.1034	2.397	0.0165 *
syncretism*age	-0.2218	0.1253	-1.770	0.0767 .

```
anova(null_model, syn_model, syn+age_model, syn*age_model)
```

	AIC	BIC	logLik	deviance	Chisq	p value
null	241.30	247.84	-118.65	237.30		
syn	237.21	247.02	-115.61	231.21	6.0896	0.01360 *
syn+a	235.92	248.99	-113.96	227.92	3.2977	0.06938 .
syn*a	234.63	250.97	-112.31	224.63	3.2911	0.06966 .

## Appendix E: IceCASE corpus

The IceCASE corpus was computed with frequency information based on data from GJEUM, the Einarsdóttir et al. 2019 language sample corpus, the Sigurjónsdóttir (2007) corpora and the Icelandic data available in CHILDES (Strömqvist et al. 1995).

Case annotation: N = nominative, A = accusative, D = dative, G = genitive.

Role annotation: A = agent, P = patient, T = theme, G = goal, E = experiencer

Animacy annotation: A = animate, I = inanimate

Parentheses denote optionality, underscores separate arguments.

Two annotations appearing simultaneously signals optionality.

Motion applies to the direct object argument.

Verbs with dative arguments are bolded.

Homonyms are split by letters.

verb	translation	rank	freq	case	animacy	role	reflexive	motion
gera	do	1	3170	N_(A)	A_(I)	A_(P)	none	0
koma	<b>come</b>	<b>2</b>	<b>2893</b>	<b>N</b>	<b>AI</b>	<b>AT</b>	<b>D</b>	<b>1</b>
segja	<b>say</b>	3	<b>2276</b>	<b>N_(D)_A</b>	<b>A_(A)_I</b>	<b>A_(G)_T</b>	<b>none</b>	<b>0</b>
sjá	see	4	1990	N_(A)	A_(AI)	E_(T)	none	0
geta	can	5	1434	N_(A)	AI_(I)	T_(T)	<b>none</b>	0
vita	know	6	1376	N_(A)	A_(I)	E_(T)	none	0
fá	<b>get</b>	7	<b>1211</b>	<b>N_(A)</b>	<b>AI_AI</b>	<b>AG_(T)</b>	<b>D</b>	<b>0</b>
vilja	want	8	1168	N_(A)	A_AI	E_(T)	<b>none</b>	0
halda	think	9	1104	N_(A)	A_(I)	E_(T)	none	0
mega	can	<b>10</b>	936	N_(A)	AI_I	T_(T)	<b>none</b>	0
leika	<b>play</b>	11	<b>897</b>	<b>N</b>	<b>A</b>	<b>A</b>	<b>D</b>	<b>0</b>

<b>taka</b>	<b>take</b>	<b>12</b>	<b>856</b>	<b>N_(A)</b>	<b>A_(I)</b>	<b>A_(T)</b>	none	<b>0</b>
<b>þurfa</b>	need	13	842	N_(A)	AI_AI	ET_(T)	<b>none</b>	0
<b>finnast</b>	<b>feel</b>	<b>14</b>	<b>807</b>	<b>D_(N)</b>	<b>A_(I)</b>	<b>E_(T)</b>	none	<b>0</b>
<b>heyra</b>	hear	15	743	N_(A)	A_(I)	E_(T)	<b>none</b>	0
<b>setjast</b>	sit down	<b>16</b>	741	N	A	A	none	0
<b>muna</b>	remember	17	712	N_(A)	A_(I)	E_(T)	<b>none</b>	0
<b>detta</b>	fall	<b>18</b>	708	N	AI	T	none	1
<b>borða</b>	eat	19	656	N_(A)	A_(I)	A_(P)	<b>none</b>	0
<b>tala</b>	speak	<b>20</b>	485	N	A	A	none	0
<b>lesa</b>	read	21	432	N_(A)	A_(I)	A_(T)	<b>none</b>	0
<b>ná</b>	<b>reach, get</b>	<b>22</b>	<b>423</b>	<b>N_(D)</b>	<b>A_(AI)</b>	<b>A_(T)</b>	<b>D</b>	<b>0</b>
<b>skoða</b>	examine	23	399	N_A	A_AI	A_T	none	0
<b>kunna</b>	know	<b>24</b>	330	N_(A)	A_(I)	A_(T)	<b>none</b>	0
<b>leita</b>	<b>search</b>	<b>25</b>	<b>329</b>	<b>N</b>	<b>A</b>	<b>T</b>	<b>D</b>	<b>0</b>
<b>sofa</b>	sleep	<b>26</b>	310	N	A	T	none	0
<b>labba</b>	walk	27	310	N	A	A	<b>none</b>	1
<b>gefa</b>	<b>give</b>	<b>28</b>	<b>307</b>	<b>N_(D)_A</b>	<b>A_(A)_AI</b>	<b>A_(G)_T</b>	none	<b>0</b>
<b>horfa</b>	watch	29	302	N	A	A	<b>none</b>	0
<b>kaupa</b>	<b>buy</b>	<b>30</b>	<b>296</b>	<b>N_A</b>	<b>A_I</b>	<b>A_T</b>	<b>D</b>	<b>0</b>
<b>sitja</b>	sit	31	292	N	A	A	none	0
<b>kíkja</b>	peek	<b>32</b>	289	N	A	A	<b>none</b>	0
<b>bíða</b>	wait	33	271	N	A	A	none	0
<b>klæða</b>	dress	<b>34</b>	265	N_A	A_AI	A_P	<b>none</b>	0
<b>lúlla</b>	sleep	35	264	N	A	T	none	0
<b>kalla</b>	call	<b>36</b>	256	N_(A)	A_(AI)	A_(T)	<b>none</b>	0
<b>langa</b>	<b>want</b>	<b>37</b>	<b>253</b>	<b>AD_A</b>	<b>A_(I)</b>	<b>E_(T)</b>	none	<b>0</b>
<b>búa</b>	<b>make</b>	<b>38</b>	<b>245</b>	<b>N_A</b>	<b>A_I</b>	<b>A_P</b>	<b>D</b>	<b>0</b>

<b>passa</b>	care for	39	228	N_(A)	A_(AI)	A_(T)	<b>none</b>	0
<b>syngja</b>	sing	<b>40</b>	223	N_(A)	A_(I)	A_(T)	none	0
<b>byrja</b>	start	41	213	N_(A)	AI_(I)	AT_(P)	<b>none</b>	0
<b>reyna</b>	try	<b>42</b>	212	N	A	A	none	0
<b>sýna</b>	<b>show</b>	43	<b>208</b>	<b>N_(D)_ (A)</b>	<b>A_(A)_ (I)</b>	<b>A_(G)_ (T)</b>	<b>none</b>	<b>0</b>
<b>keyra</b>	drive	<b>44</b>	199	N_(A)	A_(AI)	A_(T)	none	0
<b>skipta</b>	<b>change</b>	45	<b>191</b>	<b>N_(D)</b>	<b>A_(I)</b>	<b>A_(PT)</b>	<b>none</b>	<b>0</b>
<b>hlaupa</b>	run	<b>46</b>	182	N	A	A	none	1
<b>nota</b>	use	47	174	N_A	A_I	A_T	<b>none</b>	0
<b>loka</b>	<b>close</b>	<b>48</b>	<b>173</b>	<b>N_(D)</b>	<b>A_(I)</b>	<b>A_(P)</b>	none	<b>0</b>
<b>hjálpá</b>	<b>help</b>	49	<b>172</b>	<b>N_(D)</b>	<b>AI_(A)</b>	<b>AT_(G)</b>	<b>none</b>	<b>0</b>
<b>hoppa</b>	jump	<b>50</b>	168	N	A	A	none	1
<b>læra</b>	learn	51	167	N_(A)	A_(I)	A	<b>none</b>	0
<b>standa</b>	stand	<b>52</b>	166	N	AI	AT	none	0
<b>lita</b>	color	53	164	N_(A)	A_(I)	A_(P)	<b>none</b>	0
<b>drekka</b>	drink	<b>54</b>	163	N_A	A_I	A_P	none	0
<b>meiða</b>	hurt	55	153	N_A	AI_A	ATE_P	<b>none</b>	0
<b>hætta</b>	<b>quit</b>	<b>56</b>	<b>148</b>	<b>N_(D)</b>	<b>A_(I)</b>	<b>A_(T)</b>	none	<b>0</b>
<b>opna</b>	open	57	147	N_(A)	A_(I)	A_(P)	<b>none</b>	0
<b>gleyma</b>	<b>forget</b>	<b>58</b>	<b>139</b>	<b>N_(D)</b>	<b>A_(AI)</b>	<b>E_(T)</b>	<b>D</b>	<b>0</b>
<b>æfa</b>	practice	59	137	N_A	A_I	A_T	none	0
<b>finna</b>	<b>find</b>	<b>60</b>	<b>136</b>	<b>N_A</b>	<b>A_AI</b>	<b>E_T</b>	<b>D</b>	<b>0</b>
<b>vinna</b>	win	61	136	N_(A)	A_(AI)	A_(PT)	<b>none</b>	0
<b>spila</b>	play (music)	<b>62</b>	133	N_(A)	A_(I)	A_(T)	none	0
<b>hitta</b>	meet	63	128	N_A	A_A	A_T	<b>none</b>	0
<b>laga</b>	repair	<b>64</b>	120	N_A	AI_AI	AT_P	none	0

geyma	keep	65	119	N_A	A_AI	A_T	none	0
snúa	twist	66	118	N_(D)	AI_AI	AT_T	none	0
skrifa	write	67	115	N_(D)_A	A_(A)_I	A_(G)_T	none	0
klippa	clip	68	115	N_A	A_I	A_P	none	0
vakna	wake up	69	114	N	A	T	none	0
moka	dig	70	113	N_(AD)	A_(I)	A_(PT)	none	0
byggja	build	71	112	N_(A)	A_(I)	A_(P)	D	0
vanta	need	72	111	AD_A	A_I	E_T	none	0
prófa	test	73	109	N_(A)	A_I	A_T	none	0
elta	follow, chase	74	107	N_A	A_AI	A_T	none	1
renna	slip	75	106	N_(D)	AI_(I)	AT_(T)	D	1
veiða	hunt	76	104	N_(A)	A_(A)	A_(P)	none	0
teikna	draw	77	102	N_(A)	A_(I)	A_(T)	none	0
hringja	call, ring	78	101	N	AI	AT	none	0
klifra	climb	79	97	N	A	A	none	1
liggja	lie (down)	80	96	N	AI	T	none	0
hjóla	bike	81	96	N	A	A	none	1
bíta	bite	82	96	N_(A)	A_(AI)	A_(P)	none	0
ýta	push	83	95	N_(D)	A_(AI)	A_(T)	none	0
pissa	pee	84	90	N_(D)	A_(I)	A_(T)	none	0
slökkva	turn off	85	89	N_(A)	A_(I)	A_(P)	none	0
hlusta	listen	86	89	N	A	A	none	0
henda	throw	87	88	N_D	A_AI	A_T	none	0
búaB	live	88	87	N	A	A	none	0
fljúga	fly	89	87	N_(D)	AI_(I)	AT_(T)	none	1
raka	shave	90	79	N_A	A_AI	A_P	none	0

festu	fix	91	79	N_A	A_AI	A_P	none	0
kúka	poop	92	78	N_(D)	A_(I)	A_(T)	none	0
gráta	cry	93	78	N	A	A	none	0
synda	swim	94	76	N_(A)	A_(I)	A_(T)	none	1
nenna	be bothered	95	76	N_(D)	A_(I)	E_(T)	none	0
kyssa	kiss	96	76	N_A	A_AI	A_P	none	0
haldaB	hold	97	72	N_(D)	AI_(AI)	AT_(T)	none	0
heimsækja	visit	98	67	N_A	A_AI	A_G	none	1
leyfa	allow	99	66	N_(D)_(A)	A_(A)_(I)	A_(G)_(T)	none	0
velja	choose	100	65	N_A	A_AI	A_T	D	0
sækja	fetch	101	65	N_A	A_AI	A_T	none	1
spyrja	ask	102	65	N_(A)_(G)	A_(AI)_(I)	A_(G)_(T)	none	0
róla	swing	103	65	N	A	A	D	1
rífa	rip	104	64	N_A	AI_AI	A_P	none	0
skilja	understand	105	63	N_(A)	A_(AI)	E_(T)	none	0
brotna	break down	106	63	N	AI	T	none	0
þýða	mean	107	62	N_(N)	AI_(AI)	T_(T)	none	0
þvo	wash	108	62	N_(AD)	A_(AI)	A_(PG)	none	0
kveikja	turn on	109	62	N_(A)	A_(I)	A_(P)	none	0
eigaB	own	110	61	N_A	A_AI	A_T	none	0
gá	check	111	61	N	A	A	none	0
sofna	fall asleep	112	60	N	A	T	none	0
missa	drop	113	60	N_(A)	A_(AI)	A_(T)	none	0
leggja	lay	114	60	N_(A)	A_(AI)	A_(T)	none	0
dansa	dance	115	60	N	A_(I)	A_(T)	none	1
biðja	ask	116	60	N_A	A_A	A_T	none	0

skjóta	shoot	117	58	N_(AD)	A_(AI)	A_(PT)	none	0
bera	carry	<b>118</b>	58	N_A	A_AI	A_T	none	1
týna	lose	119	57	N_D	A_AI	A_T	none	0
mála	paint	<b>120</b>	56	N_(A)	A_(I)	A_(P)	none	0
stela	steal	121	55	N_(D)	A_(I)	A_(T)	none	0
klára	finish	<b>122</b>	54	N_(A)	A_(I)	A_(P)	none	0
brjóta	break	123	53	N_A	A_(I)	A_(P)	none	0
skora	score	<b>124</b>	52	N_(A)	A_(I)	A_(P)	none	0
sigla	<b>sail</b>	125	<b>52</b>	<b>N_(D)</b>	<b>A_(I)</b>	<b>A_(T)</b>	none	<b>0</b>
fletta	<b>turn page</b>	<b>126</b>	<b>52</b>	<b>N_(D)</b>	<b>A_(I)</b>	<b>A_(T)</b>	none	<b>0</b>
draga	drag	127	52	N_A	AI_AI	A_T	none	0
bursta	brush	<b>128</b>	50	N_(A)	A_(AI)	A_(P)	none	0
fela	hide	129	49	N_A	AI_AI	AT_T	none	0
baka	bake	<b>130</b>	49	N_(A)	A_(I)	A_(P)	none	0
fikta	fiddle	131	48	N	A	A	none	0
klappa	<b>pet</b>	<b>132</b>	<b>47</b>	<b>N_D</b>	<b>A_AI</b>	<b>A_G</b>	none	<b>0</b>
líta	look	133	47	N	AI	AT	none	0
flytja	move (house)	<b>134</b>	47	N_(A)	A_(AI)	A_(T)	none	1
bjarga	<b>save</b>	135	<b>47</b>	<b>N_D</b>	<b>AI_AI</b>	<b>A_G</b>	none	<b>0</b>
takast	<b>succeed</b>	<b>136</b>	<b>46</b>	<b>D_N</b>	<b>A_I</b>	<b>E_T</b>	none	<b>0</b>
smíða	build	137	46	N_(A)	A_(I)	A_(P)	none	0
sleppa	<b>let go</b>	<b>138</b>	<b>46</b>	<b>N_D</b>	<b>A_AI</b>	<b>A_T</b>	none	<b>0</b>
lenda	land	139	46	N	AI	T	none	1
færa	move	<b>140</b>	45	N_A	A_AI	A_T	none	0
stökkva	jump	141	44	N	A	A	none	1
hugsa	think	<b>142</b>	44	N_(A)	A_(I)	E_(T)	none	0

skiljaB	leave	143	42	N_A	A_(AI)	A_(T)	none	1
þurrka	dry	144	42	N_AD	A_AI	A_PG	none	0
éta	eat	145	42	N_(A)	A_(AI)	A_(P)	none	0
breyta	change	146	42	N_D	AI_(AI)	AT_(T)	none	0
greiða	comb	147	41	N_AD	A_AI	A_PG	none	0
telja	count	148	40	N_(A)	A_(AI)	A_(T)	none	0
passaB	fit	149	39	N_(D)	AI_(AI)	T_(T)	none	0
hverfa	disappear	150	39	N	AI	T	none	1
elda	cook	151	39	N_(A)	A_(I)	A_(P)	none	0
safna	collect	152	38	N_D	A_I	A_T	none	0
meina	mean	153	38	N_A	A_AI	A_T	none	0
setja	put	154	37	N_A	A_AI	A_T	none	0
skemma	damage, ruin	155	37	N_(A)	A_(AI)	A_(P)	none	0
lána	lend	156	37	N_(D)_ (A)	A_(A)_ (AI)	A_(G)_ (T)	none	0
kenna	teach	157	37	N_(D)_ (A)	A_(A)_ (I)	A_(G)_ (T)	none	0
berja	hit	158	37	N_(A)	A_(A)	A_(P)	none	0
þekkja	know	159	37	N_A	A_AI	E_T	none	0
ganga frá	put away	160	37	N	A	A	none	0
hrista	shake	161	36	N_A	A_(AI)	A_(T)	none	0
hella	pour	162	35	N_D	A_(I)	A_(T)	none	0
gista	sleep over	163	35	N	A	A	none	0
sparka	kick	164	34	N_(D)	A_(I)	A_(T)	none	0
skafa	scrape	165	34	N_(A)	A_(I)	A_(P)	none	0
hlæja	laugh	166	33	N	A	A	none	0
athuga	check	167	33	N_(A)	A_(I)	A_(T)	none	0
slá	hit	168	32	N_AD	A_(AI)	A_(PT)	none	0

perla	bead	169	32	N_(A)	A_(I)	A_(P)	none	0
stinga	sting	170	31	N_AD	A_(AI)	A_(PT)	none	0
steikja	fry	171	31	N_A	A_I	A_P	none	0
sjúga	suck	172	31	N_A	A_I	A_P	none	0
mæla	measure	173	31	N_A	A_AI	A_T	none	0
toga	pull	174	30	N_(A)	AI_(AI)	A_(PT)	none	0
kanna	check	175	30	N_A	A_I	A_T	none	0
eyðileggja	ruin	176	30	N_A	AI_I	A_P	none	0
drepa	kill	177	30	N_A	AI_A	A_P	none	0
rugla	confuse, mix	178	29	N_(AD)	A_(AI)	A_(PT)	none	0
prumpa	fart	179	29	N	A	A	none	0
prjóna	knit	180	29	N_(A)	A_(I)	A_(P)	none	0
kasta	throw	181	29	N_D	A_AI	A_T	none	0
elska	love	182	29	N_A	A_AI	E_T	none	0
reka	drive away	183	28	N_A	A_AI	A_T	none	0
deyja	die	184	28	N	A	T	none	0
velta	roll	185	27	N_(D)	AI_(AI)	AT_(T)	none	1
skottast	run around	186	27	N	A	A	none	1
púsla	do puzzle	187	27	N_(AD)	A_(I)	A_(PT)	none	0
hanga	hang	188	27	N	AI	AT	none	0
fylgja	follow	189	27	N_(D)	AI_(AI)	T_(T)	none	1
drífa	hurry	190	27	N_A	A_A	A_T	none	0
binda	bind	191	27	N_A	A_AI	A_P	none	0
sýnast	seem	192	26	D_N	A_I	E_T	none	0
stríða	pick on	193	26	N_D	A_A	A_P	none	0
spjalla	chat	194	26	N	A	A	none	0

<b>læðast</b>	sneak	195	26	N	A	A	<b>none</b>	1
<b>kubba</b>	play with blocks	<b>196</b>	26	N_(A)	A_(I)	A_(P)	none	0
<b>öskra</b>	scream	197	25	N	A	A	<b>none</b>	0
<b>lyfta</b>	<b>lift</b>	<b>198</b>	<b>25</b>	<b>N_D</b>	<b>A_AI</b>	<b>A_T</b>	none	<b>0</b>
<b>lemja</b>	hit	199	25	N_A	A_AI	A_P	<b>none</b>	0
<b>keppa</b>	compete	<b>200</b>	25	N	A	A	none	0
<b>bora</b>	drill	201	25	N_(A)	A_(I)	A_(P)	<b>none</b>	0
<b>beygja</b>	bend	<b>202</b>	25	N_(A)	AI_(AI)	AT_(P)	none	1
<b>ákveða</b>	decide	203	24	N_(A)	A_(I)	A_(T)	<b>none</b>	0
<b>stíga</b>	step	<b>204</b>	24	N	A	A	none	0
<b>snýta</b>	<b>blow (nose)</b>	205	<b>24</b>	<b>N_D</b>	<b>A_A</b>	<b>A_P</b>	<b>none</b>	<b>0</b>
<b>sleikja</b>	lick	<b>206</b>	24	N_A	A_AI	A_P	none	0
<b>skola</b>	rinse	207	24	N_A	A_AI	A_P	<b>none</b>	0
sauma	sew	<b>208</b>	24	N_(A)	A_I	A_P	none	0
<b>lýsa</b>	<b>describe</b>	209	<b>24</b>	<b>N_D</b>	<b>A_AI</b>	<b>A_T</b>	<b>none</b>	<b>0</b>
<b>banka</b>	knock	<b>210</b>	24	N	A	A	none	0
<b>súpa</b>	sip	211	23	N_(A)	A_(I)	A_(P)	<b>none</b>	0
<b>stoppa</b>	stop	<b>212</b>	23	N_(A)	AI_(AI)	AT_(T)	none	0
<b>skella</b>	<b>slam</b>	213	<b>23</b>	<b>N_D</b>	<b>AI_AI</b>	<b>A_(T)</b>	<b>none</b>	<b>0</b>
<b>senda</b>	<b>send</b>	<b>214</b>	<b>23</b>	<b>N_(D)_A</b>	<b>A_(A)_AI</b>	<b>A_(G)_T</b>	none	<b>0</b>
<b>príla</b>	climb	215	23	N	A	A	<b>none</b>	0
<b>pakka</b>	<b>pack</b>	<b>216</b>	<b>23</b>	<b>N_(D)</b>	<b>A_(I)</b>	<b>A_(T)</b>	none	<b>0</b>
<b>mæta</b>	<b>meet, show up</b>	217	<b>23</b>	<b>N_(D)</b>	<b>A_(AI)</b>	<b>A_(T)</b>	<b>none</b>	<b>1</b>
<b>klóra</b>	<b>scratch</b>	<b>218</b>	<b>23</b>	<b>N_AD</b>	<b>A_(AI)</b>	<b>A_(PG)</b>	none	<b>0</b>
<b>hræða</b>	scare	219	23	N_A	AI_A	A_P	<b>none</b>	0

<b>gangaB</b>	<b>go (well)</b>	<b>220</b>	<b>22 ND</b>	<b>AI</b>	<b>TE</b>	none	<b>0</b>
gangaC	walk	221	22 N	A	A	<b>none</b>	0
<b>tína</b>	<b>pick</b>	<b>222</b>	22 N_A	A_I	A_T	none	0
<b>svara</b>	<b>answer</b>	223	<b>22 N_(D)_D</b>	<b>A_(A)_I</b>	<b>A_(G)_T</b>	<b>none</b>	<b>0</b>
<b>ráða</b>	<b>decide</b>	<b>224</b>	<b>22 N_(D)</b>	<b>A_(I)</b>	<b>A_(T)</b>	none	<b>0</b>
hvíla	rest	225	22 N_A	A_AI	A_P	<b>none</b>	
<b>banna</b>	<b>ban</b>	<b>226</b>	<b>22 N_(D)_A</b>	<b>AI_(A)_I</b>	<b>A_(G)_T</b>	none	<b>0</b>
<b>veifa</b>	<b>wave</b>	227	<b>21 N_(D)</b>	<b>A_A</b>	<b>A_G</b>	<b>none</b>	<b>0</b>
<b>sulla</b>	<b>spill</b>	<b>228</b>	<b>21 N_(D)</b>	<b>A_I</b>	<b>A_T</b>	none	<b>0</b>
<b>skreyta</b>	decorate	229	21 N_(A)	A_AI	A_P	<b>none</b>	0
<b>rétta</b>	<b>hand</b>	<b>230</b>	<b>21 N_(D)_A</b>	<b>A_(A)_AI</b>	<b>A_(G)_T</b>	none	<b>0</b>
<b>leiða</b>	hold hand	231	21 N_A	A_A	A_T	<b>none</b>	
<b>gelta</b>	bark	<b>232</b>	21 N	A	A	none	
<b>fæða</b>	give birth	233	21 N_(A)	A_(A)	A_(P)	<b>none</b>	
<b>hittaB</b>	hit	<b>234</b>	20 N_(A)	A_(AI)	A_(P)	none	0
<b>virka</b>	function	235	20 N	AI	T	<b>none</b>	0
<b>trufla</b>	disturb	<b>236</b>	20 N_(A)	AI_(A)	A_(P)	none	0
<b>sópa</b>	<b>sweep</b>	237	<b>20 N_(AD)</b>	<b>A_(I)</b>	<b>A_(PT)</b>	<b>none</b>	<b>0</b>
<b>stækka</b>	grow up	<b>238</b>	20 N	AI	T	none	0
<b>skera</b>	cut	239	20 N_A	AI_AI	A_P	<b>none</b>	0
<b>hljóma</b>	sound	<b>240</b>	20 N	AI	T	none	0
<b>hengja</b>	hang	241	20 N_A	A_AI	A_T	<b>none</b>	0
<b>borga</b>	<b>pay</b>	<b>242</b>	<b>20 N_(D)_A</b>	<b>A_(A)_I</b>	<b>A_(G)_T</b>	none	<b>0</b>
<b>aka</b>	<b>drive</b>	243	<b>20 N_D</b>	<b>A_I</b>	<b>A_T</b>	<b>none</b>	<b>0</b>
<b>mana</b>	dare	<b>244</b>	19 N_A	A_A	A_T	none	<b>0</b>
<b>kela</b>	cuddle	245	19 N	A	A	<b>none</b>	<b>0</b>

eignast	acquire	246	19 N_A	A_AI	T_(T)	none	0
þykjast	pretend	247	18 N	A	A	none	0
vaða	wade	248	18 N	A	A	none	0
springa	explode	249	18 N	A	T	none	0
skriða	crawl	250	18 N	A	A	none	0
raða	sort	251	18 N_D	A_AI	A_T	none	0
líða	feel	252	18 D	A	E	none	0
líma	glue	253	18 N_A	A_AI	A_T	none	0
kitla	tickle	254	18 N_A	A_A	A_P	none	0
hreyfa	move	255	18 N_A	A_AI	A_T	none	0
grína	kid	256	18 N	A	A	none	0
baula	boo	257	18 N	A	A	none	0
klappaB	clap	258	17 N_(D)	A_(I)	A_(T)	none	0
æla	vomit	259	17 N	A	A	none	0
vesenast	struggle	260	17 N	A	A	none	0
spinna	spin	261	17 N	A	A	none	0
pota	poke	262	17 N	A	A	none	0
lepja	sip	263	17 N_A	A_I	A_P	none	0
blása	blow	264	17 N_(A)	A_(I)	A_(P)	none	0
verja	defend	265	16 N_(A)	A_(AI)	A_(T)	none	0
smella	click	266	16 N_(D)	A_(AI)	A_(T)	none	0
smakka	taste	267	16 N_(A)	A_(I)	A_(P)	none	0
hrópa	scream	268	16 N_(A)	A_(I)	A_(T)	none	0
gægjast	peak	269	16 N	A	A	none	0
brosa	smile	270	16 N	A	A	none	0
þora	dare	271	15 N_(AD)	A_(I)	A_(T)	none	0
útskýra	explain	272	15 N_A	A_I	A_T	none	0

væla	whine	273	15 N	A	A	none	0
vekja	wake	274	15 N_A	A_A	A_P	none	0
strauja	iron	275	15 N_(A)	A_(I)	A_(P)	none	0
spóla	rewind	276	15 N	AI	AT	none	1
skoppa	bounce	277	15 N	AI	T	none	0
plata	fool	278	15 N_(A)	A_(A)	A_(T)	none	0
mynda	form	279	15 N_(A)	A_(AI)	A_(T)	none	0
mata	feed	280	15 N_A	A_A	A_P	none	0
hræra	mix	281	15 N_(D)	A_(I)	A_(P)	none	0
hrinda	push	282	15 N_(D)	A_(A)	A_(T)	none	0
grilla	grill	283	15 N_(A)	A_(I)	A_(P)	none	0
enda	end	284	15 N_(A)	AI_(I)	AT_(P)	none	0
þrifa	clean	285	14 N_(AD)	A_(AI)	A_(PG)	none	0
vinka	wave	286	14 N_(D)	A_(A)	A_(G)	none	0
teygja	stretch	287	14 N_(A)	AI_(AI)	A_P	none	0
snerta	touch	288	14 N_A	AI_AI	A_P	none	0
máta	try on	289	14 N_(A)	A_(I)	A_(P)	none	0
læsa	lock	290	14 N_(AD)	A_(AI)	A_(P)	none	0
knúsa	hug	291	14 N_A	A_A	A_P	none	0
fjúka	blow away	292	14 N	AI	T	none	1
dæma	judge	293	14 N_(A)	A_(AI)	A_(T)	none	0
bregða	startle	294	14 D	A	E	none	0
bjóða	invite	295	14 N_(D)_ (A)	A_(A)_ (I)	A_(G)_ (T)	none	0
bila	break down	296	14 N	AI	T	none	1
tapa	lose	297	13 N_(D)	A_(AI)	A_(T)	none	0
rugga	rock	298	13 N_D	AI_AI	A_(T)	none	0

reykja	smoke	299	13	N_(A)	A_(I)	A_(P)	none	0
ljúka	finish	300	13	N_D	AI_(I)	AT_(P)	none	0
geispa	yawn	301	13	N	A	A	none	0
dingla	ring	302	13	N_(D)	A_(I)	A_(T)	none	0
batna	get better	303	13	D	A	E	none	0
þefa	smell	304	12	N	A	A	none	0
stynja	moan	305	12	N	A	A	none	0
slitna	tear	306	12	N	I	T	none	0
skila	return	307	12	N_(D)_ (D)	A_(A)_ (AI)	A_(G)_ (T)	none	0
kynna	introduce	308	12	N_A	A_AI	A_T	none	0
klessa	crash	309	12	N	A	A	none	0
kinka	nod	310	12	N_(D)	A_(I)	A_(T)	none	0
hneppa	button	311	12	N_(D)	A_(I)	A_(T)	none	0
haldaC	hold event	312	11	N_A	A_(I)	A_(P)	none	0
látaB	act out	313	11	N	A	A	none	0
sveifla	swing	314	11	N_D	A_(AI)	A_(T)	none	0
skíra	name	315	11	N_(A)_ (AD)	A_(AI)_ (I)	A_(G)_ (T)	none	0
ropa	burp	316	11	N	A	A	none	0
rifna	rip	317	11	N	AI	T	none	0
mjólka	milk	318	11	N_(A)	A_(A)	A_(P)	none	0
leira	mold	319	11	N_(A)	A_(I)	A_(P)	none	0
hvísla	whisper	320	11	N_(D)_ (D)	A_(AI)_ (I)	A_(G)_ (T)	none	0
heyja	harvest	321	11	N	A	A	none	0
grípa	catch	322	11	N_A	A_AI	A_T	none	0
fyrirgefa	forgive	323	11	N_(D)	A_(A)	A_G	none	0
bráðna	melt	324	11	N	I	T	none	0
baða	bathe	325	11	N_A	A_AI	A_PG	none	0

trúa	believe	326	10 N_(D)	A_(AI)	A_(T)	none	0
stjórna	command	327	10 N_(D)	AI_(AI)	A_(T)	none	0
sjóða	boil	328	10 N_(A)	AI_(I)	AT_(T)	none	0
rúlla	roll	329	10 N_(D)	AI_(I)	AT_(T)	none	0
ræna	rob	330	10 N_AD	A_AI	A_PT	none	0
reima	tie (shoelaces)	331	10 N_(A)	A_(I)	A_(P)	none	0
naga	gnaw	332	10 N_A	A_I	A_P	none	0
lækna	heal	333	10 N_A	A_A	A_P	none	0
hósta	cough	334	10 N	A	A	none	0
fylla	fill	335	10 N_A	A_I	A_P	none	0
flýja	flee	336	10 N	A	A	none	0
átta	realize	337	9 N_A	A_A	A_T	none	0
troða	stuff	338	9 N_D	A_AI	A_T	none	0
svæfa	put to sleep	339	9 N_A	A_A	A_P	none	0
sprauta	inject	340	9 N_AD	A_AI	A_PT	none	0
skamma	scold	341	9 N_A	A_A	A_P	none	0
selja	sell	342	9 N_(D)_ (A)	A_(A)_ (AI)	A_(G)_ (T)	none	0
lofa	promise	343	9 N_(D)_ (D)	A_(A)_ (I)	A_(G)_ (T)	none	0
grafa	dig	344	9 N_A	A_AI	A_P	none	0
fatta	realize	345	9 N_(A)	A_(I)	A_(T)	none	0
dúkka	appear	346	9 N	AI	T	none	0
drasla	trash	347	9 N_(A)	A_(I)	A_(P)	none	0
verpa	lay eggs	348	8 N_D	A_(I)	A_(T)	none	0
vaska	wash	349	8 N_(A)	A_(I)	A_(P)	none	0
tjalda	camp	350	8 N	A	A	none	0
smyrja	butter	351	8 N_AD	A_(AI)	A_(PT)	none	0

minna á	remind	352	8	NAD_A	A_AI	AE_PT	none	0
merkja	mark	353	8	N_A	A_AI	A_P	none	0
losa	release	354	8	N_A	A_AI	A_PT	none	0
leigja	rent	355	8	N_(A)	A_(AI)	A_(T)	none	0
lauma	sneak	356	8	N_D	A_AI	A_T	none	0
kjósa	vote	357	8	N_(A)	A_(AI)	A_(T)	none	0
háttá	get ready for bed	358	8	N_(A)	A_(A)	A_(P)	none	0
hrapa	fall	359	8	N	AI	T	none	0
hlýja	warm	360	8	N_D	AI_A	A_P	none	0
hita	warm	361	8	N_A	AI_I	A_P	none	0
gifta	marry	362	8	N_AD	A_A	A_TG	none	0
galdra	make magic	363	8	N_(A)	A_(AI)	A_(P)	none	0
flýta	hurry	364	8	N_D	A_(AI)	A_(T)	none	0
dúlla	doodle	365	8	N	A	A	D	0
bullá	talk nonsense	366	8	N	A	A	none	0
benda	point	367	7	N_(D)	A_(A)	A_(G)	none	0
vona	hope	368	7	N_A	A_I	E_T	none	0
stöðva	stop	369	7	N_A	AI_AI	A_T	none	0
spranga	jump	370	7	N	A	A	none	0
skíta	shit	371	7	N	A	A	none	0
skína	shine	372	7	N	I	T	none	0
skemmta	entertain	373	7	N_D	A_A	A_P	none	0
saga	saw	374	7	N_(A)	A_(AI)	A_(P)	none	0
mala	grind	375	7	N_A	AI_I	A_P	none	0
losna	get away	376	7	N	AI	T	none	1

<b>kvikna</b>	turn on	377	7 N	I	T	none	0
<b>kosta</b>	cost	<b>378</b>	7 N_A	I_I	T_T	<b>none</b>	<b>0</b>
<b>hækka</b>	rise	379	7 N_(A)	AI_(A)	AT_(T)	none	0
<b>hugga</b>	comfort	<b>380</b>	7 N_A	A_A	A_P	<b>none</b>	<b>0</b>
<b>hnerra</b>	sneeze	381	7 N	A	A	none	0
<b>fiska</b>	fish	<b>382</b>	7 N	A	A	<b>none</b>	<b>0</b>
<b>falla</b>	fall, fail	383	7 N	AI	T	none	0
<b>virða</b>	examine, respect	<b>384</b>	6 N_A	A_AI	A_T	<b>none</b>	<b>0</b>
<b>villast</b>	get lost	385	6 N	A	T	none	0
<b>vaxa</b>	grow	<b>386</b>	6 N	AI	T	<b>none</b>	<b>0</b>
<b>tækla</b>	tackle	387	6 N_A	A_A	A_P	none	0
<b>tosa</b>	pull	<b>388</b>	6 N_(A)	A_(AI)	A_(T)	<b>none</b>	<b>0</b>
<b>sussa</b>	shish	389	6 N	A	A	none	0
<b>stýra</b>	<b>conduct</b>	<b>390</b>	<b>6 N_(D)</b>	<b>A_(AI)</b>	<b>A_(T)</b>	<b>none</b>	<b>0</b>
<b>strjúka</b>	<b>pat, flee</b>	391	<b>6 N_AD</b>	<b>A_AI</b>	<b>A_PG</b>	none	0
<b>skrúfa</b>	screw	<b>392</b>	6 N_A	A_I	A_T	<b>none</b>	<b>0</b>
<b>ryksuga</b>	vacuum	393	6 N_(A)	A_(I)	A_(P)	none	0
<b>reikna</b>	calculate	<b>394</b>	6 N_(A)	A_(I)	A_(P)	<b>none</b>	<b>0</b>
<b>pússa</b>	polish	395	6 N_A	A_I	A_P	none	0
<b>nudda</b>	<b>massage</b>	<b>396</b>	<b>6 N_AD</b>	<b>A_AI</b>	<b>A_PT</b>	<b>none</b>	<b>0</b>
<b>lækka</b>	lower	397	6 N_(A)	AI_I	AT_(T)	none	0
<b>lifa</b>	live	<b>398</b>	6 N	A	T	<b>none</b>	<b>0</b>
<b>kýla</b>	punch	399	6 N_A	A_AI	A_P	none	0
<b>kveðja</b>	say goodbye	<b>400</b>	6 N_(A)	A_AI	A_T	<b>none</b>	<b>0</b>
<b>kroppa</b>	pick	401	6 N_A	A_I	A_P	none	0
<b>hvolfa</b>	<b>invert</b>	<b>402</b>	<b>6 N_D</b>	<b>A_I</b>	<b>A_T</b>	<b>none</b>	<b>0</b>

<b>hlýða</b>	<b>behave</b>	403	<b>6 N_(D)</b>	<b>A_(A)</b>	<b>T_(G)</b>	none	<b>0</b>
<b>herma</b>	imitate	<b>404</b>	6 N	A	A	<b>none</b>	<b>0</b>
<b>hefjast</b>	begin	405	6 N	I	T	none	<b>0</b>
<b>giska</b>	guess	<b>406</b>	6 N	A	A	<b>none</b>	<b>0</b>
<b>endurtaka</b>	repeat	407	6 N_A	A_I	A_T	none	<b>0</b>
<b>dreyma</b>	<b>dream</b>	<b>408</b>	<b>6 AD_(A)</b>	<b>A_(AI)</b>	<b>E_(T)</b>	<b>none</b>	<b>0</b>
<b>djóka</b>	joke	409	6 N	A	A	none	<b>0</b>
<b>afgreiða</b>	serve	<b>410</b>	6 N_(A)	A_(AI)	A_(PT)	<b>none</b>	<b>0</b>