

Frequency of use and analogical change in Korean nouns: A case against abstract representation

Yoonjung Kang
 SUNY Stony Brook
yikang@optonline.net

12/5/02

1

1. Deriving morphologically complex words in Generative Phonology

• A morpheme which has more than one non-suppletive alternant has a unique phonological representation in its lexical entry (**Underlying Representation**).

- Underlying Representation : English plural suffix

'dog':	/dɒg/
'cat':	/kæt/
'bush':	/bʊʃ/

• Word Formation Rules combine morphemes to produce morphologically complex words.

- 'dogs': /dɒg + z/ 'cats': /kæt + z/ 'bushes': /bʊʃ + z/

• Various surface pronunciations of a morpheme are derived through interaction of rules or constraints.

/dɒg+z/	AGREE (VOI)	IDENT (VOI)	/kæt+z/	AGREE (VOI)	IDENT (VOI)	/bʊʃ+z/	*SIBSIB	DEP (V)
ɔ̃ dɒgz			kætz	*!		bʊʃz	*!	
dɒgz	*!	*	ɔ̃ kæts		*	ɔ̃ bʊʃɛz		*

3

Abstract

It has been a standard assumption in generative phonology that the lexicon is a list of morphemes and that morphologically complex words are derived through Word Formation Rules. Under this view, a given morpheme has a unique Underlying Representation (UR) and its surface alternation is derived through interaction of phonological rules or constraints. However, the abstract and arbitrary character of UR has been a recalcitrant problem. Moreover, with the introduction of constraints that regulate the relations between surface forms of morphologically related words in Optimality Theory, the motivation for UR is substantially weakened, leading to proposals to do away with UR. Furthermore, there is converging evidence from psycholinguistics and historical linguistics that even morphologically complex words are treated differently depending on their frequency of use, suggesting that at least some complex words are stored in the lexicon as such. In particular, evidence shows that an analogically motivated sound change is more likely to affect infrequent words than frequent words. In this paper, I present additional evidence from Korean that even regularly inflected words undergo analogical changes at different rates depending on their frequency of use, lending further support to a more concrete view of lexical representation.

2

2. Fallacy of "economy of storage" argument: Jaeger 1986, Bybee 2001

- Economy of Storage argument
 - Memory representation is simple; access is very complex.
 - "Human storage space for memorizing words is at a premium so that every word must be memorized in a maximally economical form in which redundant (predictable) properties are eliminated...it is to our obvious advantage to memorize the rule rather than to clutter up our memory with these redundant facts (Halle and Clements 1983)."
- Detailed and Redundant storage
 - Memory representation is very complex; access is relatively simple.
 - "The indications from neuropsychology and psychology are that, instead of storing a small number of primitives and organizing them in terms of a large number of rules, we store a large number of complex items which we manipulate with comparatively simple operations (Ladefoged 1972)."
 - Redundant (predictable) acoustic cues are important in identification of speech sounds (Ohala and Ohala 1995, Miller 1994).

4

- Redundant (phonetic) features may show lexically specific behavior (Steriade 2000: English flapping).
- Distribution of sounds may depend on predictable features (Vennemann 1974: stress and reduced vowel in Icelandic).
- Phonetic variations may affect different lexical items differently (Bybee 1976: [er]-[r] variation in English).
- Exemplar model of speech perception and categorization (Johnson 1997)
- Hearers not only correctly identify words or utterances produced by different speakers but also correctly identify the voice of different speakers they have heard before.
- All perceived tokens of words and phrases are recorded in memory and they are organized into categories according to their similarity with one another.

5

4. Storage of regular, morphologically complex words: Bybee 2001

- Evidence indicates that even regularly inflected words may be stored in memory and that frequency can affect the strength of their representations.

3.1. Evidence from psycholinguistics

- Stemberger and MacWhinney (1986, 1988)
 - Speech error in formation of the past tense of regular English verbs
 - 10-low frequency verbs and 10 high-frequency verbs
 - Twice as many errors on low-frequency verbs than on high-frequency verbs
- Alegre and Gordon (1999)
 - Lexical decision on English past tense forms
 - Shorter reaction time for high-frequency words than for low-frequency words
 - Significant frequency effect in reaction time for high-frequency words
 - No frequency effect in reaction time for low-frequency words

7

3. Cross-derivational faithfulness constraints

- Flemming (1995), Benua (1997), Kenstowicz (1996, 1997), Burzio (1994, 1996, 1998, 2000), Kager (1999), Steriade (1999, 2000)
- Phonetic similarity between morphologically related words can be handled by cross-derivational faithfulness constraints which regulate the relation between surface forms of words without positing a common abstract representation (UR).
- Cyclic effects can be handled by cross-derivational faithfulness constraints without having to derive morphologically complex words on-line.
- Paradigm uniformity effects—similarity between words that are morphologically related but do not share derivational history—are better handled by constraints that directly relate surface forms of these words than derivational theory.

6

- Hare *et al.* (2001)

- Dictation of orally presented words that represented two homophones, one a past tense form and one a monomorphemic words (ex. *allowed vs. aloud*)
- Subjects tend to select the homophone that was the most frequent, even when that form was a regular past tense form.

- Sereno and Jongman (1997)

- Lexical decision on regularly inflected English nouns (singular and plural forms)
- Speed of recognition was influenced by the frequency of the particular inflected form, not by the frequency of the singular form or of the noun in general.

8

3.2. Evidence from language change

- Sound changes affect low-frequency words and high frequency words differently (Schuchardt 1885; Phillips 1983, 1984, 2001; Bybee 1985, 2001; Bybee and Hopper 2001).
- Phonetically motivated changes affect high-frequency words first.
 - Phonetically motivated changes progress with each use of the word. Therefore, the more frequent a word is put to use, the more advanced the change is.
 - o Fidelholtz (1975): vowel reduction in English
 - o Hooper (1976): schwa deletion in English
 - o Phillips (1980): Raising of /a/ before nasals in Old English
 - o Rhodes (1996): [t] flapping in English
 - o Phillips (2001): [str] → [ʃtr] in English

5. Background in Korean phonology

• PHONEME INVENTORY OF KOREAN

a) Consonants

p, p ^h , p'	t, t ^h , t'	k, k ^h , k'	
	ts, ts ^h , ts'		
s, s'			h
m	n	ŋ	
	l/r		

• Affricates are considered alveolar (cf. H. Kim 1997).

b) Vowels

i	ɨ	u
e	ʌ	o
(æ)	ɑ	

- Analogically motivated changes affect low-frequency words first.
 - High-frequency words form strong mental representations and resist change motivated by analogy to other forms.
 - o Bybee and Slobin (1982): regularization of past tense in English
 - o Corbett *et al.* (2001): regularization of noun paradigm in Russian
 - o Smith (2001): loss of *be* auxiliary for the resultative/perfect in English
 - o Bybee (2001): loss of liaison in French
- The frequency effect on lexical diffusion is found in regularly inflected words
 - Phonetically motivated change
 - o Bybee (2000): deletion of final t/d in English
 - a higher rate of deletion for high-frequency words
 - Analogically motivated change
 - o frequency effect also found among regular past tense form ???
- Korean provides a case where an analogical change affects regularly inflected words differentially depending on their frequency of use.

• CODA NEUTRALIZATION (Lee 1972; Huh 1975)

<15C> p, p ^h , (p'), β	→ p	<16C>→ p, p ^h , (p')	→ p
t, t ^h , (t'), h	→ t	t, t ^h , (t'), h	→ t
ts, ts ^h , (ts'), s, (s'), z	→ s	ts, ts ^h , (ts'), s, s'	→ t
k, k ^h , (k')	→ k	k, k ^h , k'	→ k

• In 15C, the contrast between t and s was maintained in coda position.

• AFFRICATION IN DERIVED ENVIRONMENTS (a.k.a. palatalization) (H. Kim 1997)

t, t ^h → ts, ts ^h / __](h)	
kaf ^h -a 'is the same'	kats ^h -i 'together' cf. kaf ^h -im 'being the same'
tot-a 'to rise'	hæ-tot ^h -i 'sun rise' tot-im 'rising'

• The regular affrication applies only before a **high front** vowel.

• INTERSONORANT VOICING OF PLAIN OBSTRUENTS (S. Jun 1993)

6. Overview of change in noun-final obstruents

(Kwak 1984; Choi 1986; Ko 1989; H. Kang 1994; M. Kim et al. 1997)

• NOUNS FOR WHICH CHANGE TO /s/ IS COMPLETED

: Many nouns that used to end in /t/ or /ts/ in 15C now end in /s/.

ex. 'nail'	ACC.	LOC.
15C:	<i>mot-ol</i>	<i>mot-e</i> mot#
21C:	<i>mos-il</i>	<i>mos-e</i> mot#

a) t > s: *pus* 'brush', *kas* 'hat', *nas* 'sickle', *mos* 'nail'

b) t > ts > s¹: *pas* 'friend', *t'is* 'meaning', *kos* 'place'

c) ts > s: *ius* 'neighbor', *kis* 'collar', *kus* 'exorcism', *soykos* 'ice pick',
oyas 'plum'

cf. In Present day Korean, no noun ends in /t/ before a vowel-initial suffix.

¹ Some dialects still show /ts/ and /ts/ for 'meaning' and 'place' (Kwak 1984).

• LOANWORDS

: Borrowed words ending in a coronal stop in isolation are realized with final /s/ before a vowel-initial suffix.

	NOM.	ACC.	LOC.
'cut'	<i>k'as-i</i>	<i>k'as-il</i>	<i>k'as-e</i> k'at#
'David'	<i>teipis-i</i>	<i>teipis-il</i>	<i>teipis-e</i> teipt#

• LACK OF ACCURATE DATA

- The innovative /s/ or /ts/ pronunciations are not equally available in all contexts for all nouns.
- Except for a couple of production data (Kwak 1984, Choi 1986, H. Kang 1994), no systematic quantitative study is available regarding the availability of these pronunciations in different contexts.
- Production experiments with a limited number of repetitions cannot provide a full picture of the subtle differences in well-formedness.
- Quantitative data from gradient well-formedness judgments are needed. See below.

• ALL /ts/ /ts^h/-FINAL NOUNS SHOW AN OPTIONAL VARIANT WITH FINAL /s/.²

ex. 'milk' ACC. LOC.

tsats-il ~ *tsas-il* *tsats-e* ~ *tsas-e* *tsat#*

a) *pis* 'debt', *nats* 'day', *tsats* 'milk', *tsots* 'male genitalia'

b) *pis^h* 'light', *nats^h* 'face', *k'ots^h* 'flower', *tats^h* 'trap', *suts^h* 'coal', *tsits^h* 'sail'
yuts^h 'a traditional game played with sticks'

• ALL /t^h/-FINAL NOUNS SHOW OPTIONAL VARIANTS WITH FINAL /ts^h/ AND /s/

ex. 'pot' ACC. LOC.

so^h-il ~ *so^h-il* ~ *sos-il* *so^h-e* ~ *so^h-e* ~ *sos-e* *so#*

pyat^h 'warmth of the sun', *mit^h* 'bottom', *pat^h* 'field', *k'it^h* 'end', *p'at^h* 'red bean', *so^h* 'pot', *kat^h* 'outside' *kyat^h* 'side'

* /ts^h/ is expected before *i*-initial suffixes due to the regular affrication but is an innovation before other suffixes.

² In some dialects, 'debt' and 'day' have been changed to *pis* and *nas* (Kwak 1984).

7. Change by analogy: KO 1989

• 16C~: Coda neutralization of s and t

	s → t/Coda		
	15C	os#	os-e
ex. 'clothes'	ot#	ot#	os-e
'nail'	mot#	mot-e	mot-e

* Contrasts among all coronal obstruents are neutralized in coda position.

• Salience of isolation form

- Nouns can stand alone in Korean and the speakers seem to identify the isolation form as a "base". cf. Verbs cannot stand alone.
- Evidence for salience of isolation form of nouns

o Paradigm leveling based on the isolation form (cf. Kenstowicz 1996)

	NOM.		
ex. 'knee'	mulip#	mulip ^h -i	mulip-i
'chicken'	tak#	talk-i	tak-i
'tree'	namo#	namk-i	namu-ka

o Orthographic convention of 17C: a noun and a suffix tend to be written into separate syllables while a verb and a suffix tend to be written reflecting resyllabification.

• Rule inversion

- Due to the salience of isolation form, speakers reanalyze the newly introduced alternation between t and s as a change from t (isolation form) to s (suffixed form).

• High type-frequency of /s/-final nouns

- Among nouns that end in a coronal obstruent, /s/-final nouns were and continue to be the most common.
 - The change of noun-final coronal obstruents to /s/ is by analogy to other /s/-final nouns.

Cf. Hayes (1997)'s Anticorrespondence constraints: the language learners sometimes "create relatively ad hoc, language-specific constraints... seiz[ing] upon generalizations that are *statistically useful, albeit imperfect* [emphasis original], in an attempt to improve her ability to guess the unknown inflected forms of known stems."

- English loanwords with final [t] are realized with final [s] before a vowel-initial suffix. But, [ts] is not attested.

teibit# teibis-i *teibits-i 'david, (NOM.)'

- Synchronically, according to my survey, some inflected forms of t^h-final nouns show a low rating for ts^h pronunciation but nevertheless show a high rating for s pronunciation.

h^hat^h-e ~ h^hats^h-e ~ h^has-e 'red bean, LOC.'

Average rating	2.5	1.1	3.5
----------------	-----	-----	-----

• Although fricativization of affricates (ts, ts^h > s) is in general a natural change, intervocalic ts in Korean is phonetically voiced [dz]. Therefore, a change [dz] > [s] is still phonetically unnatural.

8. Inadequacy of phonetic explanation

- A direct change of coronal stop t, t^h to a sibilant fricative is unnatural
- Kirchner (1998): "Lenition of a coronal stop to a sibilant fricative is by no means attested in the general case; it occurs only in contexts where the release of the stop inherently has significant frication."
- A change t, t^h > ts, ts^h > s is possible, but only before a high vowel.
- H. Kim (2001): "[S]top assibilation is motivated by the brief period of turbulence that sometimes occurs at the release of a plosive into a high vowel."
- In many cases, there is no evidence that t, t^h > s change went through an intermediate affricate stage.

- No philological records of affricate stage for many nouns with t > s change:
 pus 'brush', kas 'hat', nas 'sickle', mos 'nail'

9. Data gathering

• WELL-FORMEDNESS JUDGMENTS

- 16 nouns in common use that end in a coronal obstruent other than /s/.
- | | | | | | | | | | |
|-------------------|-------------------|---------------|--------------|----------------|-----------------|--------------------------|------------|----------------|------------|
| ts : | nais | 'day', tsyais | 'milk', pils | 'debt' | | | | | |
| ts ^h : | su ^h s | 'coal', nais | 'face', pils | 'light', tsais | 'trap', k'ois | 'flower' | | | |
| t ^h : | pat | 'field', k'it | 'end', kyat | 'side', kat | 'outside', pyat | 'warmth of the sun', sot | 'pot', mit | 'bottom', p'at | 'red bean' |
- 7 vowel-initial inflectional suffixes → 3 vowel contexts.
- | | |
|----|--------------------|
| t- | (NOM. & COP.) |
| i- | (ACC, DIR., & TOP) |
| e- | (LOC.) |
- cf. iy ~ e (GEN.)³
- 112 sentences containing each noun + suffix combination (16 nouns * 7 suffixes) are made up. Each sentence is paired up with one of five target

³ The genitive suffix is rarely, if ever, used in spoken language and even in read speech, it is pronounced as [e] (H. Kang 1999). The data on genitive forms are not discussed in this paper.

pronunciations (t, t^h, ts, ts^h, s) of the final coronal obstruent, resulting in 560 entries (112 * 5).

- 8 native speakers of Korean rated the target pronunciation between 1 (worst) and 4 (Good)

ex. 엄마, 할머니[소시] 너부 무거워요.

1 2 3 4
 <Mom, pot-NOM. too heavy-HON>

• FREQUENCY COUNTS FROM ON-LINE CORPUS

- KAIST Concordance program (KCPMSTAT) (<http://morph.kaist.ac.kr/kcp/>): 13.6 million words

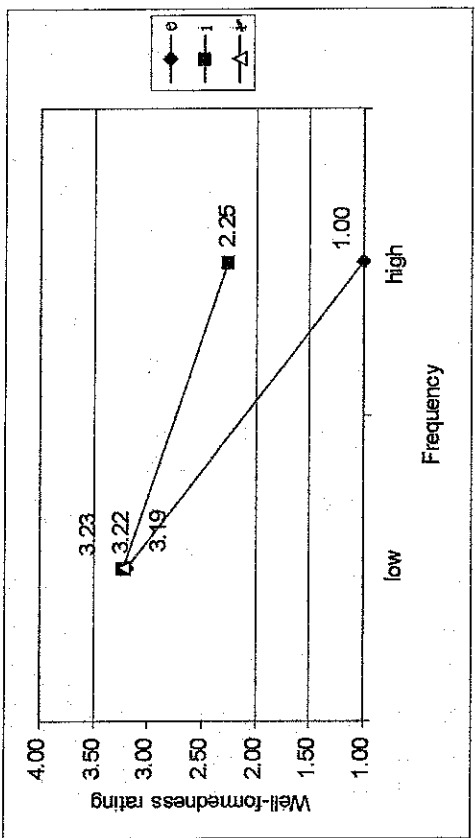
- Frequency counts for each of the 112 noun + suffix combinations are obtained from the corpus.

- NOM. and ACC. suffixes are very often omitted in casual speech but the corpus is mainly based on written texts (99%).

→ The counts for nominative and accusative forms are adjusted to 60 % and 10 % of the counts from the corpus in the following analysis.⁴

⁴ According to Cho (1981), cited in Y. Kim (1997), the nominative suffix is dropped around 40% of the time and the accusative suffix is dropped close to 90% of the time in mothers' speech to young children.

• ts → s (3 nouns)



* A complete list of words and their ratings is in the handout.

10. Change to [s]

• Ratings for [s]-final pronunciation of nouns in various contexts (all values average of 8 speakers)

• 3 factors

- "Underlying" final consonant: ts, ts^h, t^h

- Vowel of the following suffix: i, i, e

- Frequency of occurrence: low (0~67) vs. high (67~)

o Average number of occurrences of the 112 words (after adjustments in nominative and accusative forms) ≈ 67

• Overview of results

- Overall, high-frequency words show lower ratings than low-frequency words.

- In contexts where the change also happens to have natural phonetic motivation, the ratings are high across different frequency contexts and no frequency effect is found.

	Low Frequency	High Frequency	
/i/	3.23	2.25	p<0.01
/i/	3.22	NA	NA
/e/	3.19	1.00	p<0.0001
	/e/ vs. /i/: p=0.4462	/i/ vs. /e/: p<0.005	
	/i/ vs. /i/: p=0.4943		

- Phonetic motivation?

o The change is phonetically unnatural: ts is realized as [dz] due to intervocalic voicing and [dz] → [s] is a phonetically unnatural change.

- Vowel context effect

o In low-frequency words, vowel context does not make any difference in ratings.

o In high-frequency words, the apparent statistically significant vowel context effect seems to be a frequency effect in disguise.

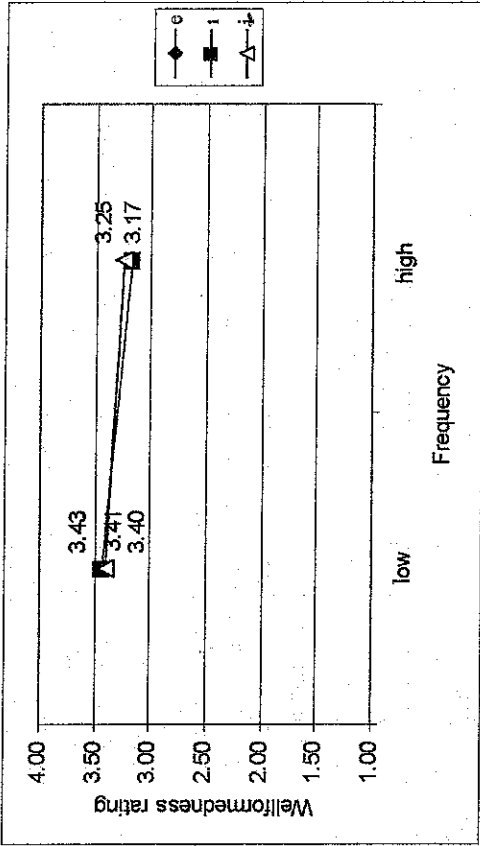
/nats/ 'day' nas-i (75): 2.25 vs. nas-e (336): 1.00

- Frequency Effect

o Statistically significant frequency effect is found for both /i/ and /e/ contexts: higher ratings for low-frequency words

o This is in line with the generalization that analogical change affect low-frequency words first.

• $\text{f}^{\text{h}} \rightarrow \text{s}$ (5 nouns)

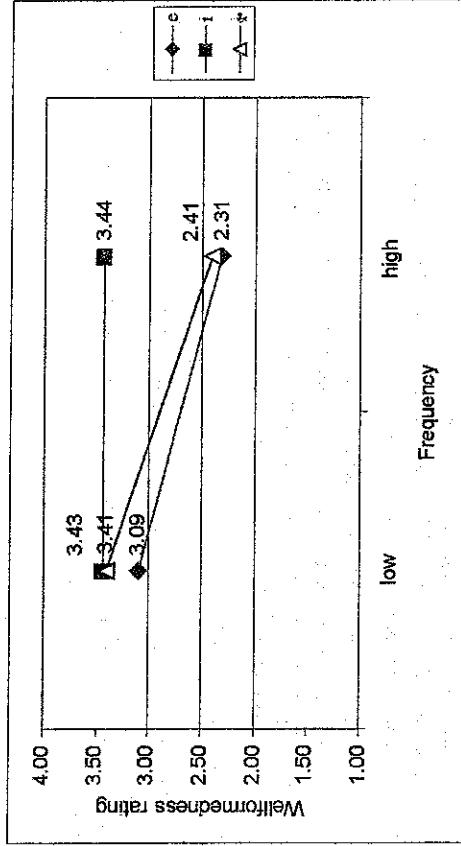


	Low Frequency	High Frequency
/i/	3.43	3.17
/h/	3.40	3.25
/e/	3.41	NA
	/i/ vs. /h/: p=0.4850	/h/ vs. /e/: p=0.4157
	/e/ vs. /i/: p=0.3328	

- Phonetic motivation?

- o $\text{f}^{\text{h}} \rightarrow \text{s}$ is a phonetically natural change (lenition) cf. Kirchner 1998, Lavoie 1999, Bybee 2001
- Vowel context effect
 - o Vowel context does not make any difference.
- Frequency Effect
 - o Overall, low-frequency words show slightly higher ratings than low-frequency words but the difference is not statistically significant.
 - o The change is analogical motivated but also the natural phonetic conditioning expedites the change even in high frequency words. The lack of frequency effect seems to be due to the ceiling effect.

• $\text{t}^{\text{h}} \rightarrow \text{s}$ (8 nouns)



	Low Frequency	High Frequency
/i/	3.43	3.44
/h/	3.41	2.41
/e/	3.09	2.31
	/i/ vs. /h/: p=0.4209	/h/ vs. /h/: p<0.005
	/h/ vs. /e/: p=0.0813	/h/ vs. /e/: p=0.3787
	/i/ vs. /e/: p=0.0717	/i/ vs. /e/: p<0.001

- Phonetic motivation?

- o A direct change of $\text{t}^{\text{h}} \rightarrow \text{s}$ is phonetically unnatural (Kirchner 1998).
- o Unlike in [e] and [i] contexts, in [i] context $[\text{f}^{\text{h}}]$ pronunciation is already established due to regular affrication. Therefore, the change ($\text{t}^{\text{h}} \rightarrow \text{s}$) is a phonetically natural change in [i] context.
- Vowel context effect
 - o Overall, the rating is highest for /i/ context and the lowest for /e/ context.
 - o Statistically significant difference is found only between /i/ vs. /e/ and /h/ in high-frequency words.
- o Only in /i/ context, the change is phonetically natural and hence the change is more advanced than in other contexts.

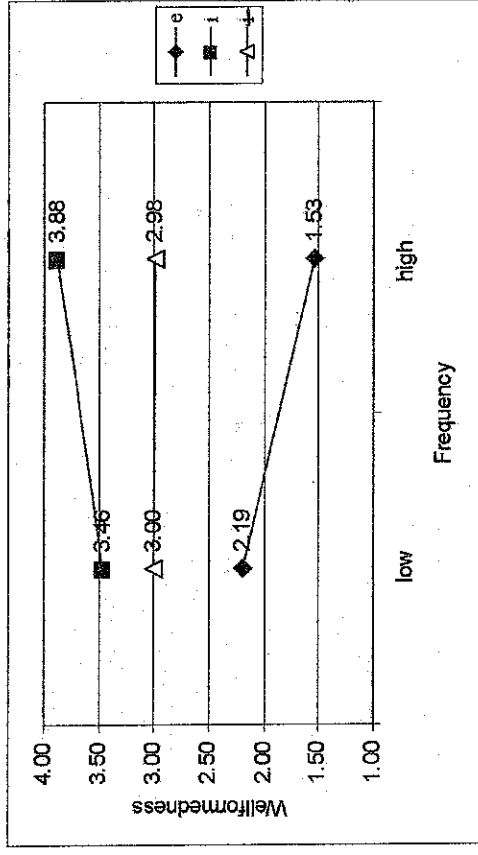
- Frequency Effect

- o [e] and [i] context: statistically significant frequency (i.e. higher ratings for low-frequency words) effect is found.
- o Again, this is in line with the generalization that analogically motivated changes affect low-frequency words first.
- o [i] context: no frequency effect is found.
- o Again, an analogically motivated change with proper phonetic conditioning shows an expedited propagation through the lexicon regardless of frequency of use.

11. Change to [tsʰ]

- tʰ-final nouns show an additional variant with [tsʰ]
- Ratings for [tsʰ]-final pronunciation of tʰ-final nouns in various contexts (all values average of 8 speakers)

• tʰ → [tsʰ] (8 nouns)



	Low Frequency	High Frequency	
/i/	3.46	3.88	p<0.0005
/ɪ/	3.00	2.98	p=0.4718
/e/	2.19	1.53	p<0.05
	/i/ vs. /ɪ/ : p<0.0001		
	/ɪ/ vs. /e/ : p<0.0005		

- Phonetic motivation?

- o Affrication is phonetically natural before [i] and [ɪ] but not before [e]. cf. H. Kim (2001): “[ʃ]top assimilation is motivated by the brief period of turbulence that sometimes occurs at the release of a plosive into a **high** **vocoid**.”

- Vowel context effect

- o Strong effect of vowel context : i > ɪ > e (high vowel > nonhigh vowel) → indicates the phonetic nature of the change.
- o Nevertheless, [e]-suffix forms nevertheless show ratings significantly higher than absolute ungrammaticality.

- Frequency effect

- o [e] context: statistically significant frequency effect (higher ratings for low frequency words) → also a sign of analogical change
- o [tsʰ] pronunciation is (marginally) available, presumably through an analogy to other forms in the paradigm.
- o [ɪ] context: no frequency effect
- o Again, analogical and phonetic motivations converge and expedite the change such that no frequency effect is visible.

- o [i] context: statistically significant frequency effect (higher ratings for low frequency words)

o This could be an indication that the change is purely phonetically motivated. However, the change seems to be already established at least for a few centuries.

- o Rather, this seems to be another side of paradigm leveling effect; the low-frequency words are affected by tʰ pronunciations found in other forms of the paradigm such that the ratings for [tsʰ] pronunciation is relatively lowered.

12. Discussion

- The current data agrees with previous psycholinguistic and historical linguistic studies in indicating that even morphologically regular words are stored in memory and the lexicon keeps track of the frequency of occurrences of these inflected words.
- A novel finding of the current study is that when analogical and phonetic motivations converge on a change, the change is further expedited such that no frequency effect is visible.
- If a sound change affects different words in a paradigm differently, as we have shown here, we expect to find a paradigm where different words reflect different stages of sound change. This is what we find in our data.

Cf. A similar observation made in Albright (2002) based on H. Lee (In preparation).

33

- Since the distribution of allomorphs is determined by a combination of factors including the frequency of occurrence of particular inflected forms, if one were to derive the surface forms from a single Underlying Representation, the rules or constraints would have to refer to information specific to individual inflected words, including the frequency of use. → Undesirable.

- If sound changes affect words in a paradigm differently, why are these 'crazy' paradigms relatively uncommon?

- OO-faithfulness constraints always linger around to ensure that a given morpheme is realized similarly across different contexts.

35

- Average ratings for 5 target pronunciations for $k'it^h$ 'end'

Frequency	NOM (i)	ACC (il)	LOC (e)	DIR (iro)	TOP (in)	COP (ita)
[t ^h]	300	25	1140	220	37	123
[t ^h]	1.13	2.63	4.00	3.50	2.75	1.13
[s ^h]	3.88	3.00	1.25	3.06	2.88	3.88
[s]	3.38	3.63	2.00	2.00	3.75	3.50
[s]	1.13	1.50	1.00	1.13	1.38	1.13
[t]	1.00	1.13	1.00	1.13	1.13	1.13

- Assuming that speakers are most likely to produce the form with the highest well-formedness rating, the paradigm will contain 4 different allomorphs of the noun.

[t^h]: $k'it^h$ -e, $k'it^h$ -iro

[s^h]: $k'is^h$ -i, $k'is^h$ -ita

[s]: $k'is$ -il, $k'is$ -in

[t]: $k'it^h$

34

13. A sketch of OT analysis

- Actual surface forms serve as input to computation.

$k'it^h$ -e → $k'it^h$ -e cf. UR-based $k'it^h$ -e → $k'it^h$ -e

$k'is^h$ -i → $k'is^h$ -i $k'it^h$ -i → $k'is^h$ -i

$k'it$ → $k'it$ $k'it^h$ → $k'it$

....

- Each lexical item projects its own series IO-correspondence constraints (Hammond 1999, Myers 1999, Burzio 2000).

IO-IDENT ([CONT]): IO-IDENT ([CONT]) $_{k'it^h}$ -e

IO-IDENT ([CONT]) $_{k'is^h}$ -i

IO-IDENT ([CONT]) $_{k'it}$

....

- A learning algorithm that is sensitive to input frequency (cf. Boersma and Hayes 2001) can learn a grammar where the ranking of the item-specific IO-

36

faithfulness constraints reflect their frequency: the more frequent the word, the higher its faithfulness constraints.

- In addition to usual OO-correspondence constraints, OO-anticorrespondence constraints (cf. Hayes 1997, Alderete 2002) that “require” alternation in paradigm are posited.

If $t/ _]_{\text{nom}} \# \rightarrow s/ _]_{\text{nom}} V_{\text{suffix}}$ ($t \rightarrow s$, for short)

- Reads: “if a noun ends in [t] in isolation form, change it to [s] before a vowel initial suffix.”
- This constraint is the underlying motivation for the change to [s] in Korean nouns.

14. Summary

- Through a thorough examination of changes occurring in noun-final coronal obstruents in Korean, the paper shows that the analogical sound change affects low-frequency words before high-frequency words.
- The results provide additional evidence that even regularly inflected words may be stored in the lexicon.
- An OT analysis which does not employ an abstract representation (UR) is suggested.
 - The lexicon stores the surface forms of words as such and they serve as input to phonological computation.
 - Item-specific IO-correspondence constraints are projected and their ranking is determined by their frequency of occurrence.
 - By interspersing these IO-correspondence constraints with a constraint that motivates the analogical change, one can simulate the pattern of lexical diffusion found in the data.

- The relative ranking between IO-faithfulness constraint that militates against change to [s] and the anticorrespondence constraint that forces the change determines whether the change occurs or not.

- High-Frequency words: IO-corr. >> Anticorr.

$k'it^h-iro$	IO-IDENT ([CONT]) $k'it^h-iro$	$t \rightarrow s$
[k'itʰ]		
$\text{e} k'it^h-iro$		*
$k'is-iro$	*!	

- Low-Frequency words: Anticorr. >> IO-corr.

$k'it^h-il$	$t \rightarrow s$	IO-IDENT ([CONT]) $k'it^h-il$
[k'itʰ]		
$k'it^h-il$	*!	
$\text{e} k'is-il$		*

→ Eventually, $k'is-il$ will form a lexical representation and $k'is-il$ will be produced from the input $k'is-il$.

Selected References

- Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge Studies in Linguistics 94. Bybee, Joan and Paul Hopper. eds. 2001. *Frequency and the emergence of linguistic structure*. Philadelphia: Benjamins.
- Choi, J. 1986. *Phonological changes in late 19C Colla dialects and its historical significance*. Northern Colla University PhD thesis.
- Hayes, Bruce. 1997. Anticorrespondence in Yiddish and its theoretical consequences, in B. Hermans & M. van Oostendorp eds. *The derivational residue in phonological Optimality Theory*. Amsterdam: John Benjamins. 175-205.
- Kang, Hee Sook. 1993. *A sociolinguistic study of phonological variations and change: with special attention to CangHung, Cennam dialect*. Cempuk University Ph.D. Thesis. [Korean]
- Kim, H. 2001. A phonetically based account of phonological stop assimilation. *Phonology* 18.1.
- Kim, Minsoo, Hocheol Choi and Murim Kim. eds. 1997. *Etymological dictionary of Korean*. Thaeaksa
- Ko, K. 1989. A new analysis of noun final change of $t \rightarrow s$. *Eneuhak* 11. [In Korean]
- Kwak, C. 1984. Spirantization of noun-final coronal consonants. *Kwikkebukmunhak* 91. [In Korean].
- Phillips, Betty. 1984. Word frequency and the actuation of sound change. *Language* 60.2.
- Phillips, Betty. 2001. Lexical diffusion, lexical frequency, and lexical analysis. In Bybee, Joan and Paul Hopper eds., *Frequency and the emergence of linguistic structure*.
- Schuchardt, Hugo. 1885 [1972]. On sound laws: against the neogrammarians.