The role of input variability for the acquisition of V2

Marc Meisezahl, Simon Kirby & Jennifer Culbertson

University of Edinburgh

Centre for Language Evolution

Collaborators

Jennifer Culbertson

Simon Kirby

German word order: V2

(1)   a.   Die Frau   *hat*   das Buch   mit Freude   gelesen
      the women has   the book   with pleasure read

   b.   Das Buch *hat* die Frau mit Freude gelesen.

   c.   Mit Freude *hat* die Frau das Buch gelesen.

   d.   Gelesen *hat* die Frau das Buch mit Freude.

   e.   Das Buch gelesen *hat* die Frau mit Freude.

   f.   Mit Freude das Buch gelesen *hat* die Frau.

subject   object   adjunct   past participle

08/06/2022

V2 diachronically

Two types of languages from diachronic perspective:

I. Languages that have retained V2 since earliest records (and even developed more strict system)
    - Most Germanic varieties

II. Languages that have lost V2 at some stage
    - English
    - Most of the Romance languages
    - Welsh

→Type II languages show steady decline of clear V2 structures

Language acquisition & change

- Lightfoot (1999, 2006) & Yang (2000): Robust attestation of evidence for V2 in learners' input necessary

- Loss of V2 in French (Yang 2000):
  - OVS, XVSO → V2; SXVO, XSVO → SVO
  - Analysis of sentences with *pro*-drop ambiguous: [X *pro* V] or [X V *pro*]
  - Roberts (1993): 5-18% VS structures, 40-52.5% SV structures in MidFr
  - More V>2 sentences than VS structures → SVO grammar

- What distribution of non-subject elements is most likely to facilitate acquisition of V2?

Overview

I. Evidence for V2

II. Corpus study

III. Artificial language learning experiment

Evidence for V2

Ideal input for learners of V2 language

- Ambiguity of SVO structures → Non-subject-initial sentences required

- Maximal variability of preverbal element (i.e. high entropy of preverbal position) and V2 without exceptions...

- ... but maximal variability of what?
  - Phrase types: NP/DP, PP, AdvP, CP etc. (Lightfoot 1999, 2006, Sitaridou 2012)
  - Grammatical functions: S, O & A (Yang 2000, 2002)

Variation and learning in the lab

The effect of variability on learning

- Facilitating effect of variability domain-general (Raviv et al. 2022)
- Goméz (2002), Goméz & Maye (2005):
  - Learning of non-adjacent dependencies by infants and adults (aXc, bXd)
  - Finding: Better learning of dependency when variability in X is higher

Variability and the acquisition of V2

- V2: X-V_{fin}
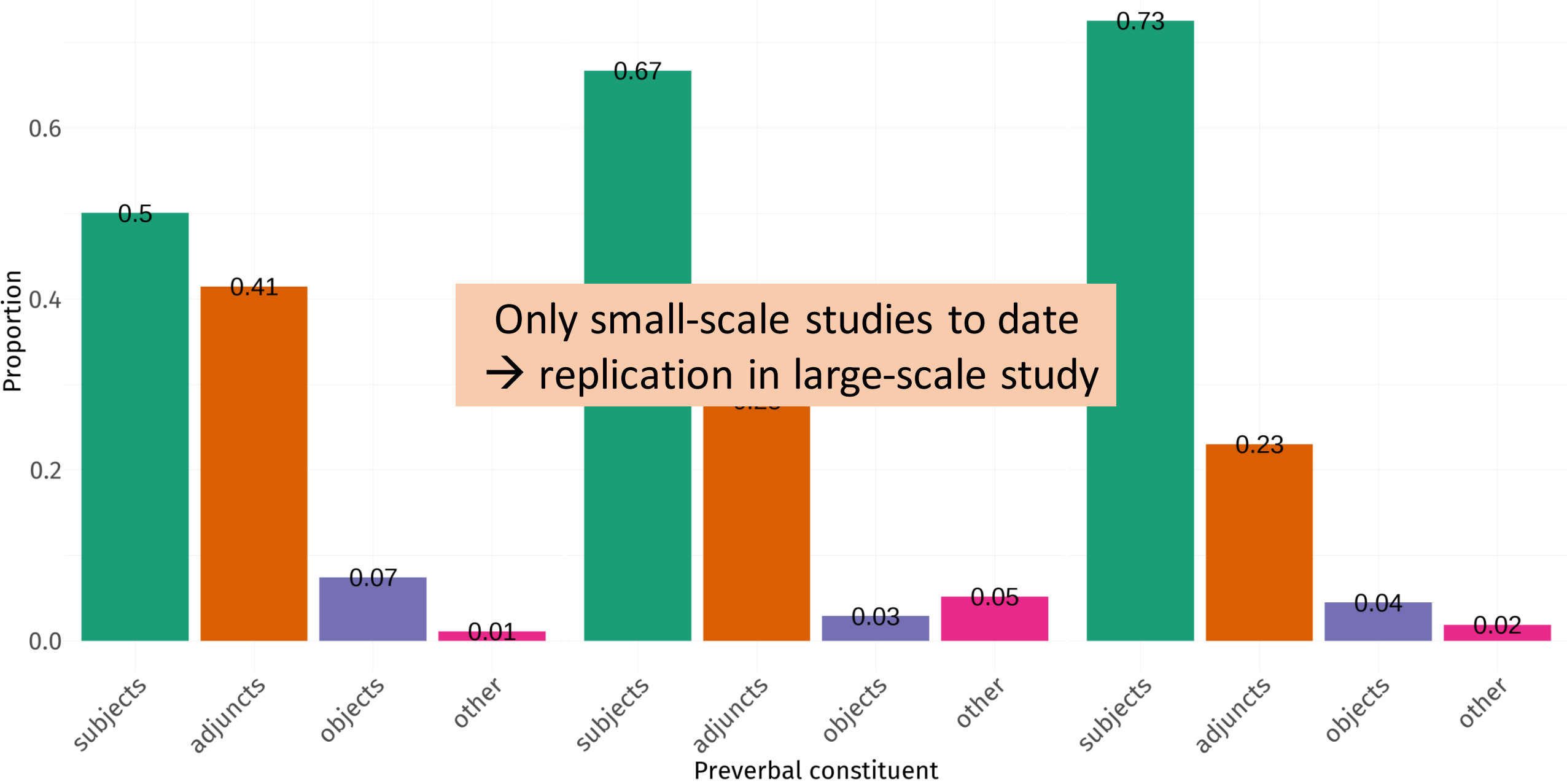- X = 1/3 S, 1/3 O & 1/3 A should result in best learning outcome

Corpus study

Proportion of different preverbal constituents

Only small-scale studies to date
→ replication in large-scale study
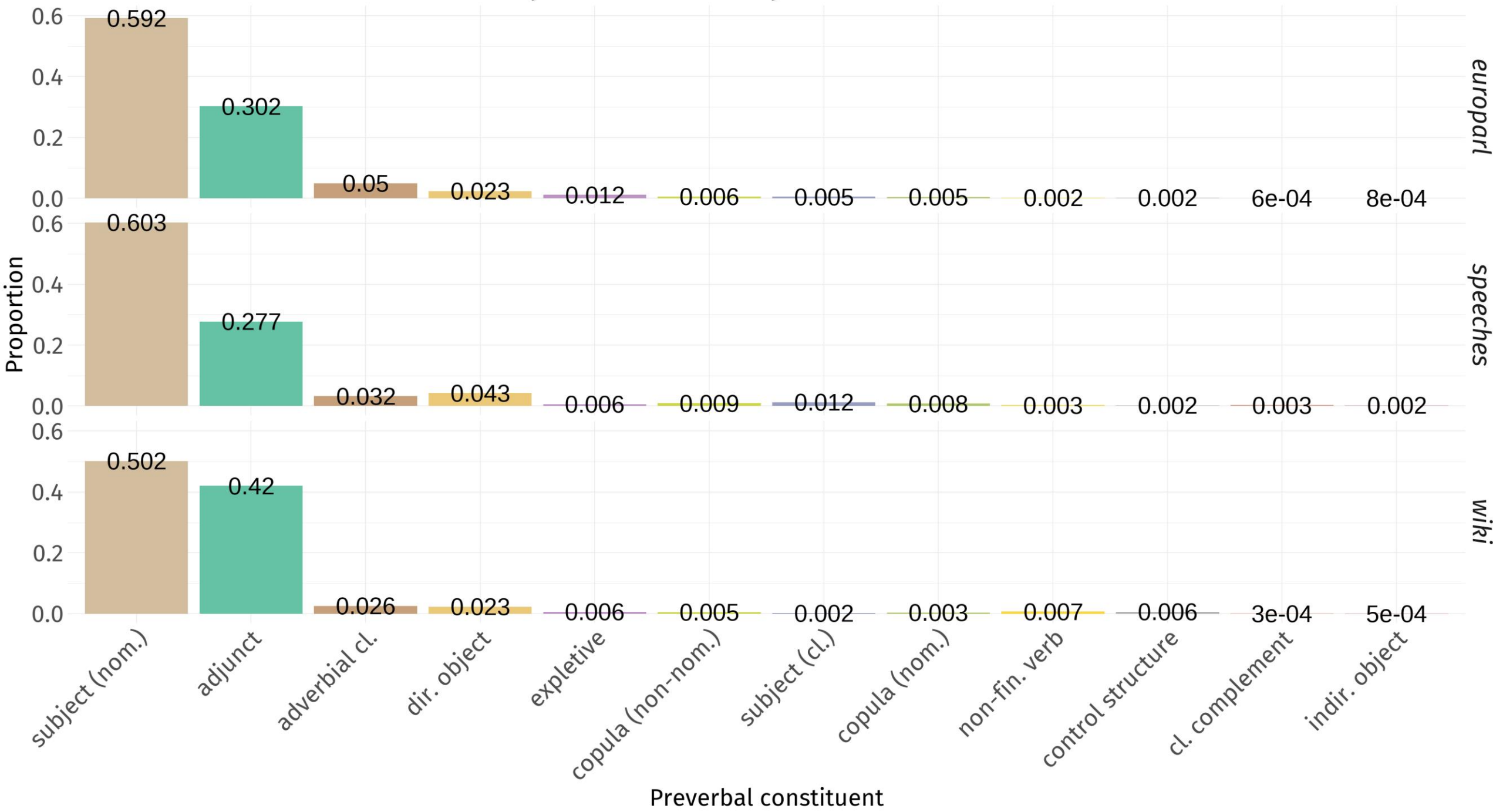
Corpus study

- German treebank *TüBa-D/DP* (de Kok & Pütz 2019):
  - Wikipedia (45.5M sentences)
  - European Parliament proceedings (2.2M sentences)
  - Speeches of German officials (619,152 sentences)
- Sentence inclusion criteria
  - Indicative V_{fin} in "prefield"
  - Minimum length of two words
  - No question or exclamation marks at the end of sentence

Proportion of different preverbal constituents

Corpus study

Discussion

- Skewed distribution persists in large scale corpora
- BUT raw frequency does not take into account the likelihood with which an element occurs in a sentence
  - $p(\text{init-S}|S) = p(\text{init-O}|O) = p(\text{init-A}|A) = \ldots?$
- Is the distribution still skewed once we take this into account?

Simulation

Procedure

- *n* V2 sentences *s*
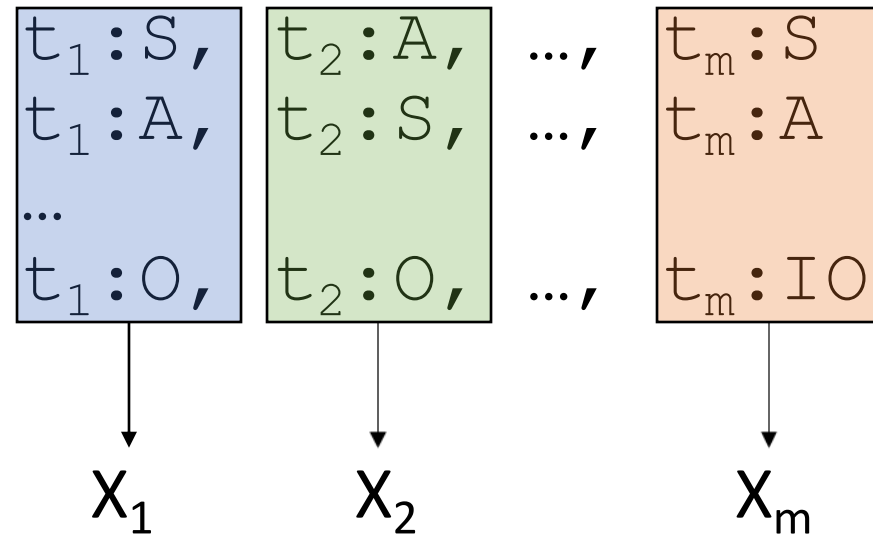
  ```
  s₁ {S, O, A}
  s₂ {S, A, A}
  …
  sₙ {S, O, IO}
  ```

→A given element is more/less often in initial position than it would be expected when outside of distributions

- *m* random init. const. for each *s*

- *m* = 10,000

Simulation

Results & Discussion

- Observed counts differ from expected/simulated counts
  - S, adv.-cl. occur MORE than expected
  - DO & IO occur LESS than expected
  - Adjuncts
    - *europarl*, *wiki*: MORE than expected
    - *speeches*: LESS than expected
- Skew persists even when base frequencies are considered

Experiment

Experiment

Hypothesis

- The learnability of a verb second (V2) grammar is conditioned on the entropy of the preverbal position

- A higher preverbal entropy entails better learning of a V2 grammar

Learning V2

- Extrapolation of the flexibility regarding the preverbal constituent to novel structures

Experiment

Preregistered

Predictions

- Participants learning a skewed V2 language should extrapolate V2 to new structures in fewer instances than participants learning a non-skewed language

- Learners of a skewed V2 language should show diminished discrimination of novel V2 and ungrammatical V3 structures compared to participants learning an unskewed V2 language

Participants

- 314 participant tested, 230 included in analysis (73.2%)
  - Uni.: 74/94
  - O-dom.: 78/118
  - A-dom.: 78/102
- Prolific
  - Self-reported US-nationals
  - Monolingual English speakers
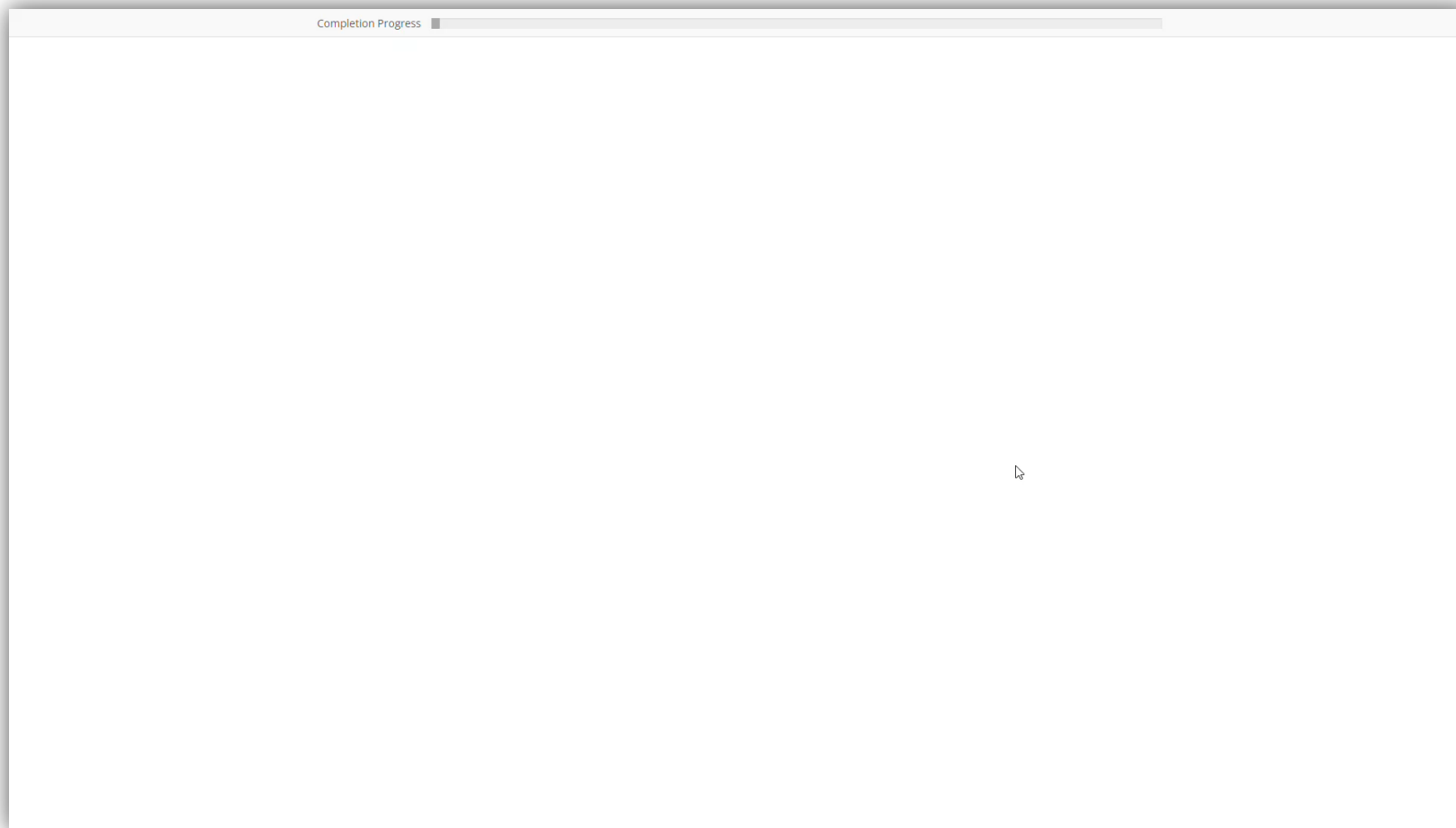 - Raised monolingually

Training phase

Materials

- Semi-artificial language
- 90 V2 sentences constructed from 30 {S, O, V, A} sets
- Uniform condition: 33.3%-33.3%-33.3%
- Skewed conditions: 60%-20%-20%

(2)   a.   The author **revises** eventually a novel in Boston.

        b.   A novel **revises** the author eventually in Boston.

        c.   In Boston **revises** the author eventually a novel.

Training phase

Training phase

Completion Progress

Form a sentence in the new English dialect with the given words
Since 2010 ___ ___ ___ ___

| brews | the witch | the potion | personally |

Reset   Submit

(or press enter)

Testing phase

Production task

- Participants are provided with scrambled English words and must form sentence in artificial language

- Seen constituent types (4 trials):
  - S, O, A (e.g. *Sophia, a carol, on Christmas*)

- Novel constituent types (4 trials each):
  - indirect objects (e.g. *to the prosecutor*)
  - complex adjuncts (e.g. *during the conflict*)

(3)   {the waiter, awkwardly, to the guest, **passes**, the saltshaker}

Testing phase

Judgement task

- Participants see V2 & V3 sentences and need to judge grammaticality of it

- Seen constituent types in initial position (4 trials each):
  - Direct objects
  - Simple adjuncts

- Novel constituent types in initial position (4 trials each):
  - Indirect objects
  - Complex adjuncts

Testing phase

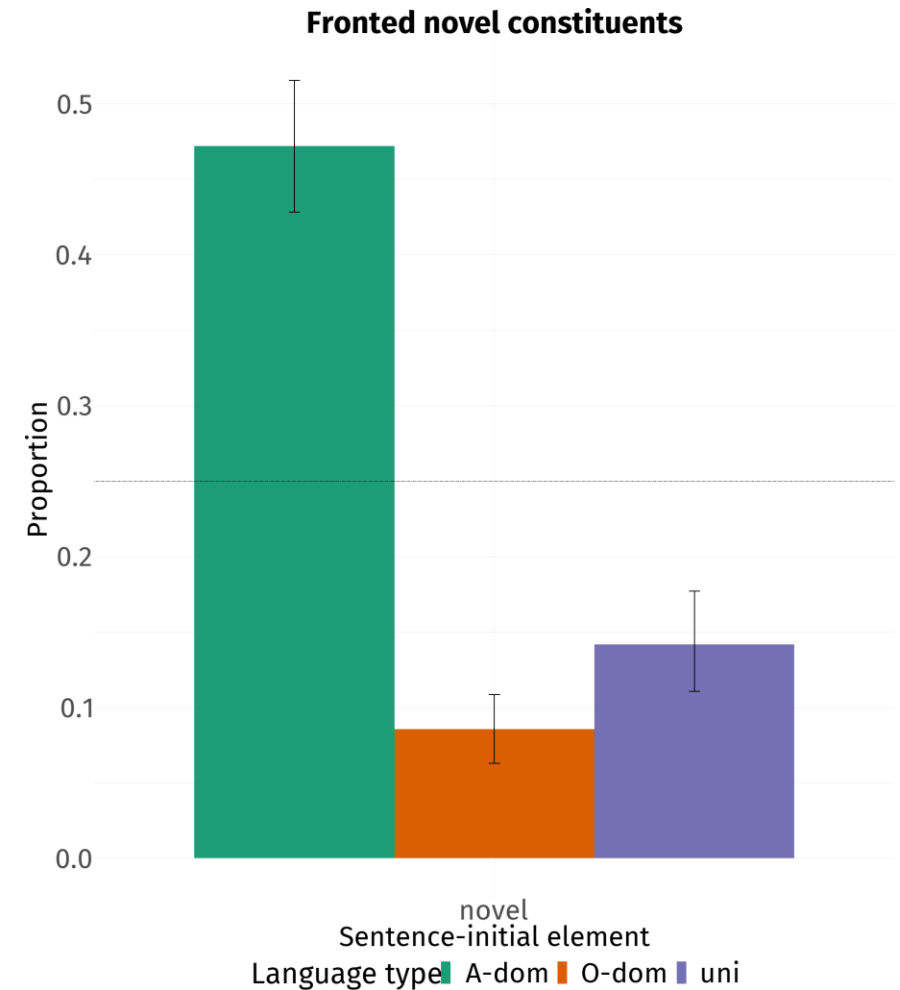(4) To the congregation **shows** the priest silently the candle.

(5) In late April **regrets** the politician openly his misconduct.

(5) To the doctor the patient **describes** precisely the pain.

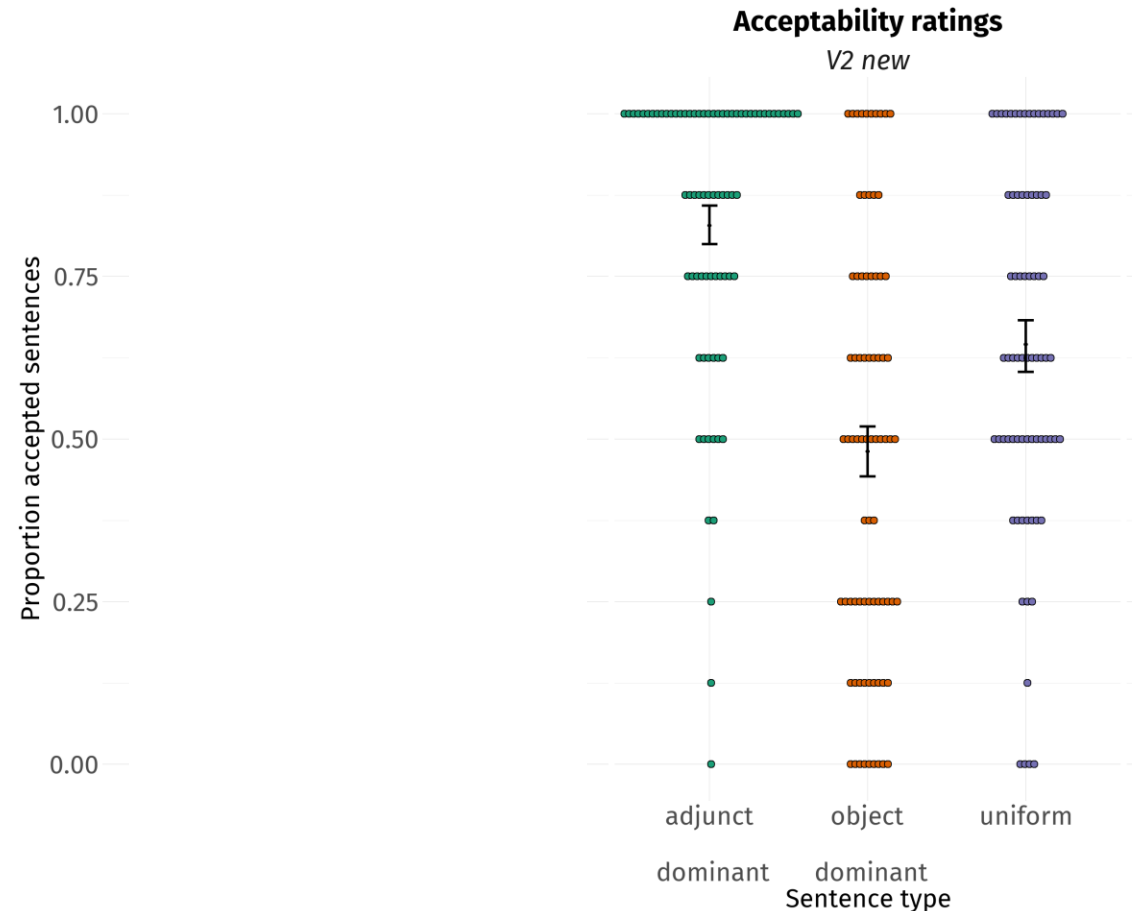(7) At the moment the referee **verifies** briefly the decision.

Results – Production

- **Prediction**: fewer novel constituents fronted in skewed condition
  - Confirmed for O-dom. but not for A-dom.
  - Apparent advantage for learners in A-dominant condition

Fronted novel constituents

Results – Judgement

- **Prediction**: Higher ratings for *V2 new* in uni. condition
  - V2-new: A-dom. > Uni > O-dom.
- **Prediction**: Better discrimination btw. *V2 new* & *V3* in uni. condition
  - Discrimination: A-dom. > Uni = O-dom.

Acceptability ratings
V2 new

Proportion accepted sentences

adjunct dominant   object dominant   uniform

Sentence type

Discussion

- V2 language easily learnable in short period

- Predictions mostly confirmed for O-dom. condition

- Participants in A-dom. condition exceed participants in uniform condition

- Why do participants in A-dom. and O-dom. condition differ?
  - More variability in A-dom. (PPs, AdvPs) than in O-dom. (DPs)?
  - Different types of violation?
  - Learning advantage through adjuncts?

General discussion

- Distributional properties affect learning outcome of V2

- Distributions of preverbal elements in natural V2 languages appear to be disadvantageous for learning

- Results support view that diminished evidence for V2 in input results in loss of V2

- Significant amount of A-initial sentences may be crucial for V2 acquisition

Literature

Bohnacker, Ute & Christina Rosén. 2008. The clause-initial position in L2 German declaratives. Transfer of information structure. *Studies in Second Language Acquisition* 30(4). 511–538.

Fodor, Janet Dean. 1998. Unambiguous triggers. *Linguistic Inquiry* 29(1). 1–36.

Gómez, Rebecca L. 2002. Variability and detection of invariant structure. *Psychological Science* 13(5). 431–436.

Gómez, Rebecca & Jessica Maye. 2005. The developmental trajectory of nonadjacent dependency learning. *Infancy* 7(2). 183–206.

Kok, Daniël de & Sebastian Pütz. 2019. TüBa-D/DP stylebook. Seminar für Sprachwissenschaft, University of Tübingen.

Lightfoot, David. 1999. *The development of language. Acquisition, change, and evolution*. Malden, MA: Blackwell.

Lightfoot, David. 2006. *How new languages emerge*. Cambridge: Cambridge University Press.

Raviv, Limor, Gary Lupyan & Shawn C. Green. 2022. How variability shapes learning and generalization. *Trends in Cognitive Sciences* 26(6). 462–483.

Sitaridou, Ioanna. 2012. A comparative study of word order in Old Romance. *Folia Linguistica* 46(2). 553–604.

Winter, Werner. 1961. Relative Häufigkeit syntaktischer Erscheinungen als Mittel zur Abgrenzung von Stilarten. *Phonetica* 7(4). 193–216.

Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 12(3). 231–250.

Yang, Charles. 2002. *Knowledge and learning in natural language*. New York: Oxford University Press.