# Lyman's Law and Limits on "Leaky" Phonotactics

Blake H. Allen

2013/04/26

# Topic

- Martin (2011) has found that phonotactic patterns strictly upheld in narrow morphological domains are also gradiently true in larger domains.
- E.g. Navajo root-affix sibilant harmony: such harmony is observed at above chance in root-root compounds.
- I demonstrate that this phonotactic "leak" is absent for the Japanese voicing restriction called Lyman's Law.
- This finding provides insight into the structural limits of the phonotactic leak phenomenon, as well the mechanism underlying phonotactic learning more generally.

# Presentation structure

1. Phonotactic "leak"
2. Lyman's Law
3. Simulation and results
4. Discussion and implications

# Summary of Martin (2011): Navajo

- Navajo sibilant harmony: within *roots* and in *root-affix* complexes.
  /ji-**s**-léé**ʒ**/ → [ji-**ʃ**-tɬéé**ʒ**]
  /ji-**s**-ti**z**/ → [ji-**s**-ti**z**]

- This restriction is **not** categorically true in *root-root* compounds:
  **tʃ**éí - **ts'**iin 'rib cage'
  **ts$^{h}$**é - **tʃ**éé? 'amber'

- Martin found, however, that harmony-violating root-root compounds were found at less than chance in Young & Morgan's (1987) dictionary of Navajo.

- His test procedure, a version of the Monte Carlo method, will be explained later in this talk.

# Summary of Martin (2011): English

- No *morpheme-internal* geminate consonants exist in English.

  *[hæ**pː**i]
  *[boʊ**zː**oʊ]

- Some morphologically derived geminates do occur, however, in *root-root* compounds like *bookcase*, *unknown*, and *solely*.

- Yet these words are also underrepresented in the CELEX English lexicon (Burnage et al. 1990 ), according to Martin's (2011) tests.

# Lyman's Law & Rendaku

- *Lyman's Law* refers to the principle that Yamato (native Japanese) morphemes in Standard Tokyo Japanese can have maximally one voiced obstruent (Ito & Mester 1993).
- This restriction manifests itself most obviously by restricting *rendaku* ("sequential voicing").

*Standard rendaku:*

| | | | | |
|---|---|---|---|---|
| te | + | **k**ami | → | te**g**ami |
| 'hand' | | 'paper' | | 'letter' |
| tokoro | + | **t**okoro | → | tokoro**d**okoro |
| 'place' | | 'place' | | 'here and there' |
| hi | + | **s**a**ɕ**i | → | hi**z**a**ɕ**i |
| 'sun' | | 'shining on' | | 'sunlight' |

(All Japanese data are from Rosen's (2003) corpus of compounds.)

# Examples of Lyman's Law

- Lyman's Law: a voiced obstruent in the second member of the compound blocks rendaku:

| | | | | | |
|---|---|---|---|---|---|
| yama<br>'mountain' | + | **ka**d**ʑ**i<br>'fire' | → | yama**ka**d**ʑ**i<br>'forest fire' | (*yamagadʑi) |
| sumi<br>'ink' | + | **ts**u**b**o<br>'jar' | → | sumi**ts**u**b**o<br>'ink bottle' | (*sumidzubo) |
| yagi<br>'goat' | + | **h**i**g**e<br>'beard' | → | yagi**h**i**g**e<br>'goatee' | (*yagibige) |

- Keep in mind that this restriction also holds for Yamato morphemes outside the rendaku environment.

# Interim Summary

- Navajo: non-local *root/root-affix* restriction $\rightsquigarrow$ *root-root*
  (sibilant harmony, long-distance)
- English: local *root* restriction $\rightsquigarrow$ *root-root*
  (dissimilation of adjacent consonants)
- What about Lyman's Law?
- Japanese: non-local *root* restriction $\rightsquigarrow$ *root-root*??
  (voicing dissimilation, long-distance)

# Monte Carlo method

- To determine whether Lyman's Law-violating root-root compounds are underrepresented in Japanese, I followed Martin's (2011) example and used the *Monte Carlo method*:

- For some large number of iterations (10,000):
  1. Put all of the observed first members of compound words into a list and randomize its order; do the same with all of the observed second members of compounds.
  2. Combine the two randomized lists pairwise into new compounds.
  3. Record the number of compounds that violate the phonotactic restriction in question.

- The mean result of all of these iterations is then compared against the observed count of compounds violating the phonotactic restriction.

# Monte Carlo method for Japanese

- For Japanese, I used a total of 90 compounds, comparable to the 97 Navajo words used by Martin (2011).

- These items were selected for having exactly two obstruents: one the last consonant of the first member of the compound, and the other the initial consonant of the second member. Again, this is equivalent to Martin's procedure.

| | |
|---|---|
| yo**k**o-**b**ue | 'transverse flute' |
| mi**s**o-**ʑ**iru | 'miso soup' |
| mi**z**u-**g**ame | 'water jug' |
| ... | |

- The observed count of such compounds violating "root-root Lyman's Law" (i.e. ...[+voi,+obs]V-[+voi,+obs]...) is **25** out of 90.

# Results

- The number of violations predicted by the Monte Carlo simulation is **25.6483**.
- This result is not significantly different from the observed count of 25, as verified by a Chi-squared test (p=0.6715).
- Therefore: there is no "leak" of Lyman's Law into the root-root domain.

# Interpreting the result

- Why would phonotactic leak occur in Navajo and English but not in Japanese?
- I propose two possible causes:
  - phonotactic leak is restricted in its phonological and/or morphological domains
  - rendaku itself is blocking the "leak"

# Martin's (2011) learning account

- Martin (2011) argues that the patterns found in Navajo and English can be attributed to properties of the phonotactic learning mechanism.
- A constraint weighting algorithm like that in Hayes & Wilson (2008) will "spread out" responsibility for observed patterns among all relevant constraints.
- If there exist both *structure-sensitive* and *structure-blind* constraints, then even patterns found only in particular structures/domains will cause its respective structure-blind constraint to increase in weight.

# Predictions of Martin's account

- In Martin's account, the only relevant distinction is between domain-specific and domain-general constraints.
- This system predicts that a phonotactic restriction in *any* domain will also be gradiently observed in *every* other domain.
- But to account for the lack of phonotactic leak in Japanese, it may be necessary to posit limitations on this phonotactic generalization mechanism.

# Domain restrictions

- The cases of Navajo, English, and Japanese are non-overlapping in pattern locality and source domain:

| | Navajo (✓) | English (✓) | Japanese (X) |
|---|---|---|---|
| *locality* | **non-local** (harmony) | local (dissimilation) | **non-local** (disharmony) |
| *source domain* | root-affix | **root-internal** | **root-internal** |

- Local phonological processes appear to be learned differently from non-local ones (Pycha et al. 2003, McMullin 2013), and may require greater representational complexity (Hayes & Wilson 2008).

- Moreover, among long-distance restrictions, root-internal ones are most common, followed by root-affix, and finally root-root (Hansson 2010). We might expect that this asymmetry is also related to the learning mechanism.

# Domain restrictions

| | Navajo (✓) | English (✓) | Japanese (X) |
|---|---|---|---|
| *locality* | **non-local** | local | **non-local** |
| *source domain* | root-affix | **root-internal** | **root-internal** |

- One explanation of the lack of phonotactic leak in Japanese is that generalizing Lyman's Law to root-root compounds would require **both**:
  - sensitivity to **non-local** phonotactics (unlike English)
  - generalization from **root-internal** phonotactics to root-root compounds, bypassing the intermediate level of root-affix phonotactic restrictions (unlike Navajo).

- This account assumes—perhaps unjustly—that:
  - local phonotactics are learned in a qualitatively different way from non-local phonotactics
  - rather than a distinction between domain-general and domain-specific constraints, phonotactic constraints refer to any level of constituency.
    (e.g. $*x\ldots x_{\text{morpheme}}$, $*x\ldots x_{\text{root-affix}}$, $*x\ldots x_{\text{root-root}}$)

# Alternative: influence from rendaku

- The domain-based account makes strong predictions, but there is no clear way to implement it in current constraint-based phonological models.
- But apart from these differences in locality and source domain between Navajo/English and Japanese, Lyman's Law is not necessarily comparable to either Navajo sibilant harmony or the English ban on morpheme-internal geminates.
- This is because of the prevalence of **rendaku**.

# The role of rendaku

*Rendaku:*

| | | | | |
|---|---|---|---|---|
| te | + | **k**ami | → | te**g**ami |
| 'hand' | | 'paper | | 'letter' |

- Rendaku (voicing of the first consonant of the second member of a compound) occurs in the majority of root-root compounds in Japanese—about 64% of eligible compounds, according to Rosen (2003).
- Unlike in Navajo and English, the tendency toward rendaku is in *conflict* with any possible dispreference for *...[+voi,+obs]V-[+voi,+obs]...* compounds.
- Because the constraint(s) that cause rendaku are certainly ranked/weighted higher than any marginal root-root disharmony-promoting constraint(s), the pressure toward rendaku nullifies the phonotactic "leak" of Lyman's Law.

# Conclusion

- Lyman's Law in Japanese does not exhibit phonotactic "leak" as do some Navajo and English phonotactics.
- Potential analyses attribute this difference to either
  1. the differences in locality and source domain between Navajo/English and Japanese
  2. the counterbalancing influence of rendaku
- The only way to resolve this explanatory dilemma is to *check for phonotactic leak in other languages*, in order to form a typology of this phenomenon.
- The analyses proposed here make predictions that help guide our choice of candidate languages for this typological survey.

# Thanks!


- Special thanks to Gunnar Hansson for introducing me to this project and discussing it with me.
- And thanks to Eric Rosen for sharing his corpus of Japanese compounds.
- Thanks as well to members of the UBC phonology group for their feedback on an earlier version.

# References

Burnage, Gavin, Rolf Harald Baayen, R Piepenbrock, & H van Rijn. 1990. CELEX: A guide for users.

Hansson, G.O. 2010. Consonant Harmony: Long-Distance Interactions in Phonology, volume 145. Univ of California Press.

Ito, J., & R.A. Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. Linguistic Inquiry 49–73.

Ito, Junko, & Armin Mester. 1993. Japanese phonology: constraint domains and struc- ture preservation. A handbook of phonological theory. Blackwell Handbooks in Linguistics Series.

Martin, A. 2011. Grammars leak: Modeling how phonotactic generalizations interact within the grammar. Language 87.751–770.

McMullin, Kevin J. 2013. Learning consonant harmony in artificial languages: Locality. Ph.D. Qualifying Paper to appear in UBC Working Papers in Linguistics.

Pycha, Anne, Pawel Nowak, Eurie Shin, & Ryan Shosted. 2003. Phonological rule-learning and its implications for a theory of vowel harmony. In Proceedings of the 22nd West Coast Conference on Formal Linguistics, volume 22, 101–114. Somerville, MA: Cascadilla Press.

Rosen, E. 2003. Systematic irregularity in Japanese rendaku: How the grammar mediates patterned lexical exceptions. The Canadian Journal of Linguistics/La revue canadienne de linguistique 48.1–37.

Young, Robert W, & William Morgan. 1987. The Navajo language: A grammar and colloquial dictionary. University of New Mexico Press Albuquerque.