1. Abstract

Successful language acquisition involves generalization, but learners must balance this against the acquisition of lexical exceptions. Examples occur throughout language. For example, English native speakers know that certain noun-adjunctive combinations are impermissible (e.g. strong winds, high prices, strong breezes, *high greenes*). Another example is the restrictions imposed by verb subcategorization, (e.g.) *gave*/*sent* the ball to him; *gave*/*sent* the ball *him*. I *donated*/*carried*/*pushed* the ball to him; I *donated*/*carried*/*pushed* him the ball. Such lexical exceptions have been considered problematic for acquisition: it learners generalize abstract patterns to new words, how do they learn that certain specific combinations are restricted? (Baker, 1979).

Certain researchers have proposed domain-specific procedures (e.g. Pinker, 1991) resolves verb subcategorization in terms of subtle semantic distinctions. An alternative approach is that learners are sensitive to distributional rules and use this information to make inferences about when generalization is appropriate (Braune, 1971).

A series of Artificial Language Learning experiments have demonstrated that adult learners can utilize statistical information in a rational manner when determining constraints on verb argument-structure generalization (Wonnacott, Newport & Tanenhaus, 2008). The current work extends these findings to children in different linguistic domains (learning relationships between nouns and particles). We also demonstrate computationally that these results are consistent with the predictions of domain-general Hierarchical Bayesian model (cf. Kemp, Perfors & Tenenbaum, 2007).

2. Background

Wonnacott, Newport & Tanenhaus (2008) (henceforth WNT) conducted a series of Artificial Language Learning experiments in which adult participants were exposed to miniature languages with two competing non-syntactic transitive verbs. Verbs in these languages were arbitrarily constrained as to whether they occurred in just one or both structures, with no semantic or phonological cues to verb-type.

WNT Central questions:
- Do learners acquire verb-specific and verb-general statistical patterns?
- What factors affect the tendency to generalize a verb to a new construction not encountered in the input?

WNT Central Findings:
- Learners acquire both verb-specific and verb-general statistical patterns, (i.e. learned the likelihood of encountering a particular structure both with a given verb and with verbs in general).
- The tendency to use a verb in a new structure was affected by:
  - Verb frequency (less likelihood of generalizing a more frequent verb).
  - Extent to which construction usage was lexically determined across the language as a whole.
- The last was particularly obvious when comparing the treatment of very low frequency ‘minimal exposure’ verbs by learners of different languages.
- WNT argued that learners were using statistical information in accordance with its utility/relevance in the past – i.e. showing rational or domain-general statistical learning.

The need for Artificial Language Learning experiments with children:
- First-language acquisition primarily occurs in early childhood.
- Language learning (first and second) is generally more successful when it begins in early childhood.

3. Experiments

Participants:
44 children recruited from Year 1 classrooms (mean age 6 years).
11 children assigned to learn each of 4 input languages below.

All monolingual native English speakers. Pseudo-random assignment matching ages across conditions.

**Language Paradigm**

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>cat, giraffe, pig, dog, cow, crocodile, mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>粒子 + 动词</td>
<td>moop = giraffe dow = cat dow</td>
</tr>
</tbody>
</table>

4. Input Languages (one for each of four groups)

**Generalist Language**
- Alternating nouns 75% dow, 25% tay
- Mixed language 1
- Alternating nouns 50% dow, 50% tay
- Mixed language 2
- Mixed language 3
- Mixed language 4

**Lexicalist Language**
- Dow-only nouns
- Lay-only nouns
- Mixed language 1
- Mixed language 2
- Mixed language 3
- Mixed language 4

Input languages differ in extent to which usage of particles is lexically determined.

Noun general usage of particles matched in lexical and generalist languages.

No noun-specific or phonological cues to noun type.

Testing Procedure : Production Test

NB – break mid testing for more sentence practice which provided the minimal exposure.

Input Languages and Test Types follow WNT. Note: new nouns explore generalization of noun general patterns (entirely novel nouns) and how learners deal with very small amount of language input (minimal exposure nouns).

Children asked to say the whole sentence. Score – use of dow versus tay.

5. Results

**Experimental Data:** Productions probabilities with familiar nouns

**In Generalist and Mixed languages usage of particles primarily influenced by noun general statistic** (t-test for both nouns). In Lexicalist language primarily influenced by noun specific statistic (t-test for both nouns).

**Minimal Exposure Nouns**
- In Generalist and Mixed languages usage of particles primarily influenced by noun general statistic (t-test for both nouns). In Lexicalist language primarily influenced by noun specific statistic (t-test for both nouns).

**Modeling data:** ‘Production’ probabilities with familiar and novel verbs.

**Summary**
Model qualitatively replicates critical aspects of human performance. (N.b. slightly more influence of ‘lexical’ constraints but model not subject to memory limitations – to be followed up?)

6. Conclusions
Like works in previous studies, children in these experiments show rational statistical learning when determining the extent of generalization. The results are consistent with a hierarchical Bayesian model capable of learning about structure variability on several levels simultaneously. Both humans and the model make inferences about the extent to which particle usage is lexically constrained. This statistic interacts with lexical frequency.

Acknowledgements
The work described in this paper was supported by Research Council and Post-Doctoral Fellowship. The first three authors’ work was supported by a grant from the Economic and Social Research Council awarded to the first author, British Academy (F2005161 2005-2007). The research was conducted in the Signal Processing Laboratory, Department of Computer Science, University of Oxford.