Defaulting effects contribute to the simulation of cross-linguistic differences in Optional Infinitive errors

Daniel Freudenthal, Julian M. Pine
Department of Psychological Sciences, University of Liverpool

Gary Jones
Division of Psychology, Nottingham Trent University

Fernand Gobet
Department of Psychological Sciences, University of Liverpool

Abstract

This paper describes an extension to the MOSAIC model which aims to increase MOSAIC's fit to the cross-linguistic occurrence of Optional Infinitive (OI) errors. While previous versions of MOSAIC have successfully simulated these errors as truncated compound finites with missing modals or auxiliaries, they have tended to underestimate the rate of OI errors in (some) obligatory subject languages. Here, we explore defaulting effects, where the most frequent form of a given verb is substituted for less frequent forms, as an additional source of OI errors. It is shown that defaulting in English tends to result in the production of bare forms that are indistinguishable from the infinitive, while defaulting in Spanish is less pronounced, and tends to result in the production of 3rd person singular forms. Dutch verb forms are dominated by the stem in corpus-wide statistics, and the infinitive in utterance-final position, suggesting defaulting in Dutch may change qualitatively across development. Defaulting is shown to increase MOSAIC's fit to English and Dutch without affecting its already good fit to Spanish, and provides a potential way of simulating the cross-linguistic pattern of verb-marking errors in children with SLI.

Keywords: Language Acquisition; MOSAIC; Optional Infinitive errors; defaulting; SLI

Introduction

Children in many languages have been shown to produce so-called Optional Infinitive (OI) errors, involving the use of non-finite verb forms in contexts in which a finite verb form is obligatory in the adult language. Thus, English-speaking children produce utterances like *He go there*, and Dutch-speaking children produce utterances like *Mama helpen* (Mummy help). While this may not be apparent from the English example (where the infinitive is a zero-marked form), the Dutch example makes it clear that the errors actually involve the use of a non-finite form – in this case the infinitive, which is marked by the infinitival morpheme *–en*. Optional Infinitive errors have attracted a great deal of attention, since they are produced in a large number of (Germanic) languages, but are relatively rare in highly inflected pro-drop languages such as Spanish and Italian.

An influential Generativist account of Optional Infinitive errors is that of Wexler (1994, 1998), who argues that language-learning children have set all the inflectional and clause structure parameters for their language from an early age but are subject to a 'Unique Checking Constraint' which prevents them from checking the D-feature of the subject DP against more than one DP. In obligatory subject languages such as Dutch and English, which require the checking of two D-features (Agreement and Tense), this results in the under-specification of Agreement or Tense and the production of an uninflected verb form. Null subject languages such as Spanish and Italian only require the checking of Tense on finite main verbs, and OI errors are therefore rare in these languages.

An alternative account of OI errors is provided by MOSAIC (Freudenthal et al. 2006, 2007, 2010). MOSAIC, which is implemented as a computational model, instantiates a view of language acquisition as input-driven learning. MOSAIC views language learning as heavily biased towards the ends of utterances and simulates OI errors as *truncated compound finites*: auxiliary/modal + infinitive constructions with a missing modal or auxiliary.

An early version of MOSAIC (Freudenthal et al. 2006, 2007) simulated (English[1]) OI errors directly through its utterance-final bias by omitting the utterance-initial modal/auxiliary from questions: *(Can) he go?* A later version of MOSAIC (Freudenthal et al. 2010, submitted) expanded on these simulations by complementing MOSAIC's utterance-final bias with a small utterance-initial bias. This version is capable of utterance-internal omission, and can produce OI errors by omitting the auxiliary or modal from declarative compound constructions: *He (can) go*. Both versions of MOSAIC have been shown to provide a good fit to the data from a range of languages (English, Dutch, German, French and Spanish). The simulations with MOSAIC show that the key aspects that determine the rates of OI errors in a language are the frequency of compound finite constructions, and the way in which these constructions pattern (see Table 1 for examples of English, Dutch and Spanish simple and compound finite constructions). Thus, MOSAIC simulates the slightly higher rates of OI errors in Dutch compared to German (two languages that are virtually identical in terms of verb placement), as a result of compound finite constructions being slightly more frequent in Dutch. It furthermore

[1] Due to the impoverished morphology of English, the presence of a 3rd singular subject is required for the identification of an OI error. In (many) other languages, the simple occurrence of a form matching the infinitive is sufficient.

simulates the far higher rates of OI errors in Dutch (and German) compared to Spanish, despite the fact that, in Spanish, compound constructions occur at rates that are similar to Dutch/German. The key aspect here is that non-finite forms in Dutch and German occur in utterance-final position, following potential complements, whereas they tend to occur in medial position in Spanish (preceding potential complement). The proportion of verbs in utterance-final position that are non-finite is therefore far higher in Dutch/German, and MOSAIC's utterance-final bias thus results in higher levels of OI errors in Dutch/German compared to Spanish.

Table 1: Simple and compound finite constructions in English, Dutch and Spanish.

Simple Finite
I eat a cookie
Ik eet een koekje (I eat a cookie)
Como una galleta ((I) eat a cookie)
Compound Finite
I want to eat a cookie
Ik wil een koekje eten (I want a cookie eat)
Quiero comer una galleta ((I) want eat a cookie)

However, one problem with the current version of MOSAIC is that it is unable to simulate the very high level of OI errors in early child English. Thus, Freudenthal et al. (2010) report that MOSAIC underestimates English OI rates by about 25 percentage points during the early stages, and suggest that a paradigm-building (cf MacWhinney 1978) or defaulting mechanism may provide an additional source of OI errors over the positional mechanism employed in MOSAIC. In particular, they point out that the impoverished morphological system of English results in a large proportion of present tense forms being indistinguishable from the infinitive. The high rate of OI errors in English children's speech may therefore partially reflect them producing the bare form as a default, and the inclusion of such a mechanism in MOSAIC may boost the model's rate of OI errors and improve the fit to the English data.

Räsänen, Ambridge and Pine (2014) provide support for this suggestion from an elicited production study in which English children were required to produce 3rd singular verb forms. Räsänen et al. found that, across 48 verbs, there was a significant negative correlation between the children's rate of 3rd singular –s production and the proportion of bare vs. 3rd singular forms for these verbs in a representative input sample. Thus, children tended to omit the 3rd singular –s most for those verbs that showed the greatest tendency to occur as bare rather than 3rd singular forms in the input.

Since defaulting effects reflect the frequency distribution of individual verbs, they are likely to differ across different languages. This suggestion is supported by data from Spanish. Aguado-Orea (2004) provides data from a Spanish child who erroneously produced (the frequent) 3rd person singular present tense in 3rd person plural present tense contexts – an error that is currently beyond MOSAIC's capabilities.

Räsänen et al. (in prep) compared elicited production data from Swedish and English children in compound (modal) and simple finite contexts. Unlike in English, the Swedish infinitive is a marked form, which is clearly distinguishable from the highest frequency finite form. Räsänen et al. found that, while English children produced OI errors in both simple and compound finite contexts, Swedish children did so only in compound finite contexts. Data from this study thus suggest that the positional effect implemented in MOSAIC operates in both Swedish and English, while the defaulting effect is much more pronounced in English where the infinitive matches the most frequent finite form.

In summary, while there is strong (cross-linguistic) evidence for OI errors being compound finite structures with missing modals or auxiliaries, children appear to produce a second type of error that contributes to OI errors in some, but not all languages.

This paper aims to extend MOSAIC's existing mechanisms for the simulation of OI errors with a lexically-specific defaulting mechanism that substitutes the most common form of a verb for less common forms where a default forms exists for a given verb. The mechanism will be tested using three languages that have different inflectional paradigms and structural properties: English, Dutch and Spanish. The key question is whether the cross-linguistic differences in the inflectional paradigms and frequencies of verb forms give rise to different patterns of defaulting and whether such a mechanism is able to boost OI levels for English simulations without affecting the good fit to other languages. A second question concerns the plausibility of utterances with defaulted forms. A key characteristic of Dutch children in the OI stage, is that, when they produce non-finite forms, they overwhelmingly place them correctly in utterance-final position (rather than in finite V2 position). Since a defaulting mechanism does not affect verb position, it could potentially result in utterances that have non-finite verb forms in finite position, a feature that is rare in child speech.

Input Analysis

In order to determine the potential effects of defaulting across the three languages, corpora of child-directed speech were analysed to derive counts for the different verb inflections. There are substantial differences between the inflectional paradigms of English, Dutch and Spanish. The English (present tense) paradigm consists of two forms: the bare form (which matches the infinitive) and the 3rd person singular. The Dutch paradigm contains three forms. The stem (1st person singular), stem+t (2nd and 3rd person singular)[2] and stem +en (all 3 plurals). Dutch plural forms match the infinitive. In Spanish, all present tense forms are different, and none of them matches the infinitive. The

[2] The –t suffix on the 2nd person singular is dropped in questions where the subject follows the verb, and the resulting form is the stem.

Spanish paradigm (including the infinitive) thus has 7 distinct forms, compared to 3 for Dutch and 2 for English.

Counts were collected from a range of speakers. For English, the adult speech directed at all (12) children in the Manchester corpus (Theakston et al. 2001) was pooled. For Dutch, the pooled data from the Groningen corpus (Bol, 1996) was used. The Spanish counts were derived from the corpora of Juan and Lucia from the Nottingham corpus (Aguado-Orea, 2004) and combined with those of the Fern-Aguado corpus.

The main aim of the input analysis was to arrive at a distribution of verb forms in contexts that require an inflected verb form. In English and Dutch, this was done by restricting the analysis to utterances containing a pronominal subject and a matching verb in an appropriate position. For Spanish, which does not require overt subjects, the morphology tier of the transcript was used. This allowed for the exclusion of imperative forms. The input analysis counted forms in interrogative as well as declarative input.

A key manipulation in the input analysis was the collection of corpus-wide counts as well as counts from utterance-final sentence fragments of differing lengths ('bins'). A central assumption in MOSAIC is that children produce increasingly long utterance-final phrases and children's knowledge of inflection is likely to reflect the changing distribution of verb forms in these increasingly long utterance-final phrases. In order to facilitate the bin analysis, the input corpora were marked in the following way. Verb forms in simple finite utterances were marked as being tensed (e.g. *he goes-tensed away*). For compound finite utterances, the non-finite main verb was marked as untensed, while the modal was marked as tensed (e.g. *he can-tensed go-untensed, does-tensed he go-untensed?*). This procedure makes it possible to identify verb forms in contexts requiring a tensed form even when restricting the analysis to utterance-final fragments.

Counts for the different verb forms were collected across the corpus as well as for utterance-final bins of increasing length. For the bin analysis, untensed forms of main verbs were counted, provided the tensed modal was not included in the bin. Thus, the utterance *does-tensed he go-untensed* contributes to the counts for *go* for bins of length 1 and 2, but not for bin 3, at which point the utterance contributes to the count for *does.* In terms of the developing child, this procedure is designed to simulate an increasing realization on the part of the child that, in modal contexts, tense and agreement are marked on the modal/auxiliary rather than the main verb. The counts generated in the input analysis were collected on a verb-by-verb basis. That is, for every English verb, the number of times it occurred as a 3rd person singular or as a form that matched the bare form was counted (past tense forms as well as progressives and participles were ignored in the counts). For Dutch, forms matching the stem, 3rd person singular and plural/infinitive were counted, and all 6 present tense forms and the infinitive were distinguished in Spanish.

The main aim of the input analysis was to determine the extent to which individual verbs showed a clear default. For present purposes, a verb was considered to have a default form when one form made up at least 65% of its occurrences. Only verbs that occurred a minimum of 5 times were included in the counts. Results for the English, Dutch and Spanish input analysis are shown in Tables 2 through 4.

Table 2: Proportion of verbs that show a default in the English input analysis for different bin sizes. Complete utterances are included in bins that exceed their length.

Bin size	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

As can be seen in Table 2, English verbs predominantly default to the bare form/infinitive. Only a handful of verbs show a default for the 3rd person singular. Examples of these verbs are *belongs*, *tastes*, *rains*, and *begins*. The proportion of verbs that show a default is affected by bin size. As bin size increases beyond 3, fewer verbs show a default. This is caused by the fact that, as bin size increases, bare forms in compound constructions disappear from the counts as they are replaced by auxiliaries and modals. However, increasing bin size does not change the nature of the default: fewer verbs show a default, but the bare form remains the dominant default form.

The analysis of the Dutch input (Table 3) shows more of a qualitative shift. For bin 1, most verbs show a default for the infinitive/plural, which occurs in utterance-final position (see Table 1). However, this effect is relatively short-lived, and, with increasing bin-size, the stem and 3rd person singular take over as defaults. By bin 10, roughly half of the verbs show a default, split between the stem (~ two-thirds) and the 3rd person singular (~ one-third). As in the analysis of English, this shift is (partly) caused by modals and auxiliaries replacing non-finite forms. However, the main reason for this shift is that the (high frequency) singular finite forms differ from the infinitive. Since these finite forms take second position, they tend to occur in larger bins.

Table 3: Proportion of verbs that show a default in the Dutch input analysis for different bin sizes.

Bin size	Inf.	Stem	3rd SG	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

Results for the Spanish analysis are shown in Table 4. The 3rd person singular is the most frequent default form across all bin sizes. While some verbs show an (initial) default for the infinitive, this effect is far less pronounced than it is in Dutch, where infinitives occur in utterance-final position (after potential complements). However, the main difference between the Spanish and Dutch analysis is that the proportion of verbs that show a default is far lower in Spanish. This difference reflects the fact that the Spanish (present tense) inflectional system is split across 6 forms, compared to 3 for the Dutch system.

Table 4: Proportion of verbs that show a default in the Spanish input analysis for different bin sizes.

Bin size	Inf.	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

Implementing defaulting in MOSAIC

The analyses reported in Tables 3, 4 and 5 suggest that defaulting is likely to result in different effects as a function of language and bin-size. In order to investigate how defaulting affects simulated child speech, MOSAIC models were trained on English, Dutch and Spanish input. The simulations were run using a new, more efficient implementation of MOSAIC, which replaces MOSAIC's discrimination net with a hash-table. The current implementation does not (yet) incorporate MOSAIC's chunking or generativity mechanism. However, the current implementation produces very similar results to those reported by Freudenthal et al. (2010), particularly at early points in development. Like the version used in Freudenthal et al., the current version complements MOSAIC's utterance-final bias with a small utterance-initial bias. This allows MOSAIC to produce incomplete utterance-final phrases as well as concatenations of utterance-initial words with utterance-final phrases. These concatenations allow for utterance-internal omission and the production of OI errors with subjects through modal omission (e.g *He* (*can*) *go*). The training procedure was similar to that employed in Freudenthal et al. (2010). For English, this involved the use of an input file that was coded for 3rd singular contexts. This allowed for the identification of English OI errors in the absence of (3rd person singular) subjects, and therefore a straightforward comparison with Dutch and Spanish, which do not require the presence of a subject to identify OI errors.

Training MOSAIC models involves feeding the input through MOSAIC several times. Each run of the model produces increasingly long utterance-final phrases, which may be associated with (short) utterance-initial phrases. Since learning is frequency-sensitive, high frequency words or phrases have a higher likelihood of being learned than low frequency ones. Output is generated from MOSAIC by producing all utterance-final phrases and their concatenations with utterance-initial words. Output from MOSAIC thus consists of a corpus of utterances, which can be directly compared to a corpus of child speech[3].

All models were trained to a stage where MOSAIC's output reached a Mean Length of Utterance (MLU) of approximately 2.0. The output was then analysed to determine the number of utterances containing a main verb that were and were not marked for finiteness. Non-finite utterances were those that only contained a verb matching the infinitive (learned from a 3rd person singular context for English), whereas finite utterances were those containing a finite main verb or auxiliary or modal. Freudenthal et al. (2010) analysed the speech of children from a number of different languages at an MLU of ~ 2.0 and report approximate rates of OI errors of .90 for English, and .75 for Dutch. OI levels for 2 Spanish children were considerably lower at .05 and .20. MOSAIC models for these languages showed OI levels of .63 (English), .65 (Dutch) and .15 (Spanish). MOSAIC thus underestimated OI levels for English by approximately 25 percentage points, and by 10 percentage points for Dutch, while Spanish scores were within the range displayed by the children. Table 5 reports the results for MOSAIC models trained on the maternal speech to two English and Dutch children (Anne, Becky, Peter and Matthijs) and one Spanish child (Juan). OI rates before defaulting are similar to those reported by Freudenthal et al. (2010). Defaulting was implemented by searching MOSAIC's output for finite (present tense) verbs. In instances where a verb showed a default and the finite form did not match this default, the finite form was replaced with the default form. Tables 6 and 7 report the OI levels (and proportion of affected utterances) after defaulting based on thresholds of 0.60, 0.65 and 0.70. Table 6 reports results based on the counts for Bin10. Table 7 reports results for Bin2, a value close to the model MLUs, and hence developmentally appropriate.

Table 5: Descriptive statistics for English, Dutch and Spanish models before defaulting.

	MLU	Prop. OI	(verbal) utterances
Anne	2.07	.59	130
Becky	1.99	.66	109
Peter	1.97	.61	513
Matthijs	2.02	.66	1008
Juan	1.97	.15	1021

As can be seen in Table 6, defaulting based on Bin10 results in a small (8-12 percentage points) increase in OI levels for English, but has (virtually) no effect on Dutch and Spanish. The proportion of utterances that is affected by the defaulting mechanism also differs across languages. For English, it varies between 18 and 28%, while it is around 5% for Dutch and 3% for Spanish.

[3] Since MOSAIC does not duplicate utterances, analyses are typically restricted to utterance types, and duplicate utterances are removed from child corpora.

Table 6: Proportion of OIs for three levels of defaulting based on Bin 10 (proportion of affected utterances in parentheses).

	Defaulting Threshold		
	.60	.65	.70
Anne	.70 (.28)	.70 (.28)	.71 (.27)
Becky	.74 (.18)	.74 (.18)	.74 (.18)
Peter	.62 (.05)	.61 (.03)	.61 (.00)
Matthijs	.66 (.06)	.66 (.05)	.66 (.03)
Juan	.16 (.03)	.16 (.03)	.16 (.02)

Defaulting based on the counts for Bin2 shows larger effects. Relative to Table 5, OI levels for English have increased by 18 to 27 percentage points, substantially more than for Bin10. OI levels for Dutch have increased by 4-5 percentage points, while Spanish OI levels are unaffected. The proportion of utterances affected by the defaulting mechanism is similar across Tables 6 and 7. The increase in OI levels for Bin2 relative to Bin10 thus does not reflect a simple increase in affected utterances. Instead, it represents a shift in the forms to which the model defaults.

Table 7: Proportion of OIs for three levels of defaulting based on Bin 2 (proportion of affected utterances in parentheses).

	Defaulting Threshold		
	.60	.65	.70
Anne	.86 (.34)	.86 (.34)	.80 (.29)
Becky	.81 (.19)	.80 (.18)	.79 (.17)
Peter	.66 (.05)	.66 (.05)	.65 (.05)
Matthijs	.70 (.07)	.70 (.07)	.70 (.06)
Juan	.16 (.04)	.16 (.03)	.16 (.03)

Error Analysis

Since verb placement in Dutch is dependent on finiteness, defaulting in Dutch has the potential to result in grammatical errors that are not typical of child speech. This possibility was investigated by examining the nature of the utterances that resulted from defaulting in Dutch (at a threshold of 0.65 and Bin2). For Peter's model, 26 utterances were affected by defaulting, all of which involved a finite form being changed into an infinitive. Eight of the utterances contained a verb in final position, and thus resulted in a stereotypical OI error. Ten utterances contained a verb in finite (V2) position – a position occupied by plural verbs. However, since these utterances did not contain a subject they were not grammatically anomalous. A further 8 utterances contained a singular subject combined with a plural verb in V2 position. This type of error is rare in child speech. However, it makes up less than 2% of the entire output for Peter's model.

A similar pattern was apparent in Matthijs's output. Out of 66 utterances affected by defaulting, 5 resulted in clear error, and the remainder was roughly equally split between subjectless plurals and stereotypical OI errors.

In summary, the analysis of the output revealed that the defaulting mechanism produced few obvious errors. At an MLU of around 2, relatively few utterances contain enough syntactic elements for defaulting to result in clear errors. Non-matching subject-verb sequences (i.e. OI errors) are attested in child speech, and verb-object sequences can be contextually, but not grammatically, inappropriate. This latter type of error is likely to remain unnoticed in situations where the context is ambiguous or cannot be recovered – a situation that is likely to occur in corpus analyses, a common method in language acquisition research. An interesting observation in this respect is made by Verhagen and Blom (2014), who argue that the rate of inflectional errors in elicited production is far higher than in spontaneous speech. Apparently, children are worse at inflection than their spontaneous speech suggests.

While the error analysis raises the question of whether defaulting in Dutch is likely to result in higher error rates at higher MLUs (when more syntactic elements are present), this is unlikely to be the case. Defaulting is likely to reduce as children grow older and their MLU increases. Larger MLUs furthermore correspond to larger bins. Fewer verbs show a clear default in larger bins, and 3rd person singular and (in particular) stem defaults become more frequent. To the extent that defaulting still occurs at high(er) MLUs, it is thus likely to involve defaulting to the stem, which is a phonologically simpler or reduced form. A relevant observation in this respect is that omission of 3rd person singular –t (or production of the stem in a 3rd person singular context) occurs in adult speech in some Dutch dialects. Defaulting at higher MLUs is therefore more likely to take the form of an error that is not uncommon in the adult language.

Conclusions

The main aim of this paper was to investigate defaulting effects as a contributing factor in children's cross-linguistic production of Optional Infinitive errors. An analysis of English, Dutch and Spanish child-directed speech corpora showed large cross-linguistic differences in the distribution of verb forms. A corpus-wide analysis showed that English verbs tend to be used in forms that match the bare form, whereas Spanish verbs tend to be used in either the 2nd or 3rd person singular form. Dutch shows a similar pattern to Spanish with most verbs showing a default for the stem or 3rd person singular. A developmentally-inspired input analysis which restricted the analysis to utterance-final phrases boosted the pattern for English. All but a handful of verbs occurred overwhelmingly in the bare form. The Dutch developmental analysis showed a shift from the stem and 3rd person singular towards the infinitive, while the Spanish analysis saw a reduction in the forms that default to the 2nd and 3rd singular forms. The input analysis thus suggested that the pattern of defaulting would differ across the three languages. English children are expected to default to the bare form. The fact that the bare form remains the default for larger bin sizes suggests that this effect may continue until late in development. Dutch children are expected to initially default to the infinitive, while they may default to the stem and 3rd person singular in later stages of development. Spanish children are expected to show relatively little defaulting, and this is likely to be restricted to 2nd and 3rd person singulars. Implementation of the defaulting mechanism in MOSAIC showed a large boost in OI errors for English a small increase for Dutch, and little change for Spanish. Defaulting thus has the largest effect for those languages where MOSAIC's earlier fit was poorest. However, defaulting does not just boost levels of OIs; it also allows MOSAIC to produce verb forms in contexts where it previously could not – including the provision of finite verb forms in incorrect contexts, and the production of non-finite verb forms in contexts that do not have a modal reading - a feature that is thought to be more prominent in English than in Dutch (Hoekstra & Hyams, 1998).

Importantly, rather than rendering MOSAIC's main mechanism for the production of OI errors (the utterance-final bias) superfluous, defaulting effects are strongest and most plausible when combined with the utterance-final bias.

An important consequence of these results is that they suggest a way in which MOSAIC might be extended to provide an account of atypical language development. Children with Specific Language Impairment (SLI) show slow development of language in the apparent absence of deficits outside the linguistic domain. English and German (but not Spanish) children with SLI are thought to have particular difficulty with verb inflection and have been characterized as going through an extended Optional Infinitive Stage (Rice, Wexler & Cleave, 1995). Compared to typically developing children at the same MLU, English children with SLI produce more OI errors, and they continue to do so at higher MLUs. The mechanism reported here suggests several potential causes for such differences. Defaulting effects may be more pronounced, continue for longer in children with SLI, or, if we take the bin analysis seriously, children with SLI may be slower to access larger bins. This latter suggestion has considerable overlap with a recent account by Leonard et al. (in press) who, in line with MOSAIC's biases, suggest that children with SLI may show a continued reliance on subject-verb sequences that have their origins in utterance-final input sequences.

Acknowledgements

Julian Pine, and Fernand Gobet are members of the International Centre for Language and Communicative Development (LuCiD) at the University of Liverpool, for which support of the Economic and Social Research Council [ES/L008955/1] is gratefully acknowledged. This research was supported by ESRC Grant ES/J011436/1.

References

Aguado-Orea, J. (2004). The acquisition of morpho-syntax in Spanish: Implications for current theories of development. *Unpublished doctoral thesis, University of Nottingham.*

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: Centre for Language and Cognition.

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.

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.

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*, *37*, 643-669.

Hoekstra, T. & Hyams, N. (1998). Aspects of Root Infinitives. *Lingua*, *106*, 81-112.

Leonard, L.B., Fey, M.E., Deevy, P. & Bredin-Oja, S.L. (in press). Input sources of third person singular –s inconsistency in children with and without specific language impairment. To appear in *Journal of Child Language*.

MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development*, *43*, 1-123.

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

Räsänen, S., 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.

Räsänen, S., Ambridge, B. & Pine, J.M. (in preparation). Investigating the optional infinitive phenomenon: a comparison of English and Swedish.

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.

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.

Verhagen, J. & Blom, E. (2014). Assymmetries in the acquisition of subject-verb agreement in Dutch: evidence from comprehension and production. *First Language*, *34*, 315-335.

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. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, *106*, 23-79.