Conditioned Variation: Children Replicate Contrasts, not Parental Variant Rate

Iris E. Nowenstein  
*University of Iceland*

Anton K. Ingason  
*University of Iceland*

Joel Wallenberg  
*University of York*

Follow this and additional works at: https://repository.upenn.edu/pwpl

**Recommended Citation**  
Available at: https://repository.upenn.edu/pwpl/vol28/iss2/14

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/pwpl/vol28/iss2/14  
For more information, please contact repository@pobox.upenn.edu.
Abstract
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 and Newport 2005, Smith et al. 2009, Cournane and Pérez-Leroux 2020). As has been extensively documented, children overgeneralize and regularize both consistent (Marcus et al. 1992) and inconsistent (Hudson Kam and Newport 2005) input. 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 and White 2019). The literature therefore shows both regularization and matching, but under different circumstances. In this paper, we argue for a third scenario and present a case where children neither regularize nor match their caregiver. Instead, they replicate the systematic contrasts they encounter and regularize within matched conditions. This is what happens in the acquisition of Icelandic Dative Substitution (DS), a stigmatized but widespread instance of grammatically conditioned morphosyntactic variation. We investigated DS in 99 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 results show that caregivers’ general DS rate did not predict the rate at which their children selected DS, regardless of age. On the other hand, when analyzing the data within conditioning factors, we found that children replicate the contrasts present in their caregivers’ speech, both at the group and individual level, and that this was in part dependent on age.
Conditioned Variation: Children Replicate Contrasts, not Parental Variant Rate

Iris Edda Nowenstein, Anton Karl Ingason, and Joel Wallenberg*

1 Introduction

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 and Newport 2005, Smith et al. 2009, Courmane and 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 and Newport 2005) input. These 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, Courmane 2019, Courmane and Pérez-Leroux 2020, Hall and 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 and White 2019).

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 and 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 saliency of the variable (e.g. Smith et al. 2009, Hendricks et al. 2018).

In this paper, we argue for a third scenario and present a case where children neither regularize nor match their caregiver. Instead, they replicate the systematic contrasts they encounter. This is what happens in the acquisition of Icelandic Dative Substitution (DS), a stigmatized but widespread instance of grammatically conditioned morphosyntactic variation. We investigated DS in 99 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 results show that caregivers’ general variant rate did not predict the rate at which their children selected DS, regardless of age. On the other hand, when analyzing the data within conditioning factors, we found that children replicate the contrasts present in their caregivers’ speech, both at the group and individual level, and that this was in part dependent on age. This is to some extent in line with previous studies (Smith et al. 2009, Hendricks et al. 2018, Courmane 2019, Hall and Maddeaux 2020, Repetti-Ludlow and MacKenzie 2022), and has implications for studies on the role of specialization (Wallenberg 2019) and the dynamics of variation in individual (Tamminga et al. 2016).

2 Background

2.1 Icelandic Subject Case Variation

Variation in subject case is one of the most researched topics of morphosyntactic change in Insular Scandinavian (e.g., Jónsson and Eythórsson 2005). As mentioned before, in the present study we focus on the most common form of variation in subject case in Icelandic: Dative Substitution (DS), also known as Dative Sickness (‘þágufallsýki’) in the prescriptivist discourse. It is a relatively stable but stigmatized variant which spread at the end of the 19th century (but see Viðarsson 2009 for examples from Old Icelandic) and has been extensively studied diachronically and in a series of

* We would like to thank all participants as well as audiences at NWAV49 and Rask 2022. The research was supported with grants from the Icelandic Research Fund and the UI Doctoral Fund.
large surveys (e.g. Svaavarsdóttir 1982, Jónsson 1997-1998, Jónsson and Eythórsson 2005, Barðdal 2011, Thráinsson 2013). Studies indicate that children’s rate of DS is associated with their parents’ socio-economic status (Svaavarsdóttir 1982, Eythórsson and Jónsson 2005). The change consists of dative being used instead of the original subject case for a number of psych verbs, as in (1):

(1)  **Hana** langar í epli. → **Henni** langar í epli.  
*her.ACC wants in apple. her.DAT wants in apple.*

‘She wants an apple.’

Most verbs with which DS occurs have traditionally accusative marked experiencer subjects, but two common DS verbs (*hlakka* ‘look forward to’ and *kviða* ‘be anxious about’) also appear in the nominative, which is by far the most frequent case for subjects in Icelandic (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 (non-nominative) subjects have certain characteristics, in addition to not triggering subject-verb agreement (the verb is in a 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 (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 in a novel verb (Guðmundsdóttir et al. 2019). Nonetheless, children do acquire DS (Sigurðardóttir 2002, Nowenstein 2017) and associate dative subjects more than nominative ones with experiencers in novel verb tasks (Nowenstein et al. 2020).

The large-scale surveys mentioned before (Svaavarsdóttir 1982, Eythórsson and Jónsson 2005, 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 (Svaavarsdóttir 1982, N = 202) and 2001 (Eythórsson and Jónsson 2005, N = 845) and 14–15-year-olds in 2006-2007 (Svaavarsdóttir 2013, Thráinsson 2013). For these four verbs, there is an increase of 6.5% (from 27.4% to 33.9%) between the results of Svaavarsdóttir (1982) and Eythórsson and Jónsson (2005) but the DS rate then drops to 20% in the study with 14–15-year-olds which was administered 5 years later (see Svaavarsdóttir 2013). This is in line with the Hall and Maddeaux (2020) “two steps forward, one step back” approach in which children innovate and then retract 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, Svaavarsdóttir 2013, Inason 2015). DS is not only specialized (Wallenberg 2019) along a continuous stylistic dimension (s-conditioning) but also along a categorical internal (grammatical) dimension (i-conditioning) (Tamminga et al. 2016). In this paper, we focus on two instances of conditioning along the internal grammatical dimension: the Person-Specific Retention (PSR) and syncretism effects (see Inason 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, see Nowenstein 2017). This has often been attributed to prescriptivism and self-correction but intra-speaker variation which is conditioned in this particular way can be found in informal contexts and child language (Nowenstein 2017). The second type of conditioning we investigate has not been confirmed until now; we find that dative is selected more often when the nominative and accusative are syncretic. This is shown in Table 1, with two inflectional paradigms which were used as stimuli in the study described in Section 3. *Stelpurnar* (‘the girls’) has syncretic nominative and accusative forms while *strikarnir* (‘the boys’) has distinct forms for each case. Here it is important to note that accusative-dative syncretism is also common in Icelandic, but the nominative and dative never pattern together without the syncretism including accusative also. 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.

Under such an interpretation, the syncretism effects might be the result of processing pressure in variation, while the PSR would be an example of how inconsistent input is regularized and
mated onto existing features.

<table>
<thead>
<tr>
<th></th>
<th>syncreric</th>
<th>non-syncreric</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominative</td>
<td>stelpurnar girls.the NOM/ACC</td>
<td>strákarnir boys.the NOM</td>
</tr>
<tr>
<td>accusative</td>
<td>stelpurnar girls.the NOM/ACC</td>
<td>strákana boys.the ACC</td>
</tr>
<tr>
<td>dative</td>
<td>stelpunum girls.the DAT</td>
<td>strákunum boys.the DAT</td>
</tr>
<tr>
<td>translation</td>
<td>the girls</td>
<td>the boys</td>
</tr>
<tr>
<td>DS rate</td>
<td>→ more DS</td>
<td>→ less DS</td>
</tr>
</tbody>
</table>

Table 1: Lack of syncretism and nominative-accusative syncretism in Icelandic subject case and possible relationship to Dative Substitution rate.

We have established that Icelandic Dative Substitution is an example of change in progress where grammatically conditioned intra-speaker variation is widespread, and now turn to the question of how children react to the patterns present in their input.

2.2 Acquiring Conditioned Variation

As already reviewed in Section 1, previous work shows both matching and regularization when variation is present in learners’ input, often within the same study. Whether matching or regularization is observed seems to depend on a number of factors. The pioneering work of Smith and colleagues (Smith et al. 2009) shows that correlations between variant rate in child-caregiver dyads may depend on age, the variable type, conditioning complexity and social saliency. 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 and Newport 2009, Hudson Kam 2015). This is the third option we want to emphasize in the current work, 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, Durham & Fortune (2009) found a significant relationship between children and their caregivers for the house-choose 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 and 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 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, we describe the methods used to investigate whether this is the case for children acquiring Icelandic Dative Substitution.

3 Methods

The data used in the present study were collected within the MoLiCoDiLaCo-project at the University of Iceland (PIs: Sigríður Sigurjónsdóttir and Eiríkur Rögnvaldsson), where the aims of the project were not to investigate the main questions of the present paper but to analyze the presence
of English in the Icelandic language community and evaluate its possible effect on the development of Icelandic in a broad sense (Sigurjónsdóttir and Nowenstein 2021). This contributes to the study being possibly underpowered within conditions despite a high number of participants. 101 children aged 3–13 (see Table 2) in a case production forced-choice task. 95 caregivers additionally participated in a task containing grammaticality judgments as well as forced-choice items. Using these data, we were 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).

<table>
<thead>
<tr>
<th>Age</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>16</td>
<td>4</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td>9</td>
<td>8</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 2: Number of participants by age.

In the children’s forced choice task, stimuli were presented orally and in written form with a blank (2) and options (e.g. the ones in Table 1).

(2) _________ langar út.
        wants out.
‘________ want to go outside’

Seven DS items were presented, two of them testing the PSR with first and third person non-syncretic pronouns with the same verb, and two of them testing syncretism effects with third person plural full NPs (also one verb). Caregivers were presented with 13 DS items in written form, four of them used to test the PSR (two verbs) 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 (3)), they were first transformed into a binomial accept/reject variable.

(3) 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 ideal, DS is a low-frequency variable and therefore not the ideal candidate for corpus-based research either, particularly when the conditioning factors discussed in 2.1 are also under investigation. Additionally, we 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.

4 Results

We 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 before moving on to the results within conditioning factors at the group and individual level.

4.1 Variant Rate

We begin with the rate of original case (nominative or accusative) – as opposed to the innovative dative – and the correspondence between the children and their caregivers. In Figure 1, we can for example see whether the caregivers with the highest rate of original case use are also associated with the children who select the original case the most. 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 with an age-caregiver rate interaction predicting the children’s rate shows no significant effects ($F (3, 80) = 0.5436, p = 0.6539$, adjusted $R^2 = -0.01677$). The children do not match their caregivers’ usage rate as measured by our proxy, which was computed across conditions, and this is not affected by age as can also be visualized in Figure 1.

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 use the original case more than their children, as would be expected with 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, 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 present in our younger participants. 7 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 nor original case use.

To summarize, when we look 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 children acquire variation instead of regularizing. We additionally see 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 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.

4.2 Person-Specific Retention

Starting with the PSR, Figure 2 shows a similar pattern emerging in the caregivers and children. 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 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.

We conducted a nested comparison (Likelihood Ratio Test) using mixed effects logistic regression models (lme4 in R, Bates et al. 2015). Random intercepts for participants and items were included in every model, as well age and an age-pronoun type interaction for the 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 ($\chi^2 (1) = 23.9, p < 0.001$). Additionally, age improves the children’s model significantly ($\chi^2 (1) = 7.1, p < 0.01$) but the age-pronoun type interaction does not ($p = 0.74$). As can be seen in figure 3, the PSR is clearly present from a young age. The difference between the youngest and oldest children in our experiment is therefore 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, we 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 4 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.

Figure 3: Dative selection in the first and third person singular, by age group (children).

4.3 Syncretism Effect
We now move on to the syncretism effects results. As for the PSR, Figure 4 shows that the syncretism effect is present in both groups. 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.
Figure 4: Dative selection/acceptance with/without syncretism on the subject, by group (children and caregivers).

We also conducted a nested comparison of mixed effects logistic models for the syncretism effects. For the adults, we observe a significant improvement to the fit of the model by 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 pronoun-age interaction ($p = 0.07$) were significant.

Figure 5: Dative selection with/without syncretism on the subject, by age group (children).

The visualization in figure 5 indicates that the lack of a significant age-syncretism interaction might be due to a power issue, as the syncretism effects seem to be acquired later. 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 Summary and Discussion

The main contributions of the paper can be put into three different categories. Theoretically (1), we aimed to investigate how children acquire grammatically conditioned case marking variation in function of caregiver language within and across conditions and how this might be dependent on age. The results showed that when acquiring the grammatically conditioned case marking variation of Icelandic Dative Substitution, children neither regularize nor match their parents’ variant rate. Instead, we observe systematic contrast replication at the group and individual level, or regularization within matched conditions. As discussed in Section 2.2, this is in line with results from previous work (e.g. Smith et al. 2007, Hall and Maddeaux 2020, Repetti-Ludlow and MacKenzie 2022). Empirically (2), 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 had previously not been investigated in Icelandic. DS therefore shows robust i-conditioning along a categorical dimension with s-conditioning also present and supposedly acquired in specific contexts (Tamminga et al. 2016). This might point towards stabilizing variation within specialized variants (Wallenberg 2019), although we also observe results which are in line with the “two steps forward, one step back” approach (Hall and Maddeaux 2020). Methodologically (3), we 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 and Maddeaux 2020 and Hall 2020 for elicitation tasks), but the results’ correspondence with previous work suggests that judgments and forced choice are viable methodological options, as long as the limitations of the present work are addressed, in particular the statistical power within conditions.

To conclude, we want to stress that in future work, we aim to put the present results in the context of language acquisition models. Such models need to be able to account for the emergence of productive intra-speaker variation patterns, and we consider a combination of Yang’s variational model of language acquisition (2002) and Tolerance Principle (2016) to be a promising approach.

References


Institute of Linguistics
University of Iceland
Reyjavík, 101, Iceland
trisen@hi.is
antoni@hi.is

Department of Language and Linguistic Science
University of York
York, YO10 5DD, UK
joel.wallenberg@york.ac.uk