Homomorphemic Diffusion in Japanese

Mark Irwin
Yamagata University

What is a Homomorph?

- ‘A group of morphemes, all of which are pronounced identically but which are semantically (and, in Japanese, also orthographically) distinct’
- Notation: braced majuscule, e.g. {SEKI}
- English example: {LI:} encompasses <lee>, <lea> (x2), <Leigh>, <Lee>, -<ly>, etc.
- Japanese example: {SEKI} encompasses <石> ‘stone’, <赤> ‘red’, <席> ‘seat’, etc.

Japanese Background

- Modern Japanese consists of four well-defined vocabulary strata: native, Sino- (SJ), foreign and mimetic (Itō (2002), McCawley (1968), Miller (1967), Shibatani (1990)).

- Standard phonemic analyses of Japanese (Bloch (1950), Kawakami (1977), Shibatani (1990), Takeuchi (1999), Vance (1987), etc.) include two 'unusual' phonemes: /Q/ and /N/. /Q/ indicates prolongation or gemination of the succeeding obstruent; /N/ is a nasal; both are moraic.

- Minimal pair ex.: /haka/ 'grave' /haQka/ 'mint'

- Japanese - as far back as Old Japanese - admits only open syllables or those ending in /Q/ or /N/.

Chinese Background

- Middle Chinese (MC), as reconstructed by Pulleyblank (1991) and others, had four tones: level, rising, departing and entering. The latter ended in a voiceless stop *-/p t k/.

- When MC lexemes containing entering tone morphemes were borrowed into Japanese along with sinography in three main waves between the 5th and 14th centuries CE, the final voiceless stop of entering tone morphemes accrued an epenthetic high vowel /i/ or /u/.

The ‘Alleged Irregularity’ Problem I

- Bimoraic SJ morphemes derived from entering tone MC codas in *-/k/ or *-/t/ show a clearly rule-based allomorphy, when they appear as the initial morpheme in a SJ bimorphemic compound whose second morpheme begins in /k/-. The elision of their final high vowel is accompanied by assimilation and concomitant lengthening of the preceding consonant to the initial /k/- of the following morpheme, i.e. generation of the mora obstruent /Q/.

-/ku/ + /k/- > /Q.k/:	/syaku/	+	/kiN/	>	/syaQ.kiN/
	借		金		借金
	‘borrow(ing)’		‘money’		‘debt’
-/tu/ + /k/- > /Q.k/:	/butu/	+	/kyoo/	>	/buQ.kyoo/
	仏		教		仏教
	‘Buddha’		‘teach(ing)’		‘Buddhism’
-/ti/ + /k/- > /Q.k/:	/niti/	+	/ki/	>	/niQ.ki/
	日		記		日記

The ‘Alleged Irregularity’ Problem II

- Apparently irregular allomorphic behaviour exists within bimoraic SJ morphemes ending in **-/ki/**, when they appear as the initial morpheme in a SJ bimorphemic compound whose second morpheme begins in /k/-.

-/ki/ + /k/- > /Q.k/: /seki/ + /kai/ > /seQ.kai/

石 灰 石灰

‘stone’ ‘ash’ ‘quicklime’

-/ki/ + /k/- > /ki.k/: /heki/ + /kuu/ > /heki.kuu/

碧 空 碧空

‘azure’ ‘sky’ ‘azure sky’

-/ki/ + /k/- > /Q.k/ ~ /ki.k/: /teki/ + /kaku/ > /teQ.kaku/ ~ /teki.kaku/

適 格 適格

The ‘Alleged Irregularity’ Problem: Summary

- The ‘alleged irregularity’ problem is confined only to bimoraic SJ morphemes ending in -/ki/.

- Irregularity appears even with the same initial sinograph:

石鹸	/seQ.keN/	(*/seki.keN/)	‘soap’
石国	/seki.koku/	(*/seQ.koku/)	toponym
石果	/seQ.ka/ ~	/seki.ka/	‘stone fruit’

- Phonology, accent, orthography and semantics all appear to play no role in this irregular allomorphy.

The Synchronic Corpus I

- Uses 4 major dictionaries: 2 monolingual, 1 bilingual and 1 sinographic.

- Each type was assigned a ‘rate of occurrence of /Q/’ (notated as *n*Q), indicative of its /Q/-fulness or /Q/-lessness. Initially set to zero, this was incremented by 1 point each time a token was listed in a dictionary in /Q/-ful form and decremented when listed in /Q/-less form.

The Synchronic Corpus II

- There are 16 extant SJ -{KI} homomorphs, but of these only 10 actually occured in the Corpus.

- In total, there were 269 types in the Corpus, appearing in over 574 dictionary entry tokens (an average of 2.13 dictionary entry tokens per type). The level of /Q/-generation varied hugely across homomorphs.

Synchronic Corpus Results I

Homo-morph	No. of types which are				No. of dictionary entry tokens which are			*n*Q
	/Q/-ful	*/Q/-var.*	*/Q/-less*	*Total*	*/Q/-ful*	*/Q/-less*	*Total*	
BEKI	0	0	5	5	0	7	7	-1.000
EKI	0	4	21	25	4	46	50	-.840
GEKI	1	4	13	18	13	30	43	-.395
HEKI	1	0	25	26	1	45	46	-.957
REKI	1	1	11	13	2	19	21	-.810
RIKI	0	1	2	3	3	6	9	-.333
SEKI	100	10	16	126	220	38	258	+.705
SIKI	0	1	15	16	3	32	35	-.829
TEKI	2	21	12	35	46	57	103	-.107
ZIKI	0	0	2	2	0	2	2	-1.000
Total	**105**	**42**	**122**	**269**	**292**	**282**	**574**	**+.017**

Synchronic Corpus Results II

- The only homomorph to show a strong tendency towards /Q/-generation is {SEKI}.
- {TEKI} has a strong tendency towards variation (/Q/ ~ non-/Q/ doublets).
- The remaining homomorphs are either statistically valueless ({BEKI, RIKI, ZIKI}) or exhibit a strong tendency towards /Q/-lessness ({EKI, HEKI, REKI, SIKI}).

Diachronic Corpus

- Historical sources were dictionaries, thesauri, etc. predating the 20th century and postdating the 11th (before which no orthography was used for /Q/).
- In all, 20 sources were employed, from the 1165 *Iroha Jiruishō* to Brinkley's 1896 *Japanese-English Dictionary*.
- The types contained in these 20 works were grouped into 6 consolidated historical periods: Meiji (1867-99), Late Edo (1750-1866), Early Edo (1610-1750), 16th century, Late Middle Japanese (1200-1500) and Early Middle Japanese (1100-1200).

Diachronic Corpus Results I

Period	Number of Types			
	/Q/-ful	*/Q/-var.*	*/Q/-less*	**Total**
Meiji	21	12	24	57
Late Edo			9	9
Early Edo			1	1
c. 16th century	1		13	14
Late Middle Japanese			6	6
Early Middle Japanese			4	4
Total	**22**	**12**	**57**	**91**

Diachronic Corpus Results II

- There is only one isolated case of a /Q/-ful type and no cases whatsoever of /Q/-variant types prior to Meiji.

- There are no types appearing as /Q/-less in the Synchronic Corpus which occur as either /Q/-ful or /Q/-variant types in the Diachronic Corpus.

- Based on the conclusions drawn from the two Corpora, the following sound change, beginning in the early 19th century, can be posited:

  {(C)Vki} → {(C)VQ} / ____ {k(y)V(C)(V))}

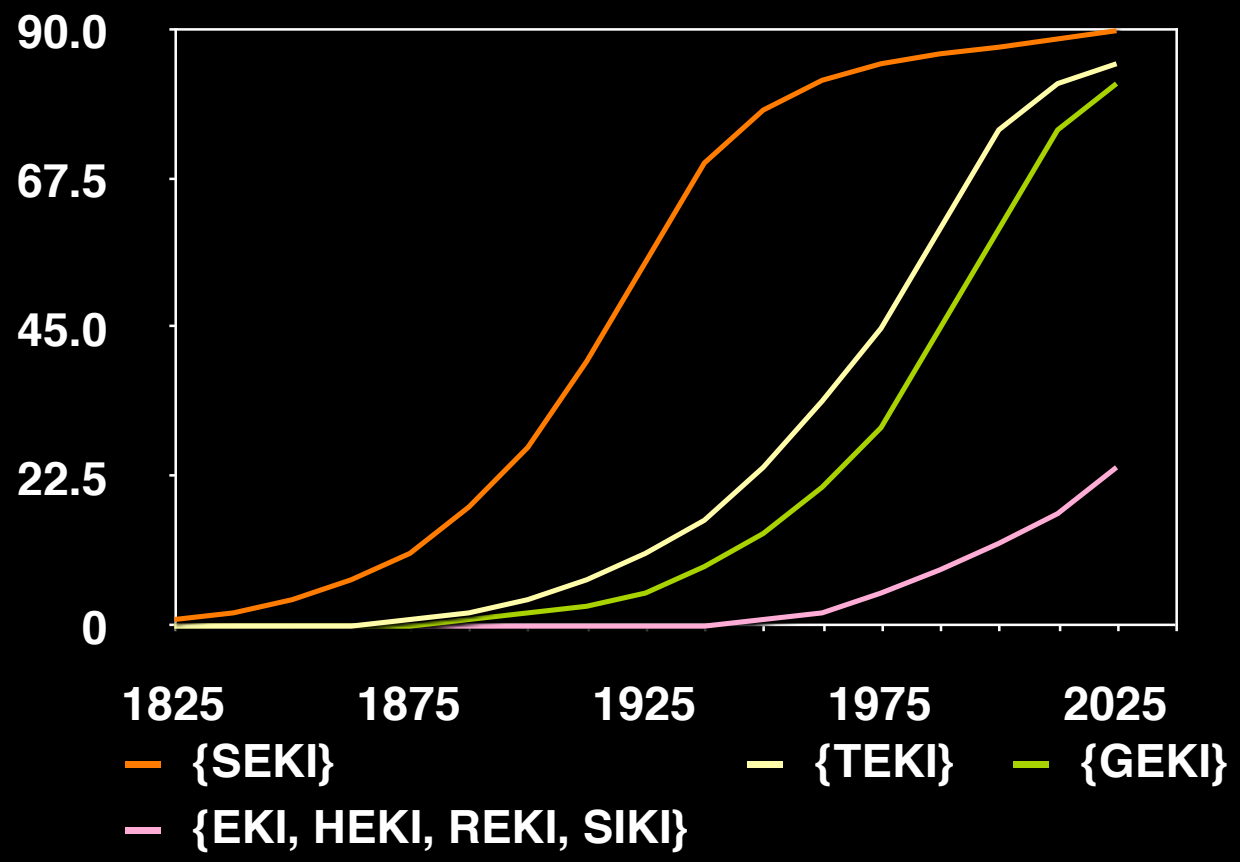
Lexical Diffusion

- Term first employed by Wang (1969), although see also Schuchardt (1885), Sturtevant (1917) and Sommerfelt (1962) for earlier, albeit nameless, models.
- Contrary to the older neogrammarian view (i.e. the 'regularity hypothesis' first exemplified by Osthoff & Brugmann (1878)), sound change is seen as 'phonetically abrupt and lexically gradual'.
- Recent significant studies include Blevins (2004:268-278), Bybee (2000, 2002), de Oliveira (1991), Krishnamurti (1998), Labov (1994:419-501), Phillips (1998, 2001) & Shen (1990).

Homomorphemic Diffusion

- When the data in the two Corpora are examined in the light of Labov's (1981) criteria for differentiating Neogrammarian sound change from lexical diffusion, it is clear we are dealing with the latter.
- However, an adequate account, and an adequate motivation, can only be found in proposing the hitherto unrecognized phenomenon of homomorphemic diffusion: Labov's (1981) criteria must be viewed under a homomorphemic, rather than a lexical or morphemic, light in order to yield the correct output.
- Entire groups of homophonous morphemes are evincing the spread of an identical sound change to a differing degree from other groups of homophonous morphemes.

Diffusion as an S-Curve



Why {SEKI}?

- Bybee (2000, 2001, 2002), Hopper (1976) and Phillips (1984, 1998, 2001) have claimed that, in cases of lexical diffusion involving a reductive sound change, higher frequency lexemes have a tendency to be affected earlier and more thoroughly as a consequence of automation in neuromotor activity brought about by repetition.

- Given that the sound change I have proposed is assimilatory and thereby reductive, we might therefore expect, if Bybee/ Hopper and Phillip's claims as well as my hypothesis on the independent status of the homomorph are both correct, that the most frequent homomorphs will exhibit the change earlier and more thoroughly.

Why {SEKI}?

Homomorph	No. types	No. dictionary entries	*n*Q
SEKI	126	258	+.705
TEKI	35	103	-.107
EKI	25	50	-.840
HEKI	26	46	-.957
GEKI	18	43	-.395
SIKI	16	35	-.829
REKI	13	21	-.810
RIKI	3	9	-.333
BEKI	5	7	-1.000
ZIKI	2	2	-1.000
r	**0.84**	**0.88**	

References

- Bloch, Bernard. (1950). Studies in colloquial Japanese IV: Phonemics. *Language* 26:86-125.
- Blevins, Juliette. (2004). *Evolutionary Phonology*. Cambridge: Cambridge University Press.
- Bybee, Joan. (2000). The Phonology of the Lexicon: Evidence From Lexical Diffusion. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based Models of Language*. Stanford: CSLI Publications. 65-85.
- Bybee, Joan. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Bybee, Joan. (2002). Word Frequency and Context of Use in the Lexical Diffusion of Phonetically Conditioned Sound Change. *Language Variation and Change* 14:261-290.
- de Oliveira, Marco Antonio. (1991). The Neogrammarian Controversy Revisited. *International Journal of the Sociology of Language* 89:93-105.
- Hopper, Joan B. (1984). Word Frequency in Lexical Diffusion and the Source of Morphophonological Change. In William Christie (ed.), *Current Progress in Historical Linguistics*. Amsterdam: North Holland. 96-105.
- Itō, Masamitsu. (2002). Goi no Ryōtekiseikaku [The Quantitative Nature of the Lexicon]. In Michiaki Saitō (ed.), *Asakura Nihongo Kōza 4: Goi, Imi [Asakura Japanese Course Vol. IV: Vocabulary and Meaning]* . Tokyo: Asakura Shoten. 29-53.
- Kawakami, Shin. (1977). *Nihongo Onsei Gaisetsu [A Survey of Japanese Phonetics]*. Tokyo: Ōfūsha.
- Krishnamurti, Bhadriraju. (1998). Regularity of Sound Change Through Lexical Diffusion: A Study of

References (cont.)

- Miller, Roy Andrew. (1967). *The Japanese Language*. Chicago: Chicago University Press.
- Osthoff, Hermann & Brugmann, Karl. (1878). *Morphologische Untersuchen auf dem Gebiet der indogermanischen Sprachen, Vol. 1*. Leipzig.
- Phillips, Betty. (1984). Word Frequency and the Actuation of Sound Change. *Language* 60:320-342
- Phillips, Betty. (1998). Word Frequency and Lexical Diffusion in English Stress Shifts. In Richard Hogg & Linda van Bergen (eds.), *Historical Linguistics 1995, Vol. II: Germanic Linguistics*. Amsterdam: John Benjamins. 223-232.
- Phillips, Betty. (2001). Lexical Diffusion, Lexical Frequency, and Lexical Analysis. In Joan Bybee & Paul Hopper (eds.), *Frequency and Emergence of Linguistic Structure*. Amsterdam: John Benjamins. 123-136.
- Schuchardt, Hugo. (1885 [1972]). *Über die Lautgesetze*. Berlin: Verlag von Robert Oppenheim.
- Shen, Zhongwei. (1990). Lexical Diffusion: A Population Perspective and a Numerical Model. *Journal of Chinese Linguistics* 18:159-200.
- Shibatani, Masayoshi. (1990). *The Languages of Japan*, Cambridge: Cambridge University Press.

- Schuchardt, Hugo. (1885 [1972]). *Über die Lautgesetze*. Berlin: Verlag von Robert Oppenheim.

- Sommerfelt, Alf. (1962). *Diachronic and Synchronic Aspects of Language*. 's-Gravenhage: Mouton.
- Sturtevant, Edgar. (1917). *Linguistic Change: An Introduction to the Historical Study of Language*. Chicago: University of Chicago Press.
- Takeuchi, Leone. (1999). *The Structure and History of Japanese: From Yamatokotoba to Nihongo*, Harlow, Essex: Pearson Education.
- Vance, Timothy. (1987). *An Introduction to Japanese Phonology*. Albany: State University of New York Press.