

音韻システムは数を数えるのか？ —ライマンの法則の研究から 再現性の危機問題まで—

慶應義塾大学
言語文化研究所
川原繁人

KLS, 2023/6/10



二部構成

- 題材はともにライマンの法則

- 第一部：理論的な問題

「音韻システムは数を数えるのか？」

- 第二部：研究規範的な問題

再現性の危機問題（replication crisis） & データ開示制度（open science）

第一部

理論的な問題

- 音韻システム（文法システム）は、3以上数えていいのか？
- 伝統的な答え： no.
- McCarthy & Prince (1986: 1):

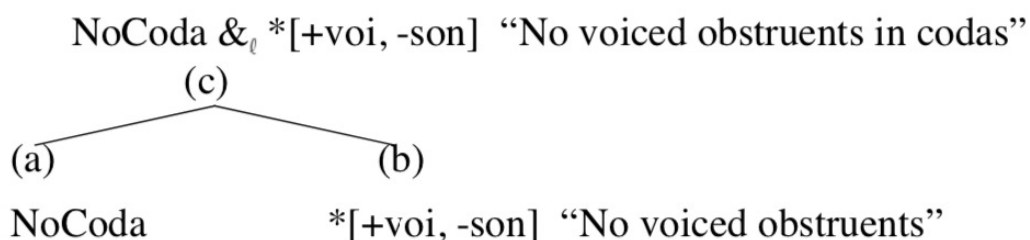
Consider first the role of counting in grammar. How long may a count run? General considerations of locality, now the common currency in all areas of linguistic thought, suggest that the answer is probably ‘up to two’: a rule may fix on one specified element and examine a structurally adjacent element and no other.

M&P 1986

- あり得ない畳語化パターン（reduplication pattern）：3つの segment をコピーする：
 - [bad-badupi] (CVC)
 - [bla-bladupi] (CCV)
 - [adu-adupi] (VCV)
- 結論：文法は（3以上を）数えない
- おそらく、音韻現象だけでなく統語現象も同じ（統語現象に関しては、ご意見お待ちしております）。

Ito & Mester (2003)より

- Ito & Mester (2003) : さまざまな制約を、より一般的な制約を結合(local conjunction)して派生させる。



- 結合は再帰的であるべき？
- その理論的な予測として...

A grammar G can expand the basic constraint set Con inherited from Universal Grammar to a superset $Con_G = Con \cup \{C_1 \& C_2\}$, for $C_1, C_2 \in Con$. Expansion is potentially recursive, so that Con_G can in turn be expanded to a superset $Con_{G'}$ by adding $C_3 \& C_4$ to Con_G , for $C_3, C_4 \in Con_G$, and so on.

With local conjunction as a recursive operation, ternary (and higher) conjunction such as $(\text{No-}\phi \&_{\delta} \text{No-}\phi) \&_{\delta} \text{No-}\phi = \text{No-}\phi^2 \&_{\delta} \text{No-}\phi = \text{No-}\phi^3_{\delta}$ are formally derivable. In the example given, the third violation of $\text{No-}\phi$ would be the fatal one. No convincing evidence has been found so far that $\text{No-}\phi^3$ is ever linguistically operative separate from $\text{No-}\phi^2$, which tends to support the old idea in generative linguistics (cf. syntactic movement theory) that the genuine contrast in grammars is not “1 vs. 2 vs. 3 vs. 4 vs. . . .”, but “1 vs. greater than 1.” [NB: ϕ is a variable representing a phonological structure and δ is a variable representing a domain]

A “Radical” proposal by Paster (2019)

[PASTER] Phonology Counts

Phonology counts

Mary Paster

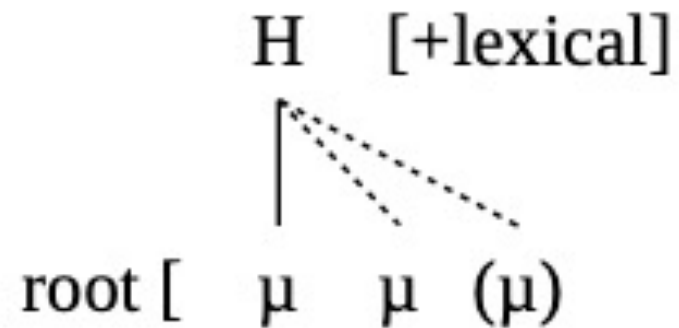
paper

[download](#)

Abstract: This paper reconsiders the widely accepted claim that phonological rules/constraints do not ‘count’ to three or more. As I discuss, there exist some phenomena that are problematic for this claim, especially in Bantu tone systems. This paper gives an overview of some phonological systems that seem to require counting. I argue that counting is more common than has been acknowledged, and that an adequate theory of phonology must account for counting.

Paster (2019)に挙げられている例

(10) H TONE TRIPLING



(32) 4TH MORA H TONE ASSOCIATION



最大エントロピー法はどのようなさ？

<i>weight:</i>	2.3	4.9	0.0	2.1	1.2	0.0	0.8	1.9	2.3	4.1	6.0			
	N _{AS} SUB	*N _C	*[_{root} m/n/ŋ	*[_{root} n/ŋ	*[_{root} ŋ	UNIF-maŋ- _{OTHER}	UNIF-paŋ-RED-	UNIF-maŋ- _{ADV}	UNIF-maŋ-RED-	UNIF-paŋ- _{NOUN}	UNIF-paŋ- _{RES}	<i>Harmony</i> (10)	<i>eHarmony</i> (11)	<i>probability</i> (8, 12)
/paŋ _{noun} + t.../														
[pa-n...] (subst.) 60.5 words (67%)	0	0	1	1	0	0	0	0	0	1	0	$1 \times 0.0 +$ $1 \times 2.1 +$ 1×4.1 $= 6.2$	$e^{-6.2}$ $= 0.0020$	$\frac{.0020}{.0020+.0008}$ $= \mathbf{0.73}$
[pan-t...] (unsubst.) 29.5 words (33%)	1	1	0	0	0	0	0	0	0	0	0	$1 \times 2.3 +$ 1×4.9 $= 7.2$	$e^{-7.2}$ $= 0.0008$	$\frac{.0008}{.0020+.0008}$ $= \mathbf{0.27}$

Lyman's Law can count only up to two

Shigeto Kawahara*, The Institute of Cultural and Linguistic Studies at Keio University, Japan, kawahara@icl.keio.ac.jp

Gakuji Kumagai, Department of English Language and Literature, Faculty of Letters, Kansai University, Japan, gakujick@gmail.com

*Corresponding author.

One long-standing question that is recurrently addressed in contemporary phonological studies is whether phonological systems can count beyond two. The traditional view is that phonological systems can count only up to two but not more (e.g., Ito & Mester, 2003; McCarthy & Prince, 1986); recently, however, some scholars have argued that phonological systems should actually be able to count beyond two (e.g., Paster, 2019; Kim, 2020). The current experiments address this general question regarding counting by studying Rendaku and Lyman's Law in Japanese. Rendaku is a morphophonological process in which the morpheme-initial voiceless obstruent of a second member of a compound becomes voiced. The application probability of Rendaku is significantly reduced if the second member already contains a voiced obstruent, a generalization that is known as Lyman's Law. Experiment 1 compared the applicability of Rendaku in nonce words which contain one voiced obstruent (e.g., [taguta]), and those which contain two voiced obstruents (e.g., [tegebi]). If Lyman's Law counts beyond two, Rendaku application is predicted to be more substantially reduced in the latter condition, as Rendaku would create morphemes that contain three voiced obstruents (i.e., [degebi]). The results show, however, that no meaningful differences were observed between the two conditions. Experiment 2 tested the recent claim that two nasal consonants may reduce the applicability of Rendaku (Kim, 2020; Kumagai, 2017).

今一度考え直してみたい。

実験1 (Kawahara & Kumagai 2023a, LabPhon)

実験1

- この「音韻システムは数えるのか？」という問題を、日本語のデータをもとに考えてみましょう。

(1) Examples of Rendaku

- a. /nise+**t**anuki/ → [nise+**d**anuki] ‘fake raccoon’
- b. /juki+**k**umi/ → [juki+**g**umi] ‘Snow Team’
- c. /hoçi+**s**ora/ → [hoçi+**z**ora] ‘starry sky’
- d. /oçi+**h**ana/ → [oçi+**b**ana] ‘dried flower’

(2) Blocking of Rendaku by Lyman’s Law

- /nise+**t**okage/ → [nise+**t**okage], *[nise+**d**okage] ‘fake lizard’
- /çito+**k**age/ → [çito+**k**age], *[hito+**g**age] ‘people’s shadow’
- /mori+**s**oba/ → [mori+**s**oba], *[mori+**z**oba] ‘cold soba’
- /çito+**h**ada/ → [çito+**h**ada], *[hito+**b**ada] ‘people’s skin’

動機：なぜこの実験を行うに至ったか？

- Kim (2019) & Kumagai (2017)：鼻音が2つあると連濁が阻止されるのでは？
- Kim (2019): 日本語話し言葉コーパス (Maekawa et al. 2004).
- Kumagai (2017): 無意味語実験.
- これらの研究によると [D...N...N]がダメ？ この制約は3つの segment を数えている.

実験1

- ライマンの法則は数えるのか？ ライマンの法則は[D...D]という構造と[D...D...D]という構造を区別するのか？
- 実験3条件:
 - (1) 有声阻害音を含まない無意味語 (e.g. [taruna])
 - (2) 有声阻害音を1つ含む無意味語 (e.g. [ta^guta]) → *[D...D]
 - (3) 有声阻害音を2つ含む無意味語 (e.g. [te^gu^bi]) → *[D...D...D]
- 和語では(3)のような単語は存在しないので、(2)と(3)が日本語の文法の中で区別されるのかは興味深い。

刺激

- 3 条件* 4 子音 * 各6 アイテム (計：72)

	0 vcd obs	1 vcd obs	2 vcd obs
/t/	[tamuma]*	[taguta]*	[tezuga]
	[tatsuka]*	[tozumi]*	[tezago]
	[taruna]*	[tegura]*	[tegubi]
	[tonime]*	[tazanu]	[taguga]
	[tekeha]*	[tegesa]	[tegozi]
	[tokeho]*	[toboφu]	[tebigi]
/k/	[kimane]*	[kidaku]	[kidabe]
	[kikake]*	[kobono]*	[kodziba]
	[kotona]	[kabomo]*	[kazido]
	[kumise]	[kedere]	[kudziba]
	[konihe]*	[kuziha]	[kezodo]
	[keharo]*	[kozana]	[kadzuba]

/s/	[semaro]*	[sebare]	[segabo]
	[sekato]*	[segeha]*	[sobogi]
	[sutane]*	[sobumo]*	[sugabi]
	[samohe]*	[sadanu]	[sobode]
	[sorise]*	[sodoka]	[sadage]
	[sateme]*	[sudaφu]	[sogebi]
/h/	[honara]*	[hobasa]*	[hogada]
	[çinumi]*	[hazuke]	[hegazu]
	[honiko]*	[hogore]*	[hedado]
	[hakisa]*	[çigiro]	[hadagu]
	[heraho]*	[φuzumo]	[çizuda]
	[çihonu]*	[hedeno]	[φubode]

実験手順

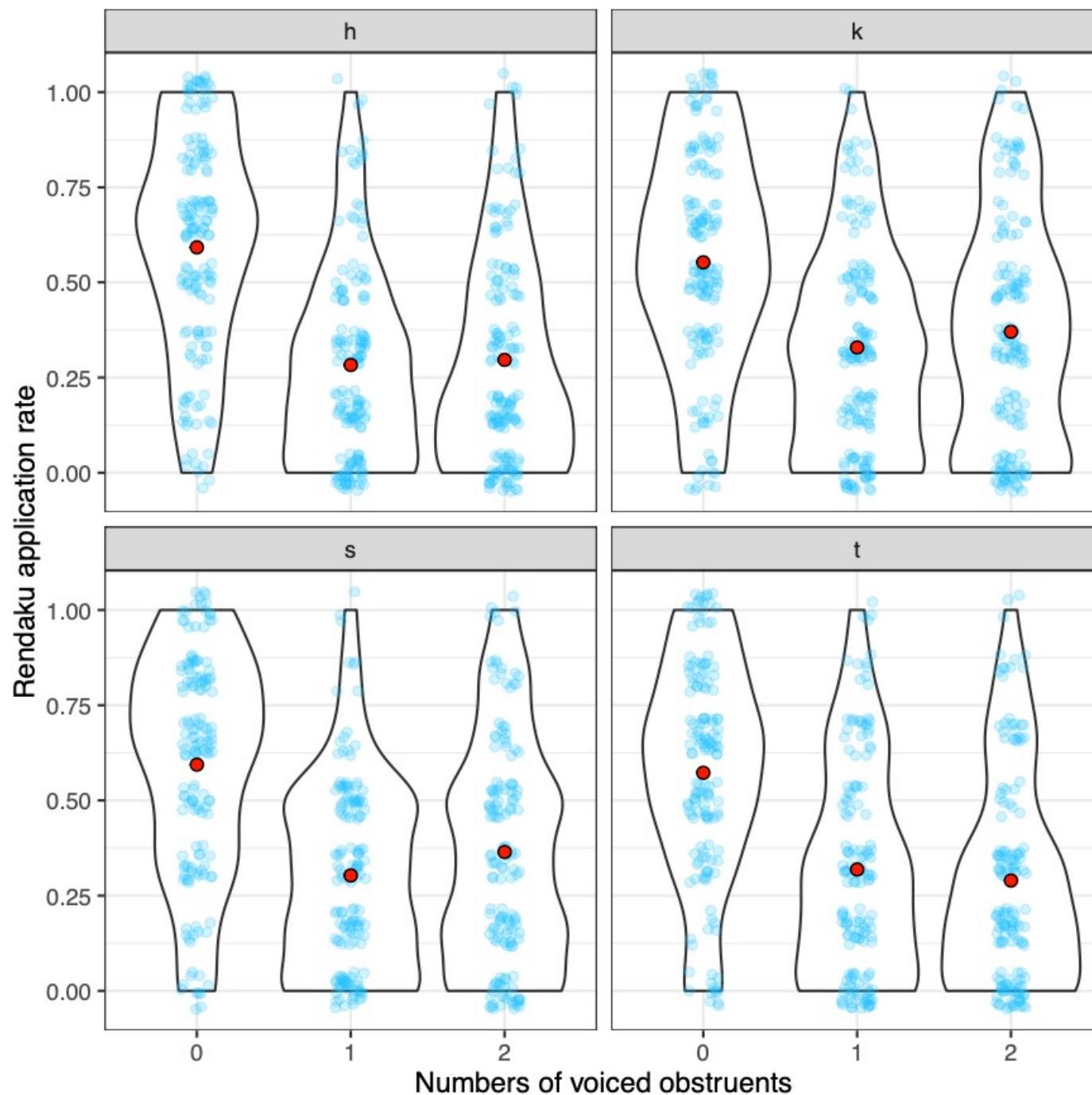
- 刺激語それぞれを「にせ」とくっつけてもらう。
- 連濁した形と連濁しない形がどちらか自然か選んでももらう。
- 例：無意味語の「たるな」という単語があったとして、「にせ」とくっつけた場合、「にせたるな」と「にせだるな」どちらがより自然に聞こえますか？

手順

- オンライン実験。149人の日本人母語話者。
- 混合ベイズアンロジスティク回帰（Baysian mixed effects logistic regression）
- 固定効果: 3条件、4子音
- 適切なランダム効果構造
- 詳細は：<https://osf.io/9qgtx> （詳細は第2部で！！）

結果

- “0” vs “1”にはっきりとした違い。
- “1” vs. 2の違いは、あまり存在しない。
- どの子音でもパターンは基本的に同じ。



回帰モデル

		β	error	95% CI
(a) intercept		-1.23	0.21	[-1.65, -0.82]
(b) sound type	/k/	0.29	0.26	[-0.22, 0.79]
	/s/	0.11	0.27	[-0.41, 0.62]
	/t/	0.20	0.27	[-0.32, 0.72]
(c) vcd obs	0 vs. 1	1.67	0.28	[1.12, 2.25]
	2 vs. 1	0.02	0.27	[-0.51, 0.53]
(d) interactions	/k/:0 vs. 1	-0.49	0.36	[-1.21, 0.21]
	/s/:0 vs. 1	-0.09	0.36	[-0.80, 0.62]
	/t/:0 vs. 1	-0.29	0.37	[-1.03, 0.43]
	/k/:2 vs. 1	0.15	0.37	[-0.56, 0.87]
	/s/:2 vs. 1	0.29	0.37	[-0.43, 1.02]
	/t/:2 vs. 1	-0.29	0.37	[-1.03, 0.45]

すごく安易な解釈
の仕方：

95% CI (Credible
Interval、信用区間)
が0を含んでいなければ、
大事な効果あり。

結論

- ライマンの法則は、 $[D...D]$ と $[D...D...D]$ を区別しない。少なくとも、その証拠は弱い。
- ベイズ統計では「どのくらいの確率で差がないか」を分析することができる（頻度統計ではできない）。52.3% of the posterior samples are contained in the ROPE of $\beta=0$ 。
- つまり、 $[D...D...D]$ が $[D...D]$ よりも悪いということとはなさそう。ただし、「完全に差がない」とは言い切れない。

理論的にはどう考えたらよいのか？

- Moreton & Smolensky (2002:306) : “*the local conjunction of C1 and C2 in D*, is a constraint which is violated whenever **there is a domain** of type D in which both C1 and C2 are violated”.
- 音韻システムが嫌うのは、違反する構造そのものではなくて、「**違反する構造を含む領域**」である。
- $[D...D...D]_{\text{domain}}$ vs. $[D...D]_{\text{domain}}$

あれ？ Kim (2019)のデータはどうなるの？

- 「日本語話し言葉コーパスにおいて鼻音が2つあると連濁しない」。実際にデータをシェアしてもらったところ.....
- すでにE2が複合語である場合 (e.g. [tate-mono])。この場合、Right Branch Condition (Otsu 1980) によって連濁は阻止されるず。
- その複合語が有声阻害音を含む場合も存在 (e.g. [tabe-mono])。この場合は「普通の」ライマンの法則で十分。
- 残った例は「好み」と「頼み」。でも、これらは連濁するのでは？
- 「かなめ」も確認されたが、これはそもそも複合語にならなそう。

じゃあ Kumagai (2017)は？

- E2にアイテムが3つしか使われていない。
- 「鼻音2つ条件」は、3つとも[XX-nama]の形。
- ちゃんと実験をやり直した方がいいよね？

ただし、[son]*2がライマンの法則を引き起こす可能性も

- 日本語の共鳴音（[+son]）の[voice]の振る舞いは、ちょっと謎。
- 共鳴音の[voice]は連濁を阻止しない (Ito & Mester 1986)。
- しかし、鼻音の[+voice]は、後続の阻害音を[+voice]に変化させられる (e.g. /kam+ta/ → [kaⁿda]).
- 共鳴音は、ラップの韻(Kawahara 2007)やダジャレ (Kawahara & Shinohara 2008)において、無声阻害音よりも有声阻害音とより韻を踏みやすい。

実験2

共鳴音1つある場合や2つある場合、連濁率を下げるのか

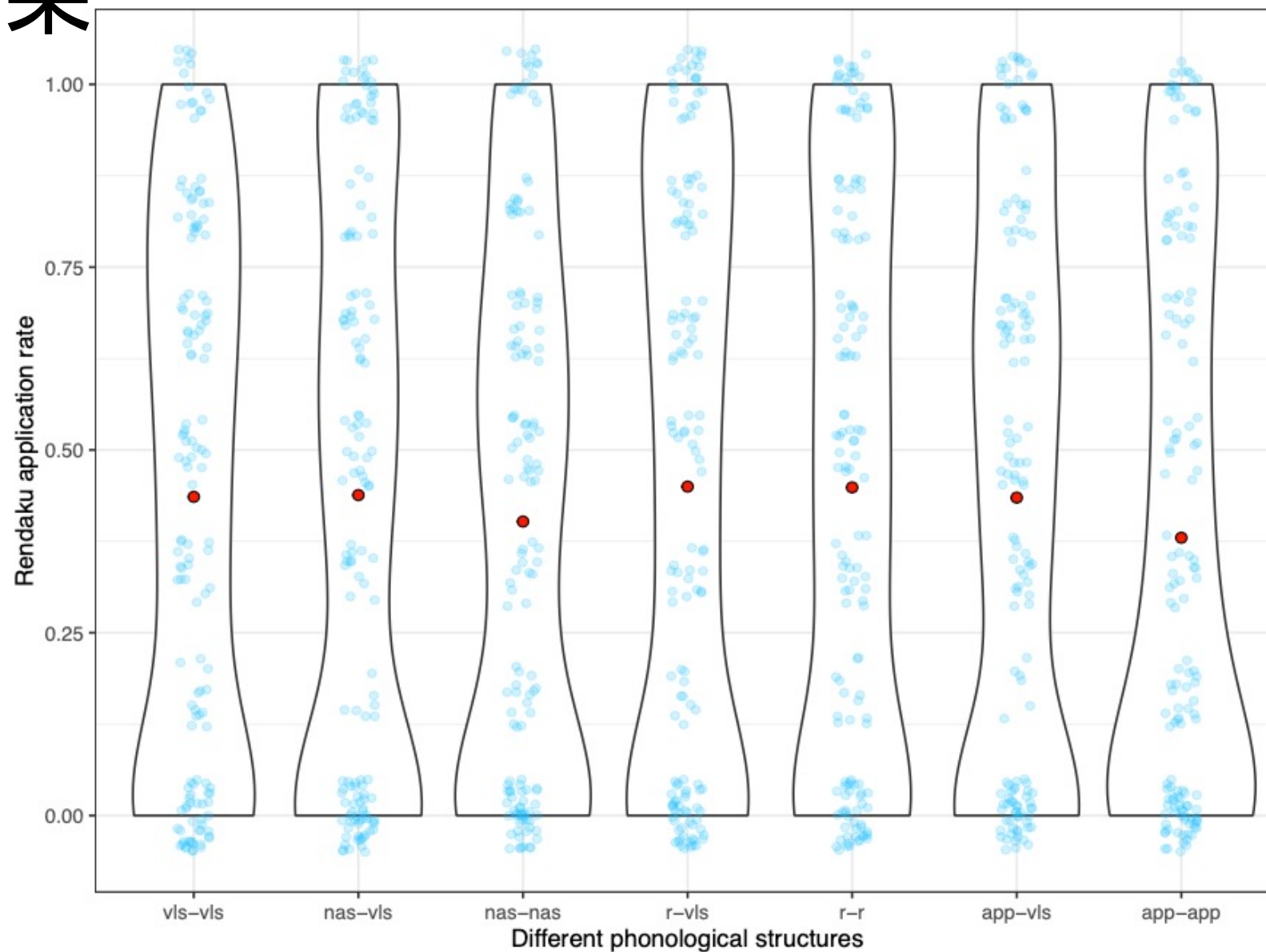
baseline

(a) [h-vls-vls]	(b) [h-nas-vls]	(c) [h-nas-nas]	
[hatosa]	[hanuta]	[hanumo]	
[hasaka]	[hanasa]	[hanama]	
[hetosa]	[henoke]	[henona]	
[hekita]	[henaso]	[henema]	
[hotaso]	[honato]	[honimu]	
[hokata]	[honika]	[honine]	
(d) [h-r-vls]	(e) [h-r-r]	(f) [h-App-vls]	(g) [h-App-App]
[harito]	[harura]	[hajuto]	[hajuwa]
[harose]	[harare]	[hawase]	[hawaja]
[herota]	[herora]	[hejata]	[hejowa]
[heresa]	[herera]	[hewasa]	[hewaja]
[horike]	[horiru]	[hojaso]	[hojuwa]
[horiso]	[horiro]	[howake]	[howaju]

- 手順などは実験1と一緒に
- 133人の日本語母語話者

実験 2 の結果

- 正直、面白い結果はなし



The model summary

- Baseline = vls-vls

- 鼻音2つある場合、ほんの少しだけ連濁率は下がるものの、そこまで下らない。

- (半母音2つだと下がる！？)

		β	error	95% CI
(a) intercept		-0.89	0.32	[-1.54, -0.25]
(b) condition	nas-vls	-0.02	0.20	[-0.42, 0.36]
	nas-nas	-0.26	0.19	[-0.63, 0.12]
	r-vls	0.10	0.20	[-0.31, 0.49]
	r-r	0.11	0.19	[-0.27, 0.49]
	app-vls	-0.02	0.20	[-0.41, 0.37]
	app-app	-0.48	0.22	[-0.91, -0.06]

第一部の結論

- 少なくともライマンの法則は数えない。
- [D...D] と [D...D...D] は同じくらいにダメ。
- *[D...N...N] といった制約も存在しない。
- Paster (2019) はどうなった？
- 「数える」とされている現象がすべて、声調など suprasegmental phonology に関わるものばかり。
- segmental phonology と suprasegmental phonology では異なるのか？
（侃々諤々の議論がなされています）。

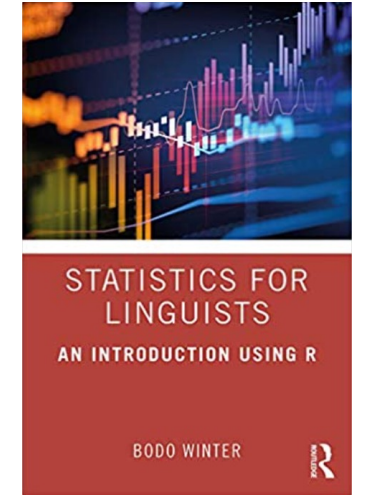
第二部

背景

- 再現性の危機問題（replication crisis）：学術論文や学会で発表された結果が、他者の研究によって再現できない問題。
- 事の発端：Bem (2011), Journal of Personality and Social Psychology. 「人間には予知能力がある」という内容。
- 論文の形式には問題がなく、統計的な有意差も示されていた。
- 他の複数の研究チームによって、再現が不可能だった。しかも、再現実験が同学術雑誌から掲載を拒否された（Ritchie et al. 2012）。

- Open Science Collaboration (2015)の大規模な調査：認知心理学の分野で、47%の結果しか再現できなかった。（ただし、再現の方法に疑義もある）。
- 同じ問題が別の分野でも指摘されている：
 - 経済学（Camerer 2016）
 - 社会科学（Camerer 2018）
 - 医学（Noseki & Errington 2017）

言語学では？



- 海外ではこの問題に関する意識が高まっている。
- **Winter (2019):** 統計の教科書だが、再現性の危機問題を大きく扱っている。
- **Linguistics (2021):** 言語学における再現性の危機問題に関する特集号
- **B4SS (2021) :** LabPhonが主催するベイズ統計のワークショップ (<https://learnb4ss.github.io/>)。頻度主義からベイズ統計への転換の主な動機は、再現性の危機問題とされていた。

再現不可能な結果の具体例

- Kluender et al. (1988): 非言語音声知覚実験。英語話者は、長い音のあとの無音区間を短いものとして知覚する。
- Fowler (1992): 逆の結果に。
- Kingston et al. (2009): Fowlerと同様、逆の結果に。
- Boroditsky (2001) : 言語が思考のパターンに影響を与える。
- Chen (2007) : 再現できず。
- 個人的にはアメリカ時代のKingstonラボで非常に苦勞をしました。また、他の例も風の噂でよく聞きます。

日本では？

- 心理学では、日本でもこの問題が認知されつつある。
- 池田 & 平石（2016）：「心理学における再現可能性危機:問題の構造と解決策」
- 平石 & 中村（2021）：「心理学における再現性の危機の 10 年—危機は克服されたのか、克服され得るのか—」

Japanese Psychological Review
2016, Vol. 59, No. 1, 3–14

心理学における再現可能性危機：問題の構造と解決策¹⁾

- ReprorducibiliTea Tokyo
(<https://twitter.com/repTeaTokyo>)

池田 功毅¹, 平石 界²

¹ 中京大学大学院 / 日本学術振興会特別研究員 PD

² 慶應義塾大学

The reproducibility crisis in psychology: Its structure and solutions

Koki IKEDA¹ and Kai HIRAISHI²

¹ Chukyo University / JSPS Research Fellow

² Keio University

ただし

- 平石先生との私信（2021/12のメールでのやりとり）
- 「心理学は再現性問題が進んでいる」との印象をお持ちのことですが、こと国内については、どこまでのものなのだろうか...と思っています。

- 川原繁人（2022）「言語学者、外の世界へ羽ばたく」の1章でこの問題について論じる。
- 平石先生との私信：再現性の問題に関心のある仲間と「他の分野ではどうなっているだろう」という話を良くしています...





大学共同利用機関法人

人間文化研究機構

[📍 アクセス](#) [✉ お問い合わせ](#) [🌐 English](#)



ご寄附はこちら
人文機構基金

[TOP](#)[機構案内](#)[研究推進・人材育成](#)[データベース](#)[イベント・展示・広報](#)[情報公開](#)[イベント・展示・広報](#)[イベント・展示情報](#)[イベント一覧](#)[展示情報一覧](#)[人文機構シンポジウム](#)[社会連携事業](#)[書籍紹介](#)

[TOP](#) / [広報](#) / 第12回 人間文化研究資源共有化研究会「人文科学におけるオープンサイエンスの課題」

第12回 人間文化研究資源共有化研究会「人文科学におけるオープンサイエンスの課題」

第12回人間文化研究資源共有化研究会 「人文科学におけるオープンサイエンスの課題」

国立国語研究所の言語資源と
オープンデータ・オープンサイエンス

第1回 SPARC Japan セミナー2019

「人文社会系分野におけるオープンサイエンス ～実践に向けて～」

国立国語研究所の言語資源と オープンデータ・オープンサイエンス

小木曾 智信

(国立国語研究所)

問題解決策のひとつ

(あくまでひとつ。万能薬ではない)

- データ開示制度 (open science initiative) : 論文とともにその分析のもととなったデータや統計スクリプトを開示する。

データ開示制度

- Garellek et al. (2020): 11名の音声学が連名でletter to the editor

Journal of Speech Sciences (9): 03-16. 2020.

Available at: <http://revistas.iel.unicamp.br/joss>

LETTER TO THE EDITOR: TOWARD OPEN DATA POLICIES IN PHONETICS: WHAT WE CAN GAIN AND HOW WE CAN AVOID PITFALLS

GARELLEK, Marc¹

GORDON, Matthew²

KIRBY, James³

LEE, Wai-Sum⁴

MICHAUD, Alexis^{5*}

MOOSHAMMER, Christine^{6,7}

NIEBUHR, Oliver⁸

RECASENS, Daniel⁹

ROETTGER, Timo B.¹⁰

SIMPSON, Adrian¹¹

YU, Kristine M.¹²

データ開示制度

- Cho (2021): Journal of Phoneticsのeditorial statement。



Journal of Phonetics
Volume 89, November 2021, 101113



Editorial

Where we are at: Impact, special
collections, open science and
registered report at the *Journal of
Phonetics*

Taehong Cho Editor-in-Chief 

データ開示制度

- 多くの国際雑誌で義務化（Language, Linguistics, Phonologyなど）

Data and referencing

Authors, associate editors, and referees are asked to pay special attention to referencing and to how data is handled in the submission. References should not only include the work that the author consulted, but if at all possible the original source as well. Data sources should be clearly referenced and, if publicly available, cited as such, e.g. with a url, or appended to the article as supplementary material. Where the data is not publicly available, (and there are often good reasons why it is not), authors should explain why it is not made available, typically in a footnote in the paper. Generally speaking, articles reporting results of empirical studies such as quantitative analyses or experimental results should allow replication by other researchers based upon the article's contents and reference, and possibly supplementary materials. Availability of field notes, for instance, clearly

- 理想的には査読時に、少なくとも論文出版時には、データをすべて開示するべし。
- 編集者としての経験：査読者によっては、「データが開示されていない」という理由で、論文を却下することもある。
- 著者としての経験：査読者がより良い統計方法で再分析してくれることもある。
- Open science framework (osf) というオンラインリポジトリが便利。

Russian Pokemonastics Experiment 2	Kumagai, Watabe, and kawahara	2023-04-25 2:17 PM
Pokemon Base Friendship	Kilpatrick, kawahara, and Ćwiek	2023-02-14 10:59 AM
Iconicity in Japanese Guessing Study	Sidhu, Winter, Woodin, and 2 more	2023-02-01 2:08 PM
Japanese baby diaper names II	Kumagai and kawahara	2023-01-31 7:08 PM
Does Lyman Count?	kawahara and Kumagai	2022-12-06 9:59 AM
Two nasals and Rendaku	kawahara and Kumagai	2022-11-29 12:44 PM
Variation, gender and perception	Sherwood, Shaw, kawahara, and 2 more	2022-11-28 12:01 PM
Limits on gestural reorganization following vowel deletion: The case of Tokyo Japanese	Shaw and kawahara	2022-11-16 11:29 AM
Random Forest Pokemon Evolution	Kilpatrick, kawahara, and	2022-10-18 8:38 AM

Does Lyman Count?	
OSF Storage (United States)	
Expt1	
lyman23-newpriors.rds	2022-11-03 08:50 AM
LymanCountResultsExpt1.pdf	2022-12-06 09:58 AM
LymanCountsVcdObs.html	2022-12-06 09:58 AM
LymanCountsVcdObs.Rmd	2022-11-03 08:49 AM
two_lymans_longformat.csv	2021-06-01 08:51 AM

ただし

- データ開示制度には、反対意見も少なくない。
- 個人的な経験。とある国際雑誌の**associate editor**会議で、大御所の先生（複数）から反対意見が。（もちろん賛成意見も）。
- 根拠：「言語学で再現性の危機が起こっているという証拠はない。」

Bodo Winter (p.c. 2022/05/19)

strictly speaking, we don't know (in an empirically based way) whether there [is a replication crisis] for linguistics. So the arguments for this existing in linguistics have to be more indirect: 1) The incentive system, the publication process, the way data are analyzed, are in no way different from psychology... so therefore it would be quite unlikely for things to be different for us... and 2) our sample sizes are even smaller than psychology, which we know is implicated with a whole suite of problems, and 3) quite a few language-relevant findings have failed to replicate...

第一部とここでつながります

- 以下の議論ではKim (2022)を議論の対象としますが、個人的な批判を意図したものではありません。
- Kim (2022, Glossa): 鼻音が2つ含まれる単語において、連濁の適用確率が下がる。例：「かもめ」「つまみ」。（話し言葉コーパスではなく）、Rendaku Database ver.4.0 (Irwin et al. 2020) に基づき「統計的な有意差」を検知。
- 鼻音を含まない場合85% (1,339/1,580)
- 鼻音を1つ含む場合80% (418/520)
- 鼻音を2つ含む場合52% (15/29)

- この主張が正しいとすれば、音韻システムは「数を数えられる」ということになる。「xxが3つあるとダメ」という制約は存在しないと考えられてきた（第一部参照）。
- このような驚きのパターンだからこそ、Glossaに掲載されたと言える。
- しかし、母語話者として、具体的な例が思いつかない.....！？論文に挙げられている例は2つだけ。データが公開されていないから、チェックしようがない！

- Kawahara & Kumagai (2023b, Glossa) : 仕方ないから、自分たちで Rendaku Corpus ver. 4.0を再分析してみよう.....
- その結果、Kim (2022)の「鼻音2つ条件」にあてはまる単語は4つのみと判明。
 - かもめ
 - つまみ
 - このみ
 - たのみ
- 「このみ」と「たのみ」は（ほぼ）確実に連濁する。

- Kim (2022)の議論は、該当する単語が4つで、2つは反対の傾向を示していた。
- 「つまみ」は2例のみ。「はなつまみ」はObject-verb compoundだから連濁しないという可能性も (Vance 2022)。「ひとつまみ」は「ひと+Y」だから連濁しないという可能性も (Irwin 2012)。
- すると、Kim の分析は「かもめ」の分析とほぼ同義では？。

- 査読時に、データが開示されていれば、査読者がこれを指摘することは容易であったはず。
- 少なくとも論文とともにデータが開示されていれば、興味を持った読者は調べることができた。
- Kawahara & Kumagai (2023b)が再分析するという手間を省くこともできた。

統計的な観点から

- Kim (2022)の統計は、ロジステック回帰を報告。「有意差はあり」とされていた。
- ただし、アイテムをランダム効果に入れていない。つまり、29個の単語があたかも別々の単語であるかのように、分析されている。
- これは間違って「有意である」という結果をだしてしまいがち（Type 1 error: Winter 2019; Winter & Grice 2021）。

- よって、統計も再分析が必要と判断。
- 「鼻音なし条件」などのデータは編集が難しいために、Kimにデータを共有するようにお願い。
- Kimはこれを拒否。以降の連絡自体も拒否。ただし、彼女は言語学者を辞めて、別の職に就いているため、拒否の理由までは不明。
- どこまでが（元）研究者の責任なのか、私も正直わかりません。この点についても議論お待ちしております。

もちろん

- 何をどこまでどうやって公開するかは議論の余地あり。
- 音声データは個人情報ですから取扱注意。
- F0やフォルマントの値などなら問題はないはず？ (Garellek et al. 2020)
- 有料のデータベースから抜き出したデータは公開できないという実情（注：Kim 2022の非開示理由はこれではない）
- 前川先生がこの問題に対して取り組んでいるようです。

広い視点から

- データを開示することで失うものはなにもない。
- 科学として進歩するためには、質の良いデータに基づいて論文を執筆する必要がある。
- でも不安も理解できる。「ミスが見つかったらどうしよう？」
- ミスは正せばよい。正してこそその信頼。
- 分野としてミスに対して寛容になるべき。

- いや、もしかしたら社会としてももっとミスに寛容になるべき？

私が娘に「100点なんて目指さないでいい」と言った理由(チラ見せ)

♡ 12



川原繁人のnote~~言語学者よ象牙の塔の外に出よう！~~

2023年5月3日 08:01



以下は7月にディスカヴァー21から出版予定の書籍のコラムの冒頭部分です。
全部公開すると出版社様に迷惑がかかってしまうので、本当に冒頭だけ。なんで先行公開させて頂いたかという、ネット上で「大学は知識を吸収するところである。知識はネットで吸収できる。だから大学はもういない」という記事を見たからです。これはちゃんと反論しないとダメだよねってことで、Twitterでつぶやいたら少し反響があったので、私の考えの背後にあったこの記事을チラ見せです！

- Replication failure = 過去の研究の否定ではない。
 - Vance (1979) = ライマンの法則の位置効果が無意味語で発見。
 - Kawahara (2012) = この効果を発見できず。
 - Ihara et al. (2009): 時代とともにこの効果が薄れている可能性を指摘。
- つまり、我々は歴史的変化を目撃しているのかも。この新たな仮説についてはただいまさらに精査中

自戒を込めて

- 正しい（理想的な？）研究態度：研究の結果として論文を執筆する。
- 現実：仕事を得るためや研究費獲得のため、大学からの要請などから論文を書くことが求められる。すると多少質の悪いデータでも「これで論文が出せるならいいや」という発想に。

自戒を込めて

- つまり「研究の結果が論文」のではなく、「論文を書くために研究をする」ことに。
- 私はこの発想の被害者でもあります、加害者でもあります。
- それでも、この雰囲気は科学の進歩には百害あって一利なし。
- データ開示制度が日本の言語研究でも広がりますように！
- 特に後半部は色々な意見があると思います。議論を楽しみにしております。