You can't get there from here: On interpreting learning experiments

Background. Experimental exposure to artificial languages has been proposed as a tool to evaluate the role of statistical learning in language acquisition (Saffran et al., 1996, et seq.). While this paradigm can be used to determine participants' ability to discriminate whether test items are in the language, it cannot be determined solely from participant performance how learners accomplished the task and whether the cues learners actually learned the desired structure using the cues given by the experimenter (cf. Perruchet and Vinter, 1998). Modeling is a necessary companion to such experiments to verify that that the proposed learning mechanisms are actually consistent with experimental results.

Proposal. We use a simple computational learning model to evaluate the claim that subjects in the experiments in Saffran 2001 learned hierarchical structures from an artificial language. We find that a learner that learns predictive dependencies between syntactic categories without any attempt to discover hierarchical information precisely predicts the observed human performance on the given task without learning the structure intended by the experimenters.

Simulation. We selected the results reported in Saffran 2001 for modeling due to the reporting of participant performance for each grammatical test administered. In that study, adult and child learners were exposed to stimuli from an artificial language generated using the context-free grammar given in Figure 1. The author claims that participants used predictive dependencies, represented as predictive relationships between syntactic categories, to learn the hierarchical structure of the language they were exposed to. Participants were tested using a two-way forced choice between a grammatical and ungrammatical item of the language (Table 2). Participants that had learned the hierarchical structure of the language would be expected to perform above chance on all tests.

However, knowledge of the hierarchical structure of the artificial language was not required to pass the administered tests. Takahashi (2009) showed that a finite state automaton (FSA) would be sufficient to pass; we explore a much more simple learning technique relying on the cues that Saffran (2001) suggest would be useful for acquiring hierarchical structure, predictive dependencies between syntactic categories.

We trained a learning model using the 18 unique syntactic category sequences generated by the grammar that the participants in the study of Saffran (2001) were exposed to examples of. The learner forms two generalizations based on the co-occurrence of syntactic categories in a sentence. It learns that a *requires* b if $p(a, b) = 1$, that is every sentence containing a also contains b. It learns that a *excludes* b if $p(a, b) = 0$, that is every sentence containing a does not contain b. Co-occurrence within the same syntactic category is defined such that if category a never occurs more than once in a sentence, a excludes a. The rules learned from exposure to the artificial language are given in Table 1.

The learner correctly responded to all but one of the sentence tests given to participants, failing the test "Rule 3" (Table 2). This tested whether subjects understood that there were two ways to expand BP that are in complementary distribution; if BP expands to E, it cannot also expand to CP and F. However, this failed test is the only test on which in both rounds of testing adult participants did not perform significantly higher than chance. If the participants had either learned the hierarchical structure of the language or learned an FSA representing the language, they would have passed this test in addition to all other tests.

Conclusion. We find that the proposed learning model provides a better fit to human performance than learning either the true hierarchical grammar or an FSA. The simplicity of the model that matches human performance brings into question the claims of Saffran (2001) that a study of that type can be used as an example of hierarchical rule learning. This finding highlights the risk of interpreting participants' general success at a task as evidence for learning the representation desired by the experimenters. Strong claims about the mechanisms of language learning must be accompanied by equally strong verification of those mechanisms and the experiments that suggest them.

Figure 1: Rules of the context-free grammar used in Saffran 2001. Optional productions are given in parentheses.

Category	**Requires**	**Excludes**
A		A
C	A	
D	A	D
E	A	E, F
F	A, C	E, F
G	A, C	

Table 1: Predictive dependency rules learned from the artificial language data.

Test	**Grammatical Item**	**Ungrammatical Item**	**Simulation Response**
Rule 1: Every sentence must contain an A word.	A C F	*C F	**Pass:** C and F each require A
Rule 2: No sentence may contain more than one A word.	A D E C	*A A D E C	**Pass:** A excludes another A
Rule 3: A BP expands to contain an E or a C but not both at once.	A D E	*A D C E	**Fail:** Learner labels *A D C E* as grammatical.
Rule 4: If there is a D word, then there must be an A word.	A D C F C	*D C F C	**Pass:** D requires A
Rule 5: If there is an F word, then there must be a C word.	A C F	*A F	**Pass:** F requires C
Rule 6: If there is a G word, then there must be a C word.	A E C G	*A E G	**Pass:** G requires C

Table 2: Forced-choice grammar tests administered in Saffran 2001 and whether the simulation correctly identified ungrammatical items.

References

P. Perruchet and A. Vinter. Parser: A model for word segmentation. *Journal of Memory and Language*, 39 (2):246–263, 1998.

J.R. Saffran. The use of predictive dependencies in language learning. *Journal of Memory and Language*, 44(4):493–515, 2001.

J.R. Saffran, R.N. Aslin, and E.L. Newport. Statistical learning by 8-month-old infants. *Science*, 274(5294): 1926–1928, 1996.

E. Takahashi. *Beyond statistical learning in the acquisition of phrase structure.* PhD thesis, University of Maryland College Park, 2009.