A morphosyntactic inductive bias in artificial language learning

Itamar Kastner
itamar@itamarkast.net
Humboldt-Universität zu Berlin

Tal Linzen
tal.linzen@jhu.edu
Johns Hopkins University

Claim

In an **artificial language experiment**: when allomorphy can be learned in morpho-syntactic or morpho-phonological terms, **learners prefer morphosyntax**.

Background: Biases in learning

- Unnatural constraints are not learned as productive patterns (Becker et al. 2011) or are dispreferred to natural ones (Hayes et al. 2009; Hayes and White 2013).
- Order of trials biases the learner (Culbertson et al. 2017; Moreton and Pertsova 2017).

Adapt the paradigm for allomorphy:

⇒ What if the input is compatible with both a phonological and a syntactic generalization? (Embick 2010; Nevins 2011 on triggers of allomorphy)

⇒ Does the learner have a bias for one of them?

Generalizations: Phonology, syntax or semantics?

Gagliardi and Lidz (2014), noun classes in Tsez:

- Learners assign noun class based on phonological cues, not semantic.
  - But not grammatical features.
  - No reference to e.g. animacy in the Tsez grammar.

Culbertson et al. (2017), artificial language:

- Learners rely on salience of the cue (semantically or phonologically).
  - But not grammatical features.
  - What's saliency? Semantic/syntactic?

⇒ What about morphosyntactic features like [PL]?

Example: Modern Hebrew

	hegdil 'enlarged'	
	SG	PL
1	hegdal-ti	hegdal-nu
2M	hegdal-ta	hegdal-tem
2F	hegdal-t	hegdal-tem
3M	hegdil	hegdil-u
3F	hegdil-a	hegdil-u

	gidel 'grew'	
	SG	PL
1	gidal-ti	gidal-nu
2M	gidal-ta	gidal-tem
2F	gidal-t	gidal-tem
3M	gidel	gidel-u
3F	gidel-a	gidel-u

Generalization 1: (Morpho-)Syntax

Lower the vowel when the subject is 1st/2nd person.

Generalization 2: (Morpho-)Phonology

Lower the vowel when the suffix is C-initial. (Kastner 2016)

Methods

Train on ambiguous input, test on held-out cases (Wilson 2006). ● N = 191/224 (Exp 1), N = 237/251 (Exp 2) after removing subjects who failed at least half the Generalization trials. ● Recruitment via Prolific Academic (prolific.ac). ● Experigen (Becker and Levine 2013), presented visually ● Training words also presented auditorially, synthesized using gTTS (Durette 2017). ● Experiment 1: Random order of items in Generalization and Test. ● Experiment 2: Counterbalanced order within Generalization and Test (four blocked lists). ● Total of ten words and four suffixes. ● Results analyzed using mixed effects logistic regression (Baayen et al. 2008; Bates and Maechler 2009).

Design

Training: Three stems and two suffixes.

(2), (5) *-u* red color
(3), (6) *-ti* and vowel lowering plural

Participants were trained on the correct form and received feedback.

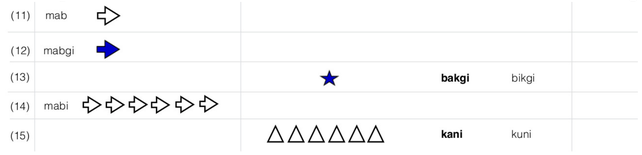
Generalization to new stems: Five new stems, same affixes.

Participants were asked to choose the correct suffixed forms and got feedback.

Test: Two new suffixes.

(11) *-i* plural
(12) *-gi* blue

Participants chose between forms with and without stem lowering.



Predictions

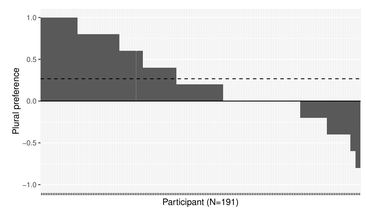
"Syntactic" generalization:	lower vowel before plural *-i* (15)	[PL]
"Phonological" generalization:	lower vowel before blue *-gi* (13)	C-initial

Results

Experiment 1: Conditions randomized within Generalization and Test.

- Main effect of Condition.
- Novel pl lowering, M = 0.546.
- Less blue lowering, M = 0.278.
- $p<0.0001$.

⇒ **Syntactic generalization is learned.**



Results: Testing in blocks

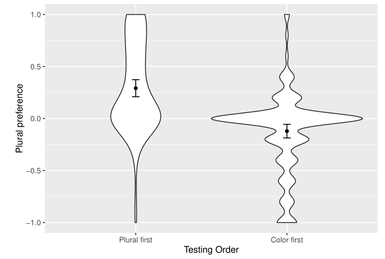
Order of learning matters (Culbertson et al. 2017; Moreton and Pertsova 2017).

⇒ **Experiment 2:** separate blocks within generalization and testing.

Condition still significant: novel pl (M=0.31) over novel blue (M=0.25), p=0.031.

Adding Order of Testing:

- Main effect of **Condition**, p=0.007.
- Main effect of **Order** (pl first vs blue first), p=0.035.
- **Interaction**, $p<0.001$.



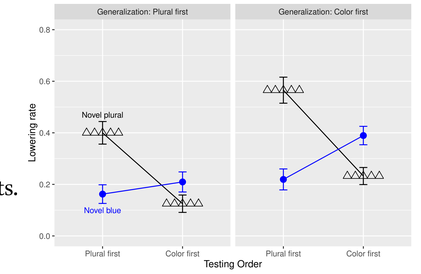
Plot shows participant preference as rate of plural lowerings minus blue lowerings.

⇒ Preference for whichever condition was tested on first.

⇒ Perhaps learners entertain both hypotheses, wait for cue to adopt one.

Results: Testing and generalization orders

- **Order of Generalization** as predictor: $p<0.004$.
- No three-way interaction.
- Error bars plot CIs within subjects.



⇒ "Color first" has no lowering; a lowering block (Plural) was closer to Test.

⇒ Hence perhaps more lowering across the board.

Discussion

- Experiment 1: **Morphosyntactic bias** in learning (contra phonological).
  - How to distinguish syntactic features from a semantic bias?
- Experiment 2: Implicit/explicit learning.
  - Ordering bias: both generalizations are entertained?
- Sensitivity to [PL] in the native language?
  - What counts as "salient"?
- Hebrew: the morphosyntactic analysis is preferable (Kastner 2016).
- Future work: beyond allomorphy.

Supported by DFG AL 554/8-1 Gottfried Wilhelm Leibniz Preis 2014 to Artemis Alexiadou.

Selected references

Becker, Ketrez and Nevins (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish devoicing neutralization. *Language* 87. **Culbertson, Gagliardi and Smith (2017).** Compteition between phonological and semantic cues in noun class learning. *JML* 92. **Gagliardi and Lidz (2014).** Statistical insensitivity in the acquisition of Tsez noun classes. *Language* 90. **Hayes, Zuraw, Spitar and Londe (2009).** Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85. **Hayes and White (2013).** Phonological naturalness and phonotactic learning. *LI* 44. **Kastner (2016).** *Form and meaning in the Hebrew verb*. NYU dissertation, lingbuzz/003028.

NELS 48, University of Iceland, October 2017