




Article

Mapping Commission Errors to Grammatical Development: A Case Study of Malayalam

Gayathri G. Krishnan , Arathi Raghunathan  and Vaijayanthi M. Sarma 

Department of Humanities & Social Sciences, Indian Institute of Technology Bombay, Mumbai 400076, India

* Correspondence: gkrishnan@iitb.ac.in

Abstract: Young children learning Malayalam use morphological categories and inflections quite productively and accurately in general. However, their utterances sometimes show the use of extra morphological material (or commission errors), revealing mismatches between adult and child grammars. In this paper, we present a survey of such errors that are observed in longitudinally collected, spontaneous speech production data of monolingual Malayalam and bilingual Malayalam–English acquiring children in order to identify both the range of commission errors and the underlying grammatical features that may have triggered them. A close analysis of the data shows us that such errors are restricted to a few grammatical loci and shed light on the specific challenges that some grammatical constraints pose for developing grammars.

Keywords: Malayalam; commission errors; early language acquisition; bilingual acquisition



Citation: Krishnan, Gayathri G., Arathi Raghunathan, and Vaijayanthi M. Sarma. 2023. Mapping Commission Errors to Grammatical Development: A Case Study of Malayalam. *Languages* 8: 29. <https://doi.org/10.3390/languages8010029>

Academic Editors: Uli Sauerland, Artemis Alexiadou and Maria Teresa Guasti

Received: 11 June 2022

Revised: 26 October 2022

Accepted: 28 December 2022

Published: 16 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Young children acquiring their first language(s) sometimes produce utterances that deviate from those that their target grammars would permit. Such productions may involve, apart from omissions of grammatical material, the substitution of grammatical items, an overgeneralisation of a default exponent in place of a specific one, or the use of exponents where none is required. In all of these contexts, the children's productions differ from those of the adults' in the same morphosyntactic contexts by including *different* affixal material and, in the last, *more* affixal material than would be found in the adult productions. These are what have been called *commission errors*. In our survey, we discuss all such errors attested in the data including substitutions, overextensions, and overproductions and treat them all as commission errors.

Malayalam, which belongs to the South-Dravidian subgroup of the family, is one of the 22 scheduled languages of India and deemed a classical language. It is a head-final language with rich agglutinative morphology (Asher and Kumari 1997). The verbs and the nouns carry much of the grammatical information in an utterance, with verbs bearing tense, aspect, and mood inflections but not overt subject–verb agreement, and nouns bearing case and number inflections with gender generally being an inherent lexical feature, though it can be overtly marked as well. Previous research on the acquisition trajectories of Dravidian languages in general include Lakshmanan (2006); Lakshmi Bai (2004); Raghavendra and Leonard (1989); Sarma (2014); Usha Rani and Sailaja (2004), and Gayathri (2019); Girija Devi (1972); Leela (2016); Raghunathan (2021) specifically for Malayalam. However, there are no systematic discussions of commission error patterns in the literature of Dravidian languages.

An analysis of the errors in child productions permits us to understand more fully the grammatical properties of the target languages and the difficulties they may present to the young children. A study of such errors, dominantly in western languages, have contributed to various aspects of our understanding of lexical access and the differences between irregular and regular morphology (for example, the *dual-route model* (Marcus

1995; Pinker and Prince 1994) and the *single-route model* (Daugherty and Seidenberg 1994; Elman et al. 1996; Rumelhart and McClelland 1986)) and, more broadly, the maturational properties of grammar (for example, the *full competence hypothesis* (Borer and Wexler 1987; Poeppel and Wexler 1993; Rizzi 1993), which tries to account for the presence or absence of certain grammatical features).

One of the goals of this paper is to present a survey of the different kinds of commission errors that we find in early Malayalam grammars. We do not, however, discuss omission errors which are also found in the data (Krishnan et al. 2022) but point out that the omissions also centre on similar grammatical features underscoring the idea, and the second goal of the paper, that the complexity of these properties is the reason for the deviations. Young learners of Malayalam produce morphological material, especially inflections, with great ease and their overall performance conforms to expected adult patterns. While the ratio of errors in the productions to the numbers of correct uses is quite small, the errors are not accidental or performance missteps, but generally follow from incomplete learning. The errors are found across different lexical and grammatical categories, although the number of observed error tokens varies across the categories. The commission errors described here typically target either affixal material, with one affix being substituted for another or being inserted where an affix is not required, or stems, with one stem substituting for another. Such productions are judged ungrammatical by adult native speakers. Though less frequent cross-linguistically (Becker and Ud Deen 2020; Pierce et al. 2013), between the two kinds of errors, commission errors can shed particular light on both the underlying grammatical constraints and the specific challenges they pose for developing grammars. Such errors may also indicate the direction of grammatical shifts in the language (Lightfoot 1999; Pinker 1999).

We look at both monolingual and bilingual data in order to present a fuller picture of the types of commission errors that are attested in early Malayalam. On the one hand, it is of interest that many of the errors occur in both sets of data, which suggests that access to another grammar/language does not alter the Malayalam acquisition trajectory and the difficulties that the children may face. On the other hand, some error types are only attested in one or the other set of data and may allow us to take into consideration the context of acquisition and whether the errors are Malayalam-centred or follow from broader issues that children acquiring any language may face. Although we do not present the bilingual children's English errors, the differences between their productions in English and those in Malayalam are informative, and will come up in our discussion. We must note that the data analysed here are longitudinally collected, spontaneous speech productions and not item-targeted experimental data. Hence, the errors discussed here are not uniformly present across all children and some errors are only attested once. Nonetheless, the discussion in this paper should help us uncover the grammatical loci that present challenges to early grammars. The ages at which these features are fully acquired remains an open question.

We also analyse the structural context of each error to discover what the error type can tell us about the features of child and adult grammars themselves. While it is conceivable that an error was just that, an error of performance, or that the tokens may be limited by the 'serendipitous' nature of spontaneous production data, we think that there is much to be inferred from the observed errors. Experimental/elicitation studies are needed to focus more fully on specific errors and to inform us of their spread and frequency in Malayalam child language.

The paper is structured as follows. We begin in Section 2 with an overview of the children in the study, as well as the methods of data collection and analysis. We then present our detailed survey of the commission errors in the data, together with a discussion of each error type in Section 3. In Section 4, we offer a discussion of the errors and our observations as to why these errors may be present. We present the conclusions in Section 5.

2. Subjects and Method

There were five children in the study, of whom two, a girl and a boy, were monolingual Malayalam learners and three, two girls and a boy, were bilingual Malayalam and English learners. The children belonged to middle-class families and their parents were educated. While Malayalam was the only language used at home and in the speech community at large for the monolingual children, the bilingual children heard both Malayalam and English at home. The monolingual data were collected at Alappuzha, Kerala. The bilingual data were collected at three different places – Bangalore, India; Dubai, U.A.E; and Missouri, U.S.A. Prior to the start of the study, a language experience questionnaire was used to determine the acquisition contexts and the degree of exposure that the children were likely to receive to the language(s) they were learning. The data collection was approved by the IIT Bombay Ethics Committee: IITB-IEC/2016/001—date of approval 1 March 2016, and IITB-IEC/2016/002—date of approval 16 March 2016.

The monolingual children did not belong to multilingual households. Although the parents were college-educated, the language of the household was solely Malayalam and the parents did not use English utterances with the children. Of course, Malayalam has a number of lexical borrowings (from Sanskrit, English, Arabic etc.), but the borrowings are typically phonologically accommodated to Malayalam in monolingual use. Both the children were recorded in their homes as they interacted with the researcher or, occasionally, with parents or other caregivers with each session being between 30–50 minutes long. Materials such as puzzles, toys, and picture books were used to support interactions with the researcher.

In the bilingual contexts, children heard both English and Malayalam roughly equally. The parents used full English utterances with the children and were not merely mixing in some English lexemes. The utterance contexts were used to fix the language context in our analysis. If the child was addressed in English, then the context was treated as English and if the child was addressed in Malayalam, the context was treated as Malayalam. The utterance contexts were quite clear across the data. All three bilingual children received similar inputs from their parents. Their different geographical locations did not limit their exposure to either English or Malayalam, although their overall linguistic experiences did vary for the following reasons. One child lived with extended family for a little while who brought more Malayalam to the learning context, another heard more English during the day at daycare, and the third heard some Kannada and Tamil from caregivers with no overall perceptible impact barring the presence of a few lexical items. These external factors did not differentially impact the development trajectories either for English or for Malayalam. Two of the bilingual children, A_b and D_b , were recorded by their parents and one child, E_b , was recorded by the researcher.

A summary overview of the data including the number of productions, the number of lexemes, the Mean Length of Utterance (MLU) ranges, the number of sessions, and the ages of recordings is provided in Table 1. As is usually the case in studies that use such spontaneous data, we excluded from our analysis nonwords, filler sounds, nursery rhymes and songs, and verbatim repetitions.

Table 1. Data overview.

Child	A_m	H_m	A_b	D_b	E_b
Ages	1;9–2;10	2;3–3;0	2;0–3;0	1;9–2;8	1;10–2;11
Number of sessions	26	18	18	20	25
MLU range Malayalam	1.9–7	3–6.3	1.4–3.2	2.7–4.5	1.2–3.2
MLU range English	-	-	2.2–4.3	1.4–3	1.1–3.6
Malayalam utterances	3673	1072	649	837	1365
Malayalam words	9953	2809	652	1328	1488
English utterances	-	-	1355	705	3178
English words	-	-	2803	769	5010

MLU values have been calculated on the basis of the number of morphemes in the utterances. The bilingual children have lower values of MLUs compared to the monolingual children. This is not unusual (Ezeizabarrena and Fernandez 2018; Hoff et al. 2014), since the exposure of the bilingual children to each of the two languages is less than the exposure that the monolingual children receive to their single language. This difference is reset to comparable values in the early school years, but the children in this study were not yet at an age for school, and we cannot say definitively when they reach this parity. There are also some child-centric variations, with A_m and E_b being more loquacious than the others, though all the children show similar overall competencies.

Adapting Cazden (1968), we took a particular grammatical item to be in productive use at the age of recording when the children use the item, where its use is obligatory, at least 90 percent of the time and in three out of four consecutive recordings. We needed to use a four sample window, since some recording sessions carried zero instances of a grammatical feature of interest. We used ELAN (Wittenburg et al. 2006) to transcribe, organise, and code the production data with layers to code the context of utterance, the parental input, and different levels of grammar. An incorrectly produced grammatical item, i.e., a production that had overt grammatical material that deviated from target adult outputs, was coded as an error of commission based on the grammaticality judgements of the two coders who were native speakers of Malayalam and who were familiar with the dialects being used by the families. Intercoder reliability determined using Cohen's kappa (κ) was 0.94. No attempts were made by either the parents or the researchers to correct the children when they had errors in their utterances.

3. Data and Analysis

In general, similar errors are found in the Malayalam productions of both the monolingual and the bilingual children, which suggests that the specific challenges posed by the Malayalam grammatical system are common to both developmental trajectories. The potential access to another linguistic system (here, English) which may not have the specific grammatical properties of Malayalam, or which may have alternate expressions for the same property or even other enabling or, conversely, more complex mechanisms does not seem to impact the choices that the children are seen to make. We must point out that the English productions of the bilingual children show 'deviations' such as the use of *root infinitives* as has been observed with monolingual English acquiring children, but their Malayalam productions do not show those properties. The two systems, then, follow independent development trajectories. We focus, here, on only the Malayalam productions and the specific grammatical features within which the errors can be located. However, some error types are only found in one or the other kind of data and, equally interestingly, in one or the other of the children offering evidence for idiosyncratic grammatical choices. The token frequencies of the error types that we observe in both sets of data are given in Table 2. We discuss the different kinds of commission errors in the data with examples in the following subsections. For ease of presentation and clarity, the errors have been grouped into nominal morphology errors that include case, number and pronominal stem errors, and verb morphology errors that include verb stem and past tense inflection errors. Errors that are less frequent are dealt with separately.

Table 2. Token frequencies of correct instances vs commission errors in the transcripts.

Error Type	Monolingual		Bilingual		
	Correct	Commission Error	Correct	Commission Error	
Nominal	Accusative marking	181	6	45	2
	Dative on subjects	360	7	134	0
	Plural marking	44	2	3	1
	Pronominal suppletion	20	75	91	4
Verbal	Participial stem	2598	28	1152	9
	Past tense marking	911	7	326	3

3.1. Nominal: Case Errors

Malayalam nouns bear case, number, and gender morphology. Case assignment ranges over the nominative, accusative, dative, genitive, sociative, locative, and instrumental. While the nominative is the unmarked or bare form, the other case forms are overtly marked. Further, the accusative, dative, genitive, and sociative cases also show phonologically conditioned allomorphy between a nasalised and a non-nasalised suffix variant (1).

(1) Case marking in Malayalam.

case	<i>puuca</i> ‘cat’	<i>kuyil</i> ‘cuckoo’
nominative	puucca	kuyil
accusative	puuccay-e	kuyil-ine
dative	puuccay-kki	kuyil-ini
genitive	puuccay-uḍe	kuyil-inte
sociative	puuccay-ooḍi	kuyil-inooḍi
locative	puuccay-il	kuyil-il
instrumental	puuccay-aal	kuyil-aal

Typically, the nominative signals the subject, the accusative the object, the genitive possession, the locative location, the dative the indirect object and so on, but the map between grammatical roles and case, since case is a selectional property of predicates, is not entirely unique (Asher and Kumari 1997). For example, subject cases in Malayalam include both the nominative and the dative. While the nominative is the default subject case, a range of predicates assign the dative case to their subjects (Nizar 2010). The dative case also signals a wider range of semantic relations including possession, location, experiencer, recipient, and benefactive. Among the case markers, it is the genitive, locative, and instrumental that show biuniqueness, i.e., have one-to-one correspondence between meaning and form, compared to the dative and the accusative case exponents.

Malayalam, as has also been discussed in the literature for, among others, various Romance languages, Hebrew, Persian and Hindi, shows differential object marking (DOM) (Aissen 2003; Bossong 1991; Butt 1993; Comrie 1989). The hierarchy runs along the animacy scale [+human] > [−human, +animate] > [−animate] and can be correlated with obligatory and overt nominal morphology. Only [+animate] objects are overtly marked for accusative case and [−animate] objects are uniformly unmarked (2), and only [+animate] nouns in the subject position are assigned the dative case by virtue of the predicate’s semantics and the applicability of the experiencer theta role (3).

- (2) amma kuṭṭi-kki paal/puuccakkuṭṭi-ne koḍuṭṭu
 mother.NOM child-DAT milk.ACC/kitten-ACC give.PST
 ‘The mother gave milk / a kitten to the child.’
- (3) *kaseera-kki/puucca-kki viḥannu
 chair-DAT/cat-DAT be hungry.PST
 ‘The chair/cat was hungry.’

Commission errors of case affixes typically involve the accusative and dative exponents. In these errors, the affixes are overtly produced in contexts that do not require them

in the adult grammar, i.e., in violation of the animacy hierarchy. Such errors also involve the use of the nominative or the bare form of the noun in the subject position, instead of the dative. Across both these case categories, the commission errors not only target specific argument positions, but also show violations of case-specific constraints, as we will see in greater detail in the following subsections where we discuss object and subject marking errors.

3.1.1. Object Marking Errors

Commission errors with objects, in both acquisition contexts, involve the use of the accusative affix with [–animate] nouns that do not carry such marking in the adult grammar as in (4)–(6).

- (4) * **booli-ne** mee-cci tar-aŋam (A_m, 2;5.16)
 ball-ACC buy-PTCP give-MOD
booli meedji-cci tar-aŋam [expected form]
 ball.ACC buy-PTCP give-MOD
 ‘(You) should buy the ball (for me).’
- (5) * **tein-e** oŋdaakki taa (H_m, 2;11.18)
 train-ACC make.PTCP give.IMP
trein oŋdaakki taa [expected form]
 train.ACC make.PTCP give.IMP
 ‘Make me the train.’
- (6) * **bukki-ne** vaayikki (A_b, 2;1.25)
 book-ACC read.IMP
bukki vaayikki [expected form]
 book.ACC read.IMP
 ‘Read the book.’

The accusative affix is additional morphological material on the noun. The utterances indicate that the children know that these nominals are ‘objects’ in the syntactic structure, though their errors suggest that the accusative–animacy correlation that determines DOM is as yet incomplete. It is also important to note that the allomorph that is applied is appropriate to the noun in question (*booli-ne* ‘ball’ vs *tein-e* ‘train’) and the errors lie in the presence of such overt marking and not in the selection of the particular variant.

3.1.2. Subject Marking Errors

In continuation with the preceding, we find a second type of commission error involving the accusative exponent. Here, the subject nominals are incorrectly assigned the (structural) accusative case instead of the expected (structural) nominative case. This kind of error is found, again, in both the monolingual (7) and the bilingual data (8).

- (7) * baag-ili **kokki-ne** oŋd-oo (A_m, 2;5.16)
 bag-LOC crane-ACC have-Q
 baag-il **kokki** oŋd-oo [expected form]
 bag-LOC crane.NOM have-Q
 ‘Is there a crane inside the bag?’
- (8) * ii **ceettan-e** eedji hoospittal-eyykk-aa poo-ŋ-ee (D_b, 2;8.6)
 this older brother-ACC which.Q hospital-LOC-be.PRS go-PRS-NMLZ
 ii **ceettan** eedji hoospittal-eyykk-aa poo-ŋ-ee [expected form]
 this older brother.NOM which.Q hospital-LOC-be.PRS go-PRS-NMLZ
 ‘Which hospital is this older brother going to?’

The subject nouns *kokki* ‘stork’ and *ceettan* ‘older brother’ that bear the accusative are [–human, +animate] and [+human, +animate], respectively, features that are high in the animacy hierarchy and which require overt case marking. The predicate *oŋdi* ‘have’ in (7) (translated as *have* but the meaning is closer to *be* or *exist*) and the predicate *pook-* ‘go’, are

both unaccusative predicates (Levin and Hovav 1995). These sentences, therefore, have the underlying syntactic structure in (9) with a *derived* subject which is in fact the direct object of the verb that is assumed to raise from within vP to Spec, TP (Burzio 1986), in the adult and, we assume, child grammars.

- (9) [DP_i [t_i v]_{vP}]_{TP}

There has been much discussion in the literature on the acquisition of unaccusative predicates and A-movement more generally (Babyonyshev et al. 2001; Costa and Friedmann 2012; Fox and Grodzinsky 1998; Lorusso et al. 2005), and we find that the children successfully use unergative and unaccusative predicates together with nominative subjects as in (10)–(12).

- (10) kokki paraṅṅi-pooy-i (A_m, 2;5.16)
stork.NOM fly.PTCP-go-PST
'The stork flew (away).'
- (11) ḍuvaava inṅṅu (D_b, 1;9.10)
Dhruv baby.NOM sit.PST
'Baby Dhruv sat (down).'
- (12) kaali pooy-i (E_b, 2;5.2)
leg.NOM go-PST
'Leg got hurt.'

The children also produce postverbal subjects with unaccusatives as shown in (13)–(15). This has also been noted in Hebrew (Friedmann 2007) but we do not discuss these properties further here.¹

- (13) ṭaaṭe viiṭ-um naan (H_m, 2;6.1)
below fall-NOM I.NOM
'I will fall down.'
- (14) pooy-i vimaanam (A_m, 2;0.3)
go-PST plane.NOM
'The plane went.'
- (15) veenikk-um iḍi (E_b, 2;6.5)
hurt-FUT this.NOM
'This will hurt.'

The accusative marking errors in (7) and (8) are instances of incorrect exponent selection and commission errors. Such accusative marking on the subject nominals is only visible with unaccusative predicates. What is of interest are the *derived* nature of the subject, the animacy features of the nouns (*kokki* 'stork' and *ceṭṭan* 'brother'), and the overt (accusative) case that has been applied, indicating a possible conflict between DOM requirements and the object origin of the DPs versus their subject positioning.

Other subject case errors involve the dative case. In fact, most of the commission errors with the dative exponent target the subject position, as can be seen in the examples in (16) and (17).

- (16) ***ṅaani** ṅanaṅṅ-ee (H_m, 2;8.16)
I.DAT get wet.PST-EMPH
ṅaan ṅanaṅṅ-ee [expected form]
I.NOM get wet.PST-EMPH
'I got wet.'
- (17) ***ammu-ni** ooṭṭooy-ii pooy-i (A_m, 2;5.16)
ammu-DAT auto rickshaw-LOC go-PST
ammu ooṭṭooy-il pooy-i [expected form]
ammu.NOM auto rickshaw-LOC go-PST
'Ammu went in an auto.'

In addition to these, where the dative marker is used instead of the nominative in the monolingual data, we also find one instance in the bilingual data of the nominative being assigned to the subject instead of the required dative. Although the nominative is an unmarked or bare form, we can tell from utterances with pronominal subjects that this is a nominative case commission error rather than an omission of the dative case, since the first-person pronoun in Malayalam uses a suppletive stem in the non-nominative cases (18).

- (18) ***jaan** maḍi (E_b, 2;6.19)
 I.NOM enough
eni-kki maḍi [expected form]
 I-DAT enough
 ‘It’s enough for me.’

As we can see, children switch between the two permissible subject cases, the nominative and the dative. Other than the use of the accusative with animate subjects of unaccusative verbs, we find no other case affix being applied to nominals in the subject position. We may recall from Table 2 that the overall accuracy of use of nominative and dative cases on subjects is high. Dative subjects are a language specific feature of Malayalam (and other Dravidian languages). It is also an areal feature of the region (Verma and Mohanan 1990). Nominative and dative subjects are, thematically, agents and experiencers, respectively, and this is made transparent when we consider predicates that can be used with both types of subjects with correspondingly different semantics as in (19) and (20).

- (19) jaan ḍaahi-ccu [volition/agency]
 I.NOM thirst-PST
 ‘I thirsted (for something).’
- (20) eni-kki ḍaahi-ccu [non-volition/experiencer]
 I-DAT thirst-PST
 ‘I was/felt thirsty.’

A closer look at the utterances in (16) and (17) shows that the errors occur again with the unaccusative predicates *ṅanay-* ‘get wet’ and *pook-* ‘go’. This is of particular interest, since a number of dative experiencer predicates are unaccusatives (Belletti and Rizzi 1988). Further, the predicate *pook-* ‘go’ signals movement, which is one of the key semantic features of dative subject predicates when they instantiate the ‘goal’ theta role that naturally implicates movement *towards*. Seen together with the preceding discussion on accusative marked subjects, we find the same properties of animacy (*jaani* and *ammuuni*), obligatory case requirement, and derived subjects, converging here as well. The nominative-dative commission errors in (16)–(18) suggest that the grammars of the young learners include the specific features of Malayalam grammar such as the use of dative subjects and the case affix allomorphy, even while they produce the errors which underscore some difficulty with applying the animacy hierarchy.

3.2. Nominal: Plural Marking Errors

Malayalam differentiates between the singular count noun (a bare nominal) and the plural count noun. Like the accusative marking on objects and the dative marking on subjects, plural marking is also crucially defined by animacy specifications (21). Malayalam employs the same three-way contrast of the lexical features [animate] and [human] to mark plurality, but differentiates between [+human] nouns and the [+animate] unlike the accusative which groups all animates together. Thus, nouns with the lexical features [+human] take the plural affix *-maar* (e.g., *caaran-maar* spy-PL ‘spies’), nouns with the lexical features [+animate, –human] take the plural affix *-ka* (e.g., *puucca-ka* cat-PL ‘cats’), and [–animate] nouns may *optionally* be left unmarked (e.g., *aari meeḥa* six table.PL ‘six tables’). We may note that unlike the accusative case affix where allomorph selection was determined by the phonological shape of the bases with the same animacy properties, the allomorphs in the plural are lexically conditioned. Further, while the accusative

exponent is obligatorily excluded with [–animate] nouns (except in rare cases of ambiguity), the plural marker is *optional*, especially if a numeral adjective precedes the noun. These properties are more clearly visible when we consider plural marking on borrowed words (e.g., [+human] *tiicar-maar* ‘teachers’ and *qooktar-maar* ‘doctors’, but [–human, +animate] *hippopottaamassu-ka* ‘hippopotamuses’ and [–animate] *keesi/keesu-ka* ‘cases’ or *bessi/bessu-ka* ‘buses’ (Asher and Kumari 1997, p. 251)).

(21) Overt plural marking in Malayalam.

[+animate]	[+human]
ki i-ka [bird-PL] ‘birds’	ka an-maar [thief-PL] ‘thieves’
pa fuk-ka [cow-PL] ‘cows’	amma-maar [mother-PL] ‘mothers’
iicca-ka [fly-PL] ‘flies’	ku an-maar [dwarf-PL] ‘dwarves’
aa qu-ka [goat-PL] ‘goats’	va ran-maar [groom-PL] ‘grooms’
pu li-ka [leopard-PL] ‘leopards’	ce ecci-maar [elder sister-PL] ‘elder sisters’

Plural marking in Malayalam is a fairly complex phenomenon with a number of exceptions.² Asher and Kumari (1997) also propose that the phonological shape of the stem has a role to play in the affix selection, but we find that this is not easily generalisable. The same phonological shape surfaces with different suffixes (e.g., *amma-maar* mother-PL ‘mothers’, but *virid^ha-ka* old woman-PL ‘old women’). Some plurals of feminine nouns also freely alternate between *-maar* and *-ka* (e.g., *paṭṭi-maar/ka* wife-PL ‘wives’ and *ṇaḍi-maar/ka* actress-PL ‘actresses’). Nouns bearing the gender affix *-an*, whether human or not, also require the suffix *-maar* (e.g., *puruṣan-maar* male.PL ‘males’ and *kurukk-an-maar* fox-M-PL ‘foxes’), as do feminine nouns ending in [a] (e.g., *b^haarya-maar* wife-PL ‘wives’). Sometimes [+human] nouns may be moved down the animacy hierarchy and use the suffix *-ka* (e.g., *aanuṅ-ṅa* ‘men’, *peṇṇuṅ-ṅa* ‘women’, and *kutti-ka* ‘children’) (Asher and Kumari 1997, p. 249). As we said, [–animate] nouns are generally unmarked though it isn’t ungrammatical to mark them and the affix used is the same as for [+animate] nouns. These patterns show that the plural marker both employs the animacy hierarchy (though differently from the accusative) and also lacks biuniqueness. The child must, then, learn the specific plural marker for a number of individual nominal stems, which contributes to the overall difficulty in acquiring plural morphology. Nominal morphology is on the whole made more complex for all the preceding reasons.

Not surprisingly, then, of the nominal inflectional categories, plural morphemes are both least frequently produced and least productively used in the data. This remains the case until quite late. Plural morphology is also prone to systematic omission errors. A quick comparison with the bilingual children’s English utterances is informative in this context. The bilingual children produce the English number affixes (-s and -z) without difficulty, despite the allomorphy, even as they uniformly avoid number marking in their Malayalam productions.³

We find commission errors with both the plural morphemes, *-maar* and *-ka*, in the data. The monolingual child H_m incorrectly uses the plural morpheme *-maar* with the numeral adjective *oru* ‘one’ (22). The bilingual child D_b uses the plural marker *-ka* in the context of a singular demonstrative pronoun *iḍi* ‘this’.

- (22) * *oru ceetṭan-maar-aa* (H_m, 2;11.1)
 one brother-PL-be.PRS
oru ceetṭan-aa [expected form]
 one brother-be.PRS
 ‘(They) are elder brothers.’
- (23) * *iḍu-ka-um* *pariykk-atte* (D_b, 2;5.20)
 this-PL-CNJ pluck-OPT
iḍ-um *pariykk-atte* [expected form]
 this-CNJ pluck-OPT
 ‘Shall (I) pluck this too?’

Additionally, the monolingual child A_m is seen to extend the marker *-ka* to the uncountable, inanimate noun *b^{fi}akṣaṇam* ‘food’ (24) which should not bear number morphology.⁴

- (24) * ii **baccaṅ-ṅa**l okke $\bar{t}\bar{i}\bar{n}\bar{n}\bar{i}$ ii valud_ṅ-aay-it_ṅi kuu_l-ii poo-y-it_ṅi
 these food-PL all eat.PTCP these big-become-PRF school-LOC go-PTCP-PRF
 paḍiykk-aṅam (A_m , 2;8.16)
 study-MOD
 ii **b^{fi}akṣaṇam** okke $\bar{t}\bar{i}\bar{n}\bar{n}\bar{i}$ valud_ṅ-aay-it_ṅi skuu_l-il poo-y-it_ṅi
 these food all eat.PTCP big-become-PRF school-LOC go-PTCP-PRF
 paḍ^{fi}iykk-aṅam [expected form]
 study-MOD
 ‘After growing bigger by eating all these foods, (I) want to go to school and study.’

To summarise, the children’s errors involve the incorrect production of the [+human, +plural] marker *-maar* and the [−human, +animate, +plural] marker *-ka* in singular contexts, and the use of the *-ka* with an inanimate and noncount noun. It is not entirely clear if in the first error (22) the child is using the plural marker to signal either the *humanness* and/or *animacy* features rather than the number feature. In the second error (23), there is a mismatch between the singular demonstrative pronoun and a plural affix, but the presence of the conjunction *-um* implies that there is more than *one* thing that is being plucked, which may have driven this error. In the third error (24), the child is indeed indicating plurality (additionally with the universal quantifier), but does not adhere to the rule’s animacy requirement. In so far as an affix is permitted with inanimate nouns, it would be *-ka* and, to that extent, the child’s selection of the variant is correct.

3.3. Nominal: Pronominal Stem Suppletion Errors

Within nominal morphology, we find, in addition to the preceding, pronominal stem errors as well. In Malayalam, the first person, singular pronoun *jaan* (nominative) is replaced by a suppletive stem *en(n)-* in all the other case affixed forms, yielding *enne* I.ACC, *ente* I.GEN, *eniykki* I.DAT, *ennooḍi* I.SOC, and *ennil* I.LOC as given in (25). The second- and third-person pronouns do not show such suppletion and use a uniform stem *ṅin-* ‘you’ and *avan* ‘he’ / *ava* ‘she’ / *aḍi* ‘it’, respectively.

- (25) Suppletion in 1SG pronoun in Malayalam.

Case	Form	Gloss
nominative	jaan	I.NOM
accusative	enn-e	I.ACC
dative	eni-kki	I.DAT
genitive	en-te	I.GEN
sociative	enn-ooḍi	I.SOC
locative	enn-il	I.LOC
instrumental	enn-aal	I.INS

The utterances of the young children show overregularisation where the nominative base is used instead of the suppletive 1SG stem as shown in the following examples from the monolingual data, as in (26) and (27), and the bilingual data, as in (28).

- (26) * **ṅaan-e** ṅookk-un_ṅni ammee (H_m , 3;0.2)
enn-e ṅookk-un_ṅnu ammee [expected form]
 I-ACC look-PRS.PROG mom.VOC
 ‘Mom, (it) is looking at me.’
- (27) * **ṅaa-ni** miṭṭaayi meeṅam (H_m , 2;5.18)
eni-kki miṭṭaayi veeṅam [expected form]
 I-DAT chocolate.ACC want.MOD.PRS
 ‘I want chocolate.’

- (28) * **naan-iite** bukk-il vaɾəlm uŋɖi (D_b, 2;7.17)
en-te bukk-il vaɾəlm uŋɖi [expected form]
 I-GEN book-LOC violin.NOM have.PRS
 ‘I have a violin in my book.’

While these stem errors are attested in both sets of data, the monolingual child H_m produces them regularly, as shown in Figure 1. Among the other monolingual and bilingual children, the use of the 1SG forms is less frequent for two primary reasons. Malayalam is a topic-drop language, and arguments (including subjects) are often dropped. Therefore, the probability of the child using a nominative 1SG pronoun in their productions is low. Further, some of the children use their names to refer to themselves and avoid the personal pronoun all together. We do not know whether this is because the Malayalam pronoun with the suppletive stem is more challenging, or because the children are reflecting their parents’ use of their names, or because pronoun use and reference-switching is itself a complex phenomenon since children across languages show this behaviour (Morgenstern 2012; Smolík and Bláhová 2021).

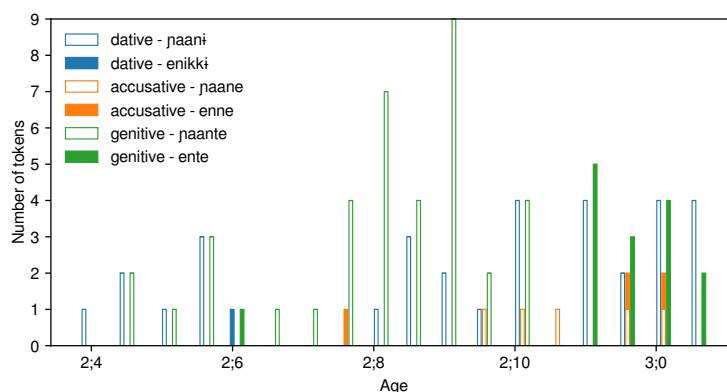


Figure 1. Pronominal stem errors in H_m's data.

In H_m's productions, the pronominal stem errors show an interesting pattern in the context of our discussion. H_m overextends the stem form *jaan* to systematically produce the nonadult forms *jaane* I.ACC, *jaani* I.DAT, and *jaante* I.GEN instead of *enn-e* I-ACC, *eni-kki* I-DAT, and *en-te* I-GEN, respectively. Interestingly, unlike in the other categories where errors are far fewer than correct usages, it is the nonsuppletive pronominal form that is more frequently found in H_m's data, with 75 such tokens compared to 19 tokens of the correct forms with the suppletive stem. What is of particular interest is that the errors to correct usage ratios are not uniform across the three case markers and do not remain so. As we can see in Figure 1, the errors with the dative case on the nonsuppletive stem persist throughout the period of recording. Only once, at the age of 2;6.1, did the child produce the correct form *enikki*. In all his other dative case productions he uniformly produces the nonsuppletive form.⁵ In contrast, the accusative marker occurs variably with and without the suppletive stem even towards the end of the study. In fact, we can see an overlapping use of correct and incorrect forms at 2;11.18 and 3;0.2 in the figure. Finally, while the genitive marker begins with such substitution errors, *jaante*, the child completely switches over to the correct suppletive form *ente* as he progresses in age. To summarise, we find variations across the three markers with the genitive case being fully correctly produced towards the end, the accusative case showing mixed use, and the dative case remaining a challenge. The full acquisition of the correct genitive form does not seem to automatically impact the other two case forms. We return to this pattern in Section 4.

3.4. Verbal: Switching of Participial Stems

Verbs in Malayalam carry tense, mood and aspect features, but not those for subject-verb agreement. The main verb can occur in one of two nonfinite shapes, the *bare stem* with the present and future markers, with modal affixes and in the infinitive, and the *past*

participle which is used in serial or multiverb constructions, or with the perfective as shown for the verb *pook-* 'go' in (29) and (30), respectively.

- (29) Use of bare stem.
Root *pook-* 'go'
 Present *pook-unṅṅu* 'is/are going'
 Future *pook-um* 'will go'
 Modal *pook-aam* 'can go'
 Infinitive *pook-aan* 'to go'
- (30) Use of the past participial form.
Root *pook-* 'go'
 Perfective *pooy-iṭṭi*
 Serial verb *pooyi-kaṭṭiṅṅu* go-finished 'gone'

Children are found to use the nonfinite, past participial verb form where the nonfinite bare verb stem is required. These errors are recorded in both the monolingual data, as in (31) and (32), and the bilingual data, as in (33) and (34).

- (31) * *pattiy-e picc-aan aatṭoottṭi puliy-e keetti*
 dog-ACC catch.PTCP-INF inside leopard-ACC made enter.PTCP
viṭ-tu (A_m , 2;6.15)
 send-PST
pattiy-e piḍiykk-aan akatṭoottṭi puliy-e keetti
 dog-ACC catch-INF inside leopard-ACC made enter.PTCP
viṭ-tu [expected form]
 send-PST
 '(I) sent the leopard inside to catch the dog.'
- (32) * *ṅaani pacc-aṅam* (H_m , 2;3.28)
 I.DAT pluck.PTCP-want.MOD
enikki pariyykk-aṅam [expected form]
 I.DAT pluck-want.MOD
 'I want to pluck (it).'
- (33) * *ṅookki amm-e puuttṭi vecc-um* (A_b , 2;9.20)
 look.IMP mom-ACC lock.PTCP keep.PTCP-FUT
ṅookki ammay-e puuttṭi vekk-um [expected form]
 look.IMP mom-ACC lock.PTCP keep-FUT
 'Look, (I) will lock (my) mother.'
- (34) * *ṅaan ceyd-aam* (E_b , 2;7.20)
 I.NOM do.PTCP-MOD
ṅaan ceyy-aam [expected form]
 I.NOM do-MOD
 'I can do (it).'

From the sample utterances in (31)–(34), we can see that the substitutions occur only where nonfinite verbs are needed, before modals, the future marker or the infinitival, and never where finite verbs are needed. In other words, while the children are employing the incorrect nonfinite form, it is in accordance with the overall syntactic requirements on *finiteness*. Thus, these are errors in the selection of an appropriate nonfinite stem. Children have to learn the individual contexts in which the participial stems are used versus those in which the bare stem is needed, which complicates the acquisition process.

3.5. Verbal: Past Tense Marking Errors

Apart from the selection of the nonfinite verb form, a second locus of commission errors with verbs is the past tense morphology. Malayalam past tense marking is a morphologically complex phenomenon (unlike the present and the future) and is expressed as one of two exponents *-i* and *-Tu*, with the latter ranging over many surface alternants given

the morphophonological contexts. While *-i* is the default affix (like *-ed* in English) and is sensitive to the phonological weight of the verb stems, the morphosyntactic specifications for *-Tu* require verb-specific, lexical information such as in-/transitivity and agentivity. These lexical features of a verb were more transparent diachronically and are becoming increasingly opaque in the morphosyntax (Krishnan and Sarma 2023). Some of the reasons for the opacity of morphological information include *morphological levelling* where phonologically similar verbs become inflectionally similar by analogy or where stem alternations within a single verb paradigm are eliminated, and *semantic shifts* that result in mismatches between the extant morphosyntactic features and the changed semantics (Hock 2009). It is sufficient here to note the opacity, and we find that the children resort to the default marker *-i* in their early productions.

We find past tense commission errors in both the monolingual and the bilingual productions, where *-i* is used instead of *-Tu*. In essence, the children are producing a tense exponent as needed, but choose the default affix over the specific one, and never the converse. Further, such errors are not specific to a single item, but are seen across several examples. Both the monolingual children show this error with the verb *koll-* ‘kill’, as in (35) and (36). In an earlier acquisition study of Malayalam (Girija Devi 1972), the author reports the use of **koll-i* for *koṅṅu* by a child in that study as well. Additionally, the monolingual child A_m also overgeneralises the marker *-i* to another geminate-ending verb *koll-* ‘hit’ (37). The bilingual child, A_b also overextends *-i* to a stem requiring *-Tu* (38). This strategy of resorting to the *-i* affix is also readily visible in adult productions, especially with borrowed lexemes.

- (35) * aḍiccu **koll-i** (A_m , 2;4.18)
 aḍicci **koṅṅu** [expected form]
 beat.PTCP kill.PST
 ‘Killed (the fly).’
- (36) * caanam cooṅ-e icci **koll-i** (H_m , 2;10.7)
 saad^hanam fooṅin-e iḍicci **koṅṅu** [expected form]
 thing phone-ACC hit.PTCP kill.PST
 ‘(That) thing killed the phone.’
- (37) * bukki vaaycc-appam kaṅṅ-ii **koḷḷ-i** (A_m , 2;3.16)
 bukki vaaycc-appam kaṅṅ-il **koṅḍu** [expected form]
 book read.PST-ADV eye-LOC hit.PST
 [ammu] bukki vaayicc-appoo [b^haḍṛay-uḍe] kaṅṅ-il koṅḍu [full form]
 [Ammu] book read.PST-ADV [Bhadra-GEN] eye-LOC hit.PST
 ‘The book hit (Bhadra) in the eye when (Ammu) read it.’
- (38) * **ṭaṅukk-i** (A_b , 3;0.20)
ṭaṅuttu [expected form]
 cold.PST
 ‘Got cold.’

3.6. Other Notable Errors

In Sections 3.1 to 3.5, we provided a survey of commission errors produced by the children across nominal and verbal forms including errors of case, number, pronominal stem, nonfinite verbal stems and past tense morphology. While there are only a few tokens of each type, they occur across children and across both kinds of data and allow us to evaluate the grammatical loci that are challenging in the language. However, we also find a few more errors in the monolingual data. These errors have a more restricted distribution, occurring in a single child’s production or even in a single token, but are of interest, and we discuss them in the following.

3.6.1. Nominal: Dative Instead of Genitive for Possession

In A_m ’s productions, we find the dative case being substituted for an expected genitive marked possessive nominal (39). While the dative case can encode the semantics of

possession (for example, with the verb *uṇḍi* ‘have’), it is never used to mark the possessor within a DP in the adult grammar.

- (39) Mother: aar-uḍe kaar-aa
 who-GEN car-be.PRS
 ‘Whose car?’
 A_m: *ammuu-**ni** (2;2.2)
 ammu-DAT
 ammuu-**nte** [expected form]
 ammu-GEN
 ‘Ammu’s (car)’

This case error does not reference animacy unlike the others, but points to the somewhat complex semantics of the dative case and the many thematic roles with which it can be aligned.

3.6.2. Verbal: Alternations

Verb alternations in Malayalam are signalled either via phonological changes, especially to the coda consonant, or via the affixation of the *-kk* suffix (40) (Krishnan and Sarma 2023).

- (40) Valence change in Malayalam.

Valence changing process	Intransitive	Transitive
Phonological change	muṇṇ- ‘sink’	mukk- ‘sink’
	kuumb- ‘join/close’	kuupp- ‘join/close’
	iḷak- ‘stir’	iḷakk- ‘stir’
	kuuṭ- ‘increase’	kuutt- ‘increase’
	iri- ‘sit’	iritt- ‘seat’
	viiṭ- ‘fall’	viiṭt- ‘fell’
<i>kk</i> -affixation	ṇanay- ‘get wet’	ṇana- kk - ‘water’
	aṇay- ‘go out’	aṇay- kk - ‘extinguish’
	ooḍ- ‘run’	ooḍi- kk ‘drive’

Both processes are seen in intransitive–transitive alternations. Additionally, morphological causatives can be built by the same *-kk* affixation, which can recurse once. With recursed *-kk-**kk*** sequences, the Obligatory Contour Principle (OCP) causes the first *-kk* to dissimilate to *-pp* (Killimangalam and Michaels 2006; Sadanandan 1999) (e.g., *ciri**kk***- ‘laugh’ – *ciri**ppikk***- ‘make laugh’).

One type of commission error that we find in the monolingual child A_m’s transcripts involves the overgeneralisation of the affix *-kk* to stems that would normally undergo phonological changes to signal the intransitive–transitive alternation, as can be seen in (41) and (42). Though the affix is incorrectly applied, the OCP application is appropriate, as is clearly visible in the sequence *oṇa-pp-ikk-* in (42). The monolingual child H_m and the bilingual children all produce many such alternating verbs (exhibiting both strategies), but this kind of commission error is only attested in A_m’s productions.

- (41) *batteccii-ḍe amma **ipp**-uṇṇ-oṇḍi (A_m, 2;4.2)
 b^hiḍra^hcecciy-uḍe amma **iritt**-uṇṇ-oṇḍi [expected form]
 bhadra sister-GEN mom sit.TR-PRS.PROG-have.PRS
 ‘Bhadra’s mom is seating (her).’
- (42) *oṇ**appikk**-aṇḍa (A_m, 2;8.16)
 oṇ**arṭṭ**-aṇḍa [expected form]
 wake.TR-MOD.NEG
 ‘Don’t wake.’

Nonconcatenative verb alternations are restricted to a smaller set of verbs compared to the larger class of verbs that undergo *-kk* affixation. There seems to be no particular phonological, morphological or semantic reason for this preference. For example, the verb

muyy- ‘sink_{intr}’ has the transitive counterpart *mukk-* ‘sink_{tr}’ formed via denasalisation of the cluster *-yy*, while a similar verb *mayy-* ‘become dull’ has the transitive counterpart *mayy-ikk-* ‘make dull’, formed via affixation. While **makk-* is phonologically possible, it does not exist in the language. The errors involving such lexical items as in (41) and (42) are, then, overregularisations where the less restrictive process is used by the monolingual child instead of the less frequent, phonological modifications. Such preferences are not dissimilar to the selection of *hanged* or *dived* as the past tense forms in English instead of the ablauted *hung* or *dove* for the verbs *hang* and *dive*, respectively. The variants that we find in the child’s productions cannot be reduced to instances of baby talk or idiosyncratic phonological alternations, since the changes are specific (*-kk* to *-pp*), show an overall increase rather than a reduction in word length, and the change is appropriate to the utterance context.

3.6.3. Verbal: Overextension of the Agentive Affix, *-kk*

As we have just seen in Section 3.6.2, Malayalam uses the agentive exponent *-kk* to derive verb alternants (e.g., *pott-* ‘break_{intr}’ – *pottikk-* ‘break_{tr}’, *maray-* ‘cover_{intr}’ – *maraykk-* ‘cover_{tr}’). The monolingual child A_m uses the affix to produce the transitive form of a verb instead of using its intransitive form as the adult would. In this utterance context, she is talking about her fallen tooth, that she has herself pulled out.

- (43) * ammu-nte palli **potticci** pooy-i (A_m , 2;3.1)
 ammu-GEN tooth break.TR.PTCP go-PST
 ammu-nte palli **potti** pooy-i [expected form]
 ammu-GEN tooth break.INTR.PTCP go-PST
 ‘Ammu’s tooth broke.’

This utterance is a creative deviation from the adult usage. The adult verb equivalent in Malayalam that corresponds to the (active) removal of a tooth is *parikk-* ‘pluck’ and not *pott-* ‘break’. The latter is more appropriate to the interpretation ‘the tooth broke’. However, the context of the utterance tells us that the child is in fact talking about such active removal (by her) of the tooth and not about it being broken off.⁶ The agentive affix, *-kk*, changes the verb from the intransitive, unaccusative *pott-* ‘break’ to the (agentive) transitive *pottikk-* ‘break’. The object *palli* ‘tooth’ in the utterance (43) is expected to occur with the unaccusative verb *pott-* ‘break’. However, there is an implicit agent which can be understood from the context. The child essentially overextends the agentive morpheme to this verb root (which is also followed by another unaccusative predicate *pook-* ‘go’) to produce a transitive verb. While the operation itself is not ungrammatical and the output is a correct variant of the verb, its use in this context is incorrect. The child can also be seen to successfully make the concomitant morphophonological changes to the stem followed by the insertion of the agentive affix *-kk*. She also produces the past participle form of *pottikk-* correctly, which is *potticci*. It is clear that the child’s competence includes a knowledge of verb alternations in Malayalam, as well as the use of *-kk* affixation for that purpose. It is likely, given that her emphasis is on her own agency in the action of tooth removal, that she uses a grammatical strategy available in the language to make the agent morphologically explicit. Such creative commission errors allow us to see the morphology in action and is more informative of the grammatical processes than even the correct uses themselves.

3.6.4. Adjectival: Overextension of the Affix, *-a*

A final error that we find in the data is the use of the marker *-a* which is the adjectival suffix in Malayalam. It is typically added to verbal participles to derive (deverbal) adjectives that are most frequently used in relative participle constructions (e.g., *vaadiya puuvi* ‘Wilted flower/The flower that is wilted.’). In nominal (N-N) compounds, the first member of the compound provides an implicit adjectival/modifier reading to the head N (Malayalam compounds being right-headed), but the modifying N is never marked with the adjectival suffix (e.g., *va:ji-ppuuvi* / **va:ia-ppuuvi* way-flower ‘wayside flower’) unlike the deverbal adjectives.

In one utterance by H_m, we find an overextension of the suffix in a compound. *komban-aana* ‘tusker-elephant’ is a compound noun where the noun (with a male gender suffix) *komban* ‘tusk’ modifies the noun *aana* ‘elephant’ to denote a specific set of elephants. H_m uses the adjectival suffix *-a* on the first nominal member of such a compound producing the ungrammatical form **komban-a* (44). With this overextension, the child is explicitly marking the adjectival function, in a manner that is similar to the use of the agentive marker to mark implicit agents.

- (44) * **komban-ay-aana** (H_m, 3;0.2)
 tusker-ADJ-elephant
komban-aana [expected form]
 tusker-elephant
 ‘a male elephant’

While the errors listed in this section are few, they reveal both that the children’s grammars incorporate these morphological processes and that the children are able to use them creatively even if the outcome is not congruent to those of the adult’s.

4. Discussion

In the preceding section, we surveyed the different types of commission errors produced by young learners of Malayalam. The full range of errors are summarised in (45). While some of these errors are seen to occur repeatedly in the data, across children in both the acquisition contexts, and to persist over time, others are only attested a few times and a few, only once as we have seen in the previous section.

- (45) Summary of commission errors in early Malayalam.

	Type of error	Expected	Example
Nominal	Accusative on inanimate object	no exponent	*[ein-e]/[rein/ ‘train’
	Accusative on subject	nominative	*[kokki-ne] /kokki/ ‘stork’
	Dative subject	nominative	*[kokki-ni] /kokki/ ‘stork’
	Nominative subject	dative	*[naan] /enikki/ I-DAT
	Dative possessive	genitive	*[ammuu-ni]/ammuu-nte/ ‘Ammu’s’
	Plural marking	no exponent	*[baccay-ŋa] /b ^β akŋam/ ‘food’
	Pronominal suppletion error	suppletive stem	*[naan-te] /ente/ I.GEN
Verbal	Nonfinite verb forms switch	bare stem	*[ceyɔ-aam] /ceyy-aam/ ‘can do’
	Past tense allomorph switch	past tense affix <i>-Tu</i>	*[koll-i] /ko- ^{nnu} / ‘killed’
	Verb alternations with affix	coda increment	*[oŋa-pp-ikk-] /oŋatt- / ‘wake _{tr} ’
	Agentive affix	intransitive	*[poɪi-kk] ‘break _{tr} ’ /poɪi- / ‘break _{intr} ’
Adjectival	Adjectival affix	noun	*[komban-ay-aana]/komban-aana / ‘tusker elephant’

The monolingual and bilingual children’s productions in the data show that errors (of omissions and commissions) are far fewer (except for H_m’s pronominal stem errors) compared to the correct usages, which is a remarkable achievement. Children also show productive use of various affixes (see Table 2). While the few errors of commission we discuss here show deviation from adult patterns, children’s utterances show that their grammars are congruent to adult grammars in most respects and, particularly, in marking higher functional categories like Tense. Thus, the children’s productions converge on full competence (Poeppl and Wexler 1993). More interestingly, the errors are limited to the options made available to them by their target language and are restricted to certain grammatical loci.

We can map the different kinds of commission errors that we have discussed above to larger grammatical patterns, as summarised in Table 3. *Incorrect selection* is one pattern. We see this in the children’s case assignment choices targeting different arguments (accusative and dative for nominative subjects, nominative for dative subjects, dative for genitive, accusative for inanimates, plural affix selection and marking) and in the selection of the nonfinite form (participle for the bare stem). With the participial errors, we see that children do not switch between finite and nonfinite categories, but switch between the

available nonfinite stems. The errors suggest that while their grammars have the underlying functional categories, it is the context-appropriate selection of the exponent or stem that continues to challenge despite an overall remarkable speed of acquisition.

Table 3. Grammatical patterns in child commission errors in Malayalam.

Patterns	Examples
Incorrect selection	accusative, dative for nominative, nominative for dative, dative instead of genitive, participial stem, plural marking
Overregularisation	pronominal stem suppletion, past tense marking, valence change
Overextension	overuse of agentive and adjectival suffixes

As we saw at some length in Section 3.1, Malayalam acquiring children produce several kinds of commission errors of case assignment. Specifically, we see deviations in subject and object case assignment. Subjects are assigned either the nominative or the dative case (depending on the predicate) and direct object nominals that are animate bear the accusative case overtly, while inanimate nouns are left obligatorily unmarked. With case assignment, we can recover the underlying role of the animacy hierarchy and its impact on child language.

The dative case in Malayalam, not only references animacy, but it is also one of the few suffixes in Malayalam that lacks biuniqueness. Further, it is employed in both argument (subject and indirect object) and adjunct positions (location and purpose) and mapped to several thematic roles including experiencer, possessor, goal, location, and benefactive. This increases the complexity of mapping the dative case to the arguments that require it. While the errors largely target the subject position, we also see the dative case being extended to mark a possessive nominal instead of the genitive case (39). *Possession* is part of the semantics of the dative case and, while the child uses the genitive and dative correctly in a majority of the cases and across argument types and semantic roles, the mismatch suggests that the child has still a little way to go in completely acquiring the precise contexts where the dative is required.

This kind of overextension of the dative case affix has also been noted by Lakshmi Bai (2004) in Tamil where, like Malayalam, the dative case can encode *location* among other properties. The child in that study is seen to substitute the dative case for the locative case, as in (46) and (47).

- (46) * bablu viiTT-ukku aaNTi illa (C, 2;1.12)
 Bablu house-DAT aunty is not
 ‘Aunty is not there in Bablu’s house.’ (Lakshmi Bai 2004, p. 257)
- (47) Father: rammiy-ooTa buk enka
 Rammi-GEN book where
 ‘Where is Rammi’s book?’
 R: * haalu-kku irkku (2;4.15)
 hall-DAT is
 ‘It is to the hall.’ (Lakshmi Bai 2004, p. 254)

The complexity of a rule is one of the main factors impacting successful acquisition and use of the morphological exponents. One source of such rule complexity is the use of grammatical items that lack biuniqueness and semantic transparency. Such items are seen to be relatively more difficult for children to acquire successfully (Dressler 2012). We see this both with the selection of the nonfinite stem where they have to learn the individual contexts and with the dative case.

A second factor that seems to impede full acquisition is the use of the animacy hierarchy, which is aligned with differential object marking in Malayalam. It is implicated in case (see Section 3.1.1) and number morphology (see Section 3.2). We have seen that Malayalam has two overt plural exponents and two accusative exponents as well. The animacy hierarchy [+human] > [−human, +animate] > [−animate] has a direct role in determining

the overt morphological outcomes. For accusative case marking, the nominals that are [\pm human, +animate] are overtly marked, but nominals that are [−animate] are obligatorily unmarked. For number, [+human] is distinguished from [−human, +animate]. Further, [−animate] may pattern with the latter or remain unmarked. An analysis of the errors produced in the data shows us that while the allomorphy per se is not the challenge, the children’s difficulties centre on exactly those grammatical processes that use the animacy hierarchy. This allows us to identify one prominent locus of difficulty in Malayalam grammar, which contributes to rule complexity as well. The commission errors that are attested in the data are congruent with the omission errors attested in the transcripts, which are not discussed here. We find that children frequently omit the nominal inflections whose selection is dependent on the animacy feature and their errors can be ascribed to the incomplete acquisition of the animacy hierarchy (Krishnan et al. 2022).

We find that other morphological rules that do not pay attention to these nominal features, such as the locative and the genitive case marking, are uniformly and successfully acquired. We do not suggest that animacy as an *ontological* category is difficult for children, that seems unlikely. But DOM as determined by the animacy hierarchy within the morphological rules that govern case and number assignments is a challenge in early language acquisition, and the errors that we have surveyed suggest that the property has not yet been fully acquired. While the children’s productions are quite accurate in general and the errors are not widespread, they are entirely within specific grammatical loci for specific reasons as we have just discussed.

Given the above discussion on rule complexity, we can now evaluate the pronominal stem error patterns in H_m ’s productions (Figure 1) that are unique to his data in the extent to which they are used (with incorrect forms far exceeding the correct usages) and in how they persist in his productions though other children also make similar substitutions, but transiently so. The recovery patterns that we observe in H_m ’s productions are of particular interest. The shift to the correct suppletive stem is not uniform and is in the order genitive > accusative > dative. This pattern directly reflects the *complexity* of the case assignment rules which is, dative > accusative > genitive, complexity being defined as the sensitivity of a rule to the animacy hierarchy and the properties of semantic transparency and biuniqueness. The dative is not only mapped to multiple theta roles and nominals in both the argument and adjunct positions, its assignment in the subject position with experiencer predicates is dependent on the [+human] feature of the nominal. Overt marking of the accusative case is similarly complex, whereas the genitive affix assignment is independent of the animacy hierarchy and shows biuniqueness, contributing to the relative ease of its acquisition. In this set of data, we have some direct evidence that the complexity of the rules impacts successful rule-application and complete acquisition. Interestingly, all three case affixes exhibit allomorphy with a nasalised and a non-nasalised form, as in (1), but the children demonstrate no difficulty with the use of the variants which are always appropriate to the nominal being used even while the case is being misapplied. It is also noteworthy that convergence on the suppletive base in one case form (genitive) does not automatically trigger a sweeping change across all the case forms, even though the suppletive stem shape is identical across all three. It appears that each rule has to be independently acquired, since the rules for each affix are different.

Another pattern that we find in the data, is *overregularisation* which is a well-recognised cross-linguistic phenomenon (Marcus et al. 1992). Children preferentially use the default rule/exponent instead of the specific rule/exponent. This application of the default over the specific is seen in the pronominal stem errors observed in both sets of data, and with the past tense affixes. The overgeneralisation of the past tense marking in Malayalam is analogous to past tense marking errors such as **brealed* and **singed* (Marcus et al. 1992) and the plural marking errors (Marcus 1995) produced by English learners. Pinker (1999) notes that children’s overregularisation errors indicate the direction in which morphological levelling takes place. Equally, children’s overregularisations in the data may point to possible directions in which morphological shifts are taking place in Malayalam, and also

show us the specific areas of grammar that already are or are becoming opaque and where the rules cannot be clearly induced. While this kind of levelling is a slow change, it is visible in the existence of doublets with one counterpart slowly going out of use. The coexistence of the irregular *dreamt* and *learnt* with the regular *dreamed* and *learned* exemplifies this slow change (Pinker 1999). We find similar dual past tense forms in Malayalam, with the default *-i* affixed forms gaining currency in modern usage. The verbs *vell-* ‘challenge’ and *coll-* ‘say’ form *ven̄nu* and *con̄nu*. These forms incorporate the past tense affix *-Tu* and are now archaic. We also find *velli* ‘challenged’, and *colli* ‘said’ with the default past tense affix *-i* (Rajaraja Varma 1917) which have supplanted the earlier forms. High frequency verbs like *cell-* ‘go’ and *koll-* ‘kill’ have managed to resist the levelling (like the forms of the English verb ‘be’) and we still find *cen̄nu* ‘went’ and *kon̄nu* ‘killed’ in common use instead of **celli* and **kolli*.

Similarly, with verb alternations, only a restricted class of verbs undergo nonconcatenative, phonological changes and Malayalam is slowly undergoing a morphological shift towards *-kk* affixation over phonologically effected changes. For example, the verb *para-* ‘fly’ has two transitive forms *paratt̄n̄*, the nonconcatenative form, and *para-ppi-kk-*, the concatenative form, with the latter gaining currency. The commission errors attested in the data (e.g., **ipp-* instead of *iritt̄n̄* ‘seat’ (41) and **on̄appikk-* for *on̄art̄t̄n̄* ‘wake_{tr}’ (42)) are in the same direction as this morphological shift. Overregularisation errors are often harbingers of morphological levelling.

A third pattern in the errors is the *overextension* of rules to syntactic contexts that do not require them as in the use of the agentive (43) and the adjectival (44) suffixes. In both these cases, the child’s utterances are an attempt at making the implicit explicit. Such productions are also constrained by the options available in the language’s grammar.

5. Conclusions

In this survey of the commission errors in the longitudinal monolingual and bilingual spontaneous speech production data, we sought to map the complexity of the grammatical processes to the errors and to evaluate its impact on children’s productions. We find that the commission errors also indicate the directions of morphological shifts in the language. We find that the errors, though few, make transparent the children’s knowledge of their target grammars and permit us to identify the grammatical loci that pose challenges to the young learner. We find that the use of the animacy hierarchy within morphological processes, given that the hierarchy is differentially used in nominal morphological processes, is a challenge, as is the lack of biuniqueness. In addition, Malayalam acquiring children show a tendency to prefer the general rule/exponent to the specific, as has also been established in studies of various languages. Although we did not discuss the omission errors in the data in this paper, the two combined drive home the same point, that despite their overall success, the children still have a little way to go to match the adults. The observed patterns of commission errors tell us what is left to be learned, and point to options exercised by the children while they learn the target grammars that are firmly within the grammatical choices provided to them by their target language, Malayalam.

Author Contributions: Conceptualisation, G.G.K. and V.M.S.; Data curation, G.G.K. and A.R.; Formal analysis, G.G.K. and V.M.S.; Investigation, G.G.K. and A.R.; Methodology, G.G.K.; Supervision, V.M.S.; Writing—original draft, G.G.K. and V.M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of IIT Bombay (IITB-IEC/2016/001—date of approval 1 March 2016, and IITB-IEC/2016/002—date of approval 16 March 2016).

Informed Consent Statement: Informed consent was obtained from the parents for all the subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We thank the three anonymous reviewers for their comments, questions, and suggestions and the editors of the special issue, Artemis Alexiadou, Maria Teresa Guasti, and Uli Sauerland. Any remaining errors are, of course, ours.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

- ¹ Malayalam permits ‘scrambling’ and so a correlation between postverbal subject positioning and unaccusativity is not a straightforward one.
- ² Our thanks to an anonymous reviewer for seeking greater clarification.
- ³ We have one instance in the data where the bilingual child uses an English affix on a Malayalam noun and produces *miin-z* ‘fish’.
- ⁴ While some dialects may treat this as a count noun, this is ungrammatical in the dialect that the child is learning and in standard Malayalam.
- ⁵ Although most of the nonsuppletive dative productions have the phonological shape *paani* ‘I’, there is also one instance of *paanikki* which uses the nonsuppletive base with the incorrectly matched dative allomorph. The point is that the stem selected is the same in both forms. Additionally, *paani* is a dative-marked form like the third person *avani* ‘he’, and the central vowel is not an epenthetic that is required by the phonology.
- ⁶ Our thanks to one of the reviewers for asking us about the inappropriate verb choice.

References

- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21: 435–83. [\[CrossRef\]](#)
- Asher, Ronald E., and T. C. Kumari. 1997. *Malayalam*. London and New York: Routledge.
- Babyonyshev, Maria, Jennifer Ganger, David Pesetsky, and Kenneth Wexler. 2001. The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry* 32: 1–44. [\[CrossRef\]](#)
- Becker, Misha, and Kamil Ud Deen. 2020. *Language Acquisition and Development: A Generative Introduction*. Cambridge: MIT Press.
- Belletti, Adriana, and Luigi Rizzi. 1988. Psych verbs and θ -theory. *Natural Language & Linguistic Theory* 6: 291–352. [\[CrossRef\]](#)
- Borer, Hagit, and Kenneth Wexler. 1987. The maturation of syntax. In *Parameter Setting*. Edited by Thomas Roeper and Edwin Williams. Dordrecht: Springer, pp. 123–72. [\[CrossRef\]](#)
- Bossong, Georg. 1991. Differential object marking in romance and beyond. In *New Analyses in Romance Linguistics*. Edited by Dieter Wanner and Douglas A. Kibbee. Amsterdam: John Benjamins Publishing Company, pp. 143–70. [\[CrossRef\]](#)
- Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Springer.
- Butt, Miriam. 1993. Object specificity and agreement in Hindi/Urdu. In *CLS29: Papers from the 29th Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, Vol. 1, pp. 80–103.
- Cazden, Courtney B. 1968. The acquisition of noun and verb inflections. *Child Development* 39: 433–48. [\[CrossRef\]](#) [\[PubMed\]](#)
- Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Chicago: The University of Chicago Press.
- Costa, João, and Naama Friedmann. 2012. Children acquire unaccusatives and A-movement very early. In *The Theta System: Argument Structure at the Interface*. Edited by Martin Everaert, Marijana Marelj and Tal Siloni. Oxford: Oxford University Press. [\[CrossRef\]](#)
- Daugherty, Kim, and Mark S. Seidenberg. 1994. Beyond rules and exceptions: A connectionist approach to inflectional morphology. In *The Reality of Linguistic Rules*. Edited by Susan D. Lima, Roberta L. Corrigan and Gregory K. Iverson. Studies in Language Companion Series 26. Amsterdam: John Benjamins Publishing Company, pp. 353–88.
- Dressler, Wolfgang U. 2012. On the acquisition of inflectional morphology: Introduction. *Morphology* 22: 1–8. [\[CrossRef\]](#)
- Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge: MIT Press.
- Ezeizabarrena, Maria-José, and Iñaki Garcia Fernandez. 2018. Length of utterance, in morphemes or in words?: MLU3-w, a reliable measure of language development in early Basque. *Frontiers in Psychology* 8: 2265. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fox, Danny, and Yosef Grodzinsky. 1998. Children’s passive: A view from the *by*-phrase. *Linguistic Inquiry* 29: 311–32. [\[CrossRef\]](#)
- Friedmann, Na’ama. 2007. Young children and A-chains: The acquisition of Hebrew unaccusatives. *Language Acquisition* 14: 377–422. [\[CrossRef\]](#)
- Gayathri G. 2019. Malayalam Morphosyntax: Inflectional Features and Their Acquisition. Ph.D. thesis, Indian Institute of Technology Bombay, Mumbai, India. [\[CrossRef\]](#)
- Girija Devi, A. 1972. Acquisition of language by a Malayali child. In *Proceedings of the First all India Conference of Dravidian Linguistics 1971*. Edited by Vadasery Iyemperumal Subramoniam and Elias Valentine. Kerala: University of Kerala, pp. 420–23.
- Hock, Hans Henrich. 2009. *Principles of Historical Linguistics*. Berlin: De Gruyter Mouton. [\[CrossRef\]](#)

- Hoff, Erika, Stephanie Welsh, Silvia Place, and Krystal M. Ribot. 2014. Properties of dual language input that shape bilingual development and properties of environments that shape dual language input. In *Input and Experience in Bilingual Development*. Edited by Johanne Paradis and Theres Grüter. Amsterdam: John Benjamins Publishing Company, pp. 119–40. [\[CrossRef\]](#)
- Killimangalam, Ashtamurthy, and Jennifer Michaels. 2006. The three 'ikk's in Malayalam. Unpublished MIT ms.
- Krishnan, Gayathri G., Arathi Raghunathan, and Vaijayanthi M. Sarma. 2022. Acquisition of Malayalam inflections: Complexity of morphosyntactic rules and its impact on developing grammars. *First Language*. [\[CrossRef\]](#)
- Krishnan, Gayathri G., and Vaijayanthi M. Sarma. 2023. Unlocking the verbal spine in Malayalam: Past tense is key. *Glossa: A journal of general linguistics* 8: 1–44. [\[CrossRef\]](#)
- Lakshmanan, Usha. 2006. Assessing linguistic competence: Verbal inflection in child Tamil. *Language Assessment Quarterly* 3: 171–205. [\[CrossRef\]](#)
- Lakshmi Bai, B. 2004. Acquisition of dative subject in Tamil. In *Non-Nominative Subjects*. Edited by Peri Bhaskararao and Karumuri Venkata Subbarao. Amsterdam: John Benjamins Publishing Company, Vol. 1, pp. 245–68. [\[CrossRef\]](#)
- Leela, Maya. 2016. Early Acquisition of Word Order: Evidence from Hindi-Urdu and Malayalam. Ph.D. thesis, Universitat Autònoma de Barcelona, Barcelona, Spain.
- Levin, Beth, and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge: MIT Press.
- Lightfoot, David. 1999. *The Development of Language: Acquisition, Change, and Evolution*. New York: Wiley-Blackwell.
- Lorusso, Paolo, Claudia Caprin, and Maria Teresa Guasti. 2005. Overt subject distribution in early Italian children. In *BUCLD 29: Proceedings Supplement of the 29th Annual Boston University Conference on Language Development*. Boston: Boston University.
- Marcus, Gary F. 1995. Children's overregularization of English plurals: A quantitative analysis. *Journal of Child Language* 22: 447–59. [\[CrossRef\]](#) [\[PubMed\]](#)
- Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, Fei Xu, and Harald Clahsen. 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 57: i+iii+v-vi+1-178. [\[CrossRef\]](#)
- Morgenstern, Aliyah. 2012. The self as other: Self words and pronominal reversals in language acquisition. In *Spaces of Polyphony*. Edited by Clara-Ubalina Lorda and Patrick Zabalbeascoa. Amsterdam: John Benjamins Publishing Company, pp. 57–72. [\[CrossRef\]](#)
- Nizar, Milla. 2010. Dative Subject Constructions in South-Dravidian Languages. Master's thesis, University of California, Berkeley, Berkeley, CA, USA.
- Pierce, Lara J., Fred Genesee, and Johanne Paradis. 2013. Acquisition of english grammatical morphology by internationally adopted children from China. *Journal of Child Language* 40: 1076–90. [\[CrossRef\]](#) [\[PubMed\]](#)
- Pinker, Steven. 1999. *Words and Rules: The Ingredients of Language*. New York: Basic Books.
- Pinker, Steven, and Alan Prince. 1994. Regular and irregular morphology and the psychological status of rules of grammar. In *The Reality of Linguistic Rules*. Edited by Susan D. Lima, Roberta Corrigan and Gregory K. Iverson. Amsterdam: John Benjamins Publishing Company. [\[CrossRef\]](#)
- Poeppl, David, and Kenneth Wexler. 1993. The full competence hypothesis of clause structure in early German. *Language* 69: 1–33. [\[CrossRef\]](#)
- Raghavendra, Parimala, and Laurence B. Leonard. 1989. The acquisition of agglutinating languages: Converging evidence from Tamil. *Journal of Child Language* 16: 313–22. [\[CrossRef\]](#) [\[PubMed\]](#)
- Raghunathan, Arathi. 2021. The Growth of Two Languages: Study of Malayalam-English Bilingual Language Acquisition. Ph.D. thesis, Indian Institute of Technology Bombay, Mumbai, India.
- Rajaraja Varma, A. R. 1917. *Kerala Paaniniyam*. Thiruvananthapuram: National Book Stall.
- Rizzi, Luigi. 1993. Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition* 3: 371–93. [\[CrossRef\]](#)
- Rumelhart, David E., and James L. McClelland. 1986. On learning the past tenses of English verbs. In *Parallel distributed processing: Explorations in the microstructure of cognition: Foundations*. Edited by David E. Rumelhart, James L. McClelland and the PDP Research Group. Cambridge: MIT Press, vol. 1. [\[CrossRef\]](#)
- Sadanandan, Suchitra. 1999. Malayalam Phonology: An Optimality Theory Approach. Ph.D. thesis, University of Southern California, Los Angeles, CA, USA.
- Sarma, Vaijayanthi M. 2014. Issues in the acquisition of Tamil verb morphology. In *South and Southeast Asian Psycholinguistics*. Edited by Heather Winkler and Prakash Padakannaya. Cambridge: Cambridge University Press, pp. 110–23. [\[CrossRef\]](#)
- Smolík, Filip, and Veronika Bláhová. 2021. Early linguistic reference to first and second person depends on social understanding as well as language skills: Evidence from Czech 30-month-olds. *First Language* 41: 109–25. [\[CrossRef\]](#)
- Usha Rani, A., and V. Sailaja. 2004. Acquisition of the non-nominative subject in Telugu. In *Non-Nominative Subjects*. Edited by Peri Bhaskararao and Karumuri Venkata Subbarao. Amsterdam: John Benjamins Publishing Company, vol. 2, pp. 209–22. [\[CrossRef\]](#)
- Verma, Mahendra K., and Karuvannur Puthanveetil Mohanan, eds. 1990. *Experiencer Subjects in South Asian Languages*. Stanford: CSLI Publications.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. ELAN: A professional framework for multimodality research. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*. Paris: European Language Resources Association.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.