Modelling syntactic development in a cross-linguistic context

Fernand GOBET
Centre for Cognition and Neuroimaging
Department of Human Sciences
Brunel University
Uxbridge UB8 3PH, UK
Fernand.Gobet@Brunel.ac.uk

Daniel FREUDENTHAL
Julian M. PINE
School of Psychology
University of Nottingham
Nottingham NG7 2RD, UK
Daniel.Freudenthal@psyc.nott.ac.uk
Julian.Pine@psyc.nott.ac.uk

Abstract

Mainstream linguistic theory has traditionally assumed that children come into the world with rich innate knowledge about language and grammar. More recently, computational work using distributional algorithms has shown that the information contained in the input is much richer than proposed by the nativist approach. However, neither of these approaches has been developed to the point of providing detailed and quantitative predictions about the developmental data. In this paper, we champion a third approach, in which computational models learn from naturalistic input and produce utterances that can be directly compared with the utterances of language-learning children. We demonstrate the feasibility of this approach by showing how MOSAIC, a simple distributional analyser, simulates the optional-infinitive phenomenon in English, Dutch, and Spanish. The model accounts for young children's tendency to use both correct finites and incorrect (optional) infinitives in finite contexts, for the generality of this phenomenon across languages, and for the sparseness of other types of errors (e.g., word order errors). It thus shows how these phenomena, which have traditionally been taken as evidence for innate knowledge of Universal Grammar, can be explained in terms of a simple distributional analysis of the language to which children are exposed.

1 Introduction

Children acquiring the syntax of their native language are faced with a task of considerable complexity, which they must solve using only noisy and potentially inconsistent input. Mainstream linguistic theory has addressed this 'learnability problem' by proposing the nativist hypothesis that children come into the world with rich innate knowledge about language and grammar (Chomsky, 1981; Piattelli-Palmarini, 2002; Pinker, 1984).

However, there is also strong empirical evidence that the amount of information present in the input is considerably greater than has traditionally been assumed by the nativist approach. In particular, computer simulations have shown that a distributional analysis of the statistics of the input can provide a significant amount of syntactic information (Redington & Chater, 1997).

One limitation of the distributional approach is that analyses have rarely been done with naturalistic input (e.g. mothers' child-directed speech) and have so far not been linked to the detailed analysis of a linguistic phenomenon found in human data, (e.g., Christiansen & Chater, 2001). Indeed, neither the nativist nor the distributional approach has been developed to the point of providing detailed and quantitative predictions about the developmental dynamics of the acquisition of language. In order to remedy this weakness, our group has recently been exploring a different approach. This approach, which we think is a more powerful way of understanding how children acquire their native language, has involved developing a computational model (MOSAIC; Model Of Syntax Acquisition In Children) that learns from naturalistic input, and produces utterances that can be directly compared with the utterances of language-learning children. This makes it possible to derive quantitative predictions about empirical phenomena observed in children learning different languages and about the developmental dynamics of these phenomena.

MOSAIC, which is based upon a simple distributional analyser, has been used to simulate a number of phenomena in language acquisition. These include: the verb-island phenomenon (Gobet & Pine, 1997; Jones, Gobet, & Pine, 2000); negation errors in English (Croker, Pine, & Gobet, 2003); patterns of pronoun case marking error in English (Croker, Pine, & Gobet, 2001); patterns of subject omission error in English (Freudenthal, Pine, & Gobet, 2002b); and the optional-infinitive phenomenon (Freudenthal, Pine, & Gobet, 2001, 2002a, 2003). MOSAIC has also been used to simulate data from three different languages (English, Dutch, and Spanish), which has helped us to

understand how these phenomena are affected by differences in the structure of the language that the child is learning.

In this paper, we illustrate our approach by showing how MOSAIC can account in detail for the 'optional-infinitive phenomenon' in two languages (English and Dutch) and its quasi-absence in a third language (Spanish). This phenomenon is of particular interest as it has generally been taken to reflect innate grammatical knowledge on the part of the child (Wexler, 1994, 1998).

We begin by highlighting the theoretical challenges faced in applying our model to data from three different languages. Then, after describing the optional-infinitive phenomenon, we describe MOSAIC, with an emphasis on the mechanisms that will be crucial in explaining the empirical data. We then consider the data from the three languages, and show to what extent the same model can simulate these data. When dealing with English, we describe the methods used to collect and analyse children's data in some detail. While these details may seem out of place in a conference on computational linguistics, we emphasise that they are critical to our approach: first, our approach requires fine-grained empirical data, and, second, the analysis of the data produced by the model is as close as possible to that used with children's data. We conclude by discussing the implications of our approach for developmental psycholinguistics.

2 Three languages: three challenges

The attempt to use MOSAIC to model data in three different languages involves facing up to a number of challenges, each of which is instructive for different reasons. An obvious problem when modelling English data is that English has an impoverished system of verb morphology that makes it difficult to determine which form of the verb a child is producing in any given utterance. This problem militates against conducting objective quantitative analyses of children's early verb use and has resulted in there being no detailed quantitative description of the developmental patterning of the optional infinitive phenomenon in English (in contrast to other languages like Dutch). We have addressed this problem by using exactly the same (automated) methods to classify the utterances produced by the child and by the model. These methods, which do not rely on the subjective judgment of the coder (e.g. on Bloom's, 1970, method of rich interpretation) proved to be sufficiently powerful to capture the development of the optional infinitive in English, and to do so at a relatively fine level of detail.

One potential criticism of these simulations of English is that we may have tuned the model's parameters in order to optimise the goodness of fit to the human data. An obvious consequence of overfitting the data in this way would be that MOSAIC's ability to simulate the phenomenon would break down when the model was applied to a new language. The simulations of Dutch show that this is not the case: with this language, which has a richer morphology than English, the model was still able to reproduce the key characteristics of the optional-infinitive stage.

Spanish, the syntax of which is quite different to English and Dutch, offered an even more sensitive test of the model's mechanisms. The Dutch simulations relied heavily on the presence of compound finites in the child-directed speech used as input. However, although Spanish child-directed speech has a *higher* proportion of compound finites than Dutch, children learning Spanish produce optional-infinitive errors *less* often than children learning Dutch. Somewhat counter-intuitively, the simulations correctly reproduce the relative scarcity of optional-infinitive errors in Spanish, showing that the model is sensitive to subtle regularities in the way compound finites are used in Dutch and Spanish.

3 The optional-infinitive phenomenon

Between two and three years of age, children learning English often produce utterances that appear to lack inflections, such as past tense markers or third person singular agreement markers. For example, children may produce utterances as:

(1a) *That go there**
(2a) *He walk home**

instead of:

(1b) *That goes there*
(2b) *He walked home*

Traditionally, such utterances have been interpreted in terms of absence of knowledge of the appropriate inflections (Brown, 1973) or the dropping of inflections as a result of performance limitations in production (L. Bloom, 1970; P. Bloom, 1990; Pinker, 1984; Valian, 1991). More recently, however, it has been argued that they reflect the child's optional use of (root) infinitives (e.g. *go*) in contexts where a finite form (e.g. *went*, *goes*) is obligatory in the adult language (Wexler, 1994, 1998).

This interpretation reflects the fact that children produce (root) infinitives not only in English, where the infinitive is a zero-marked form, but also in languages such as Dutch where the infinitive carries its own infinitival marker. For instance,

children learning Dutch may produce utterances such as:

(3a) *Pappa eten** *(Daddy to eat)*
(4a) *Mamma drinken** *(Mummy to drink)*

instead of:

(3b) *Pappa eet* *(Daddy eats)*
(4b) *Mamma drinkt* *(Mummy drinks)*

The optional infinitive phenomenon is particularly interesting as it occurs in languages that differ considerably in their underlying grammar, and is subject to considerable developmental and cross-linguistic variation. It is also intriguing because children in the optional infinitive stage typically make few other grammatical errors. For example, they make few errors in their use of the basic word order of their language: English-speaking children may say *he go*, but not *go he*.

Technically, the optional infinitive phenomenon revolves around the notion of 'finiteness'. Finite forms are forms that are marked for Tense and/or Agreement (e.g. *went, goes*). Non-finite forms are forms that are not marked for Tense or Agreement. This includes the infinitive form (*go*), the past participle (*gone*), and the progressive participle (*going*). In English, finiteness marking increases with development: as they grow older, children produce an increasing proportion of unambiguous finite forms.
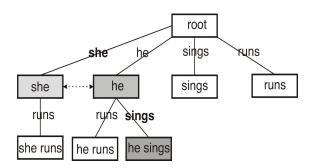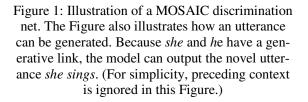
4   Description of the model

MOSAIC is a computational model that analyses the distributional characteristics present in the input. It learns to produce increasingly long utterances from naturalistic (child-directed) input, and produces output consisting of actual utterances, which can be directly compared to children's speech. This allows for a direct comparison of the output of the model at different stages with the children's developmental data.

The model learns from text-based input (i.e., it is assumed that the phonological stream has been segmented into words). Utterances are processed in a left to right fashion. MOSAIC uses two learning mechanisms, based on discrimination and generalisation, respectively. The first mechanism grows an n-ary discrimination network (Feigenbaum & Simon, 1984; Gobet et al., 2001) consisting of nodes connected by test links. Nodes encode single words or phrases. Test links encode the difference between the contents of consecutive nodes. (Figure 1 illustrates the structure of the type of discrimination net used.) As the model sees more and more input, the number of nodes and links increases, and so does the amount of information held in the

nodes, and, as a consequence, the average length of the phrases it can output. The node creation probability (NCP) is computed as follows:

$$NCP = (N / M)^L$$

where M is a parameter arbitrarily set to 70,000 in the English and Spanish simulations, N = number of nodes in the net (N ≤ M), and L = length of the phrase being encoded. Node creation probability is thus dependent both on the length of the utterance (longer utterances are less likely to yield learning) and on the amount of knowledge already acquired. In a small net, learning is slow. When the number of nodes in the net increases, the node creation probability increases and, as a result, the learning rate also increases. This is consistent with data showing that children learn new words more easily as they get older (Bates & Carnavale, 1993).

Figure 1: Illustration of a MOSAIC discrimination net. The Figure also illustrates how an utterance can be generated. Because *she* and *he* have a generative link, the model can output the novel utterance *she sings*. (For simplicity, preceding context is ignored in this Figure.)

While the first learning mechanism is based on discrimination, the second is based on generalisation. When two nodes share a certain percentage (set to 10% for these simulations) of nodes (phrases) following and preceding them, a new type of link, a *generative link* is created between them (see Figure 1 for an example). Generative links connect words that have occurred in similar contexts in the input, and thus are likely to be of the same word class. As no linguistic constructs are given to the model, the development of approximate linguistic classes, such as those of noun or verb, is an emergent property of the distributional analysis of the input. An important feature of MOSAIC is that the creation and removal of generative links is dynamic. Since new nodes are constantly being created in the network, the percentage overlap between two nodes varies over time; as a

consequence, a generative link may drop below the threshold and so be removed.

The model generates output by traversing the network and outputting the contents of the visited links. When the model traverses test links only, the utterances it produces must have been present in the input. Where the model traverses generative links during output, novel utterances can be generated. An utterance is generated only if its final word was the final word in the utterance when it was encoded (this is accomplished by the use of an *end marker*). Thus, the model is biased towards generating utterances from sentence final position, which is consistent with empirical data from language-learning children (Naigles & Hoff-Ginsberg, 1998; Shady & Gerken, 1999; Wijnen, Kempen, & Gillis, 2001).

5 Modelling the optional-infinitive phenomenon in English

Despite the theoretical interest of the optional-infinitive phenomenon, there is, to our knowledge, no quantitative description of the developmental dynamics of the use of optional infinitives in English, with detail comparable to that provided in other languages, such as Dutch (Wijnen et al., 2001). The following analyses fill this gap.

5.1 Children's data: Methods

We selected the speech of two children (Anne, from 1 year 10 months to 2 years 9 months; and Becky, from 2 years to 2 years 11 months). These data were taken from the Manchester corpus (Theakston, Lieven, Pine, & Rowland, 2001), which is available in the CHILDES data base (MacWhinney, 2000). Recordings were made twice every three weeks over a period of one year and lasted for approximately one hour per session.

Given that optional-infinitive phenomena are harder to identify in English than in languages such as Dutch or German (due to the relatively low number of unambiguous finite forms), the analysis focused on the subset of utterances that contain a verb with *he*, *she*, *it*, *this* (*one*), or *that* (*one*) as its subject. Restricting the analysis in this way avoids utterances such as *I go*, which could be classified both as non-finite and finite, and therefore makes it possible to more clearly separate non-finites, simple finites, compound finites, and ambiguous utterances.

Identical (automatic) analyses of the data and model were carried out in a way consistent with previous work on Dutch (Wijnen et al., 2001). Utterances that had the copula (i.e., forms of the verb *to be*) as a main verb were removed. Utterances that contained a non-finite form as the only verb were classified as non-finites. Utterances with an unambiguous finite form (*walks*, *went*) were counted as finite, while those containing a finite verb form plus a non-finite form (*has gone*) were classified as compound finites. The remaining utterances were classified as ambiguous and counted separately; they contained an ambiguous form (such as *bought* in *he bought*) as the main verb, which can be classified either as a finite past tense form or as a (non-finite) perfect participle (in the phrase *he bought*, the word *has* may have been omitted).

5.2 Children's data: Results

The children's speech was partitioned into three developmental stages, defined by mean length of utterance (MLU). The resulting distributions, portrayed in Figure 2, show that the proportion of non-finites decreases as a function of MLU, while the proportion of compound finites increases. There is also a slight increase in the proportion of simple finites, although this is much less pronounced than the increase in the proportion of compound finites.

5.3 Simulations

The model received as input speech from the children's respective mothers. The size of the input was 33,000 utterances for Anne's model, and 27,000 for Becky's model. Note that, while the analyses are restricted to a subset of the children's corpora, the entire mothers' corpora were used as input during learning. The input was fed through the model several times, and output was generated after every run of the model, until the MLU of the output was comparable to that of the end stage in the two children. The output files were then compared to the children's data on the basis of MLU.

The model shows a steady decline in the proportion of non-finites as a function of MLU coupled with a steady increase in the proportion of compound finites (Figure 3). On average, the model's production of optional infinitives in third person singular contexts drops from an average of 31.5% to 16% compared with 47% to 12.5% in children. MOSAIC thus provides a good fit to the developmental pattern in the children's data (not including the 'ambiguous' category: $r^2 = .65$, $p < .01$, RMSD = 0.096 for Anne and her model; $r^2 = .88$, $p < .001$, RMSD = 0.104 for Becky and her model). One obvious discrepancy between the model's and the children's output is that both models at MLU 2.1 produce too many simple finite utterances. Further inspection of these utterances reveals that they contain a relatively high proportion of finite modals such as *can* and *will* and finite forms of the dummy modal *do* such as *does* and *did*. These forms are unlikely to be used as the only verb in children's early utterances as their function is to

a: Data for Anne
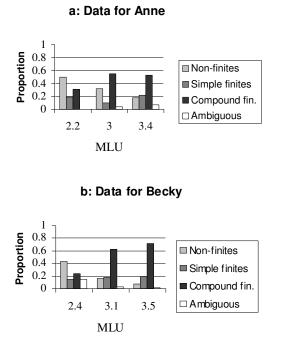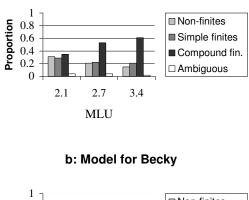

b: Data for Becky

Figure 2: Distribution of non-finites, simple finites, compound finites, and ambiguous utterances for Anne and Becky as a function of developmental phase. Only utterances with *he*, *she*, *it*, *that* (*one*), or *this* (*one*) as a subject are included.
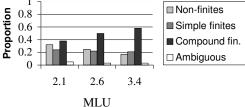
a: Model for Anne

b: Model for Becky

Figure 3: Distribution of non-finites, simple finites, compound finites, and ambiguous utterances for the models of Anne and Becky as a function of developmental phase. Only utterances with *he*, *she*, *it*, *that* (*one*), or *this* (*one*) as a subject are included.

modulate the meaning of the main verb rather than to encode the central relational meaning of the sentence.

An important reason why MOSAIC accounts for the data is that it is biased towards producing sentence final utterances. In English, non-finite utterances can be learned from compound finite questions in which finiteness is marked on the auxiliary rather than the lexical verb. A phrase like *He walk home* can be learned from *Did he walk home?*, and a phrase like *That go there* can be learned from *Does that go there?* As MLU increases, the relative frequency of non-finite utterances in the output decreases, because the model learns to produce more and more of the compound finite utterances from which these utterances have been learned. MOSAIC therefore predicts that as the proportion of non-finite utterances decreases, there will be a complementary increase in the proportion of compound finites.

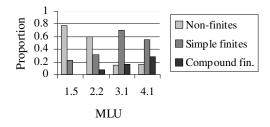## 6 Modelling optional infinitives in Dutch

Children acquiring Dutch seem to use a larger proportion of non-finite verbs in finite contexts (e.g., *hij lopen*, *bal trappen*) than children learning English. Thus, in Dutch, a very high percentage of children's early utterances with verbs (about 80%) are optional-infinitive errors. This percentage decreases to around 20% by MLU 3.5 (Wijnen, Kempen & Gillis, 2001).

As in English, optional infinitives in Dutch can be learned from compound finites (auxiliary/modal + infinitive). However, an important difference between English and Dutch is that in Dutch verb position is dependent on finiteness. Thus, in the simple finite utterance *Hij drinkt koffie (He drinks coffee)* the finite verb form *drinkt* precedes its object argument *koffie* whereas in the compound finite utterance *Hij wil koffie drinken (He wants coffee drink),* the non-finite verb form *drinken* is restricted to utterance final position and is hence preceded by its object argument: *koffie*. Interestingly, children appear to be sensitive to this feature of Dutch from very early in development and MOSAIC is able to simulate this sensitivity. However, the fact that verb position is dependent on finiteness in Dutch also means that whereas non-finite verb forms are restricted to sentence final position, finite verb forms tend to occur earlier in the utterance. MOSAIC therefore simulates the very high proportion of optional infinitives in early child Dutch as a function of the interaction between its utterance final bias and increasing MLU. That is, the high proportion of non-finites early on is explained by the fact that the model mostly produces sentence-final phrases, which, as a result of
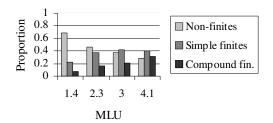
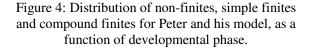Dutch grammar, have a large proportion of non-finites.

As shown in Figure 4, the model's production of optional infinitives drops from 69% to 28% compared with 77% to 18% in the data of the child on whose input data the model had been trained. In these simulations, the input data consisted of a sample of approximately 13,000 utterances of child-directed speech. Because of the lower input size, the M used in the NCP formula was set to 50,000.

a: Data for Peter

b: Model for Peter

Figure 4: Distribution of non-finites, simple finites and compound finites for Peter and his model, as a function of developmental phase.
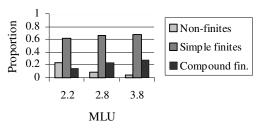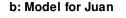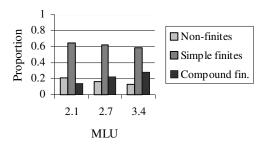
7 Modelling optional infinitives in Spanish

Wexler (1994, 1998) argues that the optional-infinitive stage does not occur in pro-drop languages, that is, languages like Spanish in which verbs do not require an overt subject. Whether MOSAIC can simulate the low frequency of optional-infinitive errors in early child Spanish is therefore of considerable theoretical interest, since the ability of Wexler's theory to explain cross-linguistic data is presented as one of its main strengths. Note that simulating the pattern of finiteness marking in early child Spanish is not a trivial task. This is because although optional-infinitive errors are much less common in Spanish than they are in Dutch, compound finites are actually more common in Spanish child-directed

speech than they are in Dutch child-directed speech (in the corpora we have used, they make up 36% and 30% of all parents' utterances including verbs, respectively).

a: Data for Juan

b: Model for Juan

Figure 5: Distribution of non-finites, simple finites and compound finites for Juan and his model, as a function of developmental phase.
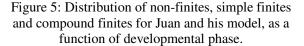
Figure 5a shows the data for a Spanish child, Juan (Aguado Orea & Pine, 2002), and Figure 5b the outcome of the simulations run using MOSAIC. The parental corpus used as input consisted of about 27,000 utterances. The model's production of optional infinitives drops from 21% to 13% compared with 23% to 4% in the child. Both the child and the model show a lower proportion of optional-infinitive errors than in Dutch. The presence of (some rare) optional-infinitive errors in the model's output is explained by the same mechanism as in English and Dutch: a bias towards learning the end of utterances. For example, the input *¿Quieres beber café? (Do you want to drink coffee?)* may later lead to the production of *beber café*. But why does the model produce so few optional-infinitive errors in Spanish when the Spanish input data contain so many compound finites? The answer is that finite verb forms are much more likely to occur in utterance final position in Spanish than they are in Dutch, which makes them much easier to learn.

8 Conclusion

In this paper, we have shown that the same simple model accounts for the data in three languages that differ substantially in their underlying structure. To our knowledge, this is the only model of language acquisition which simultaneously (1) learns from naturalistic input (actual child-directed utterances), where the statistics and frequency distribution of the input are similar to that experienced by children; (2) produces actual utterances, which can be directly compared to those of children; (3) has a developmental component; (4) accounts for speech generativity and increasing MLU; (5) makes quantitative predictions; and (6) has simulated phenomena from more than one language.

An essential feature of our approach is to limit the number of degrees of freedom in the simulations. We have used an identical model for simulating the same class of phenomena in three languages. The method of data analysis was also the same, and, in all cases, the model's and child's output were coded automatically and identically. The use of realistic input was also crucial in that it guaranteed that cross-linguistic differences were reflected in the input.

The simulations showed that simple mechanisms were sufficient for obtaining a good fit to the data in three different languages, in spite of obvious syntactic differences and very different proportions of optional-infinitive errors. The interaction between a sentence final processing bias and increasing MLU enabled us to capture the reason why English, Dutch and Spanish offer different patterns of optional-infinitive errors: the difference in the relative position of finites and non-finites is larger in Dutch than in English, and Spanish verbs are predominantly finite. We suggest that any model that learns to produce progressively longer utterances from realistic input, and in which learning is biased towards the end of utterances, will simulate these results.

The production of actual utterances (as opposed to abstract output) by the model makes it possible to analyse the output with respect to several (seemingly) unrelated phenomena, so that the nontrivial predictions of the learning mechanisms can be assessed. Thus, the same output can be utilized to study phenomena such as optional-infinitive errors (as in this paper), evidence for verb-islands (Jones et al., 2000), negation errors (Croker et al., 2003), and subject omission (Freudenthal et al., 2002b). It also makes it possible to assess the relative importance of factors such as increasing MLU that are implicitly assumed by many current theorists but not explicitly factored into their models.

An important contribution of Wexler's (1994, 1998) nativist theory of the optional-infinitive stage has been to provide an integrated account of the different patterns of results observed across languages, of the fact that children use both correct finite forms and incorrect (optional) infinitives, and of the scarcity of other types of errors (e.g. verb placement errors). His approach, however, requires a complex theoretical apparatus to explain the data, and does not provide any quantitative predictions. Here, we have shown how a simple model with few mechanisms and no free parameters can account for the same phenomena not only qualitatively, but also quantitatively.

The simplicity of the model inevitably means that some aspects of the data are ignored. Children learning a language have access to a range of sources of information (e.g. phonology, semantics), which the model does not take into consideration. Also, generating output from the model means producing everything the model can output. Clearly, children produce only a subset of what they can say. Furthermore, any rote-learned utterance that the model produces early on in its development will continue to be produced during the later stages. This inability to unlearn is clearly a weakness of the model, but one that we hope to correct in subsequent research.

The results clearly show that the interaction between a simple distributional analyser and the statistical properties of naturalistic child-directed speech can explain a considerable amount of the developmental data, without the need to appeal to innate linguistic knowledge. The fact that such a relatively simple model provides such a good fit to the developmental data in three languages suggests that (1) aspects of children's multi-word speech data such as the optional-infinitive phenomenon do not necessarily require a nativist interpretation, and (2) nativist theories of syntax acquisition need to pay more attention to the role of input statistics and increasing MLU as determinants of the shape of the developmental data.

9 Acknowledgements

References

Aguado-Orea, J., & Pine, J. M. (2002). Assessing the productivity of verb morphology in early child Spanish. Paper presented at *the IX International Congress for the Study of Child Language*, Madison, Wisconsin.

Bates, E., & Carnavale, G. F. (1993). New directions in research on child development. *Developmental Review, 13*, 436-470.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

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

Brown, R. (1973). *A first language*. Boston, MA: Harvard University Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, NL: Foris.

Christiansen, M. H., & Chater, N. (2001). Connectionist psycholinguistics: Capturing the empirical data. *Trends in Cognitive Sciences, 5*, 82-88.

Croker, S., Pine, J. M., & Gobet, F. (2001). Modelling children's case-marking errors with MOSAIC. In E. M. Altmann, A. Cleeremans, C. D. Schunn & W. D. Gray (Eds.), *Proceedings of the 4th International Conference on Cognitive Modeling* (pp. 55-60). Mahwah, NJ: Erlbaum.

Croker, S., Pine, J. M., & Gobet, F. (2003). Modelling children's negation errors using probabilistic learning in MOSAIC. In F. Detje, D. Dörner & H. Schaub (Eds.), *Proceedings of the Fifth International Conference on Cognitive Modeling* (pp. 69-74). Bamberg:Universitäts-Verlag.

Feigenbaum, E. A., & Simon, H. A. (1984). EPAM-like models of recognition and learning. *Cognitive Science, 8*, 305-336.

Freudenthal, D., Pine, J. M., & Gobet, F. (2001). Modelling the optional infinitive stage in MOSAIC: A generalisation to Dutch. In E. M. Altmann, A. Cleeremans, C. D. Schunn & W. D. Gray (Eds.), *Proceedings of the 4th International Conference on Cognitive Modeling* (pp. 79-84). Mahwah, NJ: Erlbaum.

Freudenthal, D., Pine, J. M., & Gobet, F. (2002a). Modelling the development of Dutch optional infinitives in MOSAIC. In W. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th Annual Meeting of the Cognitive Science Society* (pp. 322-327). Mahwah, NJ: Erlbaum.

Freudenthal, D., Pine, J. M., & Gobet, F. (2002b). Subject omission in children's language: The case for performance limitations in learning. In W. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th Annual Meeting of the Cognitive Science Society* (pp. 328-333). Mahwah, NJ: Erlbaum.

Freudenthal, D., Pine, J. M., & Gobet, F. (2003). The role of input size and generativity in simulating language acquisition. In F. Schmalhofer, R. M. Young & G. Katz (Eds.), *Proceedings of EuroCogSci 03: The European Cognitive Science Conference 2003* (pp. 121-126). Mahwah NJ: Erlbaum.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*, 236-243.

Gobet, F., & Pine, J. M. (1997). Modelling the acquisition of syntactic categories. In *Proceedings of the 19th Annual Meeting of the Cognitive Science Society* (pp. 265-270). Hillsdale, NJ: Erlbaum.

Jones, G., Gobet, F., & Pine, J. M. (2000). A process model of children's early verb use. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the Twenty Second Annual Meeting of the Cognitive Science Society* (pp. 723-728). Mahwah, N.J.: Erlbaum.

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

Naigles, L., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs: Effects of input frequency and structure on children's early verb use. *Journal of Child Language, 25*, 95-120.

Piattelli-Palmarini, M. (2002). The barest essentials. *Nature, 416*, 129.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Redington, M., & Chater, N. (1997). Probabilistic and distributional approaches to language acquisition. *Trends in Cognitive Sciences, 1*, 273-279.

Shady, M., & Gerken, L. (1999). Grammatical and caregiver cue in early sentence comprehension. *Journal of Child Language, 26*, 163-176.

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.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition, 40*, 21-81.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivation in child grammar. In N. Hornstein & D. Lightfoot (Eds.), *Verb movement*. Cambridge, MA: 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.

Wijnen, F., Kempen, M., & Gillis, S. (2001). Root infinitives in Dutch early child language. *Journal of Child Language, 28*, 629-660.