

Komaba campus, the University of Tokyo

Effects of intervocalic voicing on diminutive *-kko* suffixation in Tohoku Japanese



Mamoru Yakuwa¹, Chuyu Huang²

¹Aoyama Gakuin University, ²Nagoya Gakuin University

March 10, 2026

-kko suffixation in Tohoku Japanese

- (1) -kko diminutives in Tohoku Japanese (Kushibiki, 2014)
- ame-kko* ‘candy-DIM’
 - o-cha-kko* ‘HON-tea-DIM’
 - sensei-kko* ‘teacher-DIM’

Restriction (Abe, 1999)

***<...ko + -kko>**

e.g. **hanko-kko*, ‘stamp-DIM’, **sooko-kko* ‘shelf-DIM’

Effect of voicing

- When the base-final /ko/ is voiced to [go], this restriction does not apply. (Abe, 1999)
 - **Intervocalic voicing (IV): In Tohoku Japanese, /k/ and /t/ between vowels voice to [g] and [d].** (Inoue, 1968)
- (2) Compatibility of voiced /ko/ with the diminutive suffix *-kko* (Abe, 1999)
- ✓ *tago-kko* ‘octopus-DIM’
 - ✓ *hago-kko* ‘box-DIM’
 - ✓ *nogo-kko* ‘saw-DIM’

Research questions

- 1 How productive is *-kko* suffixation for /ko/-ending bases?
- 2 How do region and generation affect the pattern?
 - Abe (1999) only refers to *Hirosaki-shi* (Aomori) Japanese.
- 3 Is the /ko/ contiguity prohibition actually observed?
 - This study aims to replicate the effect experimentally.
 - Does IV affect /ko/ adjacency?
- 4 What is the underlying representation: /ko/ or /kko/?
 - How is it accounted by phonological theory?

Procedure

- Production experiment in which participants produced *-kko*-suffixed diminutives.
 - They were asked to add *-kko* for each stimulus after a brief instruction.
 - If the corresponding *-kko* diminutive form was not natural, they pronounced “I do not say it”.
- Online experiment coded with jsPsych (De Leeuw, 2015) and conducted on Cognition.run.

Participants

- 32 participants born and raised in the Tohoku region (Aomori, Akita, Iwate, Yamagata, Miyagi, Fukushima) were recruited via CroudWorks.
- Another 17 participants were recruited via snowball sampling.
 - The critical period for their native dialect was defined as up to junior high school graduation.
- Participants were also asked about their gender, age, and residence history.

Participants

- $N = 49$, $F = 25$, $M = 24$
- Mean age = 44.16
SD = 16.54
- Aomori: 6
Akita: 11
Iwate: 9
Yamagata: 5
Miyagi: 12
Fukushima: 6

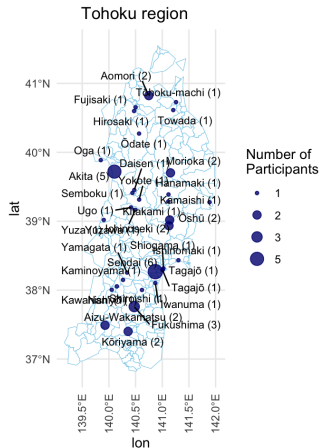


Figure: Participants' residences during critical period

Carrier sentence

- Participants were instructed to fill the blank in the carrier sentence below with a /ko/-suffixed diminutive of the given stimulus.
- Dialect-specific expressions were used to elicit dialectal features.
- The carrier sentence was taken from Huang et al. (2023).

Actual carrier sentence

_dabe. ndanda, _ da.
'It should be __, yeah, it is __.

Stimuli

- Eighteen stimulus words were prepared.
 - Each word ends with /ko/
- Half contain intervocalic /k/.
 - Intervocalic /k/ is expected to undergo voicing, whereas the others are not.
- Non-intervocalic environments
 - 1 …/N (nasal mora) ko/ as in *kinko*
 - 2 …/Q (geminate mora) ko/ as in *iburigakko*
 - 3 …/R (long vowel) ko/ as in *sooko*

Entire stimuli

Intervocalic

<i>komugiko</i>	'flour'
<i>tarako</i>	'cod roe'
<i>tabako</i>	'cigarette'
<i>taiko</i>	'drum'
<i>otyoko</i>	'a kind of cup'
<i>hudebako</i>	'pencil case'
<i>namako</i>	'sea cucumber'
<i>takenoko</i>	'bamboo sprout'
<i>neko</i>	'cat'

Non-intervocalic

<i>osinko</i>	'a kind of pickles'
<i>rakko</i>	'otter'
<i>iburigakko</i>	'a kind of pickles'
<i>kinko</i>	'vault'
<i>hiyayakko</i>	'tofu'
<i>sooko</i>	'warehouse'
<i>patiNko</i>	'pinball game'
<i>anko</i>	'bean paste'
<i>inko</i>	'parakeet'

Online experiment sample

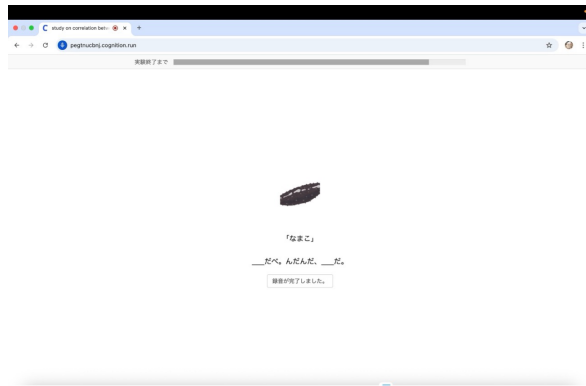


Figure: Screenshot of the experiment interface

Data cleansing

- Recorded tokens were converted from BASE64 to 16 kHz monaural WAV files using Python.
- Stimulus words that participants altered were excluded.
e.g., *tyape* for *neko* ‘cat’ *dego-duge* for *osinko* ‘a kind of pickles’
- Finally, **631 tokens** from **46 participants** were analysed.
 - Sample size: **optional stopping** (Rouder, 2014).
 - Optional stopping protocol: 95% CI ≥ 0.0 and BF ≥ 3.0 .

Raw data

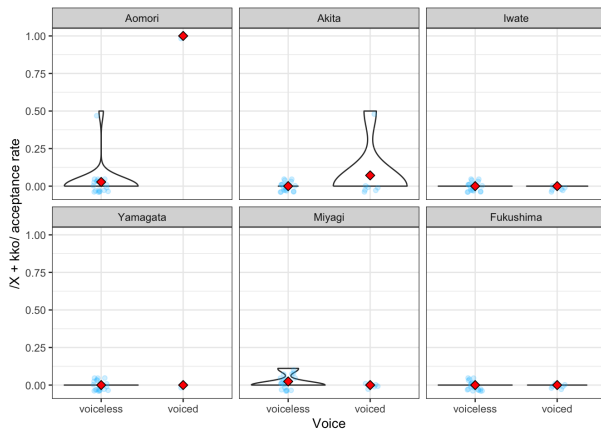


Figure: Acceptance and voice response rates by prefecture

Inferential analysis: model

● Bayesian mixed effects model

- `brms` package (Bürkner, 2017) of R (R Core Team, 2025), implemented with R Studio.
- **Prior setting**: normal (3,0) weakly informative prior (Lemoine, 2019).
- **MCMC**: 4 chains, each with 2,000 iterations, were run. The first 1,000 iterations were discarded as warmup. No thinning (`thin=1`).

Model formula

$$\underbrace{\text{acceptance}}_{\text{dependent variable}} \sim \underbrace{\text{voi_qual} + \text{pref} + \text{age_group}}_{\text{fixed factor}} + \underbrace{(1|\text{stimuli}) + (1|\text{participant})}_{\text{random factor}}$$

(not interacted)

Inferential analysis: Overall result

Parameter	Estimate	pd	lower limit of 95%CI	upper limit of 95%CI	\hat{R}
voiceless - voiced	3.55	99.10%	0.24	7.59	1.00
younger - middle	1.39	66.30%	-5.64	8.02	1.00
younger - elder	2.80	72.95%	-7.01	11.52	1.00
middle - elder	1.32	64.12%	-7.52	9.96	1.00
Aomori - Akita	-2.30	72.90%	-9.58	5.38	1.00
Aomori - Iwate	-4.30	86.08%	-12.06	3.46	1.00
Aomori - Yamagata	-3.21	77.65%	-11.52	4.90	1.00
Aomori - Miyagi	-2.06	69.67%	-9.04	5.57	1.00
Aomori - Fukushima	2.66	68.53%	-7.67	14.10	1.00
Akita - Iwate	-2.00	71.25%	-8.69	5.65	1.00
Akita - Yamagata	-0.96	59.72%	-8.62	6.65	1.00
Akita - Miyagi	0.30	54.20%	-6.76	7.35	1.00
Akita - Fukushima	5.01	82.85%	-6.85	15.25	1.00
Iwate - Yamagata	1.00	60.22%	-6.89	8.36	1.00
Iwate - Miyagi	2.29	74.17%	-4.96	9.65	1.00
Iwate - Fukushima	7.08	89.85%	-4.32	18.02	1.00
Yamagata - Miyagi	1.23	62.40%	-7.10	8.62	1.00
Yamagata - Fukushima	5.91	85.55%	-5.13	17.79	1.00
Miyagi - Fukushima	4.62	81.92%	-5.93	15.73	1.00
Participant	11.20	—	5.05	22.79	1.00
Stimuli	1.00	—	0.04	3.01	1.00

Key findings from the result

- The four chains mixed well ($\hat{R} = 1.00$).
- **Intervocalic voicing facilitates the acceptance of *-kko* suffixation.**
95% CI: [0.24-7.59], pd = 99.10%: **probably present** (Makowski, Ben-Shachar, Chen, & Lüdecke, 2019).
- Regional effects are **not supported**.
- Generational effects are also **not supported**.
 - All 95% CI ranges include 0 and pds are too low to be compelling.
- **Inter-speaker variation** (Estimate = 11.20).
- Small stimulus-level variation (Estimate = 1.00).

Identity avoidance

- UR of [kko] → /**kko**/ or /**ko**/?
- **Identity avoidance** (IA) (Yip, 1998)
Adjacency of identical segments is prohibited.
- Moraic identity avoidance in Japanese: **OCP(CV)**

Evidence 1: Sano, Shinichiro (2013)

(3) IA in *ra*-deletion (*ra-nuki*)

a. kari-**rare**-ru →[✓] [kari-re-ru] ‘can borrow’

b. ire-**rare**-ru →[?] [ire-re-ru] ‘can insert’

(4) IA in potential *sa*-insertion

a. ik-ase-ru →[✓] [ik-**as**ase-ru] ‘have somebody go’

b. tobas-ase-ru →^{*} [tobas-**as**ase-ru] ‘have somebody fly’

Identity avoidance

Evidence 2: Rendaku

(5) IA in Rendaku (Sato, Y, 1989)

- a. tobi-hi *tobi-bi 'flew-fire'
- b. kizu-tukeru *kizu-dukeru 'to hurt'

- Kawahara and Sano (2014) nonce word experiment

(6) **IA plays a role in catalyzing rendaku**

- a. ika-[ka→ga]-kaniro > ika-[ta→da]niro
- b. iga-[ka→ga]niro < ida-[ka→ga]niro

Underlying representation

- /ko-kko/: Two /ko/s are **not adjacent**. → IA is not applicable.
- Word origin: **ko** ‘child’.

Proposal: UR is /ko/

① *neko*

↓ Intervocalic voicing

② *nego*

↓ Suffixation (+*ko*)

③ *nego-ko*

↓ Gemination

④ *nego-kko*

① *neko*

↓ Suffixation (+*ko*)

② No output

Conclusion

- 1 -*kko* suffixation for /ko/-ending bases is not much productive.
- 2 Intervocalic voicing of [go] facilitates acceptance of diminutive /ko/ adjacency.
 - This supports the literature: Abe (1999).
 - Abe (1999) argues the effects of intervocalic voicing on suffixation.
- 3 Regional and generational differences were not confirmed.
- 4 Proposal: -*kko* suffixation is **opaque**.
 - **Opacity:** suffixation undergoes iff intervocalic voicing is adopted.
 - **UR:** /ko/
 - A serial derivation can account for it.

References

- Abe, T. (1999). Descriptive study on semantics of diminutive /-ko/: evidence from Tsugaru and Hirosaki dialect of Aomori prefecture [sishooji /-ko/ ni kansuru imi-kijutu: Aomori-ken tugaru-tihou, hirosakihougen o rei ni]. *Research note of sociolinguistics, Osaka University*, 1, 16–24.
- Bürkner, P.-C. (2017). *brms: An R Package for Bayesian Multilevel Models Using Stan*. (Pages: 1 Publication Title: J. Stat. Softw. Volume: 80)
- De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1–12.
- Huang, C., Matsui, S., Watabe, N., Noguchi, H., Hashimoto, A., Mizoguchi, A., & Kitahara, M. (2023). Neutralization and secondary acoustic cues of voicing contrast: A Tohoku and Tokyo Japanese production experiment. In *Proceedings of ICPHS 2023*.
- Inoue, H. (1968). Consonant System of the Tōhoku Dialect [Tōhoku-hougen no siin taikai]. *GENGO KENKYU (Journal of the Linguistic Society of Japan)*, 52, 80–98.
- Kawahara, S., & Sano, S.-i. (2014). Identity avoidance and Lyman's law. *Lingua*, 150, 71–77.
- Kushibiki, Y. (2014). Research on diminutives of "ame-kko", "ame-tyan", "Bombonica" ["amekko", "ametyan", "Bonbonica" no sishooji]. *Asia gakka nempou*, 8, 28–35.
- Lemoine, N. P. (2019). Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, 128(7), 912–928.
- Makowski, D., Ben-Shachar, M. S., Chen, S. A., & Lüdtke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in psychology*, 10, 2767.
- R Core Team. (2025). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, 21(2), 301–308.
- Sano, Shinichiro. (2013). Violable and Inviolable OCP Effects on Linguistic Changes: Evidence from Verbal Inflections in Japanese. *MIT Working Papers in Linguistics*, 66, 145–156.
- Sato, Y. (1989). Rules of accent and rendaku in compounds [hukugou-go ni okru akusento kisoku to rendaku kisoku]. *Lecture on Japanese and education of Japanese language, phenetics and phonology of Japanese (upper volume)*, 2, 233–265.
- Yip, M. (1998). *Identity avoidance in phonology and morphology*. Retrieved 2025-11-25, from <http://roa.rutgers.edu/files/82-0000/roa-82-yip-3.pdf.gz>

Thank you.

Acknowledgement

This research was supported by JSPS KAKENHI Grant Number JP25K00460.

Contact: myakuwa2005@gmail.com