

Acknowledgements

This study has been conducted under Prof. Shinobu Imai, as part of the Japanese Studies Program at Osaka University, funded by MEXT Japanese Government Scholarship.

Phrasal compounds in Japanese

Japanese phrasal compounds sometimes allow for structures where only part of the compound is modified.

- (1) [titi-no haka] -mairi father-GEN grave visiting 'visiting father's grave'
- (2) *[furui haka] -mairi old grave visiting 'visiting an old grave'
- (3) [kiree-na mati] -dukuri clean-COP town making 'construction of a clean town'
- (4)*[kyodai-na mati] -dukuri giant-COP town making 'construction of a giant town'

[3]

[7]

What licenses this?

- (1) & (2) → left-hand element of compound is a relational noun [3] [5] [7]
- (3) & (4) → (un)grammaticality of (4) contested [3] [5]
- left-hand element of the compound in (3) is a part of a **cliché** [7]
- certain verbs are semantically "transparent" in compounds [3]

The Present Study: investigates the occurrence of constructs similar to (3) and (4) to explore their acceptability as well as licensing conditions.

What is a cliché?

"memorized **units** [that] have some special situational or contextual specification; for instance, *whistle while you work* means literally what it says, but at the same time it evokes Disneyesque mindless cheerfulness through its connection with the movie *Snow White*." [4]

- (5) "whistle while you work"
- "is lexicalized or in some degree institutionalised" [2]
- (6) "May the best man win"

allows 'context-dependent innovations' [1]

- (7) "to Houdini one's way out of a closet"
- (8) Baroque flautist (a player of a Baroque flute), wooden flautist (!= a player of a wooden flute)

How productive are phrasal compounds with modifier stranding?

1. Are they always licensed by clichés?

- (8) [oto-de Koobe-o annai-suru yuniiku-na teepu] -dukuri sound-with Kobe-ACC guide-do unique-COP tape making 'creating a unique tape that guides people around Kobe through audio'

[5]

2. Are they more productive with certain verbs?

- (9) mizu -yari water giving 'watering'
- (10) *kinoo [tumetai mizu]-yari-o sita yesterday cold water giving-ACC do.PST 'I watered the plants with cold water yesterday.'

[3]

Hypothesis I: If clichés are taken to be lexicalized units, we would expect to see a lot of the same modifier-noun pairs in the data.

Hypothesis II: If clichés are taken to be units with high frequency, we would expect the frequency of the [Mod N] to correlate with its occurrence in a modifier stranding context.

Methods

A corpus study of Balanced Corpus of Contemporary Written Japanese [6] was conducted. The corpus contains 100 million words of written Japanese data from 1975-2005 and consists of three subcorpora. The following searches were performed in the corpus using online tool Chunagon [8].

Search key

X [Noun] Y constructs where

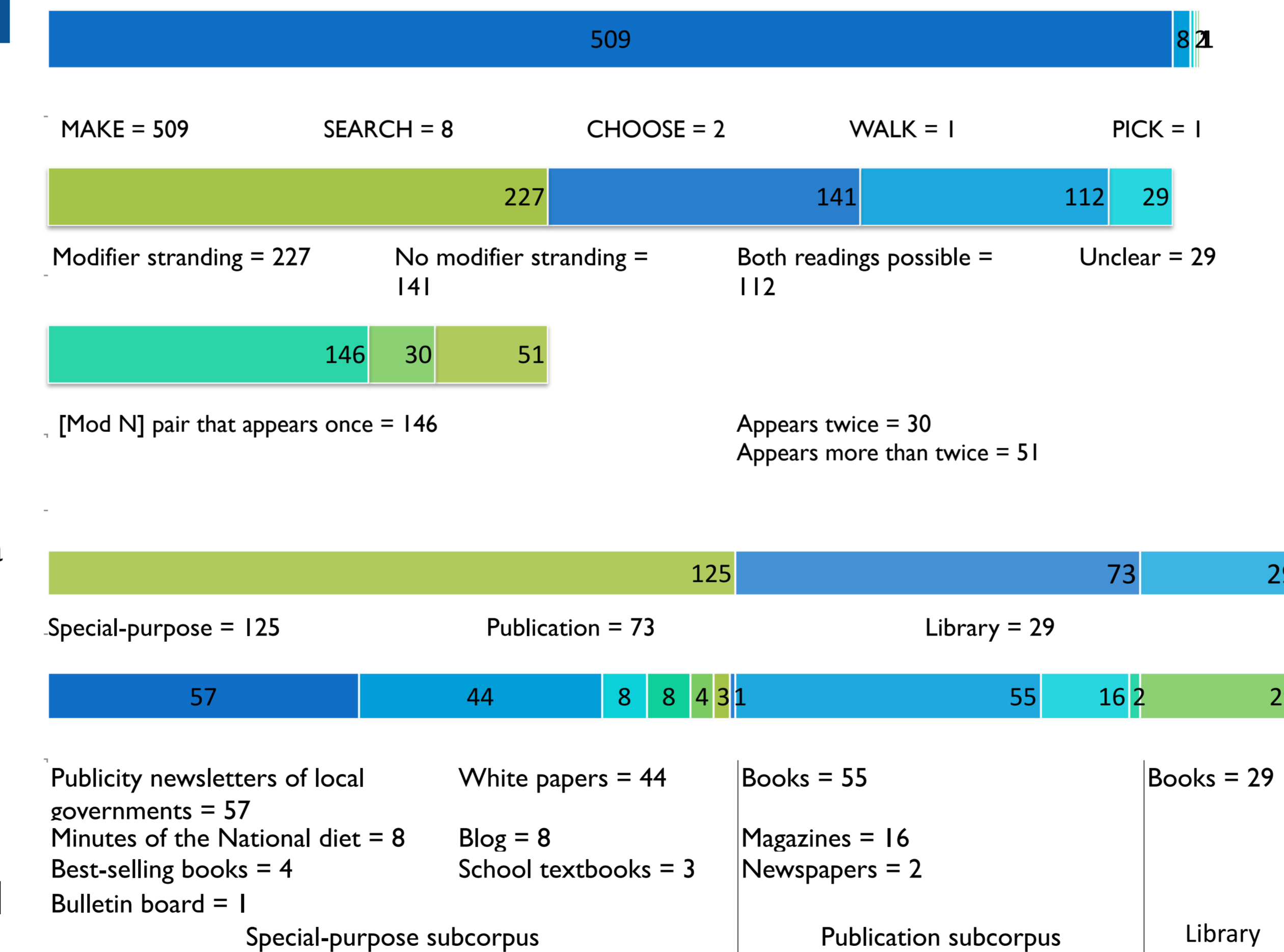
X = adjective OR adjectival noun + copula -NA

Y = nominalized forms of verbs *tukuri, sagasi, tori, aruki, erabi* (make, search, take, walk, choose)

Data

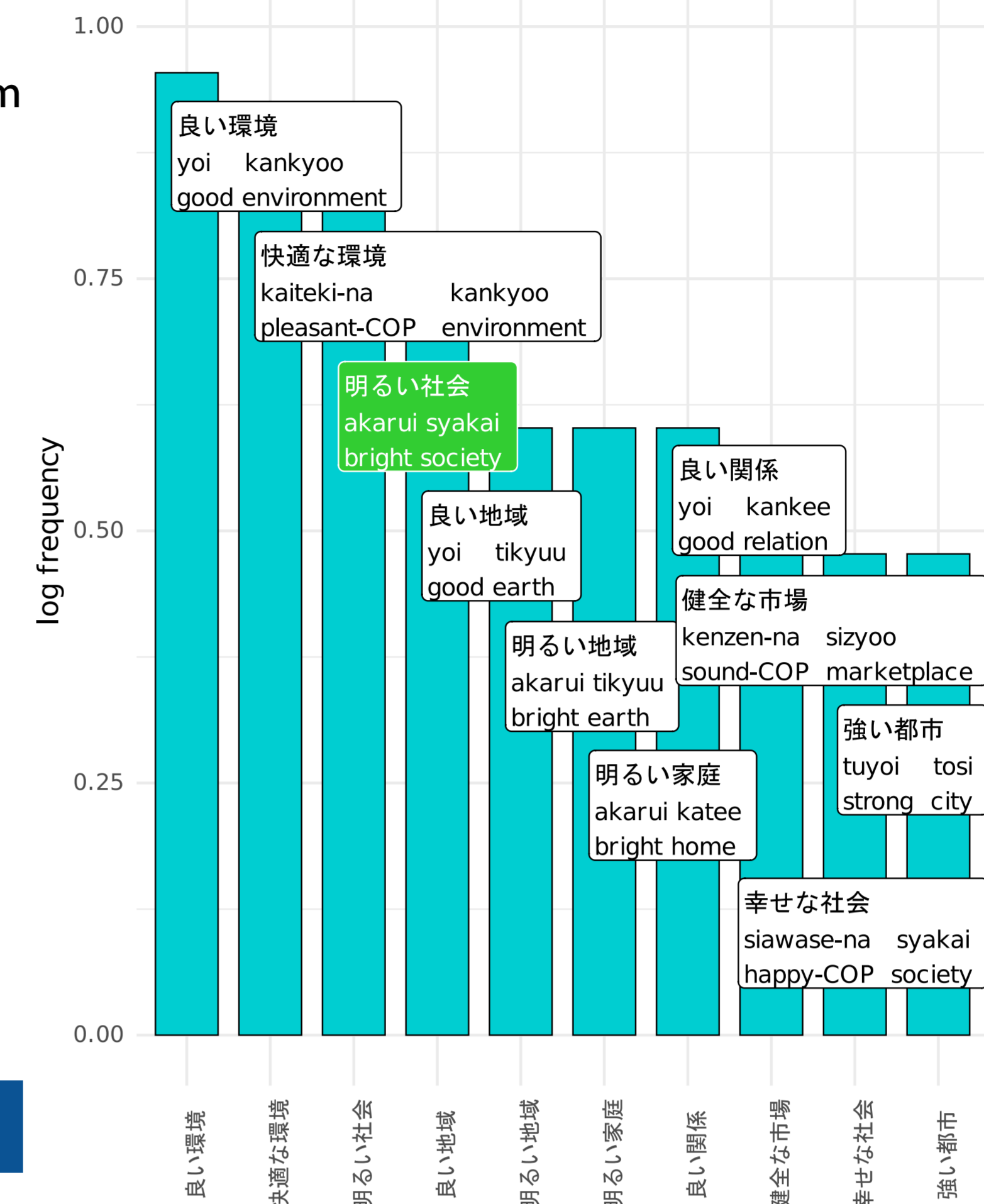
Results for key were marked by two Japanese linguists on whether adjectives modified the entire compound, only the left-hand element or if both readings were possible. Only the cases where unambiguous modifier stranding was observed were considered. An additional search was performed to extract the frequency of [Mod Noun] pairs that were found in modifier stranding in the overall corpus data.

Results

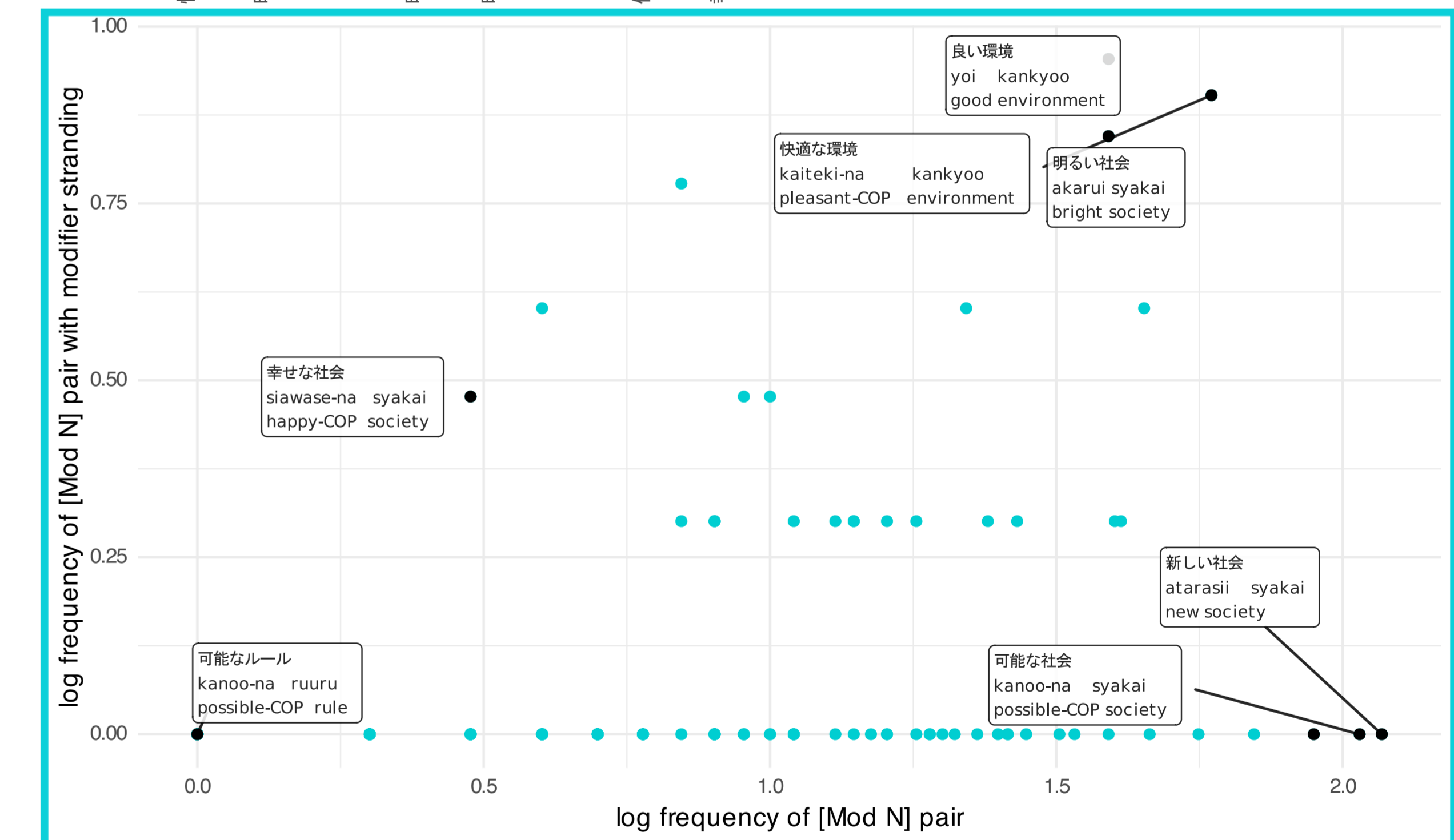


- (11) 二十一世紀を展望した地域社会の発展の核となる新たな都市づくり
[nizyuuisseeki-o tenboo-sita tikyuu-syakai-no hatten-no kaku 21st century-ACC view-do.PST world-society-GEN development-GEN core tonaru arata-na tosi] -dukuri become new-COP town making 'construction of a new town that would be the core of establishing a global community that has a 21st century point of view'

Results



- Manually marked data was processed using RStudio and package ggplot2 for data visualization [9].
- Data on verbs other than 'tukuri' (make) were omitted due to small number of samples available.
- Frequency of [Mod N] pairs were log-normalized.
- Of all the [Mod N] pairs, only one appears to clearly conform to a common definition of a cliché (given in green).



Discussion

- High ratio of cases of modifier stranding to phrasal compounds with a modifier suggest pattern is productive.
- Variety in Mod N patterns could be an indicator against the cliché account of licensing.
- Modifier stranding seems to appear mostly in contexts where it might be desirable to convey information in a compact way.
- Frequency of [Mod N] does not seem to meaningfully correlate with the frequency of [Mod N] in a modifier stranding context, suggesting a frequency-based account of modifier stranding would not be sufficient to predict the pattern.

References

[1] Bresnan, J. & Mchombo, S. A. "The Lexical Integrity Principle: Evidence from Bantu." *Natural Language & Linguistic Theory*, Vol. 13. Springer, 1995. 181-254 [2] Carstairs-McCarthy, A. *An Introduction to English Morphology*. Edinburgh University Press, 2002. [3] Imai, S. "Ku no housetu ni kansuru imironteki kousatsu", *Nihongo Nihonbunka* 24. Osaka University, 1998. [4] Jackendoff, Ray S. *The Architecture of the Language Faculty*. MIT Press, 1997. [5] Kageyama, T. *Bunpou to gokeisei*. (Hitsuji Books), 1993. [6] Kikuo Maekawa, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den. "Balanced corpus of contemporary written Japanese." *Language Resources and Evaluation* 48 (2), pp.345-371 (DOI 10.1007/s10579-013-9261-0), 2014.06. [7] Nishiyama, K. "Phrasal compounds in Japanese." *Carola Trips & Jaklin Kornfilt* (eds.), *Further investigations into the nature of phrasal compounding* (Morphological Investigations 1). Language Science Press, 2017. 149-183 [8] Teruaki Oka, Yuichi Ishimoto, Yutaka Yagi, Takenori Nakamura, Masayuki Asahara, Kikuo Maekawa, Toshinobu Ogiso, Hanae Koiso, Kumiko Sakoda, and Nobuko Kibe. 2020. KOTONOHA: A Corpus Concordance System for Skewer-Searching NINJAL Corpora. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 7077-7083, Marseille, France. European Language Resources Association. [9] Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.