Naturalness and frequency in implicit phonological learning

Youngah Do
Elizabeth Zsiga
Jonathan Havenhill

Georgetown University

LSA 90
January 7, 2016

Effect of naturalness in learning phonology

- Mixed findings in the literature wrt the role of naturalness in phonological learning
- More natural or unmarked patterns are learned more readily (Wilson, 2006; Carpenter 2010; van de Vijver, 2011)
  - palatalization before [i] $>$ [a]
  - stressed low-vowels $>$ stressed high-vowels
- No naturalness bias (Moreton & Pater, 2012; Moreton & Pertsova 2014)

Frequency effects in learning phonology

- Frequency also matters.
- Exposure to more data results in better learning.
  - When enough data is provided, unnatural alternations can be learned, overcoming any naturalness bias (Peperkamp et al., 2006; Zsiga & Boyer, in press)
- Learners rely on natural classes to generalize alternations, but also track segmental frequency (Kaplan, 2010; Albright & Do, 2013)

This study

- The role of a naturalness bias in learning phonological alternations
- The role of frequency in phonological learning
  - Learning increases with increased exposure to an alternation.
  - How do natural classes and segmental frequency contribute to generalization of alternations?
- Are there differences between implicit vs. explicit learning?
- Test case: post-nasal voicing and devoicing

Natural post-nasal voicing

- Phonetic tendency for voicing to continue into a post-nasal stop (Hayes & Stivers, 1996)
- Phonologized as *NT (Pater 1996; Hayes 1999; Hyman 2001)
- Evidenced in numerous languages
  - Indonesian (Halle and Clements 1983)
  - Chamorro (Topping 1969)
  - Malagasy (Dziwirek 1989)

*NT effects

- Typologically, many languages allow only ND: Sequence of nasal plus voiceless stop (NT) is avoided.
  - Kikuyu Postnasal Voicing (Clements 1985):
  - OshiKwanyama Voicing and Fusion (Steinbergs 1985):
- Some languages (English) allow both NT and ND, but no language will prefer NT to ND

Unnatural post-nasal devoicing?

- A counterexample: Setswana
  - Hyman 2001, Coetzee et al. 2007, 2010, Sole et al. 2010, Gouskova et al. 2011
- Voiced stops devoice in post-nasal position:

bala	'count'	mpala	'count me'
bata	'look for'	mpata	'look for me'
disa	'guard'	ntisa	'guard me'

- While voiceless stops remain unchanged:

pala	'refuse'	mpala	'refuse me'

Unnatural post-nasal devoicing?

Caveats:

- Note that Setswana has a skewed voicing distribution:
  - /b/ and /p/ contrast before all vowels
  - d occurs only as an allophone of /l/ before high vowels
    xo-bol-a ‘to rot’ bod-ile ‘rotted’
  - g does not occur
  - The alternation is therefore most frequently evidenced as b∼p

Post-nasal devoicing in Sebirwa

- Sebirwa borrowed devoicing from Setswana (Chebanne 2000).
- Doubly unnatural: devoicing of labials only (Zsiga & Boyer in press)
- Frequency may explain why only labial devoicing was borrowed

Artificial language learning experiment

- We tested whether:
  - Post-nasal voicing is learned better than devoicing. (Evidence of naturalness bias.)
  - With sufficient evidence of unnatural patterns, post-nasal devoicing can be learned as well. (Frequency can overcome naturalness bias.)
  - Voicing alternations are generalized on the basis of segmental frequency and/or natural classes

Artificial language learning experiment

- Task: learn how to create plural forms in a Martian language
- Singular $\sim$ plural pairs presented auditorily with pictures from van de Vijver et al. (2011)
- Training: 56 items presented once per participant, in random order
- Prefixal plural marker in which vowel height harmonizes with first stem vowel
  - [ubi] $\sim$ [inubi], [rasu] $\sim$ [arrasu]
- Obstruent-initial roots exhibit voicing alternations
  - [**p**abli] $\sim$ [am**b**abli], [**b**abli] $\sim$ [am**p**abli]

Languages varying naturalness and frequency of alternators

- 10 languages: 5 with postnasal voicing, 5 with postnasal devoicing

	Alternators			Nonalternators			
lg	labial	coronal	filler	labial	coronal	filler	total
1	8	0	20	8	0	20	56
2	8	4	16	8	4	16	56
3	8	8	12	8	8	12	56
4	8	12	8	8	12	8	56
5	8	16	4	8	16	4	56

- Alternating fillers: /VN+labi/ $\rightarrow$ [allabi] (liquid-initial)
- Non-alternating fillers: /VN+ipu/ $\rightarrow$ [inipu] (vowel-initial)

Implicit vs. Explicit learning

- Experiment 1: Participants instructed to figure out how plurals are formed in Martian
  - Focus is on both vowel harmony and consonant alternations
- Experiment 2: Participants instructed to figure out which prefix vowel is used with which stem
  - Vowel harmony task is made explicit, in order to reduce attention paid to consonant alternations
  - Implicit learning of consonant alternations

Test

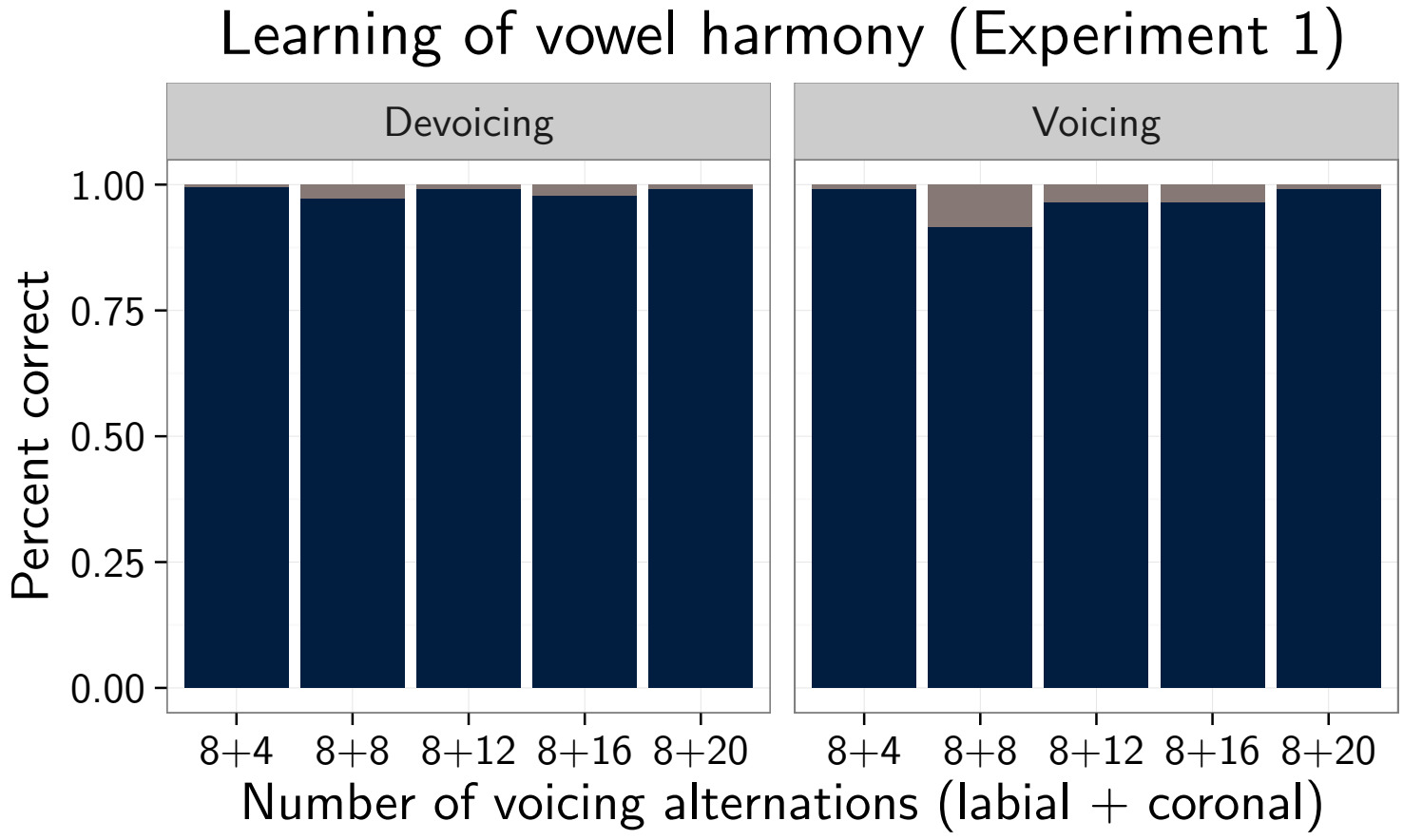
- Participants were asked to choose the correct plural form for 28 unseen words
- Test items included labial-, coronal-, and velar-initial stems
  - Velar-initial stems were not seen during training
- Four choices given for [kugri] in Natural languages:
  - Correct harmony, correct voicing: [ingugri]
  - Correct harmony, incorrect voicing: [inkugri]
  - Correct harmony, incorrect segment: [illugri]
  - Incorrect harmony, incorrect segment: [allugri]

Participants

- Experiment 1: 109 adult native speakers of English
- Experiment 2: 208 adult native speakers of English
- Criteria for exclusion:
  - Participants who did not complete entire task (training + test)
  - Participants who completed the task in under 7 minutes
- Remaining participants:
  - 105 in Experiment 1
  - 179 in Experiment 2

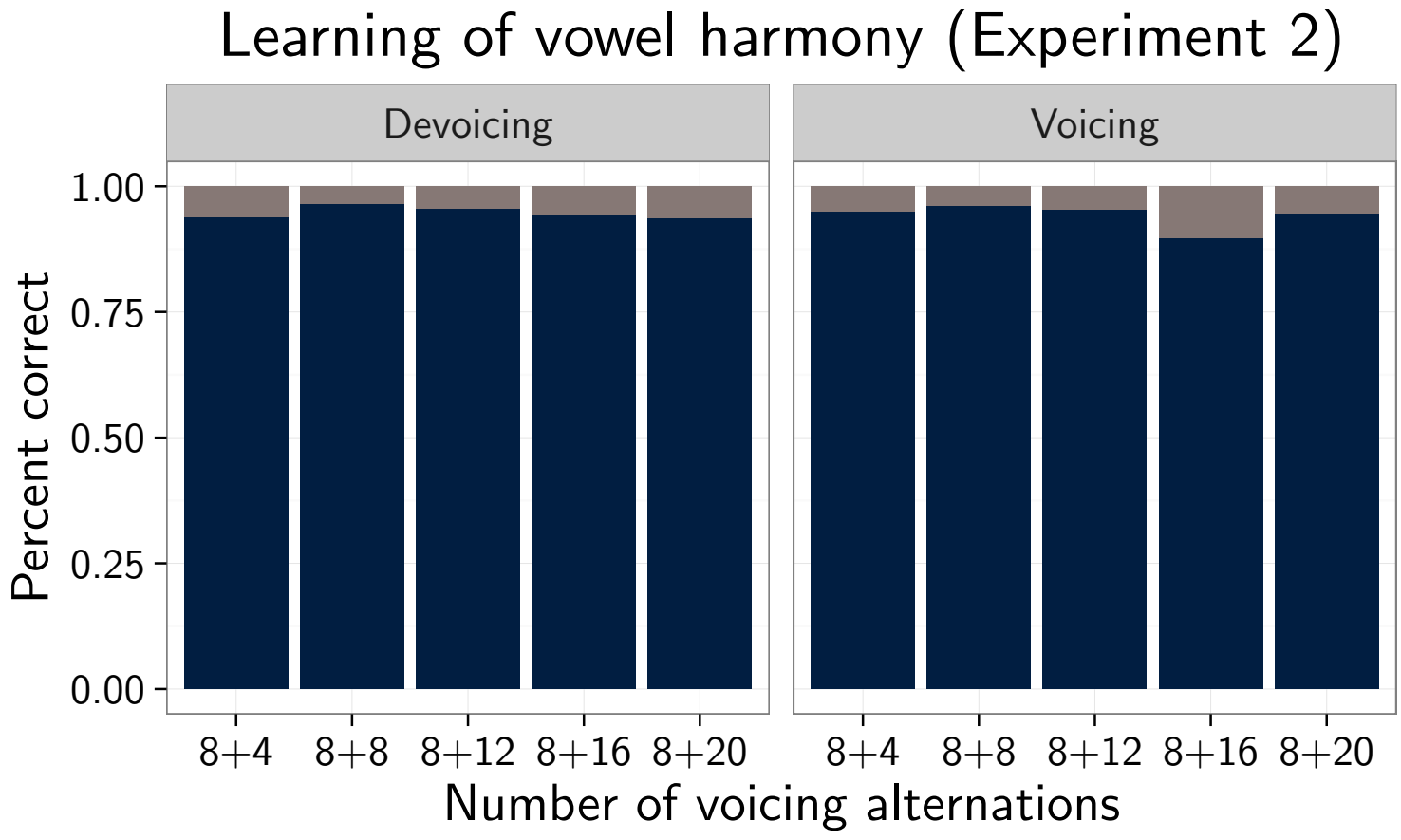
Results: Vowel Harmony

- Vowel harmony is learned equally well across natural and unnatural conditions and across frequencies:



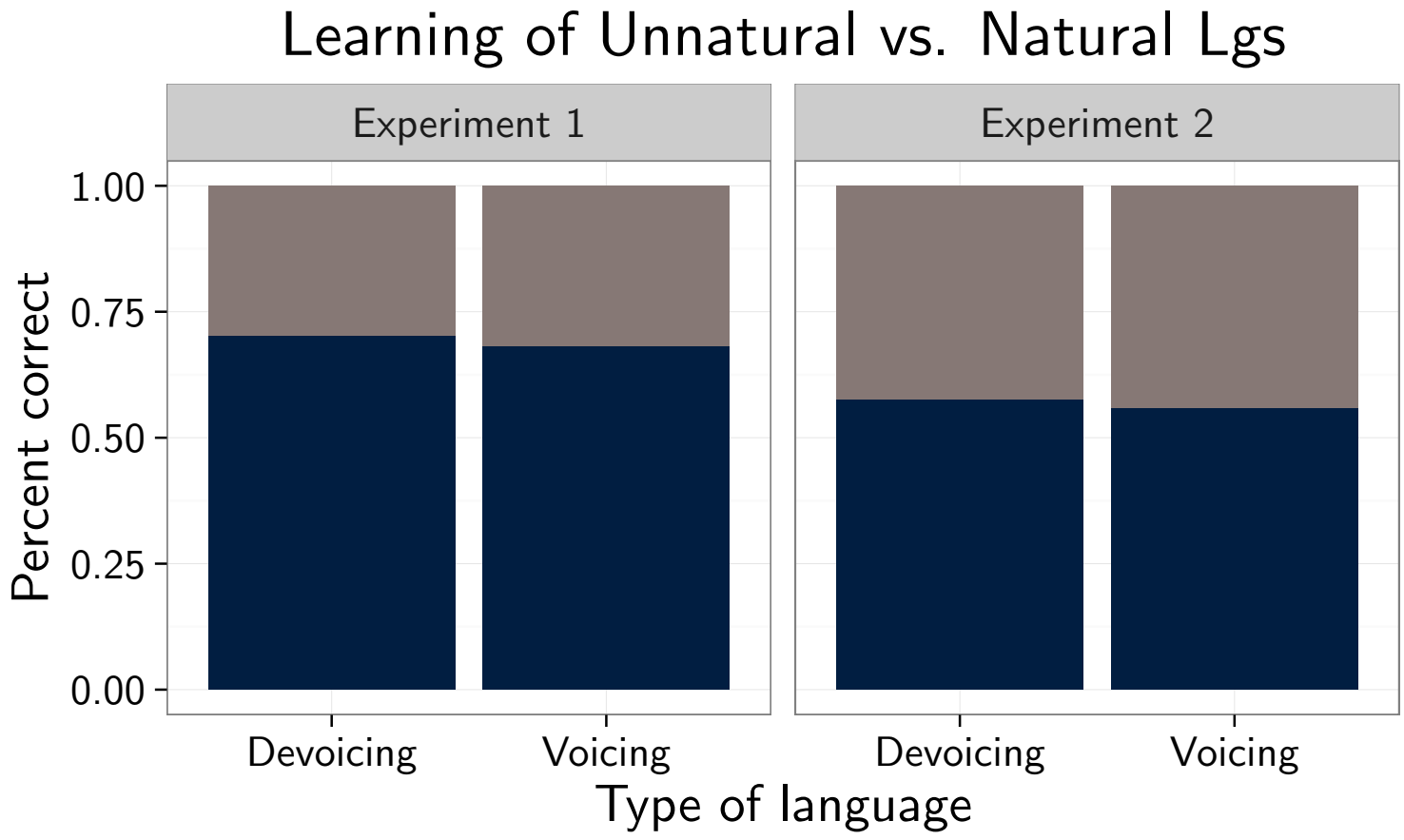
Results: Vowel Harmony

- Vowel harmony is learned equally well across natural and unnatural conditions and across frequencies:



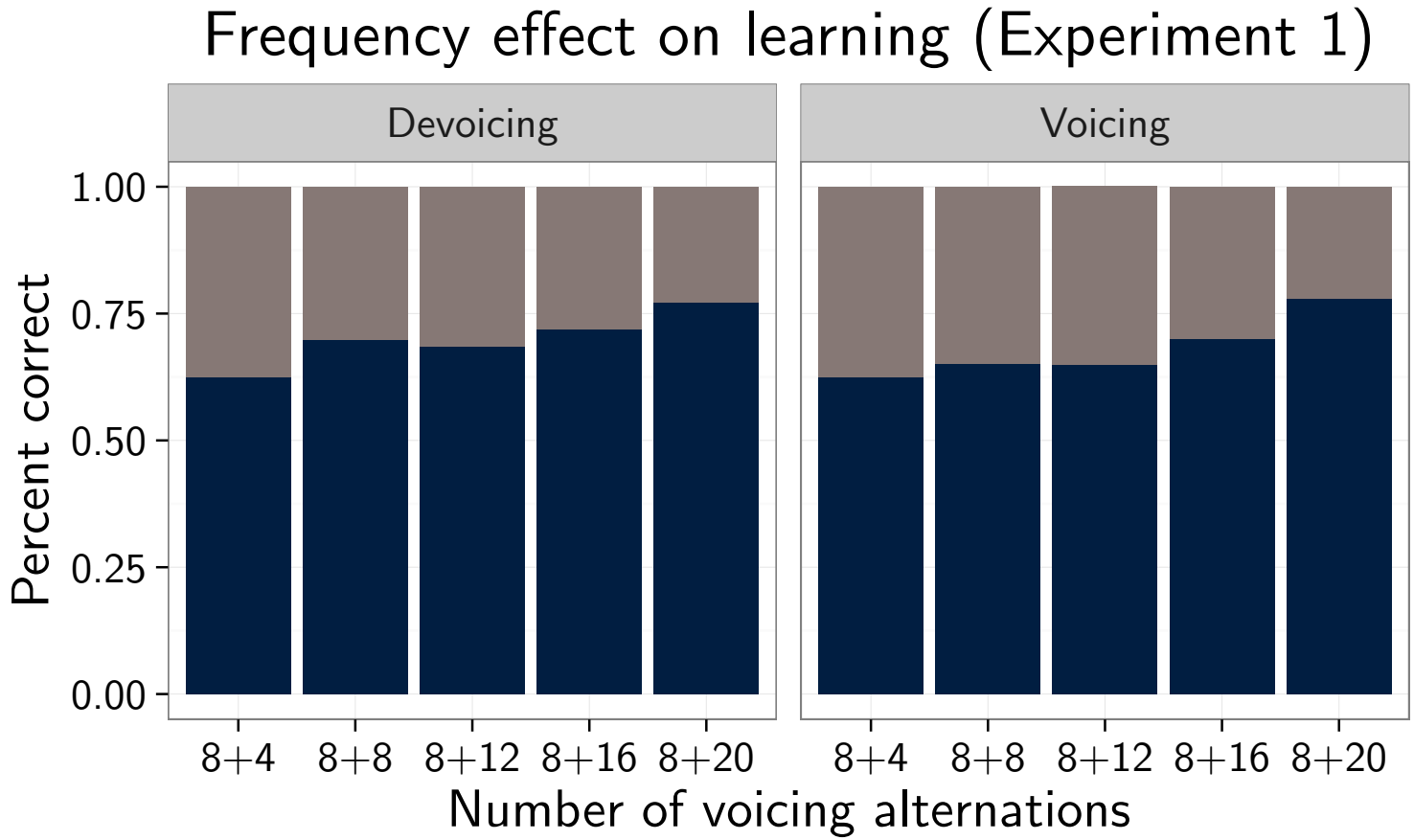
Results: Naturalness

- No significant main effect of naturalness in Experiment 1 ($p = 0.221$) or Experiment 2 ($p = 0.264$).



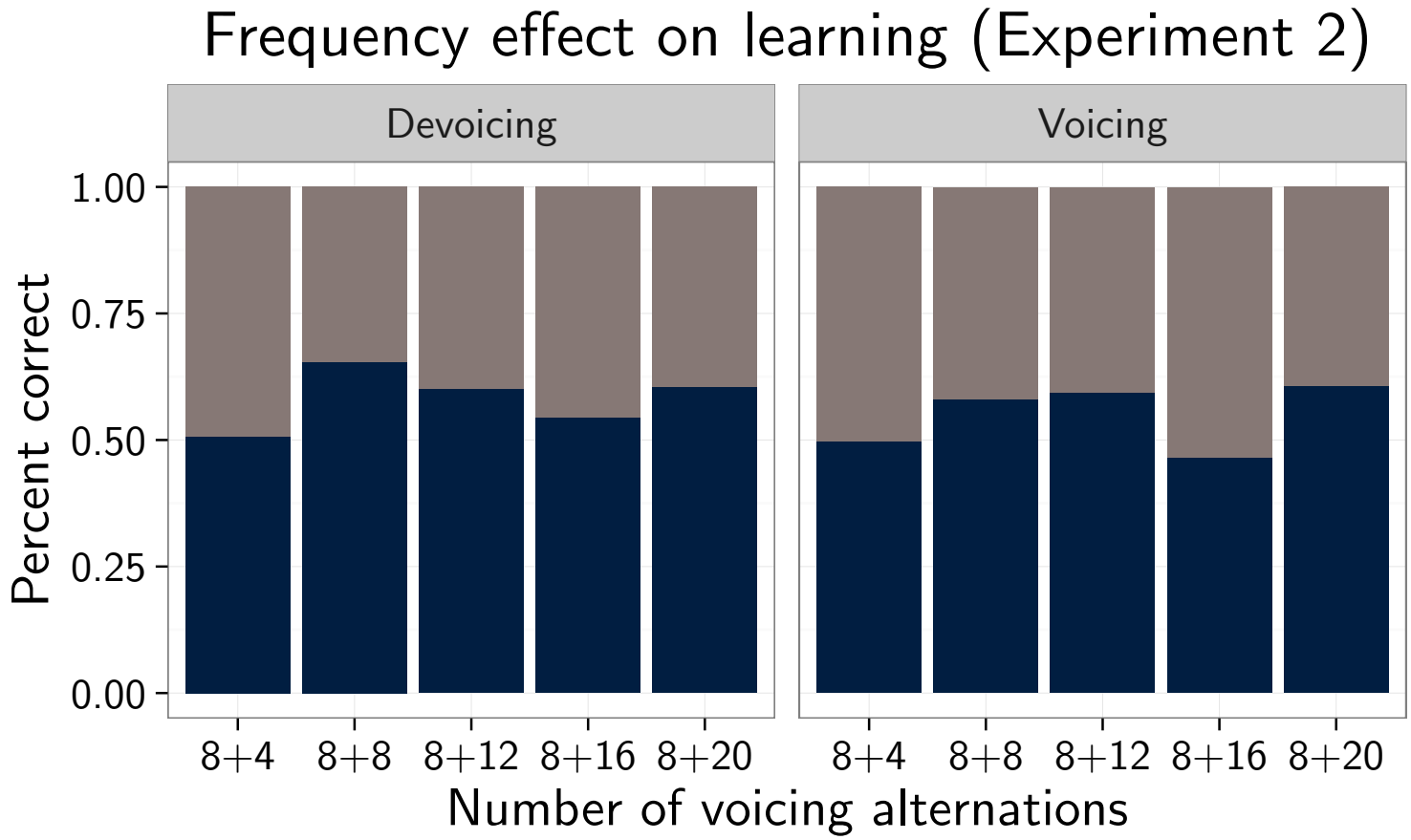
Results: Frequency

- In both experiments, learning improves with increased evidence for voicing and devoicing alternations (Experiment 1: $\beta = 0.008$, $t = 5$, $p < 0.001$, Experiment 2: $\beta = 0.003$, $t = 2.361$, $p = 0.018$)



Results: Frequency

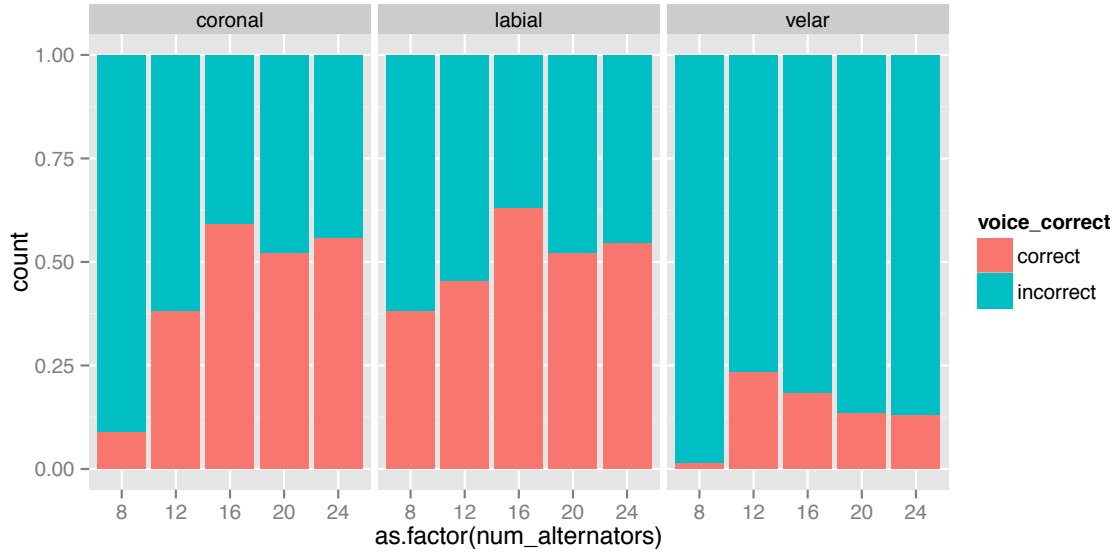
- The frequency effect is weaker in Experiment 2 (implicit learning):



- Why does frequency not play a larger role in implicit learning?

Results: Experiment 2, Natural languages

- For labials, moderate frequency effect
- For coronals, no frequency effect in languages 2-5
- For velars, accuracy is below chance in all conditions

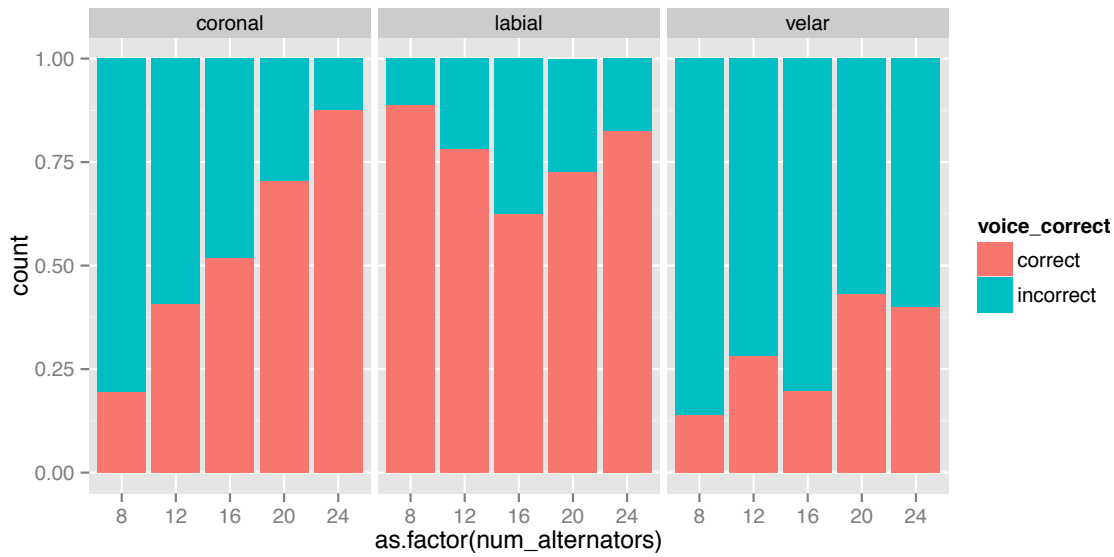


Results: Experiment 2, Natural languages

- Initial state:
  - OO-Ident[±voi] >> *NT >> *nt, *np, *nk
- Language 1 (voicing alternation only among labials):
  - With evidence for single-segment alternation, *np is promoted
  - *NT is promoted, but only minimally
  - *np >> OO-Ident[±voi] >> *NT >> *nt, *nk
- Languages 2-5 (alternation among coronals and labials):
  - With increasing evidence for coronal alternation, *nt is promoted
  - With evidence for alternation among multiple segments, general Markedness constraint promoted
  - *np, *nt >> *NT, OO-Ident[±voi] >> *nk

Results: Experiment 1, Natural languages

- For labials, high performance when participants are trained only on labial alternations.
- A drop in performance is observed as coronal alternations are introduced
- Performance increases for all segments when sufficient evidence is provided: generalization based on the natural class of voiceless stops.
- Generalization to velars is nearly signficant ($p = 0.06$)



Results: Experiment 1, Natural languages

- Initial state:
  - OO-Ident[±voi] >> *NT >> *nt, *np, *nk
- Language 1 (voicing alternation only among labials):
  - More learning of labial alternation than in Exp. 2
  - When explicit focus is given to consonant alternation, *np is promoted more than in implicit learning.
  - *np >> OO-Ident[±voi] >> *NT >> *nt, *nk

Results: Experiment 1, Natural languages

- Languages 2-3 (alternation among coronals and labials):
  - In explicit learning, *nt is promoted proportionally to frequency of coronal alternation: segment-based learning
  - No observed effect of general markedness constraint
  - Without generalization, task becomes more complex, performance on labials decreases.
  - *np >> *nt >> OO-Ident[±voi] >> *NT >> *nk
- Languages 4-5 (increased evidence for alternation)
  - General markedness constraint promoted with plentiful evidence from multiple segments
  - *NT, *np, *nt >> OO-Ident[±voi] >> *nk

Discussion

- Naturalness effect
  - Naturalness bias does not play a role in learning voicing alternation; structural simplicity may be more important (Moreton & Pater, 2012; Moreton & Pertsova 2014)
  - Postnasal voicing might not be clearly more natural than postnasal voicing
  - L1 effect: Both ND and NT are phonotactically legal in English

Discussion

- Frequency effect
  - General markedness constraint is reranked only with evidence for alternation among multiple segments
- Implicit vs. Explicit learning
  - No active role of general markedness constraint in explicit learning—learners rely on segment-specific constraints
  - Mystery: generalization to unseen segment only in explicit learning

Thank you!

Contact:

Youngah: yd79@georgetown.edu
Lisa: zsigae@georgetown.edu
Jon: jeh241@georgetown.edu