Acquisition of categorical non-adjacent dependencies in an artificial grammar

Jesse Storbeck¹, Elsi Kaiser¹, & Toben Mintz¹, 2 {jstorbenc, emkaiser, tmintz}@usc.edu
¹Department of Linguistics; ²Department of Psychology, University of Southern California, Los Angeles, CA, USA

1. Introduction

- Dependencies (between phonemes, morphemes, words, etc.) are a key feature of natural language. Two subtypes of dependencies:
  o Temporally adjacent dependencies: The cat is orange.
  o Non-adjacent dependencies (NADs): The cats down the street are orange.

- How do children and adults learn and process NADs?

- Artificial grammars are a common tool for investigating this question

- Prior work: focus on mutually-predictive (1:1) NADs (e.g., Gómez, 2002)
  o NADs like A:B; syllables A and B are mutually-predictive; X is unrelated

- But morphosyntactic NADs often involve categories, rather than discrete pairs
  o e.g., nominal NADs: that big/red/funny dog vs. those big/red/funny dogs

- Head noun agrees in number with demonstrative; adjective intervenes
  o that and dog are not mutually-predictive; are members of categories
  o This, that, these, & dog do not (dog, cat, car, ...)

- Categorically-represented NADs vs. 1:1 NADs

2. Materials & Methods

- Adults exposed auditorily to an artificial grammar with sentence structure:
  \[
  \begin{align*}
  s_1 & \rightarrow (\emptyset, \text{flanker}_{1} ) \\
  s_2 & \rightarrow (A_{1} (a), \text{slot}_{1} ) \\
  s_3 & \rightarrow (B_{1} (b), \text{slot}_{2} ) \\
  s_4 & \rightarrow (\emptyset, \text{flanker}_{2} )
  \end{align*}
  \]

- NADs are \(A_{1}(a)\) and \(B_{1}(b)\) (e.g., \(A_{1} x_{2} a_{2} f_{2}\) is “grammatical”; \(f_{2} B_{1} x_{1} a_{1} f_{1}\)
  o “teb maz perf” | “tay jeb dawk baan” | sol vip vam xodz guz | “big dafa perf”...

- Training (4 blocks of approx. 5 minutes)
  o Participants heard every “grammatical” NAD except:
    1. \((A_{2}(a), NAD\) with intervener \(x_{2}\) (\(B_{2}(b)\) NADs with intervener \(x_{2}\)
    2. All trigrams containing the NADs \(A_{1} x_{2} a_{2}\) and \(B_{1} x_{2} b_{2}\)
  o Syllables were 333 ms; 333 ms pauses between sentences
  o Each training block had 3 catch trials probing recall of most recent sentence

- Testing (4 blocks following each training with 29 items per block)
  o Participants heard and rated 3-syllable strings on 1-to-5 scale
    o Definitely heard string during training → 5
    o Definitely did not hear string during training → 1
  o Tested 50/50 mix of “grammatical” NADs and “ungrammatical” foils
    o Foils created by switching final dependent elements: \(a_{1} \leftrightarrow b_{1}, a_{2} \leftrightarrow b_{2}\)

- Test Conditions
  i. 20 trained trigrams + 20 foils
    o NADs heard during training with that particular intervener
    o Tested participants’ learning of NADs + memory of NAD-containing strings
  ii. 18 attested NADs + 18 foils
    o NADs heard during training with “new” intervener (see (1) above)
    o Trained (A1) NADs with \(x_{2}\), trained (B1) NADs with \(x_{2}\)
    o Tested participants’ learning of NADs
  iii. 20 unattested NADs + 20 foils
    o Untrained \(A_{2} x_{2}\) NADs, untrained \(B_{2} x_{2}\) NADs (see (2) above)
    o Tested participants’ generalization based on categories

3. Participants

- 47 native English speakers (30 F, 17 M)
- Undergraduates in USC Psychology courses
- 5 participants scored below chance on catch trials (6 out of 12 or below)
- They were excluded from main analysis due to insufficient attention to training stimuli
- Main analysis: \(n = 42\)
  o 28 F, 14 M
  o Catch trial performance (out of 12)
    * mean = 9.10, median = 9

4. Results

- Discriminated trained trigrams from “ungrammatical” foils
- Did not discriminate attested and unattested NADs from foils
  o 1-to-5 ratings scored based on each participant’s mean and sd in each block
  o Used linear mixed-effects regression for planned comparisons of 3 condition pairs
  o Trained trigrams rated significantly more familiar than their foils (\(t = 2.72\))
  o No such differences for attested (\(t = 0.42\) and unattested NADs (\(t = 0.92\))

- Findings contrast with results of van den Bos et al., where participants showed no learning of categorical NADs without supplementary cues during training

5. Conclusions

- Novel proof of learning category-based NADs based only on syllable co-occurrence
- Why were participants here able to learn, but failed in van den Bos et al.? (This question should be more difficult (18 discrete NADs trained vs. 4 in their task)
- Similar total training time in both experiments
- Most likely factors contributing to different experimental results:
  o Salience of utterance edges (Endress et al., 2009; Mintz et al., 2014)
  o No flankers in van den Bos et al. to vary absolute position of the NAD in sentence
  o Unnaturally slow rate of speech in previous work (Emberston et al., 2011)
  o Speech rate < 2 syllables/second, compared to 3 syllables/second in current study
- If participants learned NADs, why no discrimination in attested NAD condition?
  o Explicit learning of trigrams seems unlikely given their number (162 unique)
  o Lack of evidence for implicit learning perhaps due to consequence of foil design
  o Foils necessarily contain different component bigrams from grammatical trigrams
  o In attested NADs, grammatical contain no attested bigrams; foils contain 1
  o Adjacent statistics modulate NAD learning (Endres & Bonatti, 2007)
  o Lack of generalization may be due to insufficient overlap in (a/b) categories