A Matter of Time: Implicit Acquisition of Recursive Sequence Structures

Julia Uddén [a,b,c] (Julia.Udden@ki.se)
Susana Araujo [c,d] (smaraujo@ualg.pt)
Christian Forkstam [a,b,c] (Christian.Forkstam@ki.se)
Martin Ingvar [b] (Martin.Ingvar@ki.se)
Peter Hagoort [a,c] (Peter.Hagoort@mpi.nl)
Karl Magnus Petersson [a,b,c,d] (Karl-Magnus.Petersson@mpi.nl)

a Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
b Cognitive Neurophysiology Research Group, Stockholm Brain Institute
Karolinska Institutet, Stockholm, Sweden
c Donders Institute for Brain, Cognition and Behaviour: Centre for Cognitive Neuroimaging
Radboud University Nijmegen, the Netherlands
d Cognitive Neuroscience Research Group, Universidade do Algarve, Faro, Portugal

Abstract

A dominant hypothesis in empirical research on the evolution of language is the following: the fundamental difference between animal and human communication systems is captured by the distinction between regular and more complex non-regular grammars. Studies reporting successful artificial grammar learning of nested recursive structures and imaging studies of the same have methodological shortcomings since they typically allow explicit problem solving strategies and this has been shown to account for the learning effect in subsequent behavioral studies. The present study overcomes these shortcomings by using subtle violations of agreement structure in a preference classification task. In contrast to the studies conducted so far, we use an implicit learning paradigm, allowing the time needed for both abstraction processes and consolidation to take place. Our results demonstrate robust implicit learning of recursively embedded structures (context-free grammar) and recursive structures with cross-dependencies (context-sensitive grammar) in an artificial grammar learning task spanning 9 days.

Keywords: Implicit artificial grammar learning; centre embedded; cross-dependency; implicit learning; context-sensitive grammar; context-free grammar; regular grammar; non-regular grammar

Introduction

During the past decade, investigations of language acquisition as well as language evolution have been revitalized by the artificial grammar learning (***AGL***) paradigm which allows animals as well as children and adult humans to implicitly acquire new syntactic structures without explicit teaching, i.e., similar to the conditions for natural language development. In this context, implicit learning is a process whereby a complex, rule-governed knowledge base is acquired largely independent of awareness of both the process and product of acquisition (Reber, Walkenfeld & Hernstadt, 1991). In AGL, one separates the acquisition and the testing phase, and the paradigm consists of at least one acquisition and classification session. In the acquisition phase, participants are typically engaged in a short term memory task using an acquisition sample of sequences generated from a formal grammar. Subsequently, subjects are informed that the symbol sequences were generated according to a complex system of rules and asked to classify novel items as grammatical or not, typically with the instruction to base their classification decisions on their immediate intuitive impression (i.e., guessing based on "gut feeling"). It is a robust finding on regular grammars that subjects perform well above chance and more so after several days of learning (Folia et al., 2008; Forkstam, Elwér, Ingvar & Petersson, 2008).

Taking the perspective that some aspects of the faculty of language are shared with nonhuman animals (faculty of language in a broad sense; ***FLB***) and that other aspects are specific to human language (faculty of language in a narrow sense; ***FLN***), the quest for FLN in AGL has centered around the theoretical construct of the ***Chomsky hierarchy*** – a complexity hierarchy for formal grammars, which are divided into regular (finite state; T3), context-free (phrase-structure; T2), context-sensitive (T1), and general phrase-structure grammars (Turing-Tue; T0), and its embodiment in the recursion-only hypothesis for FLN outlined in a seminal paper by Hauser, Chomsky and Fitch (2002). For example, in a sentence such as:

The cat the rats the dog chases fear is sitting in the yard.

The recursive embedding of subordinate phrases in super-ordinate phrases introduces morphological noun-verb agreement dependencies or what we here call nested dependencies. In a recent paper (de Vries, Monaghan, Knecht, & Zwitserlood, 2008), participants were trained on such sequences following the pattern $A_1A_2A_3B_3B_2B_1$ and tested on different kinds of violations, all in one session. Critically, there was no indication of learning in the hierarchical vs. scrambled condition, where non-grammatical sequences were only violating the correspondence rules of the dependencies (i.e., $A_1A_3A_2B_3B_2B_1$ as opposed to more significant or salient violations, such as the number of repetitions of each type of constituent $A_1A_2A_3A_4B_2B_1$ where learning was present). These results replicate earlier findings showing sensitivity for gross violations (Friederici et al., 2006) but also suggest that these are most likely dependent on explicit strategies such as counting or repetition monitoring. This is also suggested by de Vries et al. (2008). They proposed that different amounts of phonological rehearsal in verbal working memory might explain the imaging results of Friederici et al. (2006) and similar studies, which has been interpreted as suggesting a specialization within Broca's complex for processing the different grammar classes. As stated in de Vries et al. (2008), it is clear that the ability to learn hierarchical embeddings in AGL still needs to be demonstrated.

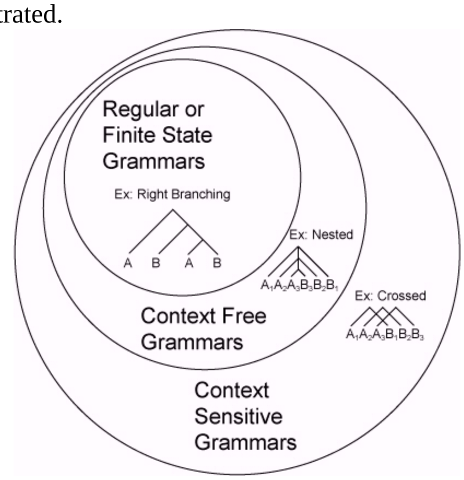

Figure 1. A Venn diagram of the first three classes in the ***Chomsky hierachy***. Informally, regular grammars are built from a collection of production rules of the form $S \rightarrow abS$ and $S \rightarrow ab$ (where lower case indicates terminal symbols and S the start symbol). It is the inclusion of the start symbol on the right hand side of the first rule that makes a grammar recursive. The context-free case allows the right hand side to involve terminal symbols around the start symbol as in $S \rightarrow aSb$ and $S \rightarrow ab$. In context-sensitive case, the left hand side has a context as in $a_1a^nSb_1b^n \rightarrow a_1a^na_{n+2}Sb_1b^nb_{n+2}$.

A variation of the nested dependencies described above is the crossed dependency pattern $A_1A_2A_3B_1B_2B_3$, famous in linguistics for being perhaps the only naturally occurring context-sensitive construction (versions exist in e.g. Dutch and Swiss German). A qualitative match between the performance of simple recurrent networks (SRN) and human perceived comprehension (investigated in Bach, Brown & Marslen-Wilson 1986) of sentences with nested dependencies and crossed dependencies was reported by Christiansen and Chater (1999). The authors argued that these results disqualify the principled Chomskian argument which, because of context-free and context-sensitive competence, language processing needs more power than the finite state architecture can provide (cf., Petersson, 2005). In fact, the SRN performance is higher on the supposedly more complex context-sensitive construction (crossed dependencies) than on context-free nested structures, thus mimicking natural language comprehension.

The crossed dependency has been relatively ignored in the empirical literature and has not yet been studied in AGL paradigms. In the present study, we dug deeply for experimental evidence for implicit learning of artificial grammars with crossed and nested dependencies. By using an AGL paradigm of extensive length (9 days) we allowed enough time needed for both generalization/abstraction processes and consolidation to take place. We minimize the influence of explicit knowledge and explicit strategies by using a preference instruction in addition to a grammaticality instruction in the test phase. In the preference version, participants are not informed about the existence of a grammar but are asked to make a preference choice for each string: like/dislike. A preference for grammaticality has been repeatedly found for finite state grammars (Folia et al., 2008; Forkstam et al., 2008) and functional neuroimaging data show the same activation pattern for preference and grammaticality classification (Folia et al., in preparation). The use of a preference classification baseline ensures that the effects observed in the classification task are due to information implicitly learnt during the acquisition phase. We used a between subject design with two groups learning two different grammars. This offers the possibility to test the robustness of learning and to compare possible differences in relation to the previously mentioned simulation and natural language results on the nested and crossed constructions.

Methods

Participants

39 right-handed healthy university students volunteered to participate in the study (28 females, 11 males, mean age±SD =21±2 years). They were all pre-screened for medication use, history of drug abuse, head trauma, neurological or psychiatric illness, and family history of neurological or psychiatric illness. All participants gave written informed consent and the study was run under the Donders Center for Cognitive Neuroimaging Experimental Approval from the local medical ethics committee at the UMC St. Radboud. 19 of the participants were exposed to a grammar with crossed agreement structure (context-sensitive) and 20 participants to a grammar with nested agreements (context-free).

Stimulus Material

We generated grammatical (G) sequences from a grammar with either a crossed (e.g., $A_1A_2B_1B_2$ or $A_1A_2A_3B_1B_2B_3$) or a nested agreement part (e.g., $A_1A_2B_2B_1$ or $A_1A_2A_3B_3B_2B_1$), of a total string length of 5-12 symbols (mean string length=10 symbols). For the regular pre- or suffix part we used the alphabet {M, N, S, V, W, R, X} and for the nested/crossed dependency part we used the alphabet {F, D, L, P}, see Figure 1. The first half of the agreement part was always taken from {F, D} and the last half from {L, P}. The crossed agreements were introduced as arbitrary agreements between the letter pairs F–L and D–P, such that if there was an F(D) in the first, second or third position of the first half, there was an L(P) in the same position in the last half of the string. The nested agreements were created by concatenating the first half with a reflection of the first half, but changed to the corresponding last half alphabet. We calculated the specific associative chunk strength (ACS) for each string in relation to the complete set of strings (Meulemans & Van der Linden, 1997). 100 strings were randomly selected and tested with respect to its ACS content in order to generate an acquisition set which was representative in terms of ACS in comparison to the complete string set and so that ACS was not significantly different between the acquisition set of nested and crossed agreements. The classification sets were derived from the remaining grammatical (G) sequences and for each of these non-grammatical (NG) sequences were derived. Agreement violations were created by keeping the structure of F's and D's in the first part and L's and P's in the second part, but violating the agreements in the first, second or third positions or in combinations of these positions. The NG sequences were selected to match the grammatical strings in terms of complete string ACS (i.e., collapsed over order information within strings). Thus, the G and NG sequences are composed of equally common bi- and trigrams in the acquisition set. Finally, test sets of 64 strings were randomly selected from the grammatical (32) and their matched non-grammatical (32) sequences in an iterative procedure; the test sets did not differ statistically in terms of ACS or string length between any condition in any test set and irrespective of nested or crossed agreements as well as independent of grammaticality status (G/NG).

Experimental Procedure

The complete experiment spanned 9 days with one implicit acquisition session per day. On day one, a baseline preference classification test was administered before the first acquisition session. On the last day, subjects performed a preference and then a grammaticality classification test.

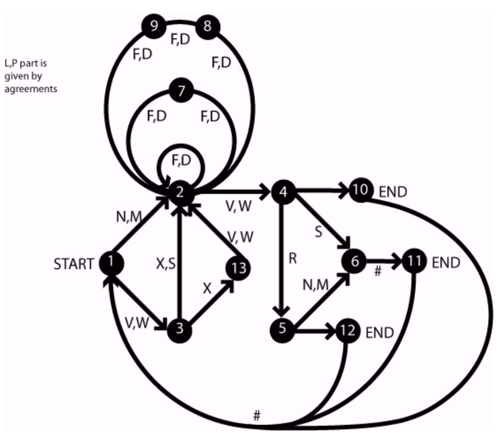

Figure 2. The transition graph used to generate the stimulus material. Grammatical strings were generated by first traversing the transition graph from the start node to the end nodes along the directions indicated by the arrows and concatenating the letters written on the traversed arrows. Nested or crossed dependencies were then imposed by changing the second half of the F/D-loops in the relevant ways. In this way, a grammatical string with crossed agreement structure, such as MFFDLLPVS, or with nested agreement structure, such as MFFDPLLVS, is generated while a non-grammatical string such as MFDFPLLVS cannot be generated in this way.

Implicit Acquisition Task

The acquisition task (~30min) was presented as a short-term memory immediate recall task to the subjects. During the acquisition task, each string was presented for 4s (whole string presentation), centrally placed on a computer screen using the Presentation software (nbs.neuro-bs.com). After the string disappeared from the screen, subjects recalled the string by self-paced typing on a keyboard. Subjects were allowed to correct themselves but no performance feedback was provided. The subjects were only exposed to grammatical examples and the presentation order of the 100 grammatical strings in the acquisition set was randomized over acquisition sessions.

Preference Classification Task

Subjects were instructed to indicate if they liked a string or not based on their immediate intuitive impression (i.e., guessing based on ''gut feeling''). Participants were told to respond as fast as possible after string onset (whole string presentation) and that there was no correct or incorrect response. The whole string was presented for 3.5s followed by an inter stimulus interval of 2.5s. The participants indicated their decision by pushing the corresponding key with their left or right index finger. The response hand was balanced across preference and grammaticality classification tests and across subjects. The presentation order of classification string of the sets was balanced across subjects.

Grammaticality Classification Task

After having finished the preference classification task, the subjects were informed about the existence of a complex system of rules used to generate the acquisition strings (but they were not informed about which the actual rules were). They were then instructed to classify novel strings as grammatical or not based on their immediate intuitive impression as correct and as fast as possible after string onset. The same classification sets were used for preference and grammaticality classification and the order of the tests were balanced over subjects.

Data Analysis

Repeated-measures ANOVAs and t-tests were used for the analysis of the data using SPSS 15 and a significance level of $P<.05$ was used. We analyzed the classification performance with endorsement rate as the dependent variable and with the factors TEST, with two levels corresponding to the baseline and preference classification; and GRAMMATICALITY, with two levels G and NG. Endorsement rate is defined as the number of strings classified as grammatical independent of their actual status, divided by the total number of recorded answers for each factor level (Meulemans & Van der Linden, 1997). For example, an endorsement rate of .3 for NG-strings means that 70% of these strings were correctly rejected. Subsequently, we analyzed the additional factor STRING LENGTH, with three levels: short (9 or fewer symbols), medium (10 symbols) and long (11 or more symbols). D-prime and response bias were calculated using standard signal detection theory (Hochhaus, 1972).

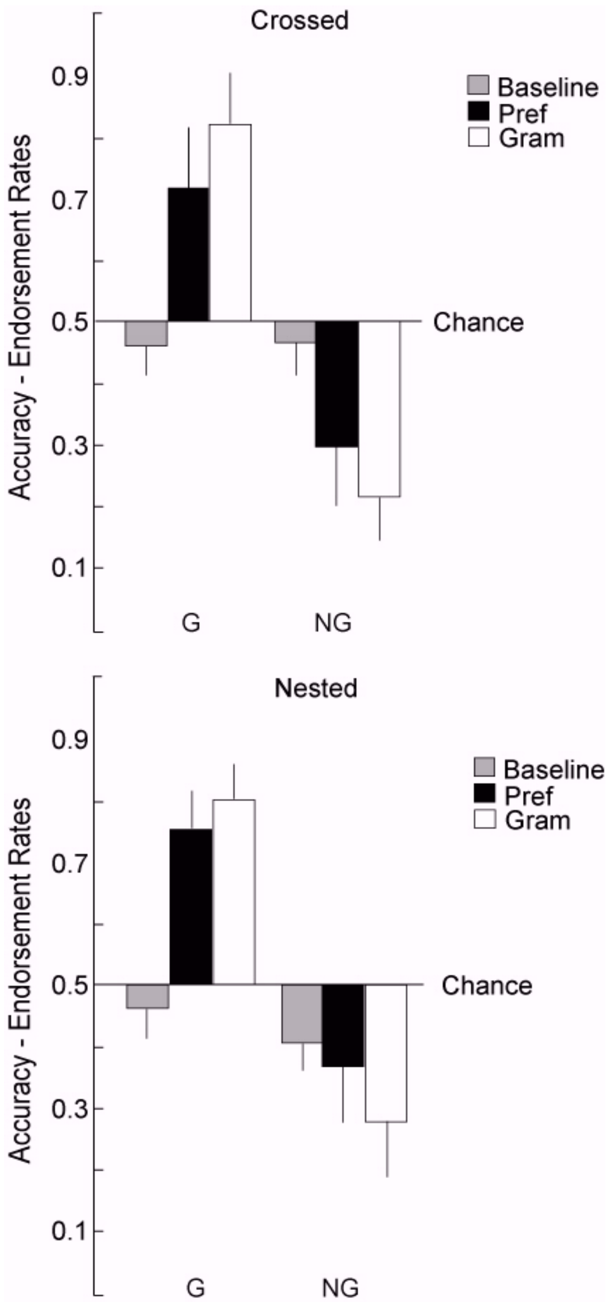

Figure 3. Classification performance in endorsement rates over the nine days of the experiment. Pref = preference classification of strings which was also used as baseline test. Gram = grammaticality classification. An endorsement rate of .3 for NG strings means that 70% of these strings were correctly rejected. Error bars indicate standard deviations.

Results

Classification Performance

Crossed dependencies**.** There was no effect of TEST on response bias (*P*=.14). There was a significant main effect of GRAMMATICALITY (*F*(1,18)=31.5, *P*<.001) and the interaction between TEST and GRAMMATICALITY was significant (*F*(1,18)=31.3, *P*<.001). The comparison between the baseline and the preference test were significant both for grammatical strings (*T*(18)=-5.5, *P*<.001, two-tailed) and agreement violations (*T*(18)=3.4, *P*<.01, two-tailed). We then compared the endorsement rates from the grammaticality classification against chance performance (.50) using a single sample, two tailed T-test. The grammatical strings were significantly different from chance (*T*(18)=8.3, *P*<.001) and well as the non grammatical strings (*T*(18)=-8.0, *P*<.001). Grammatical strings were endorsed significantly more often than NG strings in the preference (*T*(18)=6.0, *P*<.001) and in the grammaticality test (*T*(18)=9.0, *P*<.001) but not the baseline test (*P*=.87). Taken together, these results show robust implicit acquisition of the crossed syntax.

Nested dependencies**.** There was a main effect of TEST on response bias (*F*(1,19)=23.9 *P*<.001) but no significant interaction with any other experimental factor. There was a significant main effect of GRAMMATICALITY (*F*(1,19)=44.0, *P*<.001) and a significant interaction between TEST and GRAMMATICALITY (*F*(1,19)=18.8, *P*=.001). However, in the direct comparison of the preference and grammaticality test using two-tailed T-tests, we only found significant effects for grammatical strings (*T*(19)=-7.8, *P*<.001) while the effect of agreement violation was non-significant (*T*(19)=.55, *P*=.59) from the baseline to the last preference test. We then compared the endorsement rates during the grammaticality classification against chance performance (.50) using a single sample, two tailed T-test. Grammatical strings (*T*(19)=11.2, P<.001) as well as non-grammatical strings (grammaticality test 1 *T*(19)=-5.0, *P*<.001) were significantly different from chance. The grammatical strings were endorsed significantly more often than non-grammatical strings in the preference test (*T*(19)=6.0, *P*<.001) and the grammaticality test (*T*(19)=8.4, *P*<.001) but not the baseline test (*P*=.06). These results suggest two possible reasons for the weak acquisition effect between the baseline and the last preference test for the NG strings: early acquisition or preexisting bias during the baseline test, consistent with previous results (Forkstam et al., 2008). Again, the results show robust implicit acquisition of the nested syntax.

Between group effects**.** The three way interaction between test, grammaticality and group was non-significant (*P*=.34) but the two-way interaction between group and test was significant (*F*(1,37)=4.6, *P*<.05) reflecting the fact that the baseline bias was unique to the nested group and not present in the crossed group. We also analyzed the data with respect to the factor VIOLATION POSITION in the NG strings. Since the results had no clear interpretation we choose to not report them here. However, this factor turned out to explain a lot of between group variance and when we included this factor in the model, the interaction between test and group was significant ($F(1,37)=12.9$, $P<.01$). This effect was not present for the grammatical strings ($P=.56$). Without this factor, the same interaction was marginally significant ($F(1,37)=3.6$, $P=.07$). Because this result is consistent with the predicted direction for an increased correct rejections rate for crossed violations compared to nested violations, we think that these result are consistent the reported pattern of results for natural languages and supports a generalization to general sequence structures.

String length effects. There was a main effect of STRING LENGTH for both crossed ($F(2,17)=11.0$, $P=.001$) and nested strings ($F(2,18)=20.1$, $P<.001$) reflecting the fact that longer strings were more likely to be rejected independent of grammaticality. In the nested material, there was an additional interaction between STRING LENGTH and TEST ($F(2,18)=7.5$, $P=.01$), resulting from an initial bias for liking short strings and disliking long strings in the baseline test. This bias diminished during acquisition measured by the preference test on the last day.

We predicted the largest differences between grammar types on the longest string lengths. Focusing on these, the interaction between test, grammaticality and group approached significance ($F(1,37)=3.1$, $P=.09$), meaning that the crossed group showed a greater interaction between test and grammaticality compared to the nested group. This effect was related to NG strings, where there was a significant interaction between group and test ($F(1,37)=10.3$, $P<.01$) which was not present for the grammatical strings ($P>.35$).

Signal detection analysis

Crossed dependencies. D-prime varied significantly over the factor TEST ($T(18)=-4.9$, $P<.001$). D-prime was highly significant for both preference and grammaticality classification ($P<.001$) but not for the baseline preference test ($P=.89$). There was no bias in either test and the bias did not interact with TEST ($P>.15$).

Nested dependencies. D-prime varied significantly over the TEST factor ($T(19)=-2.8$, $P=.01$). D-prime was significant for both preference ($T(19)=3.2$, $P<.01$) and grammaticality classification ($T(19)=4.2$, $P<.001$) and only a trend in the baseline preference test ($P=.06$). There was also a bias trend in the baseline test ($T(19)=2.0$, $P=.06$) but not in any other tests. The bias did not interact significantly with TEST ($P>.10$).

Between group effects. There were no significant interactions between group and any other factor.

Post experimental questionnaires

The post-experimental questionnaire was distributed after the last grammaticality test. Participants accepted the following statement as true significantly better than chance: "There were correspondences between certain letters, such that if certain letters were in certain positions, this meant that there were always certain letters in certain other positions". However, when the participants were provided with the correct agreement constraints "If there was an F in the first half of the FDLP-group there had to be a L in the same position in the second half and D's corresponded to P's in the same way, as in FDFLPL" as true, they performed at chance level in terms of accepting or rejecting it. Subject who identified the correct constraints (i.e., answered yes on the last question) were not significantly better than subjects who did not.

Discussion

The present study provides strong evidence for the human capacity to implicitly acquire nested and crossed long-distance dependencies in consonant sequences. Compared to previously reported performance levels for regular grammars (e.g., Folia et al., 2008; Forkstam et al., 2008), the participants in this study performed at comparably high levels on two different types of non-regular grammars, one context-free (nested structures) and one context-sensitive (crossed structure), consistent with the theoretical analysis of Petersson (2005). Critically, participants developed a sensitivity to violations that only violate the agreement structure quantified by endorsement rates under both preference and grammaticality instructions. The preference instruction, which takes advantage of the structural mere exposure effect (cf., Folia et al., 2008; Forkstam et al., 2008), has the benefit of never making it necessary to mention the existence of an underlying generative mechanism, which also minimizes the likelihood of engaging explicit strategies (Folia et al., 2008; Forkstam et al., 2008).

A previous study of embedded (nested) recursion as Reber fragments within Reber fragments with push and pop transitions (Poletiek, 2002), reported a weak learning effect for one and two levels of nesting. De Vries et al. (2008) did not observe sensitivity to the grammaticality factor when the dependency structure was similar to the agreement constraints used here. In a series of experiments using the A^nB^n grammar without dependencies, using violations with an unequal number of repetitions of A's and B's, concluded that participants did not learn the underlying grammar (Hochmann, Azadpour, & Mehler, 2008). One reason for this pattern of results might be that subjects were not provided with enough time or exposure to grammatical items during acquisition. In these studies, subjects were only exposed during one brief acquisition session. Given the present results, it seems that more time and exposure is needed for the abstraction processes supporting implicit acquisition to take place in a measurable way. Consistent with this suggestion, the results obtained in two experiments using the A^nB^n paradigm, which showed successful acquisition in European starlings (e.g., Gentner, Fenn, Margoliash & Nusbaum, 2006) while cotton top tamarin monkeys failed to acquire the paradigm (Fitch & Hauser, 2004), might be explained by the fact that the European starlings received extensive training (~300 000 trials) while the tamarins were only given 20 min of exposure on the day preceding testing.

In the current study, the learning effect was more robust for the crossed ($F(1,18)=31.3$, $P<.001$) compared to the nested grammar ($F(1,19)=18.8$, $P=.001$). This was reflected in a significant difference between the crossed and nested group on the increase of correct rejections in preference classification compared to baseline classification (crossed>nested; *P*<.01).This is consistent with earlier results suggesting that it is easier to comprehend crossed compared to nested constructions in natural language (Bach et al., 1986), an effect that has also been captured qualitatively in simple recurrent network simulations (Christiansen & Chater, 1999). This suggests that the previously reported natural language differences between crossed and nested structures might generalize to structured sequences more generally.

Finally, we note that skepticism concerning the relevance of the Chomsky hierarchy for languages is not new (Bach et al., 1986; Christiansen & Chater, 1999; Petersson, 2005). In fact, under finite memory constraints (or finite precision computation), it is known that the Chomsky hierarchy as a complexity measure is of limited relevance. Given infinite or limitless memory resources, the Chomsky hierarchy is a memory hierarchy rather than a processing hierarchy per se. In fact, any recursive type can be instantiated in an architecture with finite memory - it simply cannot be used in an unbounded way (see e.g., Petersson, 2005). In other words, any recursive phenomena can be captured in a finite state architecture (Davis et al., 1994; Savage, 1998). However, this is not really an issue from the point of view of natural language - what are fundamental to language are long-distance dependencies and not arbitrarily long long-distance dependencies. Thus more relevant complexity measures need to be developed, for example real-time computational complexity in line with contemporary complexity theory (Savage, 1998; Petersson, 2008).

Conclusion

We have presented an implicit artificial learning paradigm based on the structural mere exposure effect which demonstrates robust implicit learning of long-distance dependencies in both context-free and context-sensitive grammars. We have enriched the ecological validity of the AGL paradigm in relation to natural syntax acquisition by investigating longer periods of learning and, by the use of preference classification, minimizing contamination by explicit processes. The results extend earlier results from natural language suggesting that Chomskian complexity theory is irrelevant for empirical research on systems with finite memory resources such as the brain.

Acknowledgments

This work was supported by Max Planck Institute for Psycholinguistics, Donders Institute for Brain, Cognition and Behaviour, Vetenskapsrådet, Hedlunds Stiftelse and Stockholm County Council (ALF, FoUU).

References

Bach, E., Brown, C., Marslen-Wilson, W. . (1986). Crossed and nested dependencies in German and Dutch: A psycholinguistic study. *Language and Cognitive Processes, 1*, 249–262.

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

Davis, M. D., Sigal, R., & Weyuker, E. J. (1994). *Computability, Complexity, and Languages*: *Fundamentals of Theoretical Computer Science* (2 ed.). San Diego, CA: Academic Press.

de Vries, M. H., Monaghan, P., Knecht, S., & Zwitserlood, P. (2008). Syntactic structure and artificial grammar learning: The learnability of embedded hierarchical structures. *Cognition*, *107*, 763-774.

Fitch, W. T., Hauser, M.D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science, 303*, 377-380.

Folia, V., Uddén, J., Forkstam, C., Ingvar, M., Hagoort, P., & Petersson, K. M. (2008). Implicit learning and dyslexia. *Ann. N. Y. Acad. Sci. 1145*, 132–150.

Folia, V., Forkstam, C., Ingvar, M., Hagoort, P., & Petersson, K. M. (in preparation)

Forkstam, C., Elwér, Å., Ingvar, M., Petersson, K. M. (2008). Instruction effects in implicit artificial grammar learning: A preference for grammaticality. *Brain Research, 1221*, 80-92.

Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. D., & von Cramon, D. Y. (2006). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cerebral Cortex, 16*, 1709-1717.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*, 1204-1207.

Hauser, M. D., Chomsky, N., Fitch, W.T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science, 298*, 1569-1579.

Hochhaus, L. (1972). A table for the calculation of d' and ß. *Psychological Bulletin, 77*, 375-376.

Hochmann, J.-R., Azadpour, M., & Mehler, J. (2008). Do Humans Really Learn A^nB^n Artificial Grammars From Exemplars? *Cognitive Science, 32*, 1021-1036.

Meulemans, T., & Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn., 23*, 1007-1028.

Petersson, K.M. (2005). On the relevance of the neurobiological analogue of the finite state architecture. *Neurocomputing, 65-66*, 825-832.

Petersson, K.M. (2008). On cognition, structured sequence processing, and adaptive dynamical systems. *Proceedings of the American Institute of Physics, Mathematical and Statistical Physics Subseries*, *1060*, 195-200.

Poletiek, F. H. (2002). Implicit learning of a recursive rule in an artificial grammar. *Acta Psychologica, 111*, 323-335.

Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and Explicit Learning: Individual Differences and IQ. *J. Exp. Psychol. Learn. Mem. Cogn., 17*, 888-896.

Savage, J. E. (1998). *Models of Computation: Exploring the Power of Computing*, Addison Wesley.