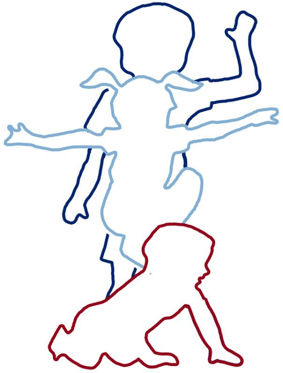
UNIVERSITY of PENNSYLVANIA

Regularizing unpredictable variation: is learning from an unreliable speaker enough?

Annalise Kendrick, Eckerd College
Kathryn Schuler, University of Pennsylvania



Children regularize variation

When children’s input contains unpredictable variation, they regularize:

- Learning a **pidgin or creole** language (e.g. Sankoff 1979)
- Learning a language from **late learners** (Mayberry & Eichen 1991, Newport & Supalla 1980)
- Experiments with natural and artificial language show regularization (e.g. Singleton & Newport 2004, Hudson-Kam & Newport 2005)

Children acquire variation

Children can learn and match the sociolinguistic variation in their input.

- Children have mastered constraints on -t/-d deletion early (Labov, 1989; Roberts 1997; Smith et al. 2009)
- Constraints on word final s-lenition in Spanish (Miller, 2013)

Why do children sometimes regularize variation and other times learn it?

Differences in the input

	Sociolinguistic (learned)	Regularization (changed)
Variation in the community	Shared	Noisy
Conditioning environment	Predictable	Not predictable
Speaker fluency	More likely	Less likely

Speaker fluency and regularization

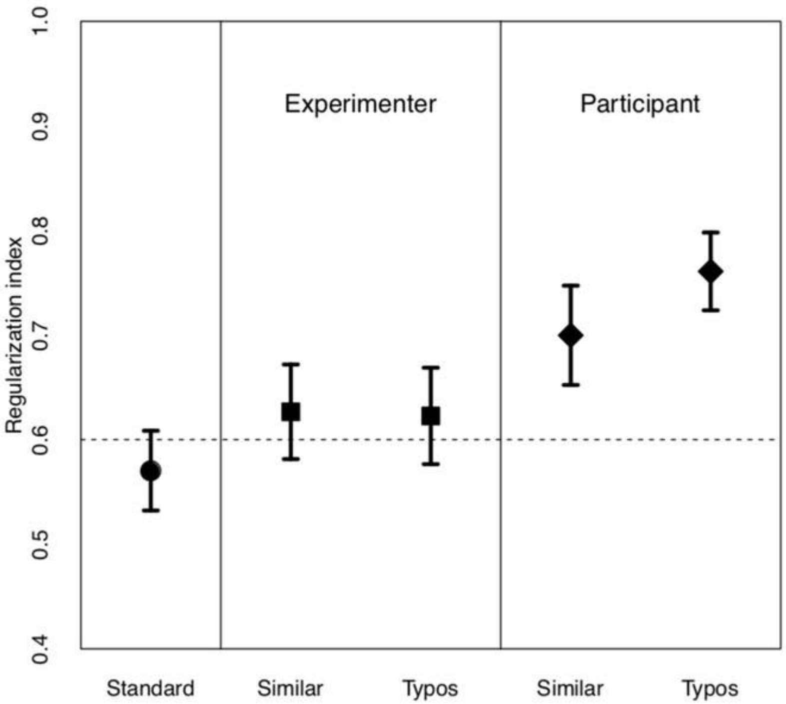
- Children are known to **regularize pidgin and creole languages** (Bickerton, 1984) **and when learning from non-native speakers** of sign language (Singleton & Newport, 2004).
- In both situations, **input is produced by non-native speakers**, who may provide **subtle linguistic cues** to indicate they are unreliable language models .
- Children may be more likely to assume variation is **noise** (e.g. mistakes by the speaker) rather than an important feature of their dialect.

Adults regularize more when they think variation is “noisy”

- Perfors (2016) argued that adults match probabilities because they **assume variation is predictable**.
- **When this assumption is challenged** — i.e. when participants are told that language input might contain mistakes — Perfors (2016) finds that **adults do regularize**
- Argues that these results **explain why children regularize** unpredictable variation in **natural language**

Perfors (2016) results

- **TASK:** Learning Written Object Labels
- **MANIPULATIONS**: Reliability of the Person Providing the Labels & Type of Marker
- **CONCLUSION:** Adults were more likely to regularize when **told the variation might be mistakes** and those mistakes look like **typos**.



Can adults use subtle linguistic cues to determine whether to learn or regularize a rule?

Experiment 1

Testing rule learning

Experiment 1
Testing rule learning

Perfors (2016)
Learning labels

DUTON

Experiment 1
Testing rule learning

Perfors (2016)
Learning labels

DUTON

Our Experiment
Plural marking rule

singular

gentif daffin

plural

gentif daffin ka

Experiment 1

Participants:

- 134 adults on prolific

Conditions:

- Two instructions conditions: **experimenter** and **participant**
- Two plural marker conditions: **distinct** and **typo**

Experiment 1

Instructions | Exposure | Test

EXPERIMENTER

PARTICIPANT

Experiment 1

Instructions | Exposure | Test

EXPERIMENTER

We are studying how people learn new languages. In this experiment, you will see pictures of farm animals paired with sentences describing them in Ackvarian.

PARTICIPANT

Experiment 1

Instructions | Exposure | Test

EXPERIMENTER

We are studying how people learn new languages. In this experiment, you will see pictures of farm animals paired with sentences describing them in Ackvarian.

PARTICIPANT

We are studying how people learn new languages **when they are given sentences from other people**. In this experiment, you will see pictures of farm animals paired with sentences describing them in Ackvarian. **The sentences actually come from a previous participant, who hard to learn Ackvarian themselves. Some participants were given a very limited time to provide descriptions, so there might be errors.**

Experiment 1

Instructions | **Exposure** | Test

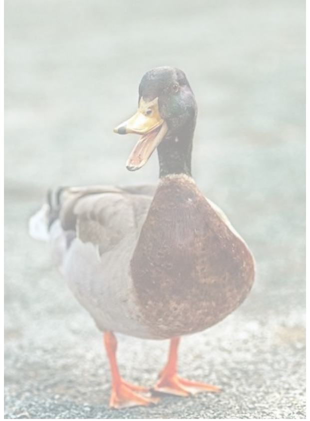
singular

gentif daffin

Experiment 1

Instructions | **Exposure** | Test

singular



gentif daffin

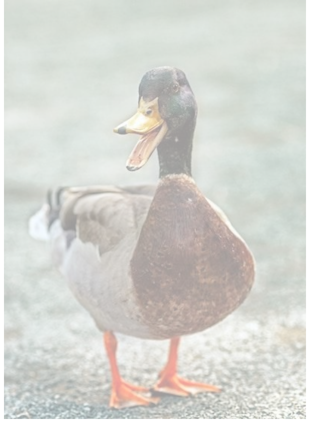
plural

gentif daffin {marker}

Experiment 1

Instructions | **Exposure** | Test

singular



gentif daffin

plural

gentif daffin {marker}

{marker}

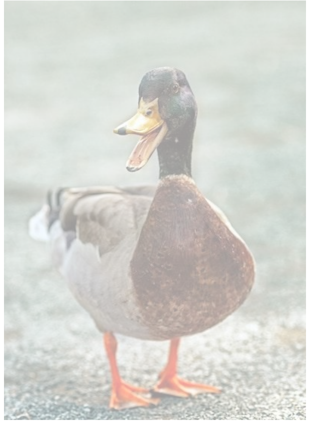
DISTINCT

ka po su ti je

Experiment 1

Instructions | **Exposure** | Test

singular



gentif daffin

plural

gentif daffin {marker}

{marker}

DISTINCT

ka po su ti je

TYPO

ka ja kq a kka

Experiment 1

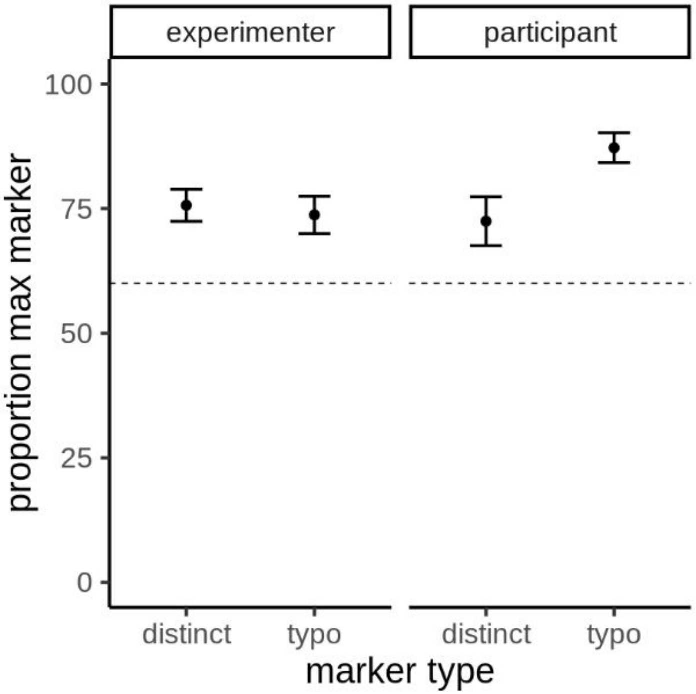
Instructions → Exposure → **Test**

Type a sentence that describes this in Ackvarian.

Experiment 1

Regularization results

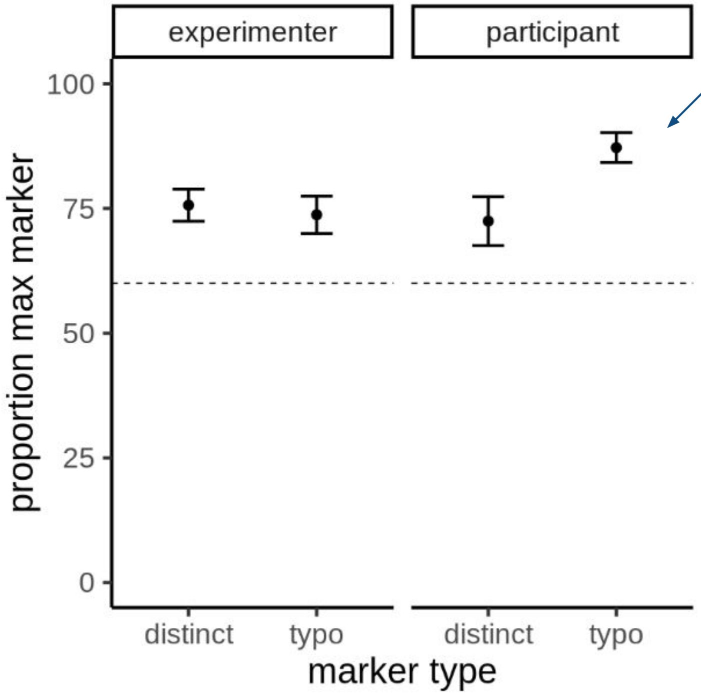
Experiment 1
Regularization results



Experiment 1

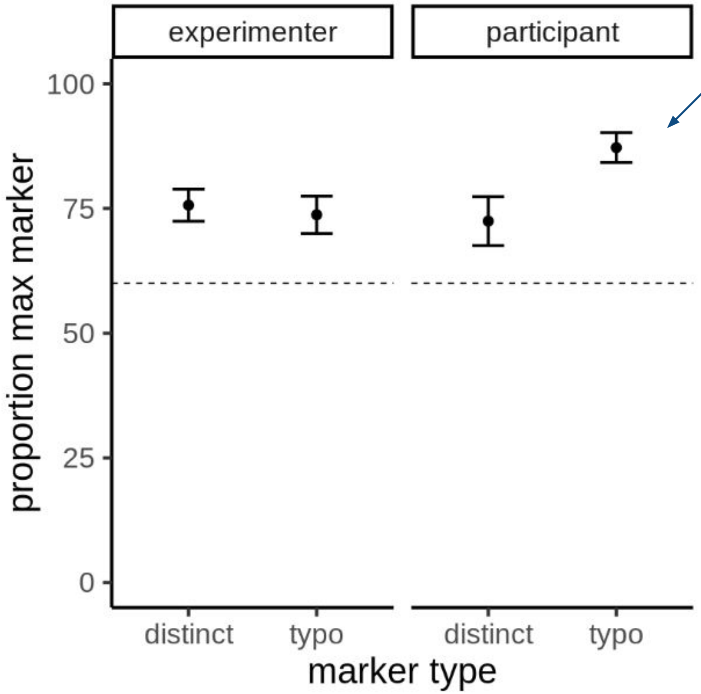
Regularization results

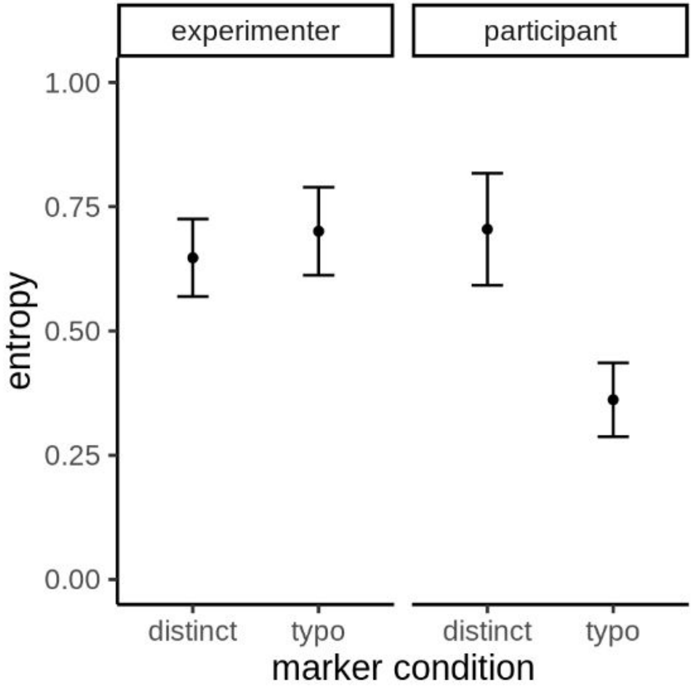
More regularization when variation can be attributed to typos



Experiment 1
Regularization results

More regularization when variation can be attributed to typos





Experiment 1

Like Perfors (2016) greater regularization when

1. Told input might be mistakes (**PARTICIPANT** condition) AND
2. Markers seemed like **TYPOS**

Unlike Perfors (2016) regularization NOT greater when told the input might be mistakes if markers were **DISTINCT**

Will participants regularize more if they know a speaker is nonnative?

Experiment 2

Participants:

- 61 adults on prolific

Conditions:

- Two instructions conditions: **native** and **nonnative**
- Only the **distinct** plural markers

Experiment 2

Instructions | Rate | Exposure | Test

NATIVE

NONNATIVE

The sentences actually come from a speaker of the language, Mary, who provided the descriptions.

Mary has lived in Ackvaria her whole life. She's been speaking Ackvarian since she was a baby. This means that Ackvarian is her native language and she almost never makes mistakes. Most of the time, she says things the right way in Ackvarian.

Experiment 2

Instructions | Rate | Exposure | Test

NATIVE

The sentences actually come from a speaker of the language, Mary, who provided the descriptions.

Mary has lived in Ackvaria her whole life. She's been speaking Ackvarian since she was a baby. This means that Ackvarian is her native language and she almost never makes mistakes. Most of the time, she says things the right way in Ackvarian.

NONNATIVE

The sentences actually come from a speaker of the language, Mary, who provided the descriptions.

Mary **just moved to** Ackvaria. She's **only** been speaking Ackvarian **for one month**. This means that Ackvarian is **not** her native language and she **makes lots of** mistakes. **Sometimes**, she says things the **wrong** way in Ackvarian.

Experiment 2

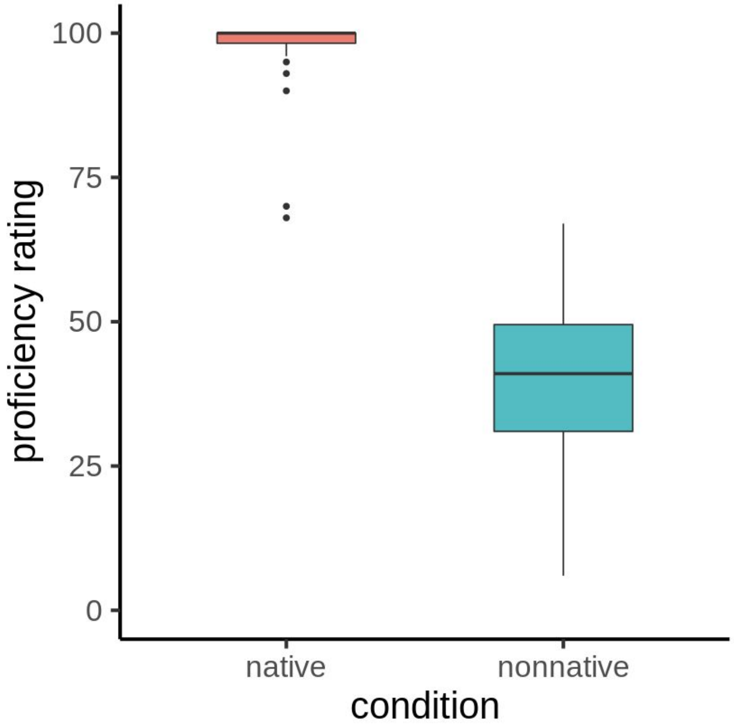
Instructions | **Rate** | Exposure | Test

Not at all | Not well | Well | Very Well

Given her experience, how well do you think Mary speaks Ackvarian?

Continue

Experiment 2
Proficiency rating results



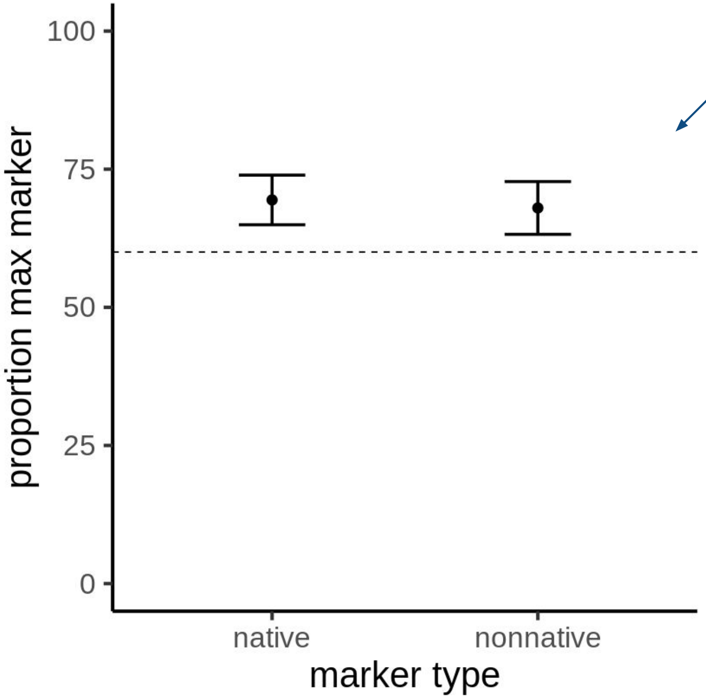
Experiment 2

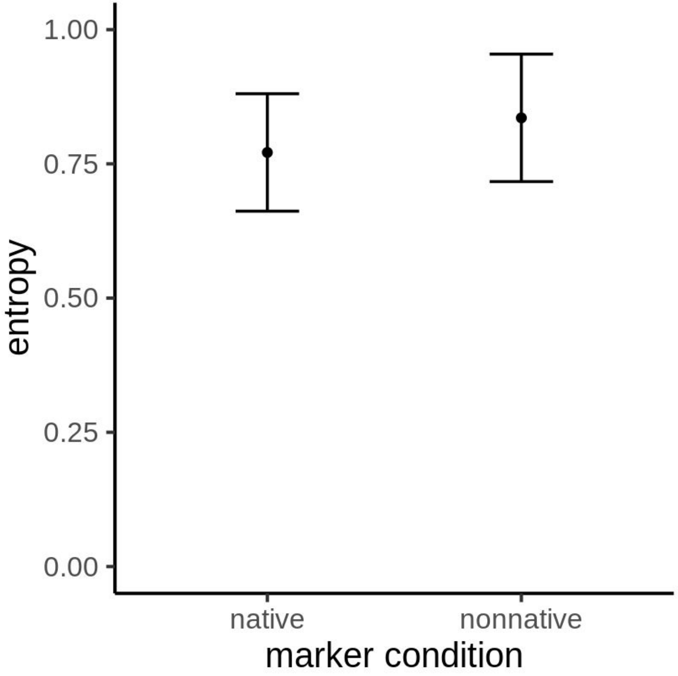
Regularization results

Experiment 2

Regularization results

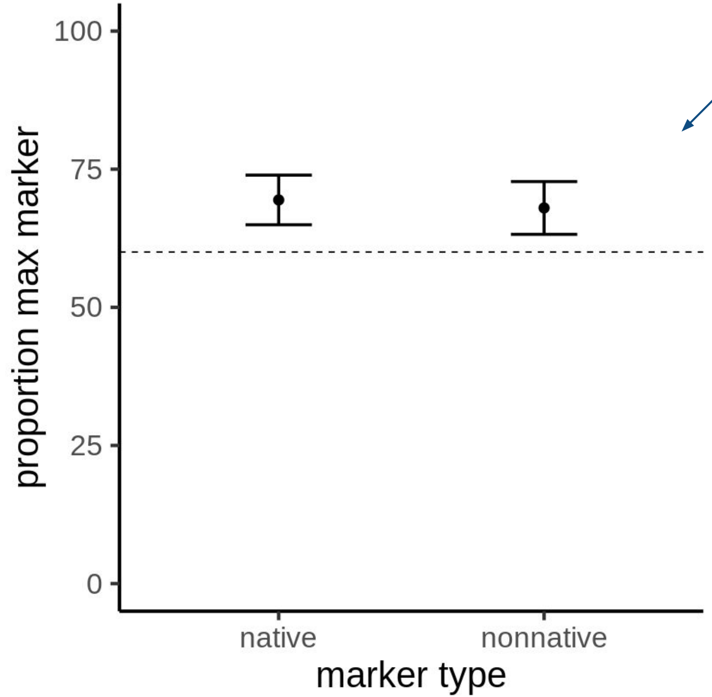
Do not regularize significantly more

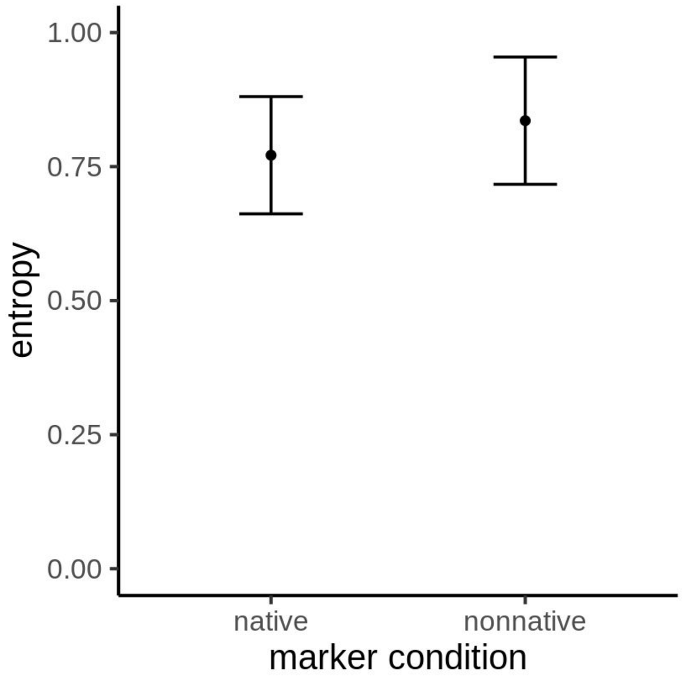




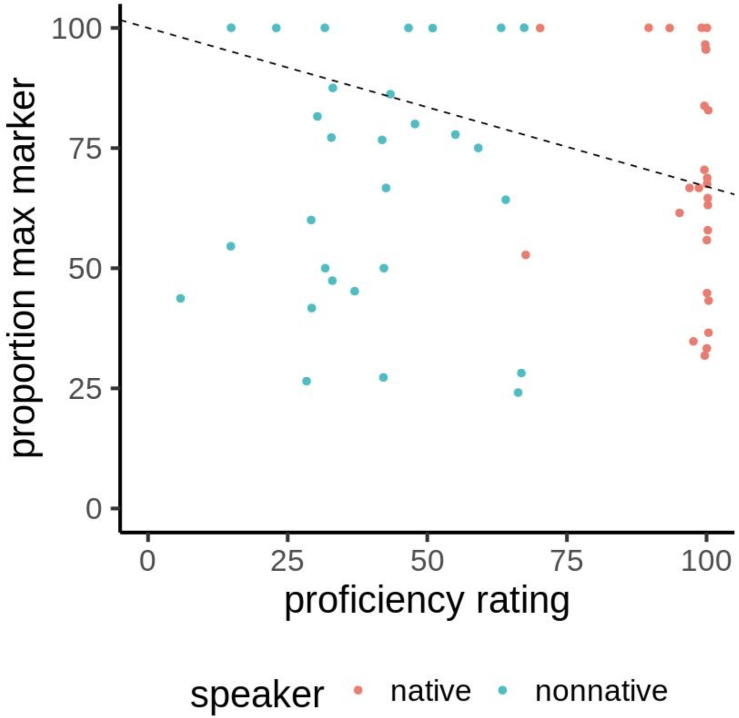
Experiment 2
Regularization results

Do not regularize significantly more





Experiment 2
Regularization by proficiency rating



Conclusion

- In **experiment 1,** adults regularized more when they were told the input might contain mistakes, but only if the markers resembled typos.
- In **experiment 2**, adults did not regularize more when learning from an unreliable language model (nonnative speaker).
- In both experiments, knowing the input contained mistakes was not enough to alter rule learning behavior.

Future directions

Future directions

- **Rule anchors** - do learners need an entrypoint to the underlying rule in order to find the rule among the noise?

Future directions

- **Rule anchors** - do learners need an entrypoint to the underlying rule in order to find the rule among the noise?
- **Red flags** - do learners need a signal *during learning* that a mistake has likely occurred?
  - Disfluencies (..uh…, …um…)
  - Facial expressions

Thank You

Thanks to the Language and Cognition Lab and the Language Evolution Lab at Penn for helpful comments and suggestions.

Questions

Aux slides

Perfors (2016) design

Cover Story

EXPERIMENTER

We are studying how people learn language. You'll see pictures of common objects with labels from an artificial language.

PARTICIPANT

The labels actually come from a previous participant, who had to learn the fake language themselves. Some were given a very limited time to provide the labels so there might be errors.

Exposure

DUTON

200 image-label pairs
10 nouns, 20x each

One main affix (60%)
Four noise affixes (10% each)

TYPO: noise affixes like typos
SIMILAR: not like typos

Test

Enter a label
40 trials, 4 per noun