Looking under the hood: Korean palatalization and derived-environment effects

Adam J. Chong
University of California, Los Angeles
ajchong@ucla.edu

1. Introduction

- Korean palatalization (e.g. Kiparsky, 1973, 1993, Iverson & Wheeler, 1988, Cho, 2001) – see Appendix A1 for historical development

(1) /t, t^h/ → [c, c^h] / __+[i, j] (*Ti)
  a. /ma**t**-i/ → [ma**c**i] 'eldest-NOM'
  b. /pa**t^h**-i/ → [pa**c^h**i] 'field-NOM'

- Palatalization is a productive process across morpheme boundaries, but does not occur within stems:

(2) /ti/ and /t^hi/ (as well as /tj/ and /t^hj/) surface faithfully within stems:
  a. /ma**t**i/ → [ma**t**i] 'knot, joint'
  b. /**t^h**im/ → [**t^h**im] 'team'

- Korean palatalization: *Morphologically Derived-Environment Effect (MDEE)*
    - Phonological processes that apply only when their conditions (environment) are met by the concatenation of two different morphemes (a.k.a. non-derived environment blocking or NDEB).
- Proven a thorny problem for phonological analyses (not reviewed here today; see Inkelas 2015 for a recent review and proposal)
- Static generalizations about the lexicon (in terms of phonotactics) mismatch with dynamic generalization that motivates alternation (Paster 2013)

- Traditional analytic assumptions:

(3) *Derived environment condition* (Revised Alternation Condition; Kiparsky 1982): Obligatory neutralization rules apply only in derived environments
  a. Trigger and target belong to separate morphemes
  b. Necessary and *sufficient* condition for alternation to apply

(4) *Static phonotactic patterns are "productive"*:
  a. Unrepaired sequences within stems (such as [ti] in Korean) are not dispreferred – sequences should be perfectly acceptable!

2. Goals for today

1. What are the static patterns in Korean as it relates to palatalization?
    - ❖ Caveat: Not going to propose a new analysis of these patterns. Taking a step-back and asking: what are the actual empirical patterns in the data?
2. How does this figure in phonotactic learning?
    a. Can you get to a markedness constraint that motivates alternations from pure phonotactic learning in Korean?
3. Implications:
    a. How does this compare to other MDEEs – very brief look at Turkish?
    b. Relation between static phonotactic and dynamic generalizations about alternations

3. Korean palatalization in more detail

3.1 Corpus investigations: NAKL

- National Academy of Korean Language corpus (NAKL; 2003)
- 53,196 commonly used Korean words (with frequencies from different print sources – e.g. newspapers)
- Includes native, Sino-Korean and loanwords (helpfully tagged)
- Initial pre-processing using Kim et al. (2002) algorithm
    - o Splits up each syllabary into constituent jamos

How many words in the entire NAKL contain coronal stop-[i, j] sequences?

CV type	Entire Lex.	Native	Sino-Kor.	Loans
[ti]	208	68	5	135
[t^hi]	167	30	4	133
[tti]	32	28	4	0
[tjV]	14	5	0	9
[t^hjV]	15	4	4	7
[ttjV]	0	0	0	0
Total	**436**	**135**	**17**	**284**

Table 1. No. of lexical items that contain /Ti/ and /Tj/ in NAKL corpus/sub-corpora

- Out of > 50,000 words in NAKL, only 436 words contain /Ti/ or /Tj/ sequences in them and ~65% of these are loanwords.

Are these sequences statistically under-represented?

- Number of times each CV sequence occurs in the corpus: **454**

	/i/, /jV/	other Vs
/t, t^h, tt/	454 (E: 5798) **O/E[1] = 0.08**	27433 (E: 22089) O/E = 1.24
other Cs	31247 (E: 25903) O/E = 1.21	93328 (E: 98672) O/E = 0.95

Table 2. Observed/Expected counts of Ti/Ci and TV/CV in entire NAKL corpus

- O/E > 1 indicates *over-representation*; O/E < 1 indicates *under-representation*
- /Ti/ and /Tj/ sequences are significantly under-represented:
  - $\chi 2(1) = 7610$, $p < 0.001$

- How does this compare with the distribution of /c, c^h, cc/?
  - Historically *t, t^h, tt > c, c^h, cc before /i, j/ across the board (Cho 2009; Lee and Ramsey 2011; see appendix A1)

	/i/, /jV/	other Vs
/c, c^h, cc/	5944 (E: 5493) **O/E = 1.08**	20474 (E: 20924) O/E = 0.98
other Cs	25757 (E: 26208) O/E = 0.98	100287 (E: 99836) O/E = 1.00

Table 3. Observed/Expected counts of CHi/Ci and CHV/CV in NAKL

- /CHi/ and /CHj/ sequences are significantly over-represented:
  - $\chi 2(1) = 56.417$, $p < 0.001$

3.1.1 Is the distribution the same across different lexical strata?

- Under-representation seems to hold strongly in the native and Sino-Korean strata of NAKL – Table 4.
- There is a statistically significant ($p < 0.0001$) under-representation in the loanword stratum of NAKL as well, although the O/E value is much higher in the loanword stratum.

[1] Observed/Expected = (Row total * Column total)/Grand Total

- Native (n = 13459), Sino-Korean (n = 36504) and loanwords (n = 3233):

	Native		Sino-Korean		Loanword	
	/i/, /jV/	Other Vs	/i/, /jV/	Other Vs	/i/, /jV/	Other Vs
/t, t^h, tt/	147 (E: 1995) **O/E = 0.07**	9770 (E: 7922) O/E = 1.23	17 (E: 3374) **O/E = 0.005**	16118 (E: 12761) O/E = 1.27	290 (E: 415) **O/E = 0.70**	1554 (E: 1428) O/E = 1.09
other Cs	8073 (E: 6225) O/E = 1.30	22864 (E: 5686) O/E = 0.93	21358 (E: 18001) O/E = 1.18	64727 (E: 68084) O/E = 0.95	1825 (E: 1700) O/E = 1.07	5728 (E: 5853) O/E = 0.98

Table 4. Observed/Expected counts of Ti/Ci and TV/CV in native, Sino-Korean words and loanwords in NAKL

Are Korean speakers just matching the statistics of English?
Given the general rate of borrowing from English, what is the rate of loaning TI words?

- English sequences /t, d, θ, ð/+/ɪ, i/ (TI) are borrowed faithfully into Korean as /ti/ or /t^hi/ (Cho 2009).
- LOANED = words that appear in both CMU Dictionary (Hayes ed. from Hayes & White 2013, CELEX freq. > 1) and NAKL loanword corpus (NAKL 1991: n = 2785) – initial estimate

	/t, d, θ, ð/+/ɪ, i/	Other CV
LOANED	241 (E: 297) **O/E = 0.81**	4379 (E: 4322) O/E = 1.01
NOT LOANED	2612 (E: 2556) O/E = 1.02	37097 (E: 37153) O/E = 1.00

Table 5. Observed/Expected counts of English words with relevant CV seqs. (loaned/not loaned against TI/CV)

- Given the general expected rate of loaning English words into Korean, the number of TI English words loaned in is significantly less than what we expect:
  - $\chi 2(1) = 12.51$, $p < 0.001$

3.2 Corpus investigations: Child-Directed Speech

How robust is the distribution? Do we see a similar distribution in Child-Directed Speech?

- Jiwon corpus (CHILDES; Ghim 2005; n = 4986)
- Mother-Child interaction: Single child (ages 2;0 – 2;3)

	/i/, /jV/	Other Vs
/t, t^h, tt/	45 (E: 299) **O/E = 0.15**	1445 (E: 1191) O/E = 1.21
other Cs	2393 (E: 2138) O/E = 1.12	8258 (E: 8512) O/E = 0.97

Table 5. Observed/Expected counts of Ti/Ci and TV/CV in Jiwon corpus

- A similarly significant under-representation as with the adult NAKL corpus:
  - $\chi 2(1) = 56.417$, $p < 0.001$

3.3 Phonotactic modeling of the Korean lexicon

Does statistical under-representation actually result in the learning of a phonotactic constraint that penalizes [Ti] and [Tj] sequences?

- UCLA Phonotactic Learner (Hayes & Wilson 2008):
  - Input:
    - Learning data: Entire corpus (including loanwords)
    - Assuming no morphological parse
    - Feature set:
      - Diphthongs (e.g. /je/ etc.) are split into a glide + vowel.
      - Assuming /c, c^h, cc/ are palatal affricates
      - Assuming that /ɛ/ and /e/ are merged to /e/ (Shin et al. 2013, Eychenne & Jang 2015)
    - Using *type* frequencies: assuming statistics counted over word types and not tokens (e.g. Pierrehumbert 2003)
  - Parameters:
    - Asked to only find bigram constraints
    - O/E accuracy threshold set at 0.3 (following most of the simulations in Hayes & Wilson (2008)).
    - No other special parameters – defaults

3.3.1 Results: NAKL

- Sanity check – what are the learned constraints with the largest weights?

No.	Constraint	Weight	Description
1.	*[+spread_glot][+word_boundary]	7.14	No aspirated stops word-finally
2.	*[+word_boundary][+nasal,+dorsal]	6.67	No word-initial [ŋ]
3.	*[+const_glot][-approx.]	6.36	No tense stops preceding non-approx. cons.
4.	*[-lateral,-syllabic][+lateral]	6.16	No non-lateral consonants before a lateral
5.	*[-syllabic][+nasal,+dorsal]	5.99	[ŋ] cannot occur following consonants
6.	*[-cons.,+front,-syll.][-labial,+high]	5.81	*[jɯ], *[ji]

- Crucially, the learner learns a constraint penalizing [Ti] and [Tj] sequences!
  - ***[-sonorant,-strident][-spread_glot,-const_glot,+high,+front]: 1.916 (ranked 55th out of 134)**
  - cf. Learner assigns this constraint a weight of 6.169 if no [Ti] or [Tj] sequences are in the corpus.
- What about including *token* frequency in training?
  - Doesn't change the result qualitatively, although the same crucial constraint gets a slightly higher weight (weight = 2.417).

3.3.2 Results: Jiwon

- Sanity check – what are the learned constraints with the largest weights?

No.	Constraint	Weight	Description
1.	*[+cont.,-lat.,-syll.][+word_bound.]	4.93	No word-final glides or fricatives
2.	*[+const_glot][+word_boundary]	4.54	No tense stops word-finally
3.	*[+spread_glot][+word_boundary]	4.40	No aspirated stops word-finally
4.	*[+const_glot][-approximant]	4.39	No tense stops preceding non-approx. cons.
5.	*[+spread_glot][-approximant]	4.31	No asp. stops preceding non-approx. cons.
6.	*[+word_boundary][+nasal,+dorsal]	4.27	No word-initial [ŋ]

- At the same O/E criterion of 0.3, the learner doesn't actually infer a constraint penalizing [Ti] and [Tj] sequences as a natural class of /t, t^h, tt/.
    - *[t^hi] = 2.71 (*[-delayed_release,+spread_glot,-labial][+high,+front,+syllabic])
    - *[tt][i, j] = 2.81 (*[+const_glot,-strident][+high,+front])
    - *[tj] = 2.49 (*[-son.,-spread_glot,-distributed][-spread_glot,-const_glot,-round,-syll.])
    - **However, [ti] and [t^hj] are not penalized.**
- Increasing the O/E criterion modestly to 0.31 though seems to get us the same *qualitative* result as with the adult NAKL corpus:
    - But the learner still does not group /t, t^h, tt/ together as a natural class, so independent constraints are doing the work.
- If the learner starts with the constraint *[-son., -strident][-consonantal, +high, +front] (i.e. *[t, t^h, tt][i, j]), it does actually assign it a somewhat comparable weight as in previous simulations (weight = 1.17), even when O/E = 0.3.

3.4 Summary: Korean

- Korean palatalization is supposedly a paradigmatic example of MDEE.
- Without much tweaking, a phonotactic learner assigns a sizable weight to a constraint penalizing [Ti] and [Tj] sequences, despite these forms actually existing in the lexicon of Korean.
    - Penalty is less robustly encoded when trained on a smaller corpus of CDS (about a tenth the size) - although qualitatively it arrives at the same result.
- Statistical under-representation does indeed translate into a *well-formedness penalty* for words with [Ti] and [Tj].

➢ Violates assumption of phonotactic "productivity" laid out in (4)!

4. Discussion

4.1 How about other MDEEs – similar or different?

- Turkish velar deletion – another parade case of MDEE (e.g. Zimmer & Abbott 1978, Lewis 1967, Sezer 1981)

(5) Suffix boundary velar deletion

a.	/bebe**k**-In/[2]	→	[bebein]	'baby-GEN'
b.	cf. /bebe**k**/	→	[bebe**k**]	'baby-NOM'
c.	/ipe**k**-A/	→	[ipee]	'cotton-DAT'
d.	cf. /ipe**k**/	→	[ipe**k**]	'cotton-NOM'

[2] Vowels in uppercase: vowels that participate in vowel harmony.

(6) Non-deletion morpheme-internally (from Inkelas 2011)

a.	/hare**k**et /	→ [hare**k**et]	'motion'
b.	/si**g**orta /	→ [si**g**orta]	'insurance'
c.	/so**k**a**k**-A/	→ [so**k**aa]	'street-DAT'
d.	/me**k**i**k**-A/	→ [me**k**ie]	'(weaver's) shuttle-DAT'

- Velar deletion **does not occur** in the following morphological/phonological environments (Sezer 1981, Inkelas 2011)
    - Verbs
    - Monosyllabic nouns
    - When the target /k/ is suffix-initial – equally morphologically-derived as /k/ being stem-final.
    - A number of lexical exceptions (mostly loanwords)
- Velar deletion seems to mostly be confined to polysyllabic nouns.
- Velar deletion **violates the Derived Environment Condition** (Inkelas 2011)
- Preliminary work (Chong *in progress*) indicates that the learner does not readily learn a *VGV constraint that motivates velar deletion.
    - **No phonotactic support for the morphologically-conditioned alternation.**

4.1.1 What of MDEEs?

- Despite structural similarities, patterns previously described together as examples of MDEE are by no means a unified phenomenon.
    - In general, once we start looking more closely at the quantitative patterns of both alternations and the lexicon, they start looking quite different.
- In both Korean and Turkish, the canonical MDEE patterns don't hold up to scrutiny, especially when one takes into consideration the assumptions laid out in (3) and (4) repeated here:

(3) *Derived-environment condition:* Turkish fails on this (Inkelas 2011, Sezer 1981).

(4) *Phonotactic "productivity":* Korean fails on this (arguably also on (3) in compounding – see Appendix A2).

- Other well-known MDEE cases are similarly murky:
    - Finnish assibilation (Antilla 2006): like Turkish, assibilation does not occur in all possible derived environments where phonological condition is satisfied.
    - Cho (2009) points out that many other MDEE patterns are similar to Korean in having exceptions that are mostly loanwords:
        - Finnish Vowel Coalescence (Anttila 2009)
        - Polish First Velar Palatalization (Lubowicz 1998)
    - It is possible then that an inspection of the lexicon of other cases would give us a similar picture as we're seeing in Korean.

4.2 MDEEs: Relation between static and dynamic generalizations

- MDEEs: loss of a static generalization, with an active dynamic one (i.e. alternation) (Paster 2013)
- BUT:
  - Korean – static phonotactic generalization is still there!
  - If offending words are rare enough, a *probabilistic* learner will nonetheless learn an, albeit weaker, phonotactic constraint.
    - If this turns out to accord with native speaker intuitions then this argues against non-probabilistic learners (e.g. Biased Constraint Demotion; Prince & Tesar 2004).
  - Turkish – there's no strong static phonotactic generalization (VGV is acceptable) but the alternation is highly morphologically conditioned.
    - It's "active" but no straightforward way to account for this without referring to morphological information or word category.
- These cases of mismatch turn out to be apparent:

	Static phonotactic generalization	**Alternation**
Korean	Yes (albeit weaker)	Yes
Turkish	No	No-*ish* (Yes but very much constrained)

- Suggests perhaps a bias for broader generalizations (see also Martin 2011), or a match between these different levels of generalizations?

5. Conclusion and future directions

- **Take home: Korean palatalization is not really a clear case of MDEE!**
- In progress: testing native Korean speakers on their well-formedness judgments of wug words with [Ti] sequences → confirm corpus and modeling work.
- In progress: Assuming a canonical MDEE language – are alternations more difficult to learn?
  - Using artificial grammar learning
- Why are [Ti] sequences tolerated in loans in Korean?
- Can we construct a typology of MDEE patterns?
- What about other kinds of generalization mismatches?
  - Derived-environment blocking (Hall 2006; Wolf 2008; Paster 2013) – mirror image of Derived-environment effects.

Thank you!

Acknowledgements to Robert Daland, Bruce Hayes, Sharon Inkelas, Karen Jesney, Stephanie Shih, Megha Sundara, Kie Zuraw and the audience at USC and UC Berkeley for discussion of various aspects of this work. Thanks also to Joo Hee Oom for help processing parts of the NAKL corpus and to Sharon Inkelas for help with the TELL corpus. This work is funded by a UCLA Dissertation Year Fellowship.

(In order to save paper, full references are available upon request)

Contact: ajchong@ucla.edu

Appendices:

A1. A brief history of Korean palatalization (Cho 2009; Lee & Ramsey 2011)

- 19th C.: Palatalization was an across-the-board sound change
  - *t, tʰ, tt > c, cʰ, cc before /i, j/
- Current mismatch has three sources:

(7) Historic monophthongization of /ɨ̯i/ sequences (circa. 19th C.) (from Cho 2009):

a. *ʌtɨ̯i > /ʌti/ 'where'

b. *matɨ̯i > /mati/ 'joint'

(8) Synchronic monophthongization of /ɨ̯i/ sequences (from Cho 2009):

a. /t*ɨ-i-ta/ → [t*ɨita ~ t*ita] 'to catch the eye'

b. /tʰɨ-i-ta/ → [tʰɨita ~ tʰita] 'to be open'

(9) Loanwords from English:

a. Eng: /stedjəm/ > Kor: /sɨtatium/ 'stadium'

b. Eng: /bɪldɪŋ/ > Kor: /piltiŋ/ 'building'

A2. Failure to palatalize elsewhere: Compounding and Prefixing

- The characterization of palatalization as a derived-environment-only rule is also problematic from the point of view of compounding.
- Palatalization does not occur across compound (or prefix) boundaries (10a, c, d) vs. (10b) (Sohn 1999, Oh 1995):

(10) Prefixes and compounds:

a. /patʰ # ilaŋ/ → [patiɾaŋ] ~ [panniɾaŋ]
field ridge 'ridge of a field'

b. cf. /patʰ + ilaŋ/ → [pacʰiraŋ]
field and 'a field and'

c. /patʰ # il/ → [patil] ~ [pannil]
field work 'field work'

d. /hot^h + ipul/ → [hotibul] ~ [honnibul] (prefix)
single comforter ‘single-layer comforter’ or ‘sheet’

- Although palatalization does not occur, another process optionally occurs: *n*-insertion followed by obstruent nasalization.
    - *n*-insertion is an independent process that occurs at compound boundaries (as well as across word and phrase boundaries) as in (11).
    - [n] is optionally inserted when the first element of the compound ends with a consonant and the second element of the compound begins with /i/ or /j/ (see Jun 2015 for a recent investigation).

(11) *n*-insertion examples in compounding more generally:
a. /com # jak/ → [comnjak]~[comjak] ‘mothball’
b. /som # ipul/ → [somnipul]~[somipul] ‘cottton sheet’

➢ Conspiracy: although palatalization doesn’t occur across the compound boundary, another repair to satisfy the markedness constraint *Ti is possible!

- So it’s unclear if faithful [ti] and [tj] sequences are also dispreferred here and the *n*-inserted forms are the preferred outputs.
- Might provide more evidence for grammatical leakage (à la Martin 2011).
    - A phonotactic constraint exerts some influence ensuring that compounds that create violations of static morpheme-internal generalizations are statistically dispreferred.
    - Even when with exceptions, a phonotactic constraint penalizing [Ti] and [Tj] sequences might still be at work here in the compounding domain.

References

Anttila, A. (2006). Variation and opacity. *Natural Language and Linguistic Theory, 24(4),* 893-944.

Antilla, A. (2009). Derived environment effects in Colloquial Helsinki Finnish. In Kristin Hanson and Sharon Inkelas (eds.), *The nature of the word: Essays in honor of Paul Kiparsky*. MIT Press: Cambridge, 433-460.

Becker, M. & Gouskova, M. (2016). Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry*, 47(3).

Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica*, 58, 129-162.

Cho, Y.-M. Y. (2009). A historical perspective on nonderived environment blocking: The case of Korean palatalization. In Kristin Hanson and Sharon Inkelas (eds.), *The nature of the word: Essays in honor of Paul Kiparsky*. MIT Press: Cambridge, 461-486.

Eychenne, J. & Jang, T.-Y. (2015). On the merger of Korean mid front vowels: Phonetic and phonological evidence. *Journal of the Korean Society of Speech Sciences, 7(2),* 119-129.

Ghim, H.-R. (2005). Korean Jiwon Corpus. CHILDES

Hall, T. A. (2006). Derived environment blocking effects in Optimality Theory. *Natural Language and Linguistic Theory, 24(4),* 803-856.

Hayes, B. (2004). Phonological acquisition in Optimality Theory: the early stages. In René Kager, Joe Pater and Wim Zonneveld (eds.), *Constraints in phonological acquisition*. Cambridge University Press: Cambridge, UK, 158-203.

Hayes, B. & Wilson, C. (2008). A maximum entropy model of phonotactics and phonological learning. *Linguistic Inquiry, 39(3),* 379-440.

Inkelas, S. (2011). Another look at velar deletion in Turkish, with special attention to the derived environment condition. In E. E. Taylan and Bengisu Rona (eds.), *Puzzles of language: Essays in honour of Karl Zimmer*.

Inkelas, S. (2015). Confidence scales: a new approach to derived environment effects. In Yuchau E. Hsiao and Lian-Hee Wee (eds.), *Capturing phonological shades*. Cambridge Scholar Press.

Inkelas, S. & C. O. Orgun (1995). Level Ordering and Economy in the Lexical Phonology of Turkish. *Language, 71,* 763-793.

Inkelas, S. & Zoll, C. (2007). Is Grammar Dependence Real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics, 45*, 133-171.

Inkelas, S., Küntay, A., C. O. Orgun & Sprouse, R. (2000). Turkish Electronic Living Lexicon (TELL): A lexical database. In M. Gavrilidou et al., eds., Proceedings of the Second International Conference on Language Resources and Evaluation. Paris: European Languages Resources Association.

Iverson, G. & Wheeler, D. W. (1988). Blocking and the elsewhere condition. In M. Hammond & M. Noonan (eds.), *Theoretical morphology: Approaches in modern linguistics.* Academic Press: San Diego. 325-338.

Keyser, S & Kiparsky, P. (1984). Syllable structure in Finnish phonology. In M. Aronoff and R. Oehrle (eds.). *Language Sound Structure*, 7–31. Cambridge, MA: MIT Press.

Kim, B., Lee, G., & Lee, J.-H. (2002). Morpheme-based grapheme to phoneme conversion using phonetic patterns and morphophonemic connectivity information. *ACM Trans. Asian Lang. Inf. Process.* 1(1), 65-82.

Kiparsky, Paul (1968). Linguistic universals and language change. In Emmon Bach and Robert T. Harms (eds.), *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, 171-202.

Kiparsky, P. (1973). Phonological representations. In Osamu Fujimura (ed.), *Three Dimensions of Linguistic Theory*. Tokyo: TEC, 1-135.

Kiparsky, P. (1982). Explanation in Phonology. Foris: Dordrecht.

Kiparsky, P. (1993). Blocking in Non-Derived Environments. In E. Kaisse and S. Hargus (eds.), *Phonetics and Phonology 4: Studies in Lexical Phonology*, 277-313. San Diego: Academic Press.

Lee, K.-M. & Ramsey, S. R. (2011). A history of the Korean Language. Cambridge: Cambridge Univ. Press.

Lewis, G. (1967). *Turkish grammar*. Oxford: Oxford University Press.

Lubowicz, A. (1998). Derived environment effects in Optimality Theory. *Lingua, 112(4),* 243-280.
Martin, A. (2011). Grammars leak: Modeling how phonotactic generalizations interact within the grammar. *Language, 87(4),* 751-770.
NAKL (2003). *Hyeondae gug-eo sayong bindo josa gyeolgwa pail [Frequently used words in modern language].* Downloadable from: http://www.korean.go.kr/
Oh, M. (1995). A prosodic analysis of non-derived environment blocking. *Journal of East Asian Linguistics, 4*, 261-279.
Paster, M. (2013). Rethinking the "duplication problem". *Lingua, 126,* 78-91.
Pierrehumbert, (2003). Phonetic diversity, statistical learning and acquisition of phonology. *Language and Speech, 46,* 115-154.
Prince, A. & Tesar, B. (2004). Learning phonotactic distributions. In René Kager, Joe Pater and Wim Zonneveld (eds.), *Constraints in phonological acquisition*. Cambridge University Press: Cambridge, UK, 245-291.
Sezer, E. (1981) The k/Ø alternation in Turkish. In G. N Clements (ed.), *Harvard Studies in Phonology*, 354–82. Bloomington: IULC.
Shin, J., Kiaer, J. & Cha, J. (2013). The sounds of Korean. Cambridge: Cambridge University Press.
Sohn, H.-M. (1999). The Korean Language. Cambridge: Cambridge University Press.
Tesar, B. & Prince, A. (2007). Using phonotactics to learn phonological alternations. *Proceedings of the Thirty-Ninth Conference of the Chicago Linguistics Society, vol. II: The panels*, 209-237.
Wolf, M. (2008). Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model. PhD Dissertation, U. Mass. Amherst.
Zimmer, K. & Abbott, B. (1978). The k/ø alternation in Turkish: Some experimental evidence for its productivity. *Journal of Psycholinguistic Research, 7(1),* 35-46.