Department of Linguistics

Learnability of a counting-involving alternation: Japanese and an artificial language

Hailang Jiang
hailang.jiang.22@ucl.ac.uk
hailangjiang.github.io
University College London

2024 LSA Annual Meeting

Summary

- Through a wug test with Japanese speakers, this study demonstrates that a lexical pattern in Japanese which ‘involves counting’ is not productive;
- Through an artificial language learning (ALL) test with English speakers, this study shows results consistent with the claim that phonology cannot count (past 2);

- This talk provides experimental evidence against counting in phonology.

Background

Japanese /g/ nasalization (VVN)

- In Tokyo (Yamanote) Japanese,

/g/	→ [g] word-initially	/geki/ [geki] ‘drama’
	→ [ŋ] word-internally	/kagami/ [kaŋami] ‘mirror’

(Ito & Mester 1996)

Japanese /g/ nasalization in compounds

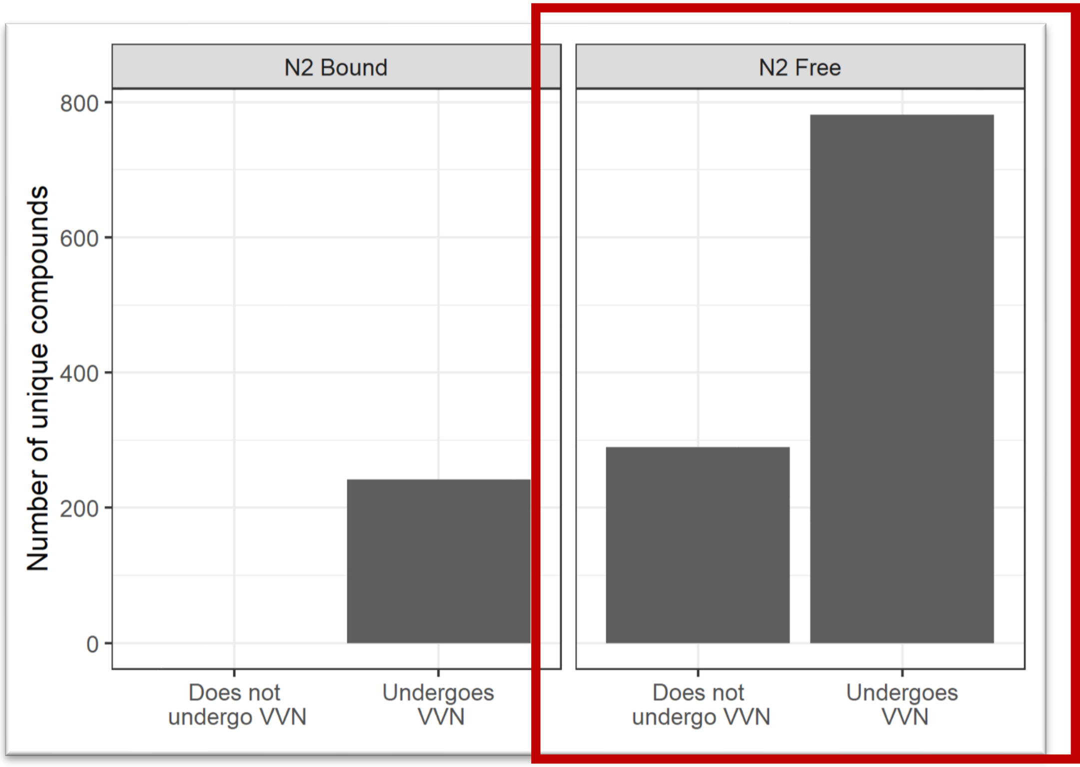
- For two-member compounds, not all /X-gY/ surface with /ŋ/.

  /doku + ga/ → [doku-ŋa]~[doku-ga]
  ‘poison moth’ free variation [g~ŋ]

  /noo + geka/ → [noo-geka] *[noo-ŋeka]
  ‘brain surgery’ one legal form [g]

  (Ito & Mester 1996; Breiss et al. 2022)

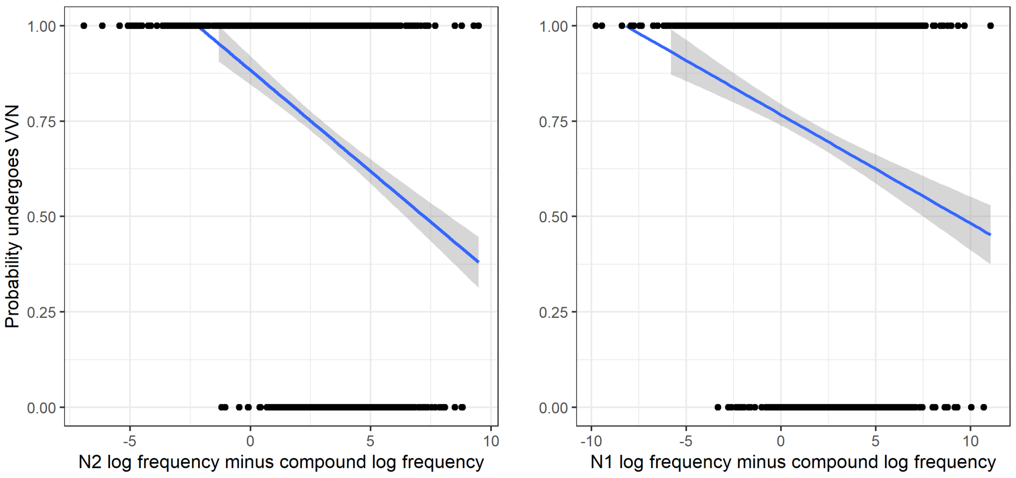
- [ŋ] is always acceptable for words with a bound second member.



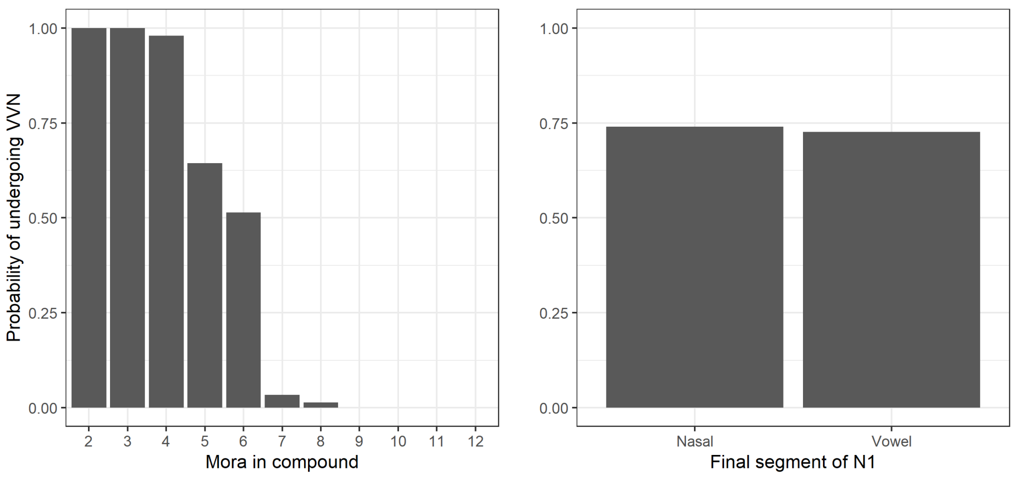
from Breiss et al. (2022)

A corpus study on /g/ nasalization

- Breiss et al. (2022) report three significant factors:



i. Relative frequency of both members (member 2 & member 1)



iii. **Mora length** of the entire compound

ii. Nasality of the preceding segment

all figures from Breiss et al. (2022)

The (un)naturalness of the three factors

i. Relative frequency of both members

→ Paradigm uniformity (Breiss et al. 2021)

ii. Nasality of the preceding segment

→ Progressive local assimilation in [nasal]

iii. Mora length of the entire compound

→ ?

(un)naturalness in the sense of Peperkamp, Skoruppa & Dupoux (2006)

second-order phonotactics (Warker & Dell 2006)

The mystery of the ‘counting pattern’

- Here, ‘counting’ is used in the sense that a exact number of a phonological units is stated in the context of a phonological rule.

  e.g., a hypothetical rule: [g] $\rightarrow$ [ŋ] / $|X_0 \_ Y_0| > 5\mu$

- Binary structure in phonology: phonology does not count to a number larger than 2 (e.g., McCarthy & Prince 1999)
- This may stem from difficulties in accessing precise information regarding the number of specific phonological units

- Counting units: syllables, moras … but never segments ☺

Counting in phonology

- Paster (2019):
  - · Phonological generalization counting to more than 2 is ***almost*** unattested
  - · There are some patterns that can only be ***analyzed*** as counting to more than 2 (e.g. grammatical tone assignment in Kuria)
  - · Even so, no pattern counts past ***4***; no similar patterns to Japanese /g/ VVN (involving counting to 5, 6, 7…) is attested
  - · Counting patterns never involve segmental features (i.e., they only condition stress & tone but never [nasal])

Experiment 1: a wug test

A wug test
design

- Wug test (Berko 1958): nonce words, to examine if speakers can productively extend attested patterns to non-existent words
- Choose a more natural form between [X-gY] and [X-ŋY]

- Participants: 30 *Tokyo* Japanese speakers from Prolific, aged between 18-65.
    - Self-report that they know the [ŋ] variant
    - 18 of them were eligible (passed the attention check, the ABX test)

A wug test (cont.)
stimuli

- 45 trials, 4 forms per trial
- Two separate members [X] & [gY] and two potential compounds [XgY] [XŋY]

  e.g., [temi] [gemo] [temigemo] [temiŋemo]
- All in Japanese orthography (hiragana), which does not distinguish [g] and [ŋ]
- Created by manipulating *two* factors: nasality of preceding seg & mora length (2-10)

e.g.,

preceding seg \ mora length	5 moras	8 moras
V (a, i, u, e, o)	dotsu'-guko'se	kasaka'so-gosoki'shi
N	no'N-ga'mehi	pehe'kiN-goro'doki

- Frequency ≡ 1, thus relative frequency was controlled
- All moras: CV, /N/ or /Q/ (/Q/ cannot end a word due to phonotactics)

A wug test (cont.)
procedure

- Consent form / audio check / instruction → 3 practice trials with real word → test

This is temi. That was my temi. The temi I saw yesterday is good. [temi] This is gemo. That was my gemo. The gemo I saw yesterday is good. [gemo]	This is temig(ŋ)emo. That was my temig(ŋ)emo. The temig(ŋ)emo I saw yesterday is good. [no audio played here] Rate the relative naturalness between two potential forms of each nonce compound word in audios. [temigemo] 1 2 3 4 5 6 7 [temiŋemo] [Next]

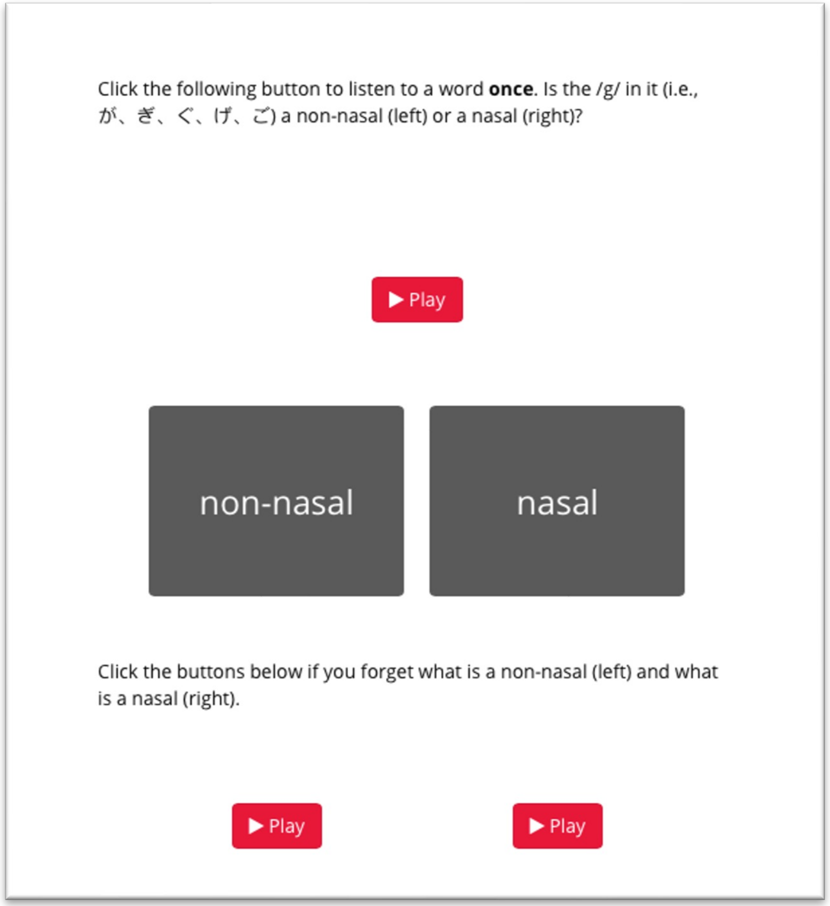
- An attention check mixed at random with trials

A wug test (cont.)
An ABX test of distinguishing [g] & [ŋ]

- 10 trials
- Audio selected from compounds in the test trials
- Could only be played **once**

- ‘Is the /g/ (i.e., が、ぎ、ぐ、げ、ご) in the audio a non-nasal or a nasal?’

- Accuracy rate < 8/10 : excluded



Results of Experiment 1

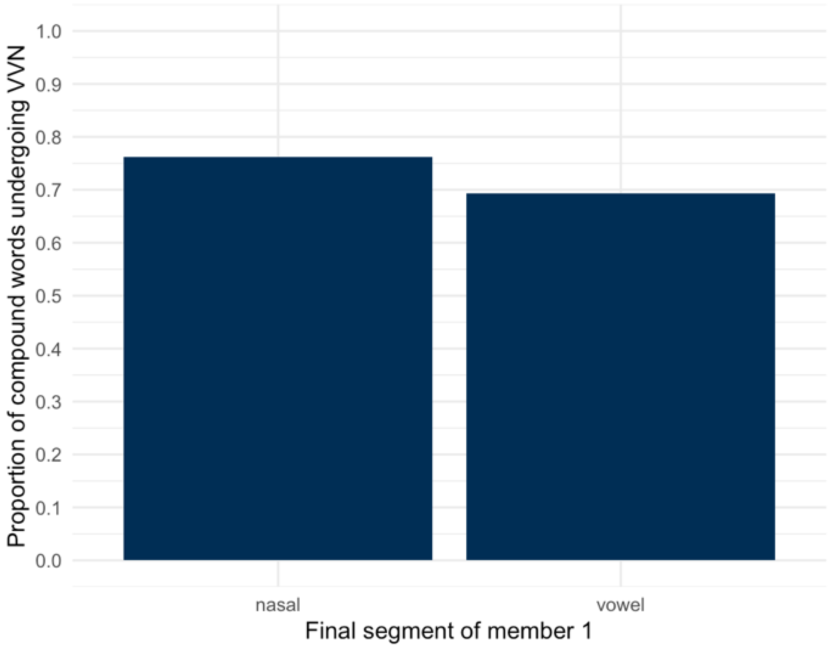
Results
Turning the 7-scale bar to a binary variable

[temigemo] 1 | 2 3 4 5 6 7 [temiŋemo]

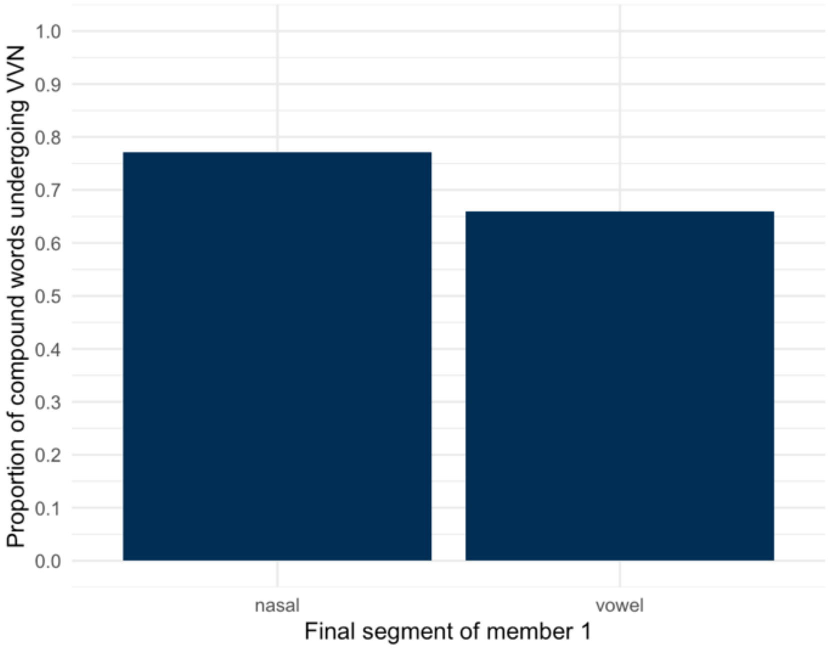
Not accept [ŋ] | Accept [ŋ]

- Transforming a variable with 7 ordinal scales into a binary variable (consistent with the corpus data)

Results
nasality of the preceding segment

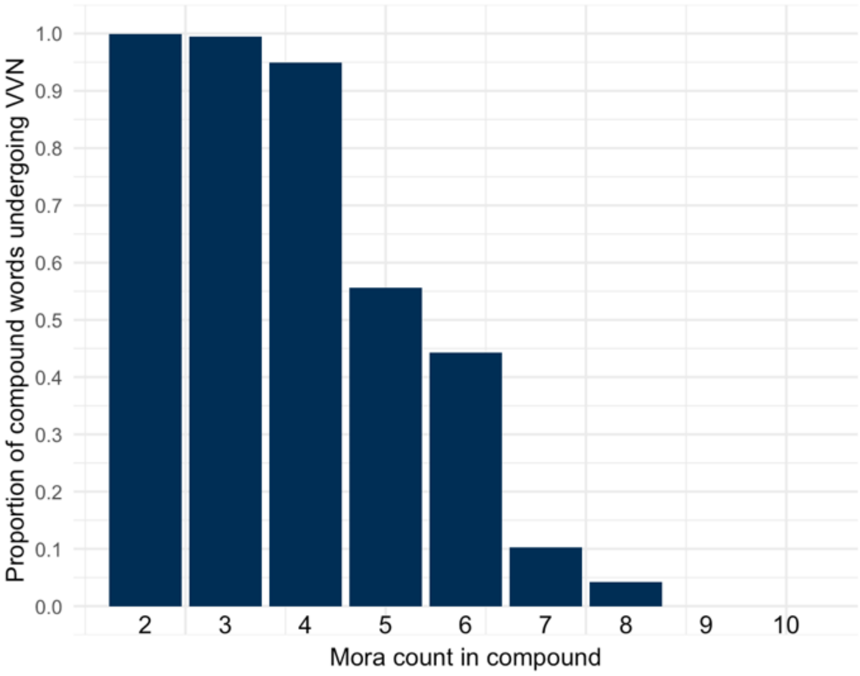


Trend in real lexicon

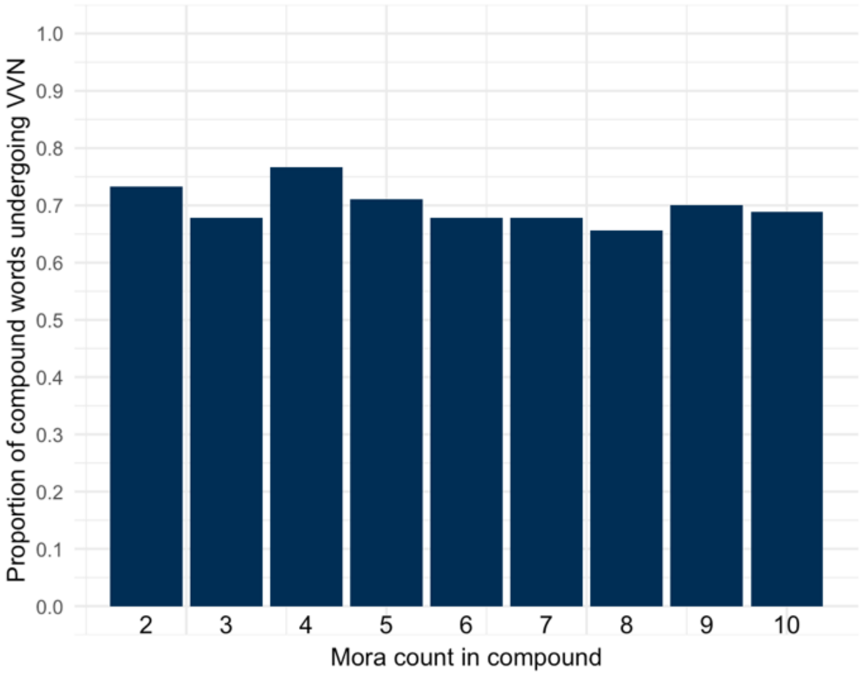


Trend in nonce words

Results
mora length



Trend in real lexicon



Trend in nonce words

- Mean value: 0.705 (real); 0.699 (nonce)
  Frequency matching (Hayes et al. 2009)

Results
Statistics

- Mixed-effect logistic model (with *max random effect*)
- Initial model: glmer(Nasalized_Response ~ nas * length + *(1 + nas + length | subject)* + *(1 | word)*, data = data, family = binomial)

- Results of fixed effects in the final model:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.4210	0.7110	1.999	0.0457 *
nasality	1.1354	0.4575	2.482	$\boxed{0.0131 *}$

- The inclusion of *length* does not significantly improve the model, according to ANOVA

Experiment 2: an ALL test

An ALL test
design

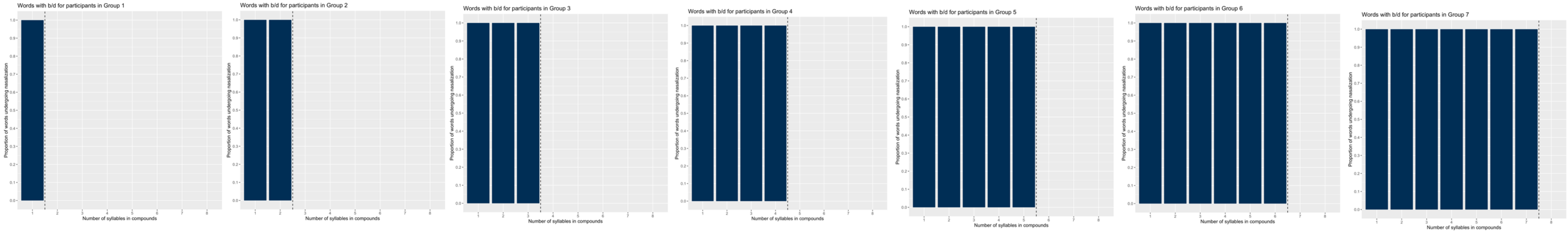
- A structure akin to Japanese VVN is designed, with a more in-depth investigation into the ‘counting’ factor ***length***
- 7 groups (10 participants each), each learning a rule with a *unique length threshold*
- The rule: /b/ (or /d/) → /m/ (or /n/) in a compound iff the compound’s length falls below the given length threshold of the group
- Same task: choose the correct form between [X-bY] and [X-mY]

- Participants: 70 American English speakers from Prolific, aged between 20-66.

  - 49 of them were eligible (passed the attention check)

An ALL test (cont.)
more on the group condition

- Different groups learned different patterns during the training phase



group 1 group 2 group 3 group 4 group 5 group 6 group 7

/b/ (or /d/) → /m/ (or /n/) in a compound
iff the compound’s length in syllables is not greater than 1, 2, 3, 4, 5, 6 or 7

An ALL test (cont.)
stimuli

- 8 lengths (1-8 syllables), 20 trials per length (160 in total)
- Among the 20 trials of each length:

  10 were focus words (containing /b/ or /d/)

  10 were filler words (not containing /b/ or /d/)
- Two separate members [X] & [bY] and two potential compounds [XbY] [XmY]

  e.g., [sapi] [butʃa] [sapibutʃa] [sapimutʃa]
- All stimuli were in the form of recordings
- All syllables: CV

  C= consonant in English inventory; no /b/ /d/ /m/ /n/ elsewhere

  V= /a/ /i/ /u/

An ALL test (cont.)
the morphological context: compounding

- The context of the nasalization rule is compounding (as in Japanese):

  [sapi] + [butʃa] = [sapibutʃa] or [sapimutʃa] (depending on the length threshold)

+

=

An ALL test (cont.)
procedure

- Training phase

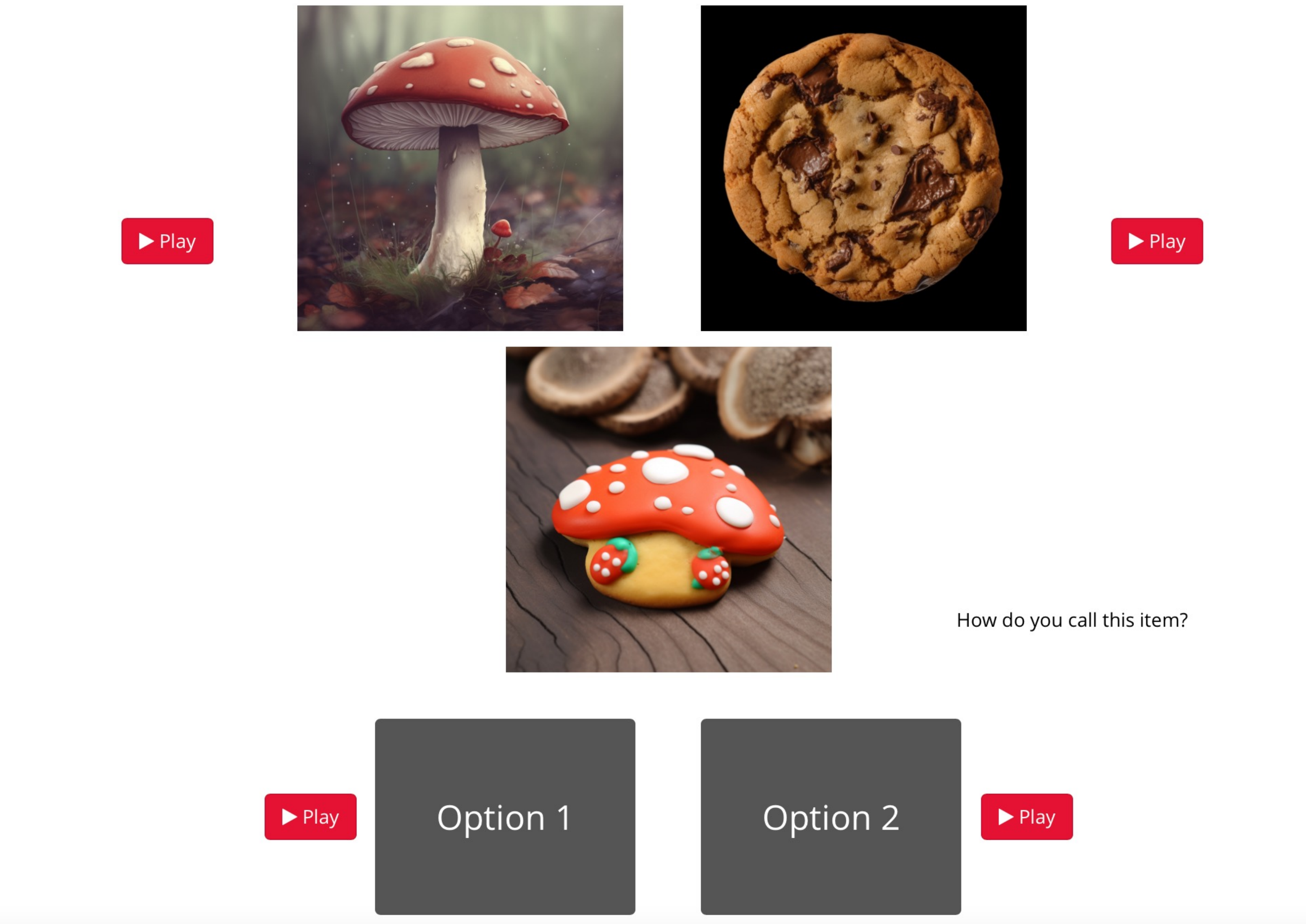
  10 trials per length (80 in total)

  feedback provided
- Testing phase

  10 trials per length (80 in total)

  no feedback

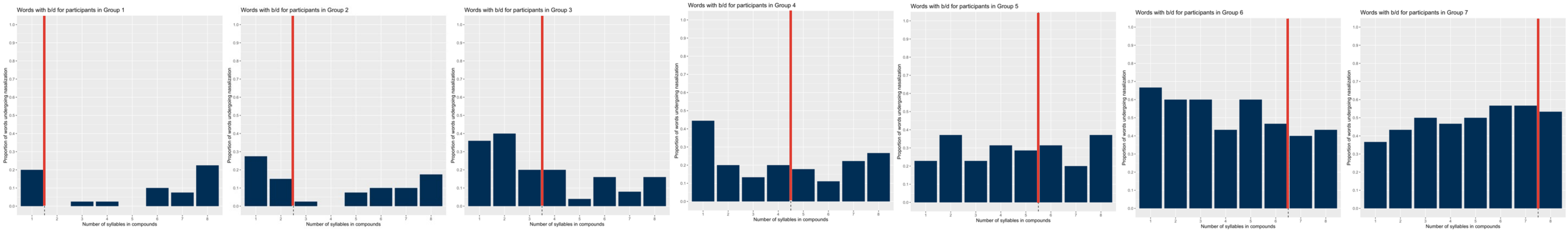
- An attention check mixed at random with trials



Results of Experiment 2

Results
no significant difference between the two sides

- Counting is allowed → Able to distinguish two sides of the length threshold as a condition of the rule, since the threshold is reflected by a number



group 1 group 2 group 3 group 4 group 5 group 6 group 7

- In each group, there is no substantial difference between both sides of the length threshold.

Results
Statistics (excl. filler words)

- Mixed-effect logistic model (with *max random effect*)
- Initial model: glmer(Nasalized_Response ~ group * side + (1 + side | subject) + (1 + group * side | item), data = filler_data, family = binomial)
- Results of fixed effects in the final model:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.2449	0.5558	-5.838	0.00000 ***
group2	0.1031	0.7411	0.139	0.88936
group3	1.2685	0.7957	1.594	0.11089
group4	0.6084	0.7403	0.822	0.41113
group5	2.2576	0.7219	3.127	0.00176 **
group6	2.3137	0.7704	3.003	0.00267 **
group7	2.2923	0.8234	2.784	0.00537 **
side	-0.3816	0.3655	-1.044	0.29653

Results

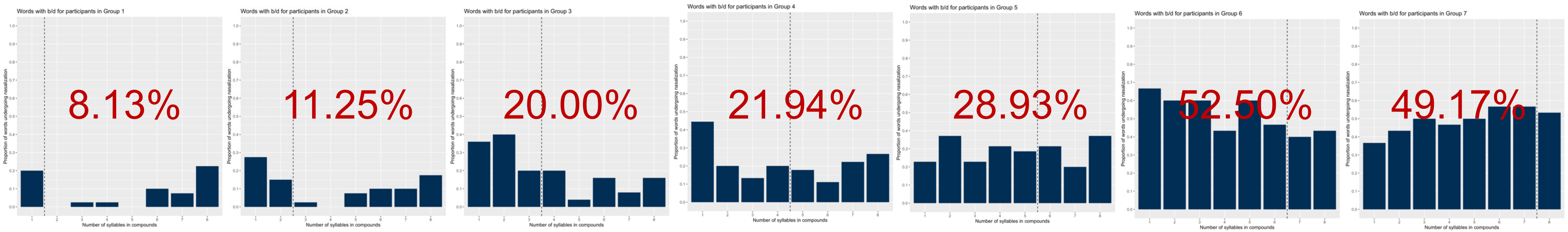
Statistics (incl. filler words)

- Mixed-effect logistic model (with *max random effect*)
- Initial model: glmer(Nasalized_Response ~ group * side * wordtype + (1 + side * wordtype | subject) + (1+ group * side | item), data = all_data, family = binomial)
- Results of fixed effects in the final model:

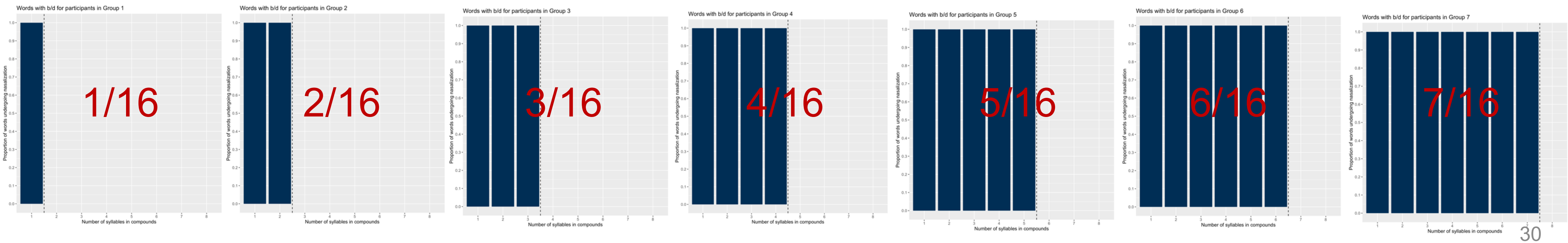
	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.46162	0.46041	-7.519	0.0000 ***
group2	0.16418	0.57476	0.286	0.7751
group3	1.30096	0.61934	2.101	0.0357 *
group4	1.06046	0.56159	1.888	0.0590 .
group5	2.29214	0.56600	4.050	0.0000 ***
group6	2.90064	0.59698	4.859	0.0000 ***
group7	2.83847	0.64302	4.414	0.0000 ***
side	-0.52396	0.33505	-1.564	0.1179
wordtype	0.08027	0.23458	0.342	0.7322
side:wordtype	0.95897	0.36442	2.631	0.0085 **

Results
frequency matching again

- Learning results (for focus words only)



- Exposure (for focus words only; filler words never nasalize)



Discussion

A learning bias against counting

- There exists a learning bias against alternations that involve ‘counting.’
- For Japanese VVN,

  · This counting-based pattern is not internalized. Unnatural patterns are disfavored and tend to be underlearned (e.g., Hayes et al. 2009)
- For ALL,

  · The learning bias can explain the indifference on two sides of the threshold

  · The results is consistent with the typological observation that counting-based patterns are rare (Paster 2019).

What is the cause of the mora-counting pattern in Japanese?

A ‘surfeit of the stimulus’ effect (Becker et al. 2011, 2012)

- Just an accidental generalization
- No synchronic explanation required
- Possibly due to a now inactive diachronic process

An alternative explanation: token frequency

- The longer a word is, the less frequently it appears. (Zipf 2013)

An alternative explanation of the ALL test: bias against source-oriented learning

- My experiment is designed to induce source-oriented (rule) learning (Kapatsinski 2012)
- The results shows a product-oriented generalization (‘learn surface forms’)

    · Even when the presentation condition is favorable for extracting rules, participants showed a strong preference for product-oriented generalizations (Kapatsinski 2012)

    · The results are biased to a certain extent by the tendency to draw product-oriented generalizations

    · A more source-oriented design may influence the results

Conclusion

- Japanese speakers internalize the natural factor of nasality of the preceding segment conditioning the tendency of /g/ undergoing nasalization; they fail to directly internalize the unnatural factor of the number of moras

- Artificial language learners did not learn the counting-based factor

- A learning bias against counting can explain the results of both experiments

References I

- Becker, M., Ketrez, N., & Nevins, A. (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language*, 84-125.
- Becker, M., Nevins, A., & Levine, J. (2012). Asymmetries in generalizing alternations to and from initial syllables. *Language*, 231-268.
- Berko, J. (1958). The child's learning of English morphology. *Word, 14*(2-3), 150-177.
- Breiss, C., Katsuda, H., & Kawahara, S. (2021). Paradigm uniformity is probabilistic: Evidence from velar nasalization in Japanese. In *Proceedings of WCCFL* (Vol. 39).
- Breiss, C., Katsuda, H., & Kawahara, S. (2022). A quantitative study of voiced velar nasalization in Japanese. *University of Pennsylvania Working Papers in Linguistics*, *28*(1), 4.
- Itô, J., & Mester, A. (1996). Correspondence and compositionality: The ga-gyō variation in Japanese phonology.
- Kapatsinski, V. (2012). What statistics do learners track? Rules, constraints or schemas in (artificial) grammar learning. *Frequency effects in language learning and processing, 1*, 53-82.

References II

- Hayes, B., Siptár, P., Zuraw, K., & Londe, Z. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language*, 822-863.
- McCarthy, J. J., & Prince, A. (1999). Prosodic morphology 1986. *Phonological theory: the essential readings*, 238-288.
- Paster, M. (2019). Phonology counts. *Radical: A Journal of Phonology*, *1*, 1-61. White, J. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition*, *130*(1), 96-115.
- Peperkamp, S., Skoruppa, K., & Dupoux, E. (2006). The role of phonetic naturalness in phonological rule acquisition.
- Warker, J. A., & Dell, G. S. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(2), 387.
- Zipf, G. K. (2013). *The psycho-biology of language: An introduction to dynamic philology*. Routledge.

Thank you!

Hailang Jiang
hailang.jiang.22@ucl.ac.uk
hailangjiang.github.io