MOSAIC+: A cross-linguistic model of verb-marking error in typically-developing children and children with Developmental Language Disorder

Daniel Freudenthal[1,3]

Fernand Gobet[2,3]

Julian M. Pine[1,3]

[1]University of Liverpool

[2]London School of Economics and Political Science

[3]ESRC International Centre for Language and Communicative Development (LuCiD)

Declaration of Interest: None

Abstract

This study extends an existing cross-linguistic model of verb-marking error in children's early multi-word speech (MOSAIC) by adding a novel mechanism that defaults to the most frequent form of the verb where this accounts for a high proportion of forms in the input. Our simulations show that the resulting model not only provides a better explanation of the data on typically developing children, but also captures the cross-linguistic pattern of verb-marking error in children with Developmental Language Disorder (DLD), including the tendency of English-speaking children to show higher rates of Optional Infinitive errors and the tendency of Dutch-, German- and Spanish-speaking children to show higher rates of agreement errors. The new version of MOSAIC thus provides a unified cross-linguistic model of the pattern of verb-marking error in typically developing children and children with DLD.

Introduction

Explaining the pattern of verb-marking error in typically-developing (TD) children and the pattern of verb-marking deficit in children with Developmental Language Disorder (DLD) is a key challenge for theories of language acquisition. Verb-marking errors are a characteristic feature of the speech of TD children. For example, in many languages, young children make errors (often referred to as Optional-Infinitive [OI] errors) in which they use infinitives and other non-finite verb forms in contexts in which a finite verb form is required. Deficits in verb-marking are a characteristic feature of DLD. For example, English-speaking children with DLD tend to produce OI errors for longer than TD children and at higher rates than controls matched for mean length of utterance (MLU). However, both the pattern of verb-marking error in TD children and the pattern of verb-marking deficit in children with DLD vary across languages.

MOSAIC (Model of Syntax Acquisition in Children) is a computational model of language learning that simulates the developmental patterning of verb-marking errors across several different languages in terms of the interaction between edge-based biases in learning and the distributional properties of the input language. MOSAIC simulates differences in the rate of OI errors in Dutch, French, German, and Spanish. However, in its current form, it cannot simulate either the very high rates of OI errors in English-speaking children or the cross-linguistic pattern of verb-marking deficit in children with DLD.

In this study, we supplement MOSAIC's basic learning mechanism with a mechanism that defaults to the most frequent form of the verb when the relative frequency of that form in the input is above a certain threshold. We investigate whether this new version of the model (MOSAIC+) provides both a better explanation of the cross-linguistic data on TD children and a means of simulating the cross-linguistic pattern of deficit in children with DLD. Our simulations show that MOSAIC+ can simulate both the very high rates of OI error in early child English and the fact that English-speaking children with DLD tend to show significantly higher rates of OI errors than MLU-matched controls, whereas Dutch- and German-speaking children do not, tending instead to show elevated, though still relatively low, rates of agreement and positioning errors.

Background Literature

The OI Phenomenon

Verb-marking errors are a characteristic feature of children's early language. For example, between the ages of 2 and 4 years, English-speaking children often make errors like (1) and (2) in which they use a zero-marked form in a context that requires a third-person singular (3sg) present-tense form (examples taken from the Manchester corpus; Theakston et al., 2001).

(1) *This go there (Anne, 2;6.4)

(2) *And the lorry go on top (Warren, 2;7.05)

Early analyses of these kinds of errors assumed that they reflect incomplete knowledge of the target inflection (e.g., Brown, 1973), or the dropping of the inflection due to performance limitations in production (e.g., Bloom, 1990). However, cross-linguistic analyses (e.g., Wexler, 1994) have shown that, in languages other than English, the equivalent errors tend to include verb forms marked with an infinitival morpheme like those in (3) and (4).

<u>Dutch</u>

(3) *Mama radio aan doen. (Peter, 2;0.7; Bol, 1996)

Mummy radio on put-INF

Mummy put radio on

German

(4) *Oma Brücke bauen. (Leo, 2;2.1; Behrens, 2006)

Grandma bridge build-INF

Grandma build bridge

Since these errors do not involve the use of a bare stem, they cannot be explained in terms of inflection drop, and this has led to the view that the pattern of verb-marking errors across languages (including the incorrect use of zero-marked forms in English) reflects the use of infinitives and other non-finite forms in finite contexts. These kinds of errors are typically referred to as Optional-Infinitive (OI) errors (Wexler, 1994), and the period during which they occur as the OI stage.

Most research on the OI stage has been conducted within a linguistic nativist framework. However, in a series of papers, we have used a computational model of language development (MOSAIC) to show that the cross-linguistic patterning of OI errors can be understood in terms of input-driven learning (see Pine et al., 2020, for a review). Below we outline the key features of MOSAIC. A more extended description of the model (including model architecture and learning mechanisms) is provided in Appendix S1 in the supporting materials.

MOSAIC

MOSAIC is an unsupervised learning model that learns from input in the form of orthographically-transcribed child-directed speech. MOSAIC gradually builds a network of words and strings of words from the input to which it has been exposed and produces output in the form of 'utterances' that become progressively longer as learning proceeds. Some of these utterances are produced by rote (i.e., have occurred as utterances or parts of utterances in the input). Others are produced generatively (i.e., by substituting distributionally similar words into frames that have occurred as utterances or parts of utterances in the input). Since the average length of MOSAIC's output increases with learning, MOSAIC can be used to simulate developmental changes in children's speech as a function of increasing MLU.

A key feature of MOSAIC is that it is subject to a strong utterance-final bias in learning. Early versions of MOSAIC (Freudenthal et al., 2006, 2007, 2009) learned entirely from the right edge of the utterance. That is, the model could only encode a word or phrase when everything that followed that word or phrase in the utterance had already been encoded in the network. MOSAIC thus built up its representation of an utterance by starting at the end of the utterance and slowly working its way to the beginning. This mechanism, which implements a recency effect in learning, can be likened to a moving window or buffer. Whenever an unknown word or word transition is encountered, the contents of the buffer are emptied, and only the most recently encountered word is left as a target for encoding. For example, when first exposed to the utterance *He goes home*, the model is only able to encode the word *home*. The word *goes* only becomes a target for encoding if the model has already encoded the word *home*, and the phrase *goes home* only becomes a target for encoding if the model has already encoded the words *goes* and *home*.

This utterance-final bias had the effect of restricting the strings that MOSAIC was able to produce to utterance-final sequences that had occurred in the input (or generative utterances based on such sequences). The current version of the model (Freudenthal et al., 2015) supplements MOSAIC's utterance-final bias with a (smaller) utterance-initial bias or left-edge learning mechanism. Left-edge learning works in a similar way to right-edge learning, except that it is anchored at the left edge of the utterance and restricted to a single word. MOSAIC now builds up its representation from both edges of the utterance, and is subject to a (small) primacy and a (larger) recency effect in learning. MOSAIC also combines the products of right- and left-edge learning by associating utterance-initial and utterance-final elements based on their co-occurrence in utterances in the input. For example, MOSAIC now represents strings such as *He go home* by learning to associate utterance-initial words such as *He* and utterance-final phrases such as *go home* based on their co-occurrence in utterances such as *He can go home*.

The addition of left-edge learning and a mechanism for associating the products of right- and left-edge learning has the effect of expanding the range of strings that the model can produce to include strings with missing utterance-internal elements. This mechanism has the potential to result in concatenations of elements with implausibly long intervening sequences such as *Jason (the boy you met at playgroup) plays football*. These are avoided by making the probability of associating utterance-initial and utterance-final elements dependent on the distance between the elements. It also has the potential to generate non-child-like concatenations such as *The (girl is going) to play football*. These are avoided by restricting concatenations to utterance-initial and utterance-final elements that are anchored at both edges of utterances in the input. That is, utterance-initial words can only be concatenated if they have also occurred in utterance-final position, and utterance-final elements can only be concatenated if the first word in the element has occurred in utterance-initial position.

MOSAIC simulates OI errors because of the way it learns from the edges of the utterance and associates the products of right- and left-edge learning. This results in the production of partial utterances that were present as utterance-final phrases in the input and concatenations of utterance-initial words and utterance-final strings. The structures in the input that give rise to OI errors are compound-finite structures: utterances that contain a finite modal or other auxiliary and an infinitive, such as the English utterance *This could go there* or the German utterance *Oma kann die Brücke bauen* (Grandma can the bridge build-INF). The truncation of utterances like these results in subjectless OI errors such as *go there* and *Brücke bauen*. The concatenation of utterance-initial words and utterance-final phrases from such utterances results in OI errors with subjects such as *This go there* or *Oma Brücke bauen* (Grandma bridge build-INF).

MOSAIC simulates the developmental patterning of OI errors because it learns to produce progressively longer utterances as a function of the amount of input to which it has been exposed. Children produce OI errors at high rates early in development and produce fewer OI errors as the length of their utterances increases. MOSAIC simulates this pattern because of the way that compound finites pattern in the relevant languages. In compound finites, the finite auxiliary precedes the infinitive. Since MOSAIC produces increasingly long utterance-final phrases, the short phrases it produces early on are likely to contain only non-finite verb forms. As the phrases MOSAIC produces become longer, finite auxiliaries start to appear, and OI errors are gradually replaced by the compound finites from which they have been learnt.

It is worth emphasising at this point that MOSAIC is a relatively simple distributional analyzer with no access to semantic information, which is clearly not powerful enough to acquire many aspects of adult syntax. MOSAIC is therefore best viewed as a simplified model of grammatical development that does not incorporate several factors that are known to affect children's language learning. Nevertheless, because of its ability to produce child-like utterances across a range of different languages, MOSAIC provides a powerful means of testing hypotheses about the relation between cross-linguistic variation in children's early language and cross-linguistic differences in the language to which they are exposed.

In an early paper, Freudenthal et al. (2007) showed that a right-edge learning model that learned OI errors from both questions and declaratives could simulate variation in the developmental patterning of OI errors across Dutch, German, and Spanish and the developmental patterning of OI errors with third-person singular (3sg) subjects in English. They also showed that the key factor was the way that MOSAIC's utterance-final bias interacted with the relative frequency of non-finite and finite verbs in utterance-final position (high in Dutch, moderately high in German, and low in Spanish, as are children's rates of OI errors in the respective languages). In a later paper, Freudenthal et al. (2009) showed that the same version of MOSAIC could simulate semantic-conditioning effects including the Modal Reference Effect and the Eventivity Constraint (the fact that in many languages OI errors tend to have a modal meaning and to be restricted to eventive rather than stative verbs), and the absence or reduced size of these effects in English. In a more recent paper, Freudenthal et al. (2015) showed that a version of the model that distinguished between declaratives and questions in the input and learned from both edges of the utterance could simulate the cross-linguistic patterning of OI errors in declaratives and Wh-questions in English, Dutch, German, and Spanish.

However, Freudenthal et al. (2010) have also shown that MOSAIC suffers from one important weakness as an account of the cross-linguistic data: it is unable to explain the very high rate of OI errors in English at low MLUs. Freudenthal et al. (2010) compare MOSAIC with Legate and Yang's (2007) Variational Learning Model (VLM) — a probabilistic parameter setting model which also has the potential to explain differences in the rate of OI errors across languages. More specifically, they investigate how well the two models predict the rate and lexical patterning of OI errors at an MLU of approximately 2 in English, Dutch, German, French, and Spanish. Their results provide support for MOSAIC's account of OI errors in the form of significant correlations between the rate at which children produce OI errors on particular verbs and the rate at which those verbs occur in compound-finite as opposed to simple-finite structures in child-directed speech in all 5 languages studied. However, they also show that, although both MOSAIC and the VLM are good at predicting differences in the rate of OI errors in Dutch, German, French, and Spanish, neither is able to predict the very high rate of OI errors in English. Freudenthal et al. therefore argue for a model of verb-marking error in which some errors reflect the use of infinitives learned from compound-finite structures in the input and others reflect a process of defaulting to the most frequent form of the verb when the target form is only weakly represented in the child's system. Such a model would predict particularly high rates of OI errors in English, where the most frequent form of the verb is usually the bare stem, and where bare-stem errors are indistinguishable from OI errors.

Supplementing MOSAIC with a Frequency-Sensitive Defaulting Mechanism

An extended version of MOSAIC that supplements the model's basic learning mechanisms with a frequency-sensitive defaulting mechanism has several potential advantages as an account of the cross-linguistic pattern of verb-marking error. First, adding some degree of frequency-sensitivity to the model's output has the potential to explain a wider range of errors, and is consistent with a wealth of evidence that frequency at a variety of levels not only increases fluency and protects items from error, but can also result in errors in which low-frequency items are replaced by higher-frequency words and sequences (see Divjak & Caldwell-Harris, 2015, and Ambridge et al., 2015, for reviews). Thus, although in many languages the most common type of verb-marking error is the use of a non-finite form in a finite context, there is evidence from more highly inflected languages that young children also make verb-marking errors in which they use the most frequent finite form of the verb in the wrong person/number context (Rubino & Pine, 1998; Aguado-Orea & Pine, 2015; Räsänen et al., 2016; Engelmann et al., 2019). A frequency-sensitive defaulting mechanism would provide a straightforward explanation of these kinds of errors.

Second, such a model has the potential to provide a better explanation of the rate of OI/bare-stem errors in early child English. Thus, because the bare stem covers 5 of the 6 cells in the English present-tense paradigm, defaulting errors in English are particularly likely to involve the use of the bare stem and, since the bare stem is indistinguishable from the infinitive, these errors will increase the rate of OI errors. In fact, there is already evidence that at least some apparent OI errors in English reflect frequency-sensitive defaulting. For example, in an elicited production study, Räsänen et al. (2014) found a significant relation between children's tendency to use bare forms of particular verbs in 3sg present-tense contexts and the relative frequency with which these verbs occur as bare forms versus third-person singular forms in finite present-tense contexts in English child-directed speech. Moreover, this result has since been replicated by Kueser et al. (2018) in a group of English-speaking children with DLD and a group of MLU-matched controls.

Third, such a model has the potential to explain the cross-linguistic pattern of verb-marking deficit in children with DLD. DLD, also referred to in the literature as specific language impairment (SLI), refers to a significant deficit in language ability that cannot be attributed to hearing loss or neurological damage (see Leonard, 2014, for a review). Although children with DLD are not a homogeneous population, deficits in verb-marking are a characteristic feature of the disorder. However, the pattern of verb-marking deficit in DLD varies across languages. Thus, English-speaking children with DLD tend to produce OI/bare-stem errors at higher rates than MLU-matched controls, even at high MLUs (Kueser et al., 2018; Rice et al., 1995). However, this effect appears to be specific to English. For example, Wexler et al. (2004), found no such differences in the rate of OI errors in their Dutch-speaking sample at MLU=3 and MLU=4, and, although Rice et al. (1997) did find an MLU-matching effect at MLU=2.66 in their German sample, this effect had disappeared by MLU=3.77.

In contrast, several researchers have found higher rates of subject-verb agreement and verb-positioning errors in Dutch and German. For example, both de Jong (1999) and Wexler et al. (2004) report elevated (though still relatively low) rates of agreement error in Dutch-speaking children with DLD; Clahsen et al. (1997) report higher rates of agreement error in

German-speaking children with DLD; and a number of researchers have reported verb-positioning errors in both Dutch (de Jong, 1999; Wexler et al., 2004) and German (Clahsen et al. 1997; Hamman et al., 1998; see Leonard, 2014, for a review). These positioning errors typically involve the incorrect use of finite forms (which are restricted to second position in Dutch and German) in utterance-final position, though the incorrect use of infinitives in second position has also been reported.

Taken together, these findings suggest that it might be possible to simulate the cross-linguistic pattern of verb-marking deficit in DLD by changing the defaulting threshold in an extended version of MOSAIC. Increasing the rate of defaulting by reducing the threshold at which defaulting errors occur is likely to increase the rate of OI/bare-stem errors in English. However, it is likely to increase the rates of agreement and verb-positioning errors in Dutch and German, where the most frequent form of the verb is likely to be a finite form that is readily distinguishable from the infinitive, and restricted to second position in main clauses.

The Current Study

The aim of the current study is to investigate whether an extended version of MOSAIC which supplements MOSAIC's basic learning mechanism with a novel defaulting mechanism provides both (a) a better explanation of the cross-linguistic data on TD children and (b) a means of simulating the cross-linguistic pattern of verb-marking deficit in children with DLD. In a first set of simulations, we investigate the extent to which adding a defaulting mechanism to MOSAIC improves the model's ability to simulate differences in the rate of OI errors in English, Dutch, German, and Spanish at MLU=2. In a second set of analyses, we investigate how this defaulting mechanism interacts with the frequency statistics of child-directed speech in the four languages to result in different levels of defaulting and different types of defaulting error. In a final set of simulations, we investigate whether increasing the rate of defaulting errors by reducing the defaulting threshold in the model allows us to simulate the cross-linguistic pattern of differences in the rate of OI, agreement, and verb-positioning errors in children with DLD relative to MLU-matched controls.

Method

In the current paper, MOSAIC+ was implemented by combining the version of MOSAIC described above and in Appendix S1 with a novel defaulting mechanism that was applied to the model's output. In this section, we first describe the way simulations are run in MOSAIC, and then the novel defaulting mechanism and the way in which defaulting rates were manipulated in the simulations that follow.

Running MOSAIC Models

MOSAIC is trained by feeding an input corpus through the model multiple times. This is necessary because the child-directed speech samples available in the languages modelled are typically not large enough to support gradual learning. Learning in MOSAIC is slow and MOSAIC initially represents just a few short utterance-final phrases. As learning proceeds, MOSAIC represents more phrases that extend further to the left of the utterance, as well as utterance-initial words, some of which have been associated with utterance-final phrases to form concatenations. Output is generated from MOSAIC by producing all the utterance-final phrases and concatenations of utterance-initial words and utterance-final phrases that it represents. Output from MOSAIC thus consists of a corpus of utterances that can be directly compared to corpora of child-directed speech. Since the average length of MOSAIC's output increases with increased training, it can also be matched to children in different stages of development based on their MLU in words.

For the current simulations, MOSAIC was trained on the child-directed speech from Anne and Becky's transcripts from the English Manchester corpus (Theakston et al., 2001), the child-directed speech from Matthijs and Peter's transcripts from the Dutch Groningen Corpus

(Bol, 1996), the child-directed speech from the German Leo Corpus (Behrens, 2006), and the child-directed speech from Juan's transcripts from the Spanish OreaPine corpus (Aguado-Orea & Pine, 2015). These are the same corpora as Freudenthal et al. (2010) used in their comparison of MOSAIC and the VLM. As in Freudenthal et al. (2010), we used versions of the English input corpora that were coded for the occurrence of verbs in 3sg contexts (e.g., *That goes-3SG there; She can go-3SG home; He is going-3SG out*). This feature allows for the identification of verbs that are learned from/produced in 3sg contexts even when no subject is present, and thus for a meaningful comparison of OI error rates in English and in the other languages, where OI errors can be readily identified even when the subject is absent.

The Novel Defaulting Mechanism

Defaulting was implemented in the model by identifying the most frequent form of each verb in a large corpus of child-directed speech in each language, and substituting this form for lower-frequency forms of the same verb in MOSAIC's output if its proportional frequency exceeded a certain threshold. Defaulting was implemented deterministically rather than probabilistically. That is, changes were always made when the proportional frequency of the relevant form exceeded the threshold, and never made when it did not. Implementing defaulting in this way does not reflect a theoretical commitment to deterministic defaulting, but rather an attempt to keep the defaulting mechanism as simple as possible to make it easier to understand the effects of manipulating the model's tendency to default across the different languages. The setting of the defaulting threshold is inevitably somewhat arbitrary. In the simulations of children's speech at MLU=2, we explore the use of values of .60, .65 and .70. These values were chosen to restrict defaulting to verb forms that made up a relatively large proportion of the relevant instances of that verb in the input while at the same time leaving scope for increasing the defaulting threshold as a function of increasing MLU in the later simulations. In our simulations of the speech of TD children and children with DLD at MLU=3 and MLU=4, we use values of .85 and .95 for the TD models and .65 and .75 for the DLD models. These values were chosen to allow us to distinguish clearly between the TD and DLD models while at the same time increasing the thresholds used in both sets of models as a function of increasing MLU.

Verb counts were collected from both declarative and interrogative input utterances. However, since MOSAIC assumes that children represent progressively longer utterance-final strings, the defaulting counts used at different MLUs were based not on corpus-wide statistics, but on utterance-final strings matched to the model's MLU. Thus, defaulting counts for models at MLU=3 were based on utterance-final strings of up to 3 words and defaulting counts at MLU=4 were based on utterance-final strings of up to 4 words. Corpus-wide statistics were also computed for purpose of comparison. This allowed us to investigate the extent to which defaulting is affected by imposing a similar utterance-final bias on the defaulting mechanism to MOSAIC's utterance-final bias in learning. To maximise the reliability of our defaulting counts, we used corpora larger than the child-specific corpora used in the actual simulations. These were, for English, the combined input for the 12 children of the Manchester corpus (Theakston et al., 2001: ~350,000 utterances); for Dutch, the combined input for the 8 children of the Groningen corpus (Bol, 1996: ~80,000 utterances); for German, the child-directed speech of the dense Leo corpus (Behrens, 2006: ~240,000 utterances) and for Spanish, the input for the two children from the OreaPine corpus (Aguado-Orea & Pine, 2015) and the 50 children from the Fern-Aguado corpus (~120,000 utterances combined). These corpora are all available in the CHILDES database (MacWhinney, 2000).

Defaulting Counts

Since defaulting is assumed to reflect competition between finite forms of the verb, defaulting counts were restricted to finite lexical verbs and, for the sake of simplicity, to present-tense verb forms. However, since in English, Dutch, and German, the infinitive is homophonous with one of the present-tense forms, infinitives were also included in the counts when they occurred in a finite utterance (e.g., *He can go there*, *Does he like that?*) and the finite auxiliary was not part of the relevant utterance-final string. This allowed us to investigate how defaulting interacts with MOSAIC's utterance-final bias in learning. In English, Dutch, and German, finite utterances were identified by searching the input for subject pronouns and a relevant verb form in an appropriate position. For Spanish, which allows null subjects, the Mor (morphology) tier of the transcript was used. This procedure allowed for the exclusion of verbs in imperative contexts.

In order to facilitate the collection of defaulting counts for utterance-final strings of different lengths, present-tense verbs in simple-finite contexts were marked as tensed (e.g., *He goes-tensed to school, They go-tensed home*), while forms in imperative contexts were left unmarked (e.g., *Go home*). This made it possible to distinguish the two forms of *go* when analysing two-word utterance-final strings – *go-tensed home* contributed to the counts for *go*, while *go home* was ignored. Infinitives that occurred in a compound-finite context (e.g., *He can go there*, *Does he go home?*) were marked as modal, while the finite auxiliary was marked as tensed (*e.g., He can-tensed go-modal there, Does-tensed he go-modal home?*). Forms marked as modal contributed to the counts for the relevant verb, provided the tensed auxiliary was not part of the relevant utterance-final string. That is, the infinitive form of *go* in *Does-tensed he go-modal home?* contributed to the counts for *go* in two- and three-word utterance-final strings, but not in four-word utterance-final strings. This procedure is designed to simulate the child's increasing sensitivity to the fact that, in compound-finite contexts, tense is marked on the auxiliary rather than on the lexical verb.

Table 1 illustrates the present-tense verb paradigm in English, Dutch, German, and Spanish using the verb *run,* which is regular in all four languages. Pronouns are not included for Spanish as Spanish is a pro-drop language. It can be seen from Table 1 that the English present-tense paradigm comprises 2 forms, one of which matches the infinitive, the Dutch paradigm comprises 3 forms, one of which matches the infinitive, the German paradigm comprises 4 forms, one of which matches the infinitive, and the Spanish paradigm comprises 6 forms, none of which matches the infinitive. Since the defaulting counts collapse across matching forms, this means that, all other things being equal, defaulting is likely to be most pervasive in English and least pervasive in Spanish, with Dutch and German falling somewhere in between. However, since defaulting is applied on a verb-by-verb basis, it is also possible that different verbs will default to different forms. For example, English verbs that tend to occur in 3sg contexts (e.g., *fits*) may default to the 3sg form.

Table 1

Present-tense paradigm for the verb *run* in English, Dutch, German, and Spanish.

	English	Dutch	German	Spanish
Infinitive	Run	Rennen	Rennen	Correr
1st singular	I run	Ik ren	Ich renne	Corro
2nd singular	You run	Jij rent[i]	Du rennst	Corres
3rd singular	She runs	Zij rent	Sie rennt	Corre
1st plural	We run	Wij rennen	Wir rennen	Corremos
2nd plural	You run	Jullie rennen	Ihr rennt	Corréis
3rd plural	They run	Zij rennen	Sie rennen	Corren

Table 2 illustrates the basic word-order patterns of the four different languages. It can be seen that finite lexical verbs usually occur before their complements in all four languages.

However, the infinitive in compound-finite structures occurs before its complements in English and Spanish, and after its complements in Dutch and German, where it is tied to utterance-final position. This results in infinitives being more common than finite forms in utterance-final position in Dutch and German, which means that, in these languages, defaulting is likely to interact with MOSAIC's utterance-final bias in learning such that the model is more likely to default to the infinitive at low MLUs (when verb counts are drawn from short utterance-final strings) and more likely to default to a finite form at high MLUs. It also means that defaulting has the potential to result in verb-positioning errors in these languages, since defaulting from a finite form to an infinitive will result in an infinitive that precedes its complement and defaulting from an infinitive to a finite form will result in a finite form that follows its complement.

Table 2

Examples of simple-finite and compound-finite constructions in English, Dutch, German, and Spanish.

	Simple-finite
English	I drink coffee
Dutch	Ik drink koffie (I drink-FIN coffee)
German	Ich trinke Kaffee (I drink-FIN coffee)
Spanish	Bebo café ((I) drink-FIN coffee)
	Compound-finite
English	I want to drink coffee
Dutch	Ik wil koffie drinken (I want-FIN coffee drink-INF)
German	Ich moechte Kaffee trinken (I want-FIN coffee drink-INF)

Spanish	Quiero beber café ((I) want-FIN drink-INF coffee)

Defaulting in MOSAIC's Output

Defaulting counts were applied to MOSAIC's output by searching the relevant output file for the occurrence of verb forms considered in the child-directed speech analysis and substituting default forms (e.g., forms that made up more than 65% of the relevant forms in the input) for non-default forms. Thus, the output utterance *he eats-3SG* was changed to *he eat-3SG* if *eat* had been identified as the default form of the verb *eat*. Likewise, the utterance *they fit there* was changed to *they fits there* if *fits* had been identified as the default form of the verb *fit*. The only exceptions to this rule were instances where an infinitive verb form occurred in an utterance with a tensed auxiliary (e.g., *can fit there*). These utterances were left unchanged even if *fits* had been identified as the default form, because they already contain a tensed form. That is, in line with the input analysis (where the phrase *can-tensed go-modal away* did not contribute to the counts for *go* if the modal verb *can* was included in the relevant utterance-final string), it was assumed that, as utterance-length increases, the child becomes increasingly sensitive to the fact that, in compound-finite contexts, tense is marked on the auxiliary rather than on the lexical verb and so does not substitute tensed verb forms for untensed verb forms in compound-finite constructions. Note that this mechanism prevents the model from producing errors such as *kann rent* and *can runs* but still allows the model to produce positioning errors in Dutch and German by substituting tensed verb forms (e.g., *drink, drinkt*) for infinitives (e.g., *drinken)* in utterances from which the tensed auxiliary is absent (e.g., *koffie drinken* → **koffie drinkt*[ii]).

Results

The present paper reports three sets of simulations and analyses. The first set of simulations (Study 1) focuses on the extent to which adding a defaulting mechanism to MOSAIC improves the model's ability to simulate differences in the rate of OI errors in English, Dutch, German, and Spanish at MLU=2. The second set of analyses (Study 2) focuses on how the novel defaulting mechanism interacts with the frequency statistics of child-directed speech in the four languages to result in different levels of defaulting and different types of defaulting error. The third set of simulations (Study 3) focuses on whether it is possible to capture the cross-linguistic pattern of differences in the rate of OI, agreement, and verb-positioning errors in children with DLD relative to MLU-matched controls by changing the model's defaulting threshold. All statistical analyses were conducted in R version 4.1.0 (R Core Team, 2021).

Study 1. Simulating Cross-Linguistic Variation in the Rate of OI Errors at MLU=2

In these simulations, we investigated the extent to which adding a defaulting mechanism to MOSAIC improved the model's ability to simulate the rate of OI errors in English, Dutch, German, and Spanish at MLU=2. This was done by comparing the output of models in the absence of defaulting with the output of models after defaulting had been applied with the defaulting threshold set to .60, .65, and .70 in each of the four languages. Output was analysed in the same way as in Freudenthal et al. (2010) by distinguishing between utterances that contained only a non-finite verb form (e.g., *rennen* in Dutch and *run-3SG*, *running-3SG* in English) and utterances that contained at least one finite verb form (e.g., *rent* and *kan rennen* in Dutch and *runs-3SG*, *can run-3SG* and *is running-3SG in English*). The dependent variable (proportion of OI errors) was the number of utterances in the first category divided by the sum of the utterances in the first and second category. Further details of the coding scheme are provided in Appendix S2 in the supporting materials.

Rates of OI errors in MOSAIC in the absence of defaulting

Table 3 shows the rate of OI errors in MOSAIC's output in the absence of defaulting (together with the corresponding rates for the children on whose input the models were trained, as reported in Freudenthal et al., 2010). It is clear from Table 3 that MOSAIC substantially underestimates the proportion of OI errors in early child English (by 30% for Anne and 38% for Becky). These results replicate those of Freudenthal et al. (2010) and show that, while MOSAIC's edge-based learning mechanism is sufficient to capture differences in the rate of OI errors across Dutch, German, and Spanish, it cannot capture the very high level of OI errors in early child English.

Table 3

Proportion of OI errors in MOSAIC's output at MLU=2 for models trained on English (Anne, Becky), Dutch (Matthijs, Peter), German (Leo), and Spanish (Juan) together with the proportions for the corresponding children.

Child	MLU	Proportion OI errors Child	Proportion OI errors Model	No. of utterances with verbs in model's output
Anne	2.04	.87	.57	99
Becky	2.00	.97	.59	60
Matthijs	2.10	.77	.65	719
Peter	2.06	.74	.65	680
Leo	1.97	.58	.57	1236
Juan	2.01	.20	.12	824

Rates of OI errors in MOSAIC with defaulting based on utterance-final statistics

Table 4 shows the rate of OI errors in MOSAIC's output after defaulting at thresholds of .60, .65, and .70 based on utterance-final words and two-word strings. These data suggest that defaulting using utterance-final statistics results in a better fit to the child data than was obtained using the previous version of the model, with differences in the defaulting threshold having little effect on the overall pattern of results.

This was confirmed by running arcsine transformations on the child and model rates reported in Tables 3 and 4 and computing Pearson correlations (with Bayes Factors). This analysis revealed a marginally significant correlation between the child and model rates for the old version of the model (r = .808, p = .052, BF = 1.84, indicating inconclusive support for the hypothesised relation), and significant correlations for each of the new versions of the model (r = .948, p = .004, r = .950, p = .004, and r = .962, p = .002) with Bayes Factors of 3.76, 3.81 and 4.23, respectively, all indicating moderate support for the hypothesised relation. A comparison of the data in Tables 3 and 4 reveals that the improvement in fit is mainly due to a substantial increase in the proportion of OI errors in English (of 34% for Anne's model and 26% for Becky's model). However, it also reflects a smaller increase in the proportion of OI errors in Dutch (of 5% for Matthijs's model and 7% for Peter's model), with the proportion of OI errors in German and Spanish being largely unaffected. These results show that combining MOSAIC's utterance-final bias with a mechanism that defaults to the most frequent form of the verb provides a better explanation of the cross-linguistic data.

Table 4

Proportion of OI errors in MOSAIC's output at MLU=2 for models trained on English (Anne, Becky), Dutch (Matthijs, Peter), German (Leo), and Spanish (Juan) input with defaulting at thresholds of .60, .65, and .70 based on utterance-final words and two-word strings. (Proportion of affected utterances in parentheses.)

Child	MLU	Proportion OI errors Child	Proportion OI errors Model Threshold=.60	Proportion OI errors Model Threshold=.65	Proportion OI errors Model Threshold=.70
Anne	2.04	.87	.91 (.34)	.91 (.34)	.80 (.23)
Becky	2.00	.97	.85 (.28)	.85 (.28)	.83 (.27)
Matthijs	2.10	.77	.70 (.08)	.70 (.08)	.70 (.08)
Peter	2.06	.74	.72 (.09)	.72 (.09)	.71 (.08)
Leo	1.97	.58	.60 (.08)	.59 (.06)	.59 (.05)
Juan	2.01	.20	.12 (.04)	.12 (.03)	.12 (.02)

Rates of OI errors in MOSAIC with defaulting based on corpus-wide statistics

Table 5 shows the rate of OI errors in MOSAIC's output after defaulting at thresholds of .60, .65, and .70 based on utterance-final strings of up to 10 words in length (effectively based on a corpus-wide analysis). These data are interesting because they suggest that defaulting using corpus-wide statistics results in a poorer fit to the child data than defaulting using utterance-final statistics. This is partly because it results in a much less pronounced increase in the proportion of OI errors in English (of 11% for both Anne and Becky's models compared to 34% and 26%, respectively, for the previous models). However, it is also because it results in a decrease in the proportion of OI errors in Dutch and German (of 16% for Matthijs's model, 13% for Peter's model and 6% for Leo's model). This reduction in the fit to the Dutch and German data is not particularly surprising since, in both languages, although the infinitive is the most common form of the verb in utterance-final position, it is not the most common form of the verb in the input as a whole. However, it does underline the need to link defaulting to the model's utterance-final bias in learning to explain the cross-linguistic data. That is, it suggests that it is necessary to assume that the same utterance-final bias that shapes the development of the model's representations also affects its sensitivity to the relative frequency of different verb forms in the input.

Table 5

Proportion of OI errors in MOSAIC's output at MLU=2 for models trained on English (Anne, Becky), Dutch (Matthijs, Peter), German (Leo), and Spanish (Juan) with defaulting at thresholds of .60, .65, and .70 based on utterance-final strings of up to 10 words in length. (Proportion of affected utterances in parentheses).

Child	MLU	Proportion OIs Child	Proportion OIs Model Threshold=.60	Proportion OIs Model Threshold=.65	Proportion OIs Model Threshold=.70
Anne	2.04	.87	.69 (.27)	.68 (.26)	.67 (.24)
Becky	2.00	.97	.67 (.33)	.70 (.30)	.70 (.30)
Matthijs	2.10	.77	.49 (.22)	.49 (.21)	.52 (.15)
Pet	2.06	.74	.52 (.17)	.52 (.18)	.54 (.12)
Leo	1.97	.58	.51 (.08)	.51 (.06)	.52 (.06)
Juan	2.01	.20	.12 (.04)	.12 (.03)	.12 (.02)

In summary, the simulations presented above show that adding a defaulting mechanism to MOSAIC allows the model to simulate the very high rate of OI errors in early child English without affecting the model's previously good fit to the data on Dutch, German and Spanish. However, they also show that this is only the case when the defaulting counts are based on utterance-final phrases matched to the model's MLU. They therefore underline the important role played by MOSAIC's utterance-final bias in explaining the developmental data.

Study 2. Defaulting as a Function of the Statistics of Child-Directed Speech in the Four Languages

In these analyses, we investigated how the defaulting mechanism interacted with the frequency statistics of child-directed speech in the four languages. This was done by setting the defaulting parameter to .65 and exploring the pattern of defaulting and the proportion of affected verbs for defaulting counts based on utterance-final strings of different lengths in the child-directed speech corpora.

Tables 6 to 9 show the results of these analyses for each of the four languages. Results are expressed as the proportion of verbs that would be subject to defaulting, based on a threshold of .65. Results are shown for utterance-final words and utterance-final strings of 2, 3, 5, and 10 words. Complete utterances are included in string sets that exceed their length. Since utterances of more than 10 words in length are rare in child-directed speech, the analysis of 10-word utterance-final strings is, in effect, a corpus-wide analysis.

Results for the English input analysis are presented in Table 6. It is clear from these data that most English verbs would default to the bare form (i.e., occur as a bare form in over 65% of tensed contexts). Fewer verbs would default to the bare form in longer strings. This is because untensed forms in compound structures contribute to the counts for short, but not for longer strings, which are more likely to include both a tensed auxiliary and an untensed lexical verb (e.g., *That might-tensed go-modal there*).

Table 6:

Proportion of verbs that would default to a particular form of the verb in English at different maximum string-lengths.

String-length	Bare Form	3rd Sg.	No default	N
1	.94	.01	.06	108
2	.96	.01	.03	181

3	.94	.02	.04	215
5	.89	.05	.07	213
10	.82	.06	.12	195

Nevertheless, even in the 10-word string analysis, more than 80% of verbs would default to the bare form. This reflects the fact that, in English, zero-marked forms like *I go* and *you go* are far more frequent than overtly-tensed forms like *he goes*, and explains why adding a defaulting mechanism to the model has such a profound effect on the rate of OI errors in English. It also suggests that allowing the model to default at high MLUs (based on statistics from longer utterance-final strings) may be an effective way of simulating the higher rate of OI errors in English-speaking children with DLD relative to MLU-matched controls. This is because defaulting based on statistics from longer utterance-final strings will increase the rate of OI errors in the model's output without affecting the model's MLU.

Results for the Dutch and German analyses are presented in Tables 7 and 8. It is clear from these data that fewer verbs would default to the infinitive in Dutch and German than would default to the bare form in English, regardless of string-length. However, it is also clear that the most common default form in Dutch and German changes as string-length increases. The infinitive (stem+en) form is the most common default for utterance-final words and two- and three-word strings, but for longer strings the 1^{st} and 2^{nd} singular (stem) in Dutch and the 3^{rd} singular (stem+t) in German are the most common defaults.

Table 7

Proportion of verbs that would default to a particular form of the verb in Dutch at different maximum string-lengths.

String-length	Infinitive (Stem+en)	1st/2nd SG (Stem)	2nd/3rd SG (Stem+t)	No default	N
1	.82	.05	.04	.09	78
2	.57	.14	.09	.20	93
3	.39	.20	.09	.31	99
5	.12	.26	.14	.49	101
10	.00	.36	.19	.46	101

Table 8

Proportion of verbs that would default to a particular form of the verb in German at different maximum string-lengths.

String-length	Infinitive (Stem+en)	1st SG (Stem+e)	2nd SG (Stem+st)	3rd SG (Stem+t)	No default	N
1	.79	.00	.02	.08	.12	168
2	.63	.03	.02	.12	.21	196
3	.49	.02	.01	.16	.31	229
5	.21	.03	.03	.21	.53	265
10	.10	.04	.04	.28	.55	283

This pattern reflects the SOV/V2 nature of Dutch and German, where non-finite forms (including the infinitive) take utterance-final position, whereas finite forms take second position and precede their complements. Non-finite forms are therefore more likely to occur in short utterance-final strings, while finite forms are more likely to occur in longer strings.

The fact that fewer Dutch and German verbs show a strong preference for the infinitive explains why the model’s defaulting mechanism has less effect on the rate of OI errors in

Dutch and German than it does in English; and the fact that the pattern of preference changes as a function of string-length explains why it is necessary to link defaulting to the model's utterance-final bias in learning in order to explain the Dutch and German data. However, these differences between Dutch and German and English also suggest that allowing the model to default at high MLUs (based on statistics from longer utterance-final strings) is likely to have a different effect in Dutch and German than it does in English. This is because, in Dutch and German, it is likely to result in the replacement of low-frequency finite forms with high-frequency *finite forms* rather than the replacement of finite forms with infinitives (as is the case in English). It may therefore provide a way of simulating the fact that, in contrast to English-speaking children with DLD, Dutch- and German-speaking children with DLD tend to produce agreement and verb-positioning errors rather than OI errors at high MLUs.

Results for the Spanish input analysis are presented in Table 9. Plural forms are not included in the table because there were no verbs that would default to a plural verb form at any string-length. Compared to Dutch and German, the pattern in Spanish is relatively stable. Most verbs would not be subject to defaulting, regardless of string-length. However, 15 to 25% would default to the 3sg form.

This pattern of defaulting is likely to result in the replacement of low-frequency finite forms with high-frequency 3sg forms rather than the replacement of finite forms with infinitives, and explains why the model's defaulting mechanism has so little effect on the rate of OI errors in Spanish. It is also consistent with the pattern of verb-marking error reported in early child Spanish, in which children tend to make agreement errors at relatively low rates, most of which involve the inappropriate use of 3sg forms.

Table 9:

Proportion of verbs that would default to a particular form of the verb in Spanish at different maximum string-lengths.

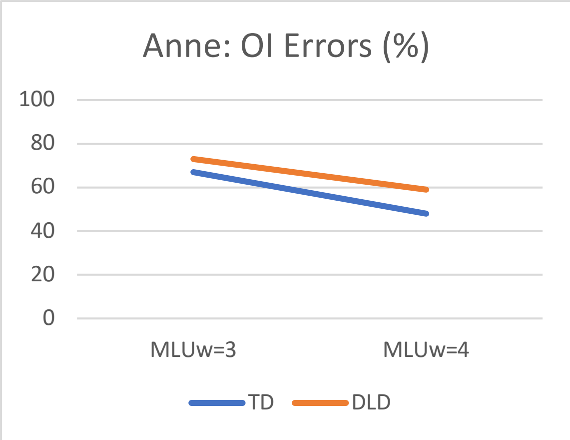
String-length	Infinitive	1st SG	2nd SG.	3rd SG	No default	N
1	.15	.02	.03	.17	.63	147
2	.10	.01	.02	.15	.71	163
3	.07	.01	.02	.21	.70	179
5	.04	.01	.02	.21	.72	189
10	.04	.01	.02	.25	.68	193

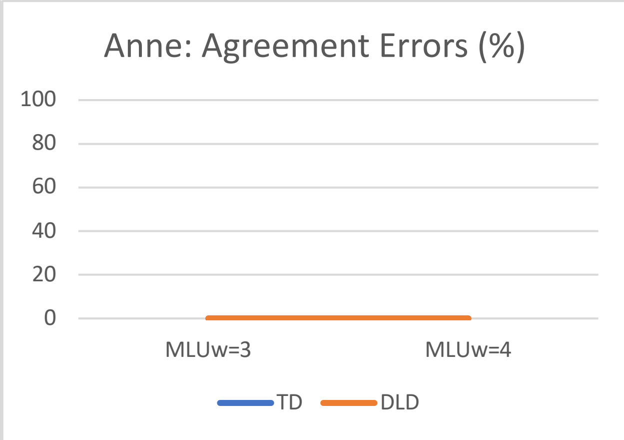
Finally, this pattern of defaulting suggests that allowing the model to default at high MLUs (based on statistics from longer utterance-final strings) would simply prolong the period during which agreement errors were made, which is broadly consistent with the data on verb-marking error in Spanish-speaking children with DLD, who tend to show slightly higher rates of agreement error relative to age-matched, but not MLU-matched controls (Bedore & Leonard, 2001).

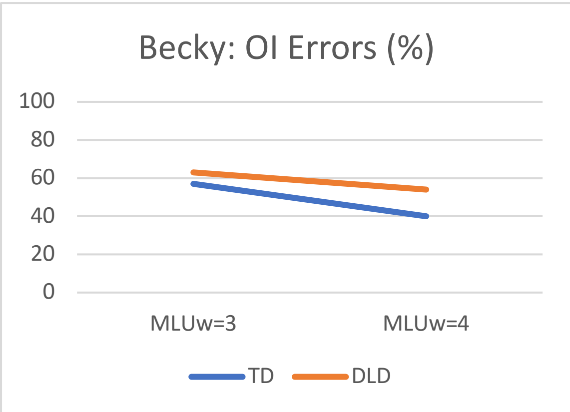
In summary, the input analyses presented above reveal that English verbs show an overwhelming preference for the bare form, which is consistent across different string-lengths. Fewer verbs show a strong preference in Dutch and German, and the preferred form changes as a function of string-length. In short strings, there tends to be a preference for the infinitive, which occurs in utterance-final position; in longer strings there tends to be a preference for finite forms, which occur in V2. Even fewer verbs show a strong preference in Spanish, but where they do, this tends to be a preference for the 3sg form. These input analyses therefore explain why defaulting has a large effect on the rate of OI errors in English, a smaller effect in Dutch and German and virtually no effect in Spanish. They also suggest that defaulting at higher MLUs will tend to increase the rate of OI errors in English, but not in Dutch, German, or Spanish, where it is likely to result in defaulting to the highest frequency finite form and hence to agreement and positioning rather than OI errors.

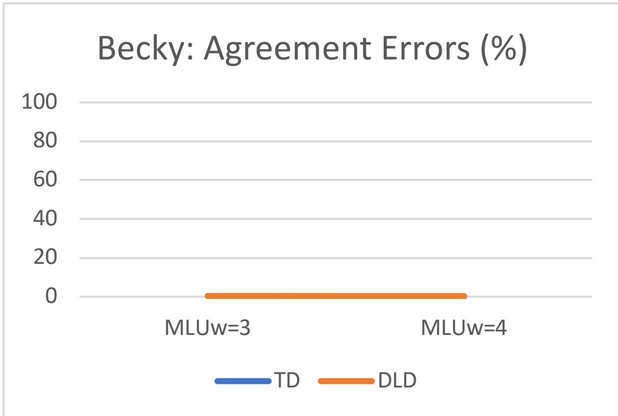
Study 3. Simulating the Cross-Linguistic Pattern of Verb-Marking Error in Children with DLD

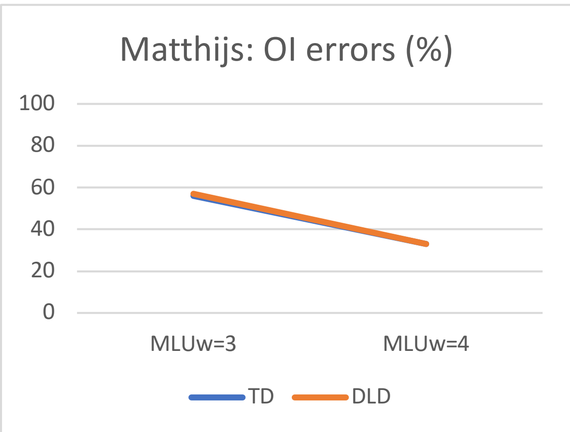
In these simulations, we investigated the model's ability to capture the cross-linguistic pattern of verb-marking error in typically-developing children and children with DLD. This was done by applying different defaulting thresholds to the output from each of the children's models at MLU=3 (based on utterance-final strings of up to 3 words) and MLU=4 (based on utterance-final strings of up to 4 words). The thresholds for the models with DLD were .65 and .75, respectively. The thresholds for the typically-developing models were .85 and .95, respectively. Since a lower defaulting threshold will result in higher levels of defaulting error, these values were chosen to result in higher levels of defaulting in the DLD than the TD models and decreasing levels of defaulting as a function of MLU. Rates of OI errors were calculated in the same way as in the first set of analyses. Rates of agreement error were calculated by identifying cases in which defaulting led to the substitution of one finite form for another (e.g., *rent* for *ren* in Dutch and *runs* for *run* in English) and dividing the number of such cases by the total number of simple-finite contexts. Rates of verb-positioning errors in Dutch and German were calculated by identifying cases in which defaulting led to the substitution of an infinitive into a finite context and cases in which defaulting led to the substitution of a finite form into an infinitival context. In each case, the denominator was the sum of the number of such cases and the number of correctly placed infinitives or finite forms.

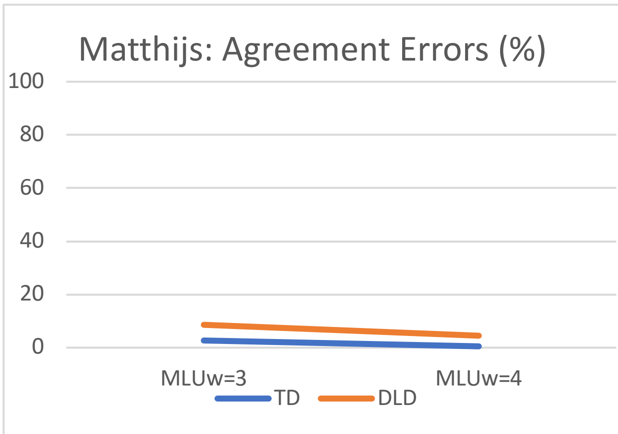

Anne: OI Errors (%)
100
80
60
40
20
0
MLUw=3
MLUw=4
TD
DLD


Anne: Agreement Errors (%)
100
80
60
40
20
0
MLUw=3
MLUw=4
TD
DLD


Becky: OI Errors (%)
100
80
60
40
20
0
MLUw=3
MLUw=4
TD
DLD


Becky: Agreement Errors (%)
100
80
60
40
20
0
MLUw=3
MLUw=4
TD
DLD


Matthijs: OI errors (%)
100
80
60
40
20
0
MLUw=3
MLUw=4
TD
DLD


Matthijs: Agreement Errors (%)
100
80
60
40
20
0
MLUw=3
MLUw=4
TD
DLD

(Panel a) (Panel b)

Figure 1. Rates of OI errors (Panel a) and Agreement errors (Panel b) for each TD model and the equivalent model with a lower defaulting threshold (DLD). Raw data are provided in Appendix S3, Table 1.

Figure 1 shows the rates of OI errors (panel a) and agreement errors (panel b) in each TD model and the equivalent model with a lower defaulting threshold (DLD). Figure 1 suggests that reducing the defaulting threshold at higher MLUs does have an effect on the rate of OI errors in the English models, resulting in increases of 6% at MLU=3 and 11 and 14% at MLU=4, but has little or no effect on the rate of OI errors in the Dutch, German and Spanish models.

These data were analysed by running a mixed-effects Poisson regression model with a random effect of modelled-child on the intercept, and fixed effects of Error (OI error, Correct), Model-Type (TD, DLD), Language (English, Non-English) and MLU (3, 4), plus all two-way interactions between Error, Model-Type and Language, and the critical three-way interaction between, Error, Model-Type and Language. All binary variables were coded as -0.5 and 0.5. All fixed effects were significant (see Table 2 in Appendix S3), including the critical three-way interaction between Error, Model-Type and Language ($\beta = 0.457$, $SE = 0.037$, $z = 12.20$, $p < .001$), which indicates that the difference in the relative frequency of OI errors versus correct utterances was greater in the English DLD vs. TD models than in the non-English DLD vs. TD models (where it was essentially zero). Importantly, a model including the three-way interaction gave a better fit than any sub-model (delta-AIC = 146.52 for the next best model)[iii].

These results confirm that reducing the defaulting threshold at high MLUs allows the model to simulate the increased rate of OI errors relative to MLU-matched controls that is seen in English-speaking children with DLD, but not seen in Dutch-, German-, and Spanish-speaking children. In the Spanish models, this is a straightforward consequence of the fact that defaulting to the infinitive is extremely rare. In Dutch and German, it reflects the fact that, although some verbs do still default from a finite to the infinitive form at high MLUs, other verbs default in the opposite direction, cancelling out any potential increase.

Figure 1 also suggests that, in contrast to the increase in OI rates in the English models, reducing the defaulting threshold at higher MLUs in the Dutch, German, and Spanish models results in increased rates of agreement error. These increases can be seen in all 3 languages at both MLU points. However, they result in relatively low overall error rates (never greater than 10%). This pattern is also consistent with the cross-linguistic literature on DLD, which reports elevated, but still relatively low, rates of agreement error in Dutch-, German-, and Spanish-speaking children.

These data were analysed by running a mixed-effects Poisson regression model on the non-English data, with a random effect of modelled-child on the intercept, and fixed effects of Error (Agreement error, Correct), Model-Type (TD, DLD), and MLU (3, 4), and the critical two-way interaction between, Error and Model-Type. All binary variables were coded as -0.5 and 0.5. All fixed effects were significant (see Table 3 in Appendix S3), including the critical two-way interaction between Error and Model-Type ($\beta = 1.600$, $SE = 0.065$, $z = 24.39$, $p < .001$), which indicates that the difference in the relative frequency of Agreement errors versus correct utterances was greater in the DLD than the TD models. Importantly, a model including the two-way interaction gave a better fit than any sub-model (delta-AIC = 778.58 for the next best model).

Finally, as noted earlier, defaulting in Dutch and German has the potential to result in positioning errors in which infinitives occur in V2 and finite forms occur in inappropriate utterance-final contexts. The rates at which such errors occur in the Dutch and German models are reported in Table 10. It can be seen from these data, that although positioning errors are relatively rare (ranging from 0 to 2.1% in the TD models and 2.0 to 6.4% in the DLD models), they are more common in the DLD than the TD models at both MLU points.

These results were analysed by running separate mixed-effects Poisson regression models with random effects of modelled-child on the intercept, and fixed effects of Error (Positioning error, Correct), Model-Type (TD, DLD), and MLU (3, 4), plus the critical two-way interaction between, Error, and Model-Type. In both models, all fixed effects were significant (see Tables 4a and 4b in Appendix S3), including the critical two-way interactions between Error and Model-Type (β = 1.816, *SE* = 0.125, z = 14.50, p < .001 and β = 1.707, *SE* = 0.133, z = 12.87, p < .001), indicating that the difference in the relative frequency of Positioning errors and correct utterances was greater in the DLD models. Importantly, models including these two-way interactions gave a better fit than any sub-models (delta-AICs = 294.91 and 28.63, for the next best models).

These results are consistent with the literature on child Dutch and German, where positioning errors are rare in TD children, particularly at high MLUs, but occur at elevated, though still relatively low, rates in children with DLD. They thus provide further support for the idea that a model in which defaulting occurs at different rates in impaired and unimpaired children provides a plausible account of the cross-linguistic pattern of verb-marking error in TD children and children with DLD.

Table 10

Rates of Positioning Errors in the Dutch (Matthijs and Peter) and German (Leo) models and the equivalent models with a lower defaulting threshold (DLD)

Model	Finites in Final Position	Finites in V2	%Error	Infinitives in V2	Infinitives in Final Position	%Error
Mathijs MLU=3	10	817	1.2	15	1396	0.1
Mathijs-DLD MLU=3	46	761	5.7	46	1381	3.2
Mathijs MLU=4	4	2164	0.2	0	1571	0

Mathijs-DLD MLU=4	54	2105	2.5	32	1550	2.0
Peter MLU=3	8	1306	0.6	33	1536	2.1
Peter-DLD MLU=3	82	1209	6.4	74	1497	4.7
Peter MLU=4	18	2983	0.6	0	1304	0
Peter-DLD MLU=4	58	2897	2.0	76	1268	5.7
Leo MLU=3	17	2881	0.6	17	2871	0.6
Leo-DLD MLU=3	79	2776	2.8	72	2850	2.5
Leo MLU=4	18	6081	0.3	3	3463	0.1
Leo-DLD MLU=4	130	6059	2.1	67	3342	2.0

In summary, modelling the verb-marking deficit in DLD in terms of an increased tendency to default to the most common form of the verb captures both the tendency of English-speaking children with DLD to produce OI errors at higher rates than MLU-matched controls and the tendency of Dutch-, German-, and Spanish-speaking children with DLD to show problems with subject-verb agreement, and Dutch- and German-speaking children to show problems with verb placement. It thus suggests that MOSAIC+ has the potential to explain both the cross-linguistic pattern of verb-marking error in TD children and the cross-linguistic pattern of verb-marking deficit in children with DLD.

Discussion

The aim of the present study was to investigate whether a model which supplements MOSAIC's basic learning mechanism with a mechanism that defaults to the most frequent form of the verb provides both a better explanation of the cross-linguistic data on TD children and a means of simulating the cross-linguistic pattern of verb-marking deficit in children with DLD. In a first set of analyses, we investigated the extent to which adding a defaulting mechanism to MOSAIC improved the model's ability to explain the cross-linguistic pattern of OI errors in TD children at MLU=2. Our results show that the addition of a defaulting mechanism allows MOSAIC to simulate the very high rate of OI errors in early child English without affecting the model's previously good fit to the data on Dutch, German, and Spanish. These findings are consistent with the idea that at least some apparent OI errors in English reflect a process of defaulting to the most frequent form of the verb (Räsänen et al., 2014; Kueser et al., 2018), and suggest that the very high rate of OI errors in English reflects the fact that, in English, but not in the other languages, defaulting tends to result in the same kind of errors as the learning of infinitives directly from the input.

In a second set of analyses, we investigated how the novel defaulting mechanism interacted with the frequency statistics of child-directed speech in the 4 languages across utterance-final strings of different lengths. The results of these analyses show that defaulting is likely to result in bare-stem errors in English and 3sg errors in Spanish regardless of string-length. However, they also show that defaulting is likely to result in different kinds of errors in Dutch and German, depending on the length of the utterance-final strings on which defaulting counts are based. Thus, defaulting based on short utterance-final strings is likely to result in OI errors and infinitives in V2, whereas defaulting based on longer utterance-final strings is likely to result in agreement errors and finite forms in utterance-final position. These findings show that defaulting can explain why English-speaking children tend to make bare-stem errors in their speech, whereas Spanish children tend to make 3sg errors. They also show why it is necessary to allow defaulting to interact with MOSAIC's utterance-final bias in learning to explain the Dutch and German data at MLU=2. The reason is that defaulting based on corpus-wide statistics would reduce the rate of OI errors in the model's output by reducing the number of infinitives and increasing the number of finite forms. Finally, they suggest that defaulting based on the statistics of longer utterance-final strings will tend to result in different patterns of error in English than in Spanish, Dutch, and German, with bare-stem errors being the most common type of error in English and agreement errors being the most common type of error in Spanish, Dutch, and German.

In a final set of analyses, we manipulated the defaulting parameter in the model in order to simulate the cross-linguistic pattern of differences in the rate of OI, agreement and verb-positioning errors in typically-developing children and children with DLD. As is clear from our results, increasing the amount of defaulting at high MLUs by lowering the defaulting threshold allows the model to simulate the higher rate of OI errors in English-speaking children with DLD and the absence of this effect in Dutch-, German-, and Spanish-speaking children. It also allows the model to simulate the increased rate of agreement errors in Dutch-, German-, and Spanish-speaking children with DLD and the increased rate of verb-positioning errors in Dutch- and German-speaking children. An important feature of these simulations is that, although significantly higher in the DLD than the TD models, the rates of agreement and positioning errors are never unrealistically high (i.e., never greater than 10%). Note that this feature of the data is a straightforward consequence of the use of a frequency-sensitive defaulting mechanism which, by its very nature, only results in defaulting errors when the target is a relatively low-frequency form of the verb. This tends to result in low overall error rates which hide higher error rates in low-frequency parts of the system. Interestingly, this is exactly the pattern of error reported in detailed analyses of the speech of children learning more highly inflected languages (e.g., Aguado-Orea & Pine, 2015; Engelmann et al. 2019).

Overall, these findings suggest that a model in which defaulting occurs at different rates in impaired and unimpaired children provides a plausible account of the cross-linguistic pattern of verb-marking error in TD children and the cross-linguistic pattern of verb-marking deficit in children with DLD. They are also consistent with a wealth of evidence that, while frequency at both the word and sequence level can increase fluency and protect items from error, it can also result in errors in which low-frequency items are replaced by higher-frequency items (see Ambridge et al., 2015, for a review). The implication is that the verb-marking deficit in DLD reflects a system that is particularly susceptible to intrusions from high-frequency items.

Limitations and Future Research Directions

It is worth noting at this point that, since different levels of defaulting are implemented in the current model by directly manipulating the defaulting parameter, our findings still leave unanswered the question of what underlying mechanism is responsible for the different levels of defaulting seen in TD children and children with DLD. One possibility that maps more or less directly onto the way defaulting is implemented in the current version of the model is that greater defaulting in DLD reflects a deficit in the ability to inhibit competition from higher-frequency forms (see McMurray et al., 2019, for an explanation of lexical deficits in DLD in terms of reduced lexical inhibition). A second possibility is that greater defaulting reflects a deficit in word-learning and paradigm-building. According to this view, greater defaulting reflects an underlying deficit in the ability to learn low-frequency forms and morphological patterns that leaves the child with DLD more susceptible to competition from high-frequency forms of the verb (see Harmon et al., in press, for an account of deficits in past-tense marking along these lines). And a third possibility is that greater defaulting reflects a deficit in the ability to process long-distance dependencies that differentiate between contexts that require lower- and higher-frequency forms. According to this view, children with DLD use the most frequent form of the verb because they have yet to distinguish between contexts that require a lower-frequency form of the verb (e.g., *Dolly sits there)* and contexts that require a higher-frequency form of the verb *(*e.g., *Does Dolly sit there?* or *We let Dolly sit there*). This leaves children with DLD susceptible to competition from higher-frequency forms of the verb for longer than TD children (see Leonard et al., 2015, for a more detailed description of this Competing Sources of Input account, and Freudenthal et al., 2021, for a model of the verb-marking deficit in DLD which shows how a deficit in the ability to take account of information in the preceding context interacts with the distributional properties of English and Spanish to result in a greater verb-marking deficit in English than in Spanish). Determining which of these mechanisms provides the most plausible account of the increased level of defaulting in DLD is clearly beyond the scope of the present study — and, given the multi-faceted nature of DLD, it is possible that all of them may have some role to play. However, it opens up a number of avenues for future research which have the potential to further increase our understanding of the factors that underlie the verb-marking deficit in children with DLD.

Conclusion

This study shows that a new version of MOSAIC that defaults to the highest-frequency form of the verb can explain both the cross-linguistic pattern of OI errors in TD children and the cross-linguistic pattern of verb-marking deficit in children with DLD. This model has several advantages over previous models of verb-marking error. First, it can explain the very high rate of OI errors in early child English. Second, it can explain why children learning languages other than English tend to make both OI and agreement errors in their speech. Third, it can explain why English-speaking children with DLD produce OI errors at higher rates than MLU-matched controls, whereas children learning other languages tend to make more agreement and positioning errors.

References

Aguado-Orea, J. & Pine, J.M. (2015). Comparing different models of the development of verb inflection in early child Spanish. *PloS One*, *10*(3), e0119613. https://doi.org/10.1371/journal.pone.0119613

Ambridge, B., Kidd, E., Rowland, C.F., & Theakston, A.L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, *42*, 239-273. https://doi.org/10.1017/S030500091400049X

Bedore, L.M. & Leonard, L.B. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *44*, 905-924. https://doi.org/10.1044/1092-4388(2001/072)

Behrens, H. (2006). The input-output relationship in first language acquisition. *Language and Cognitive Processes, 21*, 2–24. https://doi.org/10.1080/01690960400001721

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry, 21*, 491-504.

Bol, G.W. (1996). Optional subjects in Dutch child language. In C. Koster & F. Wijnen (Eds.), *Proceedings of the Groningen Assembly on Language Acquisition* (pp. 125-135). Groningen, The Netherlands: Centre for Language and Cognition.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics, 10*, 151-171. https://doi.org/10.1016/S0911-6044(97)00006-7

De Jong, J. (1999). Specific language impairment in Dutch: Inflectional morphology and argument structure. *Groningen Dissertations in Linguistics*.

Divjak, D. & Caldwell-Harris, C. (2015). Frequency and entrenchment. In E. Dąbrowska & D. Divjak (Eds.). *Handbook of Cognitive Linguistics* (pp. 53-75). Berlin: De Gruyter.

Engelmann, F., Kolak, J., Granlund, S., Szreder, M., Ambridge, B., Pine, J., Theakston, A., & Lieven, E. (2019). How the input shapes the acquisition of verb morphology: Elicited production and computational modelling in two highly inflected languages. *Cognitive Psychology*, *110*, 30-69. https://doi.org/10.1016/j.cogpsych.2019.02.001

Freudenthal, D., Pine, J.M., & Gobet, F. (2006). Modelling the development of children's use of optional infinitives in Dutch and English using MOSAIC. *Cognitive Science*, *30*, 277-310. https://doi.org/10.1207/s15516709cog0000_47

Freudenthal, D., Pine, J.M., & Gobet, F. (2009). Simulating the referential properties of Dutch, German and English root infinitives. *Language Learning and Development*, *5*, 1-29. https://doi.org/10.1080/15475440802502437

Freudenthal, D., Pine, J.M., & Gobet, F. (2010). Explaining quantitative variation in the rate of Optional-Infinitive errors across languages: A comparison of MOSAIC and the Variational Learning Model. *Journal of Child Language*, *36*, 643-669. https://doi.org/10.1017/S0305000909990523

Freudenthal, D., Pine, J.M., Aguado-Orea, J., & Gobet, F. (2007). Modelling the developmental patterning of finiteness marking in English, Dutch, German and Spanish using MOSAIC. *Cognitive Science*, *31*, 311-341. https://doi.org/10.1080/15326900701221454

Freudenthal, D., Pine, J.M., Jones, G., & Gobet, F. (2015). Simulating the cross-linguistic pattern of Optional-Infinitive errors in children's declaratives and Wh-questions. *Cognition*, *143*, 61-76. https://doi.org/10.1016/j.cognition.2015.05.027

Freudenthal, D., Ramscar, M., Leonard, L.B., & Pine, J.M. (2021). Simulating the acquisition of verb inflection in typically-developing children and children with Developmental Language Disorder in English and Spanish. *Cognitive Science*, 45(3), e12945. https://doi.org/10.1111/cogs.12945

Hamann, C., Penner, Z. & Lindner, K. (1998). German impaired grammar: The clause structure revisited. *Language Acquisition*, *7*, 193-245. https://doi.org/10.1207/s15327817la0702-4_5

Harmon, Z., Barak, L., Shafto, P., Edwards, J., & Feldman, N.H. (in press). A competition-compensation account of Developmental Language Disorder. *Developmental Science*.

Kueser, J.B., Leonard, L.B., & Deevy, P. (2018). Third-person singular -s in typical development and specific language impairment: Input and neighbourhood density. *Clinical Linguistics & Phonetics. 32*(3)*,* 232-248. https://doi.org/10.1080/02699206.2017.1342695

Legate, J.A. & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, *14*(3), 315-344.

Leonard, L.B. (2014). *Children with specific language impairment* (2nd Edition). Cambridge, MA: MIT Press.

Leonard, L.B., Fey, M., Deevy, P., & Bredin-Oja, S. (2015). Input sources of third-person singular –s inconsistency in children with and without specific language impairment. *Journal of Child Language, 42,* 786-820. https://doi.org/10.1017/S0305000914000397

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd Edition). Mahwah, NJ: Erlbaum.

McMurray, B., Klein-Packard, J., & Tomblin, J.B. (2019). A real-time mechanism underlying lexical deficits in developmental language disorder: Between-word inhibition. *Cognition*, *191*, 104000. https://doi.org/10.1016/j.cognition.2019.06.012

Pine, J.M., Freudenthal, D., & Gobet, F. (2020). Understanding the cross-linguistic pattern of verb-marking error in typically-developing children and children with Developmental Language Disorder: Why the input matters. In C.F. Rowland, B. Ambridge, A.L. Theakston, & K.E. Twomey (Eds.), *Current perspectives on child language acquisition:*

How children use their environment to learn (pp. 221-246). Amsterdam: John Benjamins. https://doi.org/10.1075/tilar.27.10pin

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Räsänen, S.H.M., Ambridge, B., & Pine, J.M. (2014). Infinitives or bare stems? Are English-speaking children defaulting to the highest-frequency form? *Journal of Child Language*, *41*, 756-779. https://doi.org/10.1017/S0305000913000159

Räsänen, S.H.M., Ambridge, B., & Pine, J.M. (2016). Comparing generativist and constructivist accounts of the acquisition of inflectional morphology: An elicited production study of Finnish. *Cognitive Science*, *40*(7), 1704-1738. https://doi.org/10.1111/cogs.12305

Rice, M.L., Noll, K.R., & Grimm, H. (1997). An extended Optional-Infinitive stage in German-speaking children with Specific Language Impairment. *Language Acquisition, 6*, 255-295.

Rice, M.L., Wexler, K., & Cleave, P.L. (1995). Specific Language Impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*, 850-863. https://doi.org/10.1044/jshr.3804.850

Rubino, R. & Pine, J.M. (1998). Subject-verb agreement in Brazilian Portuguese: What overall error-rates hide. *Journal of Child Language*, *25*, 35-59. https://doi.org/10.1017/s0305000997003310

Theakston, A.L., Lieven, E.V.M., Pine, J.M., & Rowland, C.F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, *28*, 127-152. https://doi.org/10.1017/S0305000900004608

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivation in child grammar. In N. Hornstein & D. Lightfoot (Eds.), *Verb Movement* (pp. 305-350). Cambridge: Cambridge University Press.

Wexler, K., Schaeffer, J. & Bol, G. (2004). Verbal syntax and morphology in typically-developing Dutch children and children with SLI: How developmental data can play an important role in morphological theory. *Syntax*, *7*(2), 148-198. https://doi.org/10.1111/j.1467-9612.2004.00006.x

Supporting Information:

Appendix S1: Extended description of MOSAIC

Appendix S2: Further details of OI coding scheme

Appendix S3: Mixed-effects model results tables

---

[i] In Dutch (but not German), the 2nd person singular suffix (-t) is omitted in questions and the resultant form is a bare stem, which is homophonous with the 1st person singular form and therefore boosts the frequency of this form in the input.

[ii] In fact, MOSAIC can also simulate errors like **koffie drinkt* and **Kaffee trinkt* through right-edge learning. This is because, although ungrammatical in main clauses in Dutch and German, such sequences are grammatical at the ends of utterances in subordinate clauses.

[iii] Note that, at the suggestion of a reviewer, we also ran Bayesian alternatives to all the frequentist mixed-effects Poisson regression models in the current paper using the brms package in R employing default priors. In all cases the 95% Confidence Intervals for the critical interaction terms did not cross zero (see Tables 5 to 7B in Appendix S3)