Rethinking Lexical Indexing Models: Evidence from Japanese and English

Ryan Hearn — Cornell University

The 91st Annual Meeting of the Linguistic Society of America, January 2017

The *Core-Periphery* (*CP*) Model

- Itô & Mester (2004, "I&M") divide modern Japanese lexical items into four distinct strata.
- The strata are characterized by surface adherence to a different number of the following constraints:

a. **SyllStruc**: Prevents complex onsets and codas
b. **No-DD**: No voiced obstruent geminates
c. **No-P**: Prevents nongeminate/singleton [p]
d. **No-NT**: Postnasal obstruents must be voiced

- Hierarchical behavior of the four strata, from I&M (2004: 557).

	SyllStruc	No-DD	No-P	No-NT
Yamato	✓	✓	✓	✓
Sino-Japanese (SJ)	✓	✓	✓	violated
Assimilated Foreign	✓	✓	violated	violated
Unassimilated Foreign	✓	violated	violated	violated

Origin and Behavior of Lexical Strata

- Lexical items that were borrowed during the same era of a language's history show similar phonological and phonotactic behavior.
- Kiparsky (1973): Differences between strata are gradual and hierarchical.
  - Older strata show heavier phonotactic restrictions than younger strata.

Same Underlying, Different Surface

- I&M posit multiple synchronic strata with same overall ranking of markedness constraints.
  - Each stratum defined by a separate ranking of Faith.
- Below, two separate lexical items with identical underlying forms, /pan/, result in two different surface forms since they belong to different strata.

/pan/	No-DD	Faith/ Assim	No-P	Faith/ SJ	No-NT
'bread' ☞[pan]			*		
Assimilated Foreign [han]		*!			
'group' [pan]			*!		
Sino-Japanese SJ ☞[han]				*	

- Assimilated stratum: Faith ≫ No-P ⇒ surface form [pan].
- SJ stratum: No-P ≫ Faith ⇒ surface form [han].

Conclusions

- **Alternations accounted for by indexing models are often lexical residue of earlier constraint rankings.**
- **Allowing underlying forms of lexical items to update in response to sound change eliminates the need for multiple synchronic constraint rankings.**
- **Stratal generalizations are descriptively and historically interesting, but lexical items that fall through the cracks are expected as analogy and other processes subsequently affect individual lexical items.**

Exceptions to the *CP* Model

1. Exceptions to No-NT

- Yamato *intiki* 'trickery' and *anta* 'you' violate No-NT.
  - *Anta* from *anata* via syncope, **moved from core toward periphery**.
- I&M: exceptions "undoubtedly native, but peripheral" (1995).
  - No explanation for movement toward periphery.

2. Exceptions to No-P

- Classifier *pun* 'minute' combines with numbers.
- Member of the SJ stratum, should obey No-P constraint.
- However, paradigm has leveled in fluent speech.
  - **Moved from core toward periphery of lexicon**.

Expected Compound Pronunciation	**New Compound Pronunciation**	Number Morpheme	Counter Morpheme	Meaning
ip-pun	**ip-pun**	ichi 'one'	pun	'one minute'
ni-ɸun	**ni-pun**	ni 'two'	pun	'two minutes'
san-bun	**san-pun**	san 'three'	pun	'three minutes'

- Also, why would the SJ *pun* become *bun* at all?

3. Exceptions to No-DD

- I&M's proposed adaptations for voiced obstruent-final English borrowings:
  - Assimilated Stratum → geminate voiceless obstruent.
  - Unassimilated Stratum → geminate voiced obstruent.
- Too simplistic: actually *five* different adaptation mechanisms.
- Crawford (2009): voiced geminate borrowings were most popular adaptation mechanism in oldest attestations.

What the Exceptions Tell Us

- The *Core-Periphery* model tends toward **overgenalization**, and fails when grammatical processes affect **individual lexical items** through phonological processes (syncope) or lexical processes (analogy).
- Why not just claim that these items switch strata? There is **no motivation** for them to switch strata, especially in the case of *pun*, when the nearly identical *pon* does not.

My Proposal: Lexical Update

- **A more traditional view of OT:**
  - Only one constraint ranking for all synchronic behavior.
- I&M's multiple Faith rankings correspond to the constraint rankings of different eras in a language's history.
- Hierarchical nature of strata is due to the gradual nature of sound change.
- Once constraint reranking occurs, new speakers no longer generate these forms using productive phonology.
  - They must separately store each alternation previously generated by the old constraint ranking in the lexicon.
- This **lexical update** process moves the alternation from the synchronic phonology to the mental lexicon.
- These forms can now be modified by grammatical processes without violating highly-ranked synchronic constraints.

Why the Lexicon?

- We not only allow for, but **motivate** the analogy of *pun*.
- Kiparsky (2012): "analogical change is grammar optimization, elimination of unmotivated grammatical complexity or idiosyncrasy" (p. 21)
- **h/p alternation is not motivated synchronically** .

Different Underlying, Different Surface

- Derivations for stratal data like *pan*/*han* become trivial.
- Underlying form of *han* 'group' was stored as /han/ once the reranking occurred that later allowed *pan* 'bread' to be borrowed as-is.
- When they coexisted, they had different underlying forms.

	SyllStruc	Faith	No-P
/pan/ 'bread' ☞[pan]			*
Assimilated Foreign [han]		*!	
/han/ 'group' [pan]		*!	
Sino-Japanese ☞[han]			*

Capturing Generalizations: A Case Study of English Fricatives

I. Fricative Distribution

- English voiceless fricatives contrast with voiced.
- In Old English these pairs were allophones.
  - Voiceless initially/finally, and medially voiced or geminate.
- By 1400s, these changes generated modern distribution:
  1. **Final** [ə] **deleted**, making finals phonemic.
  2. **Initial** [v] **borrowed**, making initials phonemic.
  3. **Prosodic voicing of initial** [ð] in e.g. *the*.
  4. **Geminates degeminated**.

II. Stratal Division

- Two strata are important for our purposes.
  - *Native* words: all words pre-initial/final voiced fricatives.
  - *Loan* words: all words post-initial/final voiced fricatives.
    - **Importantly**, very early Latin loans are *Native* as they show devoicing, e.g. *fan* from *vannus*.
  - Later Latin *Loan* borrowings preserved initial voicing.

III. Capturing Generalizations

- 1-4 above conspired to:
  1. Allow initial and final voiced fricatives.
  2. Degeminate geminates.
- These generalizations easily capturable by reranking Faith.
- This should be ideal to differentiate strata in *CP* model.
  - But, *Native* words like *the* are problematic.
  - Requires either ad hoc constraint just for these (unsatisfying) or ad hoc movement to *Loan* stratum (unmotivated).
- **Instead, I argue once the phonological system allowed initial voiced fricatives, prosodic processes were then free to voice *the*.**
- Without being constrained to a voiceless fricative stratum, *the* is free to develop initial voicing.

Future Considerations

- Line between phonology and lexicon must be better defined.
- We still lack a thorough understanding of analogy.

Selected References

Bermúdez-Otero, Ricardo. 2006. Phonological change in optimality theory. *Encyclopedia of language and linguistics*, **9**, 497–505.

Crawford, Clifford James. 2009. *Adaptation and transmission in Japanese loanword phonology*. Ph.D. thesis, Cornell University.

Itô, Junko, & Mester, Armin. 2004. The Phonological Lexicon. *Pages 552–564 of:* McCarthy, John J (ed), *Optimality Theory in Phonology: A Reader*. Wiley-Blackwell.

Kiparsky, Paul. 2012. Grammaticalization as optimization. *Grammatical change: Origins, nature, outcomes*, 15–51.

Contact: rdh238@cornell.edu | conf.ling.cornell.edu/ryanhearn/

Special thanks to Abby Cohn, John Whitman, and Michael Weiss, and to my colleagues in the CLC for their invaluable aid and advice.