

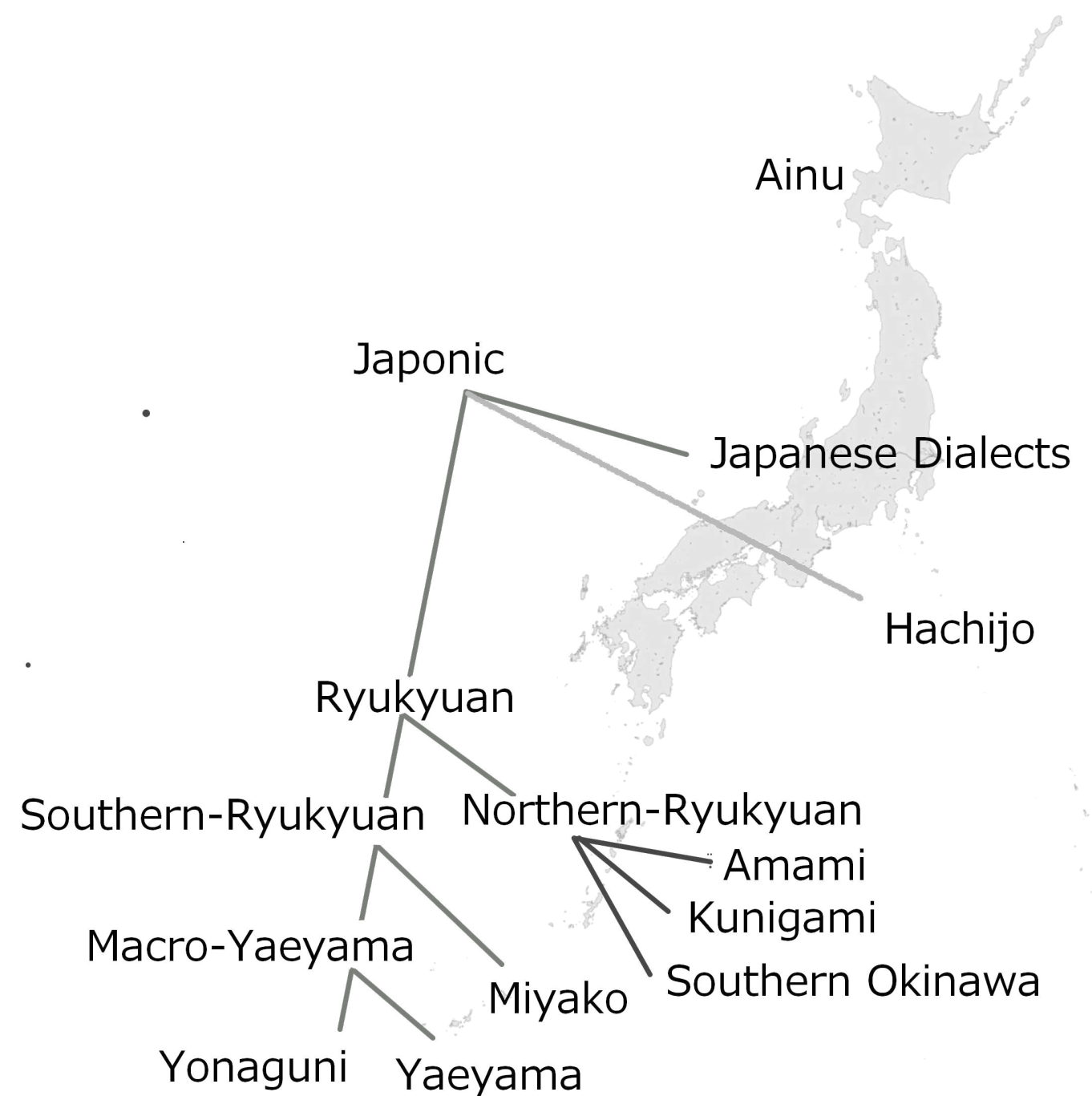
Archiving System of Endangered Languages in

Japan: A Preliminary Report

Natsuko Nakagawa, Masahiro Yamada, Kenan Celik,
Nobuko Kibe, & Yukinori Takubo

National Institute for Japanese Language and Linguistics
nakagawanatuko@ninjal.ac.jp

ENDANGERED LANGUAGES IN JAPAN



CURRENT RESOURCES

Corpus of Japanese Dialects (Kibe et al., to appear)

- 75-hour recordings from 25 places collected in 1977-1985 (to be published before 2021)
- **Searchable online both in Standard Japanese and Dialects**

Endangered Language Database

- **Online database of vocabularies and discourse**
<http://kikigengo.ninjal.ac.jp>
- 14 places across Japan (mainly Ryukyu and Hachijo)
- Approximately 9000 words & expressions with recordings

Shima Shotai Font (Ogawa et al., 2019)

- Sounds that cannot be written in Standard Japanese writing system
- **All Ryukyuan varieties are written under unified system**
- Kana-based system (phonogramic) instead of kanji-based (ideogramic)
- Makes it easier to type on PCs
- Vertical writing also possible
- Unicode registration as future task

いちぐわーちゃんや くぬいん ‘うーいーなーてい、 はてーうちだれん。
うっちい んちゃー やーらく なーちか くぬいんぬ ない ‘うーえーゆい。
ねんなん はるとう あき たーけえい ‘うーいーれゆんど。

つかいぬ みどうむぬ いーぬ ぬくりーりていーどう、…。
かぬ、 すでいみずうー かたみー、 ていんから たすき、 えー
ぼーや んめー、 んめ (ばなす)。
すでいりていーぬ

ARCHIVING & EXHIBITION SPACE (ON-GOING)

Language archiving system

- Building archiving system for researchers & community members
- Japanese-mediated (makes it possible for almost all citizens in Japan to access the system)
- **Easily depositable & immediately visible/usable**

Language museum

- **Incremental update & publish** to motivate depositors to keep depositing data
- Exhibition with coherent stories (Takubo, 2017)

FOR & WITH COMMUNITIES

Picture books w/ audio, video, & text

- Picture books for children
- The parent generation (**passive speakers**) are supposed to read for them and **practice the language**
- Audio & video available for practice

Dictionary

- Traditional: Compiled for 20+ years by a single linguist
- New: **Collected & published immediately & incrementally** by community members & linguists

Interactive ILG-ed transcript (Karaoke)

- Video or sound with transcription
- The corresponding part is played by clicking the transcription

Generation gaps

- Older gen.: knowledgeable in **language**, less in technology
- Younger gen.: knowledgeable in **technology**, in about language

REFERENCES

- Kibe, N., Sato, K., Nakanishi, T., & Nakazawa, K. (to appear) Corpus-based study of Japanese dialects: Regional differences in accusative case marking system. *Proceedings of Methods in Dialectology XVI*.
- Ogawa, S., Yamada, M., Hayashi, Y., & Ueda, H. (2019) A typeface for endangered languages, ATyPl2019 (<https://shimanomoji.site/atypi.html>).
- Takubo, Y. (2017) The Digital Museum Project for the Documentation of Endangered Languages: The Case of Ikema Ryukyuan. Vovin, A. & McClure, W. (Eds.) *Studies in Japanese and Korean Historical and Theoretical Linguistics and Beyond*, 3–12, Leiden: Brill.