Two wrongs make a right: Learnability and word order consistency

Padraic Monaghan (Padraic.Monaghan@warwick.ac.uk)
Markus Gonitzke (Markus.Gonitzke@warwick.ac.uk)
Nick Chater (N.Chater@warwick.ac.uk)
Department of Psychology, University of Warwick
Coventry, CV4 7AL, UK

Abstract

Languages often demonstrate word order inconsistencies, and such inconsistencies ought to make languages harder to acquire. We present an integrative approach exploring the relationship between learnability and word order, incorporating syntactic theory, corpus analyses and computational modelling. We focus on comparisons between English and German, and conclude that inconsistencies may be preserved in the language due to the interaction between several syntactic structures.

Introduction

Inconsistent structures in language are harder to learn than consistent structures by computational systems, whether inconsistencies are at the syntactic level (Christiansen & Devlin, 1997), or at the lexical level, in terms of grapheme to phoneme correspondences (Plaut, et al., 1996), or semantic ambiguities (Cottrell, 1986). Several languages are entirely consistent in terms of head position (e.g., the position of the verb in verb phrases), such as Japanese or Irish. However, a degree of inconsistency is present among most languages (Kroch, 2000), even if there is still a high degree of consistency (Lupyan & Christiansen, 2002; van Everbroeck, 1999). One possible contributor to learnability is case-marking, which is particularly useful in languages with relatively free word order (Lupyan & Christiansen, 2002), though this appears to be a necessary rather than a sufficient condition (Kiparsky, 1996) for the learnability of such languages.

Viewed in evolutionary terms, languages that are harder to learn are more likely to die out (Christiansen & Devlin, 1997), and given the high rate of change of languages across time, it is a significant challenge to explain how word order inconsistencies are learned within languages.

Our approach to this challenge to account for learnability of inconsistencies was to bring together syntactic theory with analyses of the frequencies of different structures in real language corpora, and combine these with computational modelling. Previous simulations of word order have largely ignored the proportions of different syntactic structures (though with notable exceptions, e.g., MacDonald & Christiansen, 2002). Through the use of real language corpora in modelling, we can increase the precision of determining the extent to which the processes of sequential learning are engaged in language processing.

This paper presents a series of corpus analyses and simulations that explore the conditions under which inconsistencies in the language are learnable. As a test case, we focus on word order in German and English. German and English are particularly interesting for comparison as they have the same root (Cable & Bough, 1993), however, they differ in several important respects: In main clauses, German has a subject-verb-object (SVO) word order, whereas in subordinate clauses, the order is subject-object-verb (SOV), as shown in Sentence 1. The sentences are subscripted with subject (S), object (O), finite verb (Vf), infinite verb (Vi), and complementiser (C) to indicate the structures. In English, in contrast, word order is SVO in both main and subordinate clauses (Sentence 2, translation of Sentence 1). German and English also differ in terms of verb position in infinite verb phrases. In German, the infinite verb is sentence final after the object (Sentence 3), whereas in English, the infinite verb occurs after the finite verb, and before the object (Sentence 4, translation of Sentence 3). In subordinate clauses, this is complicated further in German: the finite verb moves to after the infinite verb after the object (Sentence 5), but word order is the same in English (Sentence 6, translation of Sentence 5).

[1] $_{S}$Ich $_{Vf}$benutzte $_{O}$das Werkzeug $_{C}$das $_{S}$ich $_{O}$dir $_{Vf}$gab
[2] $_{S}$I $_{Vf}$used $_{O}$the tool $_{C}$that $_{S}$I $_{Vf}$gave $_{O}$you

[3] $_{S}$Er $_{Vf}$hat $_{O}$seine Meinung $_{Vi}$geändert
[4] $_{S}$He $_{Vf}$has $_{Vi}$changed $_{O}$his opinion

[5] $_{S}$Er $_{Vf}$kauft $_{O}$den Teppich $_{C}$da $_{S}$sie $_{O}$die alte $_{Vi}$zerstört $_{Vf}$haben
[6] $_{S}$He $_{Vf}$buys $_{O}$the carpet $_{C}$since $_{S}$they $_{Vf}$have $_{Vi}$destroyed $_{O}$the old one

Such differences are real-world examples in languages with the same origin that have many similarities in common. The claims we make from our synthesis of corpus analyses and computational modelling generalise to word order consistency in general. However, the two languages we have selected are of especial interest as an example as English used to have the SOV structure in subordinate clauses, but this has now changed to SVO which is consistent with main clause word order (Lightfoot, 1991). Subordinate clauses constitute only a small proportion of phrases, and so the different word order in a minority of clauses is puzzling in its persistence – the greater frequency of SVO has not overwhelmed the SOV structure. Indeed, SVO structures are as easily parsed as SOV structures in subordinate clauses in studies on German speakers (Weyerts et al., 2002). We suggest that general sequential learning behaviour, as reflected in simple recurrent networks, contributes towards preserving such inconsistencies in German word order.

In this paper, we explore three grammar fragments of German, compared to the corresponding fragment in English. We postulate that, though subordinate structures may be harder to learn in German, the occurrence of verb-final structures in main clause infinite verb phrases results in easier learning of these structures. Finite verbs in final position are rarer than infinite verbs in final position (26% compared to 74%), and the verb-final position of infinite verbs is acquired earlier than the position of the finite verb (Clahsen & Muysken, 1986). This suggests that verb-ordering in German is influenced by the occurrence of both finite and infinite verb phrases. Finally, we make predictions about the scaffolding of relatively infrequent word order inconsistencies through interaction with other, more frequent structures. We first detail how we combined corpus analyses with modelling in our comparisons between English and German.

Corpus data in modelling

MacDonald and Christiansen (2002) illustrated that the different frequencies of linguistic structures have an impact on their ease of processing. It is, therefore, extremely useful to have a representation of the relative frequencies of different structures in languages in order to make assertions about the ease of acquisition of inconsistencies that may occur within the language. Such frequency information is not usually employed in modelling syntactic structures, with models training on corpora generated from randomised proportions of grammatical rules.

There are two major influences on the ability of simple recurrent networks to learn sequential orders. The first is predictability in word order: if a noun is always followed by a verb, for example, then the verb can be more easily predicted in the grammar. If word order is inconsistent then this will result in greater difficulty of learning. If there are many branching structures then learning to correctly predict the next item will be difficult, though the model will be able to learn the transitional probabilities between elements in the sequence. Another influence on learning in simple recurrent networks is the impact of centre-embeddings in structures (Christiansen & Chater, 1999). The number of intervening lexical items between the subject noun and the verb will affect the accuracy of verb agreement, as long-distance dependencies are more difficult to learn. This is one interpretation of the results of Christiansen and Devlin's (1997) study of recursive inconsistencies: Learning difficulty relates to the distance between the subject noun and the verb (see Figure 3 in Christiansen & Devlin, 1997; see also Grüning, 2003).

We analysed two tagged corpora in order to assess the relative proportions of the different structures in the grammars. For German, we examined the NEGRA corpus (Skut et al., 1997), which is composed of approximately

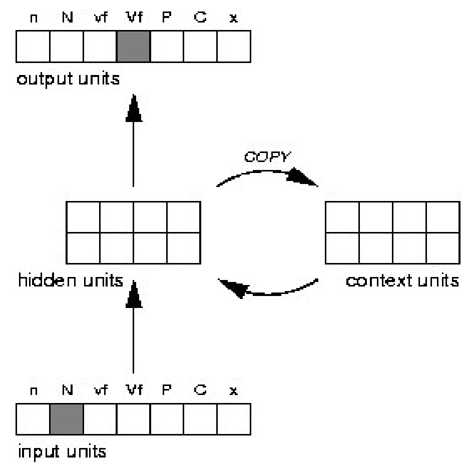

Figure 1. The simple recurrent network architecture we used in the simulations. The model is trained to predict the next word in the sentence, given the current word and the context of the previous state of the hidden units.

20,000 hand-tagged sentences from German newspapers. For English, we employed the British National Corpus (Burnard, 1995), composed of 100 million words of automatically tagged English (with an estimated error rate of 1.7%, with an additional 4.7% of words given ambiguous tags).

We derived simplified versions of the corpora by focusing on the NP, PP and VP structures. We omitted all words that modified, but did not alter the NP, PP and VP sentence structure[1]. Finally, we omitted any sentences that contained ambiguous tags, unclassified words, numerals, alphabetical letters, existential *there*, conjunctions or postpositions.

These simplifications resulted in 8,814 sentences in the NEGRA corpus, and 2,823,034 sentences in the BNC corpus which were comprised entirely of nouns, finite verbs, infinite verbs, prepositions, and complementisers[2]. Despite the differences in scale and text source, the overall proportions of different structures in the English and German corpora are approximately similar, as shown below in the more detailed analyses. Given the large number of sentences omitted from the corpora, however, the proportions given ought to be taken as a general guide only.

The corpus data were used to generate sets of sentences that were plausible approximations to the proportions of different phrase structures in English and German. The models we used were simple recurrent networks (Elman, 1990), or SRNs. SRNs are feedforward backpropagation networks with an extra layer of 'context' units that record the previous state of the network, and feedback onto the hidden layer of the model (Elman, 1990). This architecture

[1] We omitted adjectives, adverbs, articles, interjections, possessive markers, infinitive markers, and the negative particle.

[2] We also analysed the full corpora without omitting any sentences and found very similar proportions of subordinate/main clause as those reported on the cleaned-up corpora.

has been used to learn sequences in a range of tasks, including sequential learning of linguistic stimuli (Christiansen & Devlin, 1997). An example of the models used in our simulations is shown in Figure 1. This model, for the first simulation, had 7 input and output units, and eight hidden and context units. The 7 input/output units corresponded to plural/singular noun (N/n), plural/singular finite verb (V_f/v_f), preposition (P), complementiser (C), and an end of utterance marker (x). The unit corresponding to each category is indicated by the letter above the unit. The model was trained to predict the next token in a sentence, so in Figure 1, the model is presented with a singular noun, and predicts that the next element in the sentence will be a singular verb. Weights were updated using backpropagation with gradient descent of error with learning rate .1 and momentum .9.

Table 1. Grammar 1 for English and German with main and subordinate clauses, with proportions of each structure derived from BNC and NEGRA corpora.

ENGLISH	PROPORTION
S → S S-bar	7.7%
S → NP V_fP	92.3%
S-bar → C NP $V_f P_{sub}$	100%
NP → N (PP)	76.2% (23.8%)
PP → P NP	100%
V_fP → V_f (NP) (PP)	34.3% (48.5%) (17.2%)
$V_f P_{sub}$ → V_f (NP) (PP)	32.5% (46.5%) (21.0%)
GERMAN	
S → S S-bar	10.5%
S → NP V_fP	89.5%
S-bar → C NP $V_f P_{sub}$	100%
NP → N (PP)	70.4% (29.6%)
PP → P NP	100%
V_fP → V_f (NP) (PP)	11.3% (72.0%) (16.8%)
$V_f P_{sub}$ → (NP) (PP) V_f	(69.3%) (10.9%) 19.8%

Main and subordinate clauses

The first grammar that we compare between English and German consisted only of finite verb phrases in subordinate clauses. The purpose of this simulation was to test whether the grammar with SVO in main clause and SOV in subordinate clause was harder to learn than the grammar with SVO in both main and subordinate clauses. The rules of the simple grammars we compared are shown in the first column of Table 1. The grammatical rules are read as the element to the left of the arrow can be composed of the structures to the right of the arrow. So, a sentence (S) can be composed of a noun phrase (NP) followed by a finite verb phrase (V_fP). Rules can be recursive, as in the first rule, where a sentence (S) can be composed of a sentence followed by a complementiser (C) followed by a subordinate clause sentence (S-bar notates the complementiser and the subordinate clause sentence).

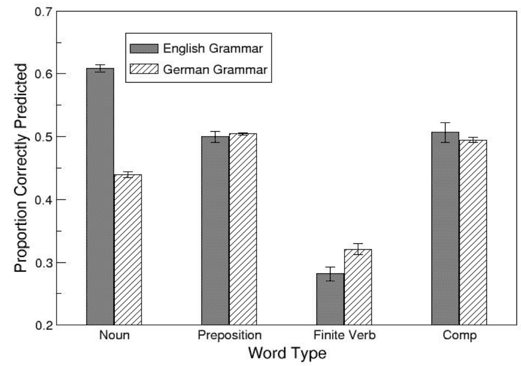

Figure 2. Proportion of words in each category correctly predicted by the model for Grammar 1 (see text for details).

Elements in parentheses indicate that these are optional, so a NP can be composed of a noun (N), or a N and a prepositional phrase (PP).

The second column of Table 1 indicates the proportions of each structure in the BNC and NEGRA corpora. So, for example, in English, 92.3% of sentences are composed of NP V_fP, and 7.7% of the time, they are composed of S S-bar. Percentages in parentheses indicate the proportion of times that the corresponding element in parentheses in the grammar occurs. So, the proportions associated with the rule V_fP → V_f (NP) (PP) are 34.3% (48.5%) (17.2%), which means that, of the occurrences of V_fP, V_fP → V_f 34.3% of the time, V_fP → V_f NP 48.5% of the time, and V_fP → V_f PP 17.2% of the time (in these cases, the rows sum to 100%).

The grammar rules in Table 1 described approximately 33.6% of the sentences in the simplified English corpus, and 42.3% of the simplified German corpus. Of particular interest from this corpus analyses were the proportions of the main clause sentences, and those composed of a main clause and a subordinate clause. In both English and German, only a small number of each are sentences composed from the rule: S → S S-bar (7.7% in English and 10.5% in German). Lightfoot (1991) has proposed that the change in English in subordinate clauses from SOV to SVO was due to the overwhelming of the rarer SOV structure by the more frequent SVO.

We generated a corpus of 1000 sentences, with branching according to the proportions of each structure that we found in the corpora for each grammar. We ran 10 simulations for each grammar, with different randomly generated corpora of 1000 sentences.

We tested the model for its ability to predict the proportion of occurrence of the next item in the training corpus when tested on a different randomly-generated corpus of 1000 sentences. A measure of mean squared error (MSE) for each word in the test corpus, as used by Christiansen and Devlin (1997), reflected the ability of the model to learn the conditional probabilities in the corpus. Lower MSE means that the model reflects the conditional probabilities in the corpus with greater accuracy.

For the German grammar, overall MSE was 0.045 (sd = 0.001) for each 'word' in the test corpus. For the English grammar, MSE was 0.013 (sd = 0.002), which was significantly lower, $t(18) = -37.30, p < 0.001$.

We also assessed the ability of the model to correctly predict the next word for each current lexical category given the sentential context. For each occurrence of the lexical category at the input (time t), we scored the proportion of times that the unit with highest activity in the output corresponded to the next word in the sentence ($t+1$). A higher proportion correctly predicted indicates that the current lexical category occurs in a more predictable context. Figure 2 shows the model's ability to correctly predict the words occurring after the different categories. Consistent with our hypothesis, the model learned the lexical category following nouns in German with greater difficulty than the category following nouns in English, $t(18) = 22.88, p < 0.001$ (adjusted for multiple comparisons), but responses for no other categories reached significance. This is due to the subject noun occurring before the verb in main clauses (SVO), and before the object noun in subordinate clauses in German (SOV). The inconsistency in word order resulted in greater difficulty in learning the sequential order of the language, even though the inconsistent subordinate clause structure in German occurs a small proportion of the time.

The rare occurrence of the SOV structure in German would put pressure on the survival of the inconsistency, and so we looked to a fuller grammar to see whether other structures may contribute towards the preservation of the subordinate clause word order. We looked firstly at differences in word order for finite and infinite verb phrases.

Table 2. Grammar 2 for English and German with main clause finite/infinite verb phrases, with proportion of each structure from BNC and NEGRA corpora.

ENGLISH	PROPORTION
$S \rightarrow NP\ V_fP$	35.7%
$S \rightarrow NP\ V_iP$	64.3%
$NP \rightarrow N\ (PP)$	60.0% (40.0%)
$PP \rightarrow P\ NP$	
$V_fP \rightarrow V_f\ (NP)\ (PP)$	34.3% (48.5%) (17.2%)
$V_iP \rightarrow V_f\ V_i\ (NP)\ (PP)$	35.6% (40.0%) (24.5%)
GERMAN	
$S \rightarrow NP\ V_fP$	50.8%
$S \rightarrow NP\ V_iP$	49.2%
$NP \rightarrow N\ (PP)$	70.0% (30.0%)
$PP \rightarrow P\ NP$	
$V_fP \rightarrow V_f\ (NP)\ (PP)$	8.5% (73.9%) (17.6%)
$V_iP \rightarrow V_f\ (NP)\ (PP)\ V_I$	10.1% (69.5%) (20.4%)

Finite and infinite verbs

The grammar fragments that we employed to assess the learnability of sentences containing finite and infinite verb phrases are shown in Table 2. In English, the infinite verb interposes between the finite verb and the object noun (shown in the main clause of Sentence 4). In German, the infinite verb is positioned at the end of the sentence (Sentence 3). Consequently, in German, the object noun is generally preceded by the finite verb. Our second simulations test the hypothesis that this results in greater difficulty in learning the sequential order of the English grammar.

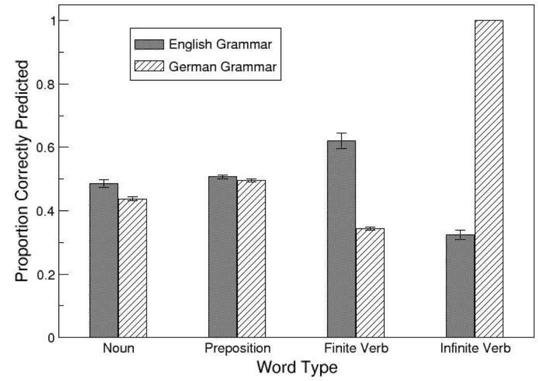

Figure 3. The model's performance on the English and German versions of Grammar 2 with finite and infinite verb phrases in main clauses.

The grammar fragments shown in Table 2 accounted for 70.1% of the BNC corpus and 71.2% of the NEGRA corpus. Infinite verb phrases constitute approximately half of all verb phrases in the German grammar, and almost 2/3rds of verb phrases in the English grammar. They therefore make a significant contribution towards learning noun-verb sequential order.

We adapted the SRN model to fit the new grammar, with units in input/output layers for plural/singular noun, preposition and end of sentence marker. For the verbs, we included a unit that was activated whenever a verb occurred in the sequence. One of three other units was also active for each verb, corresponding to whether the verb was finite and singular, finite and plural, or infinite. This was to ensure that there was some overlap in the representation of each verb. The model was trained and tested in exactly the same way as for the model trained on the first grammar. After 10 epochs of training, MSE for English was 0.038 (sd = 0.004), and 0.018 (sd = 0.005) for German.

The model trained on the German grammar learned the conditional probabilities with greater accuracy, $t(18) = 10.39, p < .001$. Figure 3 shows the model's performance on predicting each category of word for Grammar 2. The interposition of the infinite verb after the finite verb resulted in greater difficulty in learning for the English grammar. The verb-final structure of the German grammar, in contrast, was learned with relative ease. Figure 3 indicates that this was due to the 100% predictability of the end-of-sentence occurring after the infinite verb in German. In English, the infinite verb can precede a noun, a preposition or an end-of-sentence marker. The difference for infinite verbs was highly significant, $t(18) = -48.01$, $p < 0.001$. However, the difference for finite verbs was also significantly different, with English being more predictable than German, $t(18) = 10.61$, $p < 0.001$. This was due to the high frequency of the infinite verb following the finite verb in English, whereas in German the finite verb can precede a noun, a preposition, or an infinite verb. No other comparisons were significant. Finite verb phrases and infinite verb phrases are learned earlier in main clauses than word order in subordinate clauses (Clahsen & Muysken, 1986). However, children do not make word-order errors in constructing subordinate clauses in German (Meisel & Müller, 1992; Rothweiler, 1993). Does this verb-final construction in infinite verb phrases assist the acquisition of the word-order inconsistencies in the subordinate clauses in German? The next simulation tests this question.

Table 3. Grammar 3 for English and German with finite and infinite verb phrases in main and subordinate clauses, showing the proportions of each structure.

ENGLISH	PROPORTIONS
S → S S-bar	18.6%
S → NP V_fP	35.7%
S → NP V_iP	64.3%
S-bar → C NP V_fP_{sub}	25.2%
S-bar → C NP V_iP_{sub}	74.8%
NP → N (PP)	75.7% (25.2%)
PP → P NP	100%
V_fP → V_f (NP) (PP)	34.3% (48.5%) (17.2%)
V_iP → V_f V_i (NP) (PP)	35.6% (40.0%) (24.5%)
V_fP_{sub} → V_f (NP) (PP)	59.9% (28.9%) (11.2%)
V_iP_{sub} → V_f V_i (NP) (PP)	43.1% (45.5%) (27.5%)
GERMAN	
S → S S-bar	18.7%
S → NP V_fP	50.8%
S → NP V_iP	49.2%
S-bar → C NP V_fP_{sub}	45.7%
S-bar → C NP V_iP_{sub}	54.3%
NP → N (PP)	59.2% (40.8%)
PP → P NP	100%
V_fP → V_f (NP) (PP)	8.5% (73.9%) (17.6%)
V_iP → V_f (NP) (PP) V_i	10.1% (69.5%) (19.0%)
V_fP_{sub} → (NP) (PP) V_f	19.8% (69.3%) (10.9%)
V_iP_{sub} → (NP) (PP) V_i V_f	(77.0%) (19.8%) 3.3%

Finite/infinite and main/subordinate clauses

We constructed the grammar fragment of English and German that included both finite and infinite verbs in main and subordinate clauses. The grammar is shown in Table 3. This grammar fragment accounted for 90.8% of the German corpus and 86.2% of the English corpus. The model was adapted from the previous simulation by adding a unit at the input and output layers for the complementiser. Once again, 10 simulations of each grammar were trained on training sets of 1000 sentences that were randomly generated according to the general proportions of structures found in the language corpora.

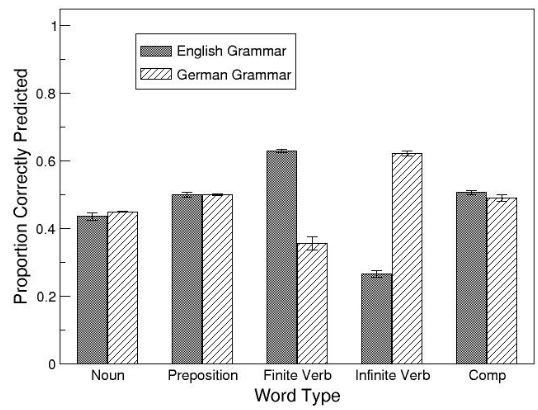

Figure 4. Proportion of correct predictions for each word category for Grammar 3 with finite and infinite verb phrases in main and subordinate clauses.

After training for 10 epochs, the model that had been exposed to the English grammar resulted in MSE = 0.028 (sd = 0.006) and the model trained on the German grammar had MSE = 0.024 (sd = 0.004). The difference was not significantly different: $t(18) = 1.76$, $p = 0.09$.

Figure 4 shows the performance of the model on predicting the word category following each word type. Though the model's performance did not differ overall between the English and the German grammar, there were differences for the different verb types. For infinite verbs, German was easier than English, $t(18) = -28.47$, $p < 0.001$. For finite verbs, English was easier than German, $t(18) = -13.67$, $p < 0.001$. The lower predictability of words following finite verbs in German is consistent with the claim that case-marking is especially useful for languages with greater variation in word order (Lupyan & Christiansen, 2002). No other comparisons were significant.

Of particular note is the absence of an effect of prediction after the noun. For Grammar 1, the sequence following a noun is easier to learn for English than German, but for Grammar 3 there is no significant difference between English and German. We interpret this as the different word order in subordinate clauses in German being rendered more easily learnable when infinite verb structures are also included in the grammar. This makes sense if one interprets the language learner acquiring one of a competing set of grammars. The verb-final structure appears to affect performance on learning the word order of subordinate clauses.

Conclusion

The models indicate that, for the learning of sequential order of nouns and verbs, SVO and SOV word order inconsistencies are alleviated by overlapping patterns of word ordering in other structures in the grammar. The verb-final structure in German infinite verb phrases results in less difference between predictions of the lexical category following a noun for English and German. We contend that the survival of the different word order in subordinate clauses in German is due, in part at least, to these verb-final constructions. The verb-final construction in subordinate clauses is rendered more easily learnable as this pattern of word order is more common in Grammar 3 than in Grammar 1. More generally, this indicates that the implications for inconsistencies in structures in languages are not sufficient alone to determine whether the language will survive: the interplay between different structures must be considered.

We have indicated the importance of melding corpus-based analyses with computational modelling in order to test hypotheses about the contributions of different word orders towards learnability of sequences in languages. Though simplified, the grammars we used to train the models were based on real corpora. We have also indicated that fine-grained analyses of the performance of SRNs, by focusing on performance by lexical category, can reveal points at which learning sequences is improved and impaired when grammars are varied.

Acknowledgments

The first and third authors were supported by a Human Frontiers of Science Program grant. The second and third authors were supported by European Commission grant RTN-HPRN-1999-00065.

References

Burnard, L. (1995). *Users Guide for the British National Corpus.* British National Corpus Consortium, Oxford University Computing Service.

Cable, T. & Baugh, A.C. (1993). *A History of the English Language*, *4th Edition*. London: Routledge.

Christiansen, M.H. & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, **23**, 157-205.

Christiansen, M.H. & Devlin, J.T. (1997). Recursive inconsistencies are hard to learn: A connectionist perspective on universal word order correlations. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*. (pp.113-118). Mahwah, NJ: Lawrence Erlbaum.

Clahsen, H. & Muysken, P. (1986). The availability of Universal Grammar to Adult and Child Learners. *Second Language Research*, **5**, 1-29.

Cottrell, G.W. (1986). A model of lexical access of ambiguous words, in S.I. Small, G.W. Cottrell & M.K. Tanenhaus (eds.) Lexical ambiguity Resolution. San Mateo, CA: Morgan Kaufman, pp.179-194.

Elman, J.L. (1990). Finding Structure in Time. *Cognitive Science*, **14**, 179-211.

Grüning, A. (2003). *Why Verb-Initial Languages are not Frequent*. MPI-MIS Preprint Series, 10. Max Planck Institute for Mathematics in the Sciences, Leipzig.

Hawkins, J. A. (1994): *A Performance Theory of Order and Constituency.* Cambridge: Cambridge University Press.

Hawkins, J. A. (1990). A parsing theory of word order universals. *Linguisitc Inquiry*, **21**, 223-261.

Kiparsky, P. (1996): The shift to head-initial VP in Germanic, in: Epstein, S., Thrainsson, H. & Peters, S. (Eds.): *Studies in Comparative Syntax, Volume 2*. Dordrecht: Kluwer.

Kroch, A. (2000). Syntactic Change, in: Baltin, M., Collins, C.: *Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell.

Lightfoot, D. (1991). *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: MIT Press.

Lupyan, G., Christiansen, M.H. (2002). Case, word order and language learnability: Insights from connectionist modeling. *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 220-225). Mahwah, NJ: Lawrence Erlbaum.

MacDonald, M.C., Christiansen, M.H., (2002): Reassessing Working Memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, **109**, 35-54.

Meisel, J.M. & Müller, N. (1992). Finiteness and Verb Placement in Early Child Grammars: Evidence from Simultaneous Acquisition of French and German in Bilinguals. In: Meisel, J. M. (Ed.). *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition.* Dordrecht: Kluwer.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S. & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, **103**, 56-115.

Rothweiler, M. (1993). *Der Erwerb von Nebensätzen im Deutschen: Eine Pilotstudie*. Tübingen: Niemeyer (Linguistische Berichte, Sonderheft 3).

Skut, W., Krenn, B., Brants, T. & Uszkoreit, H. (1997). An annotation scheme for free word order languages. *Proceedings of the Fifth Conference on Applied Natural Language Processing. Washington, DC.*

Van Everbroeck, E. (1999). Language type frequency and learnability. A connectionist approach. *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp.755-760). Mahwah, NJ: Lawrence Erlbaum.