

Probing knowledge of similarity through puns

Shigeto Kawahara
kawahara@rci.rutgers.edu

Rutgers

University of Delaware, March 20th 2009

- 1 Introduction
 - Setting up the stage
 - An example from Japanese
 - 2 Part I: Consonant correspondence
 - Introduction
 - Method
 - Results
 - Evidence for perceptual similarity
 - 3 Part II: Vocalic similarity
 - Introduction and method
 - Results
 - 4 Part III: Positional effects
 - General Introduction
 - Experiment I: Initial vs. internal
 - Experiment II: Long vs. short
 - Discussion
 - 5 General discussion
 - Summary
 - 6 Acknowledgements
- References51

Theoretical background

- Some kind of similarity seems important in shaping phonological patterns.
- Intuitively, speakers can change /A/ to [B] if A and B are “sufficiently similar” (Steriade, 2001, *et seq*).
- But what do we really mean by “sufficiently similar”?
- More formally, what is knowledge of similarity that speakers deploy to shape phonological patterns?

Perceptual similarity

- Speakers can change /A/ to [B] if A and B are **perceptually similar** (Steriade, 2001).
- Japanese voicing example (Nishimura, 2003). If there are two voiced obstruents within the same stem:

boddo → batto	<i>cf.</i> bado → * bato
baggu → bakku	<i>cf.</i> bagu → * baku
doggu → dokku	<i>cf.</i> dagu → * daku

A hypothesis (Kawahara, 2006, 2008)

- A voicing contrast is less perceptible in geminates than in singletons.
 - ▶ DD is sufficiently perceptually similar to TT.
 - ▶ D is perceptually dissimilar to T.
- There may be something special about voicing in voiced geminates.

The acoustic evidence

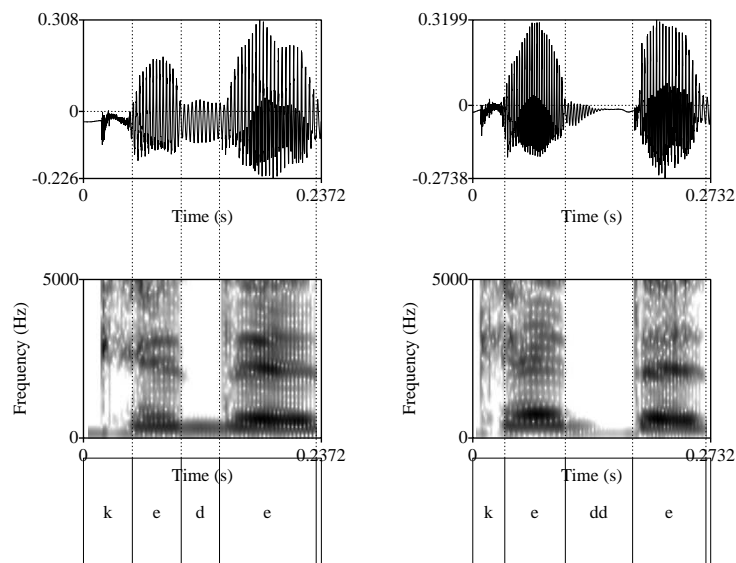


Figure: Closure voicing

The perceptual evidence (Kawahara, 2006)

- An identification experiment under noise.
- 17 native speakers of Japanese judged the voicing quality of intervocalic consonants.
- d' , which represents a perceptual distance, is calculated for voicing contrasts in singletons and geminates:
- $d'(\text{sing})=3.79$ vs. $d'(\text{gem})=.71$.

Other examples

- Nasals are more likely to assimilate in place than oral consonants (Jun, 1995): nasal pairs are sufficiently similar whereas oral pairs are “too different”.
- A voicing contrast is more likely to neutralize than other manner contrasts (Steriade, 2001): minimal pairs differing in voicing are similar enough to be interchangeable.

Summary

- The less the perceptual change, the more likely it occurs.
- In other words, speakers exert stronger grammatical pressure against a larger perceptual change.
- I attempt to provide further support for this principle from the perspective of imperfect puns (*dajare*).

Part I: Consonantal similarity (Kawahara & Shinohara, 2009)

- Now we turn to puns...
- Japanese speakers traditionally create puns by combining two similar sounding words or phrases.
- Examples: *Arumikan-no ue-no aru mikan*, *Aizusan-no aisu*, *okosama-o okosanai-de*.
- Corresponding sounds do not need to be identical, but nevertheless need to be “similar” (Cutler & Otake, 2002; Shinohara, 2004; Kawahara, 2007; Steriade, 2003; Zwicky & Zwicky, 1986).

Question and approach

- Again, what do we mean by “similar”?
- Perceptually similar? *as per* Steriade (2001)?
- A traditional ‘introspection-based’ approach does not seem appropriate to answer this question (Shattuck-Hufnagel, 1986).
- One can (only?) answer this question with a statistical/experimental approach.

Method

- We collected 2371 examples of imperfect puns from online websites as well as through elicitation.
- We defined pun domains as sequences of moras with matching vowels: **okosama**-o **okosana**ide.
- We counted a number of mismatched consonant pairs (e.g. [m-n] pair in the above example; 535 consonant pairs in total—other pairs are discussed later).

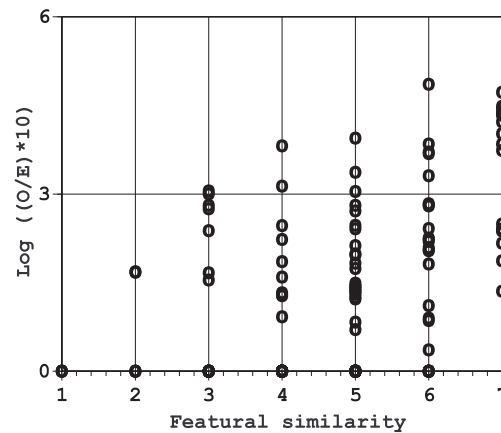
O/E ratios

- We need a measure of combinability of two consonants—the likelihood of two consonants making a pun pair.
- Suppose that an [A]–[B] pair occurs more often than an [A]–[C] pair, can we say that the [A]–[B] pair is more combinable than the [A]–[C] pair?
- Not necessarily, because [B] might simply be more frequent than [C].
- O/E values provide a nice measure. They are observed frequency relativized with respect to expected frequency.

O/E ratios cont'd

- If $O/E > 1$, the pair is observed more often than expected (overrepresentation).
- If $O/E < 1$, the pair is observed less often than expected (underrepresentation).
- To statistically establish the relation between similarity and combinability in puns, we started out with a simple version of featural similarity.
- Featural similarity is defined in terms of the number of shared feature specifications (Bailey & Hahn, 2005, and others).

Correlation



Combinability and (featural) similarity correlate with each other
($r_s = .497, p < .001$)

Evidence for perceptual similarity

Featural similarity accounts for some broad patterns, but it ultimately fails to capture some of the detailed aspects of pun pairings:

- 1 Contextual effect of [place].
- 2 Different saliency of different features.
- 3 The effect of a phonologically inert feature.
- 4 Similarity with ϕ .

Evidence I: Contextual effects

- Nasals are more likely to assimilate in place than oral consonants (Mohan, 1993).
- A [place] contrast is known to be less salient in nasal consonants than in oral consonants (Boersma, 1998; Jun, 1995; Pols, 1983; Mohr & Wang, 1968).
- The parallel in the pun pattern:

m-n: 8.85	b-d: 1.09	p-t: 1.11
	b-g: .65	p-k: 1.08
	d-g: .39	t-k: .87

Evidence I: Contextual effects

- Speakers indeed consider the [m-n] pair as more similar than any other kind of oral consonant minimal pair.
- The contribution of [place] to similarity differs depending on whether it is hosted by nasals or oral consonants.
- The featural similarity model cannot explain this observation.

Evidence 2: Saliency of different features

- A [voice] contrast is less perceptible than [nas] and [cont] contrasts (Steriade 2001 and other earlier psycholinguistic works).
- Again the parallel in pun pattern:

cont		nasal		voice	
p-φ:	5.58	b-m:	4.68	p-b:	8.51
t-s:	.90	d-n:	1.12	t-d:	7.64
d-z:	1.68			k-g:	8.03
				s-z:	11.3
				ʃ-ʒ:	6.81

Evidence 3: Effect of phonologically inert features

	Voiced obstruents	Voiceless obstruents
Paired with sonorants	63 (18.2%)	30 (6.0%)
Total	346	497

- Speakers combine sonorants more often with voiced obstruents than with voiceless obstruents.

But phonologically...

- The [+voice] feature of Japanese sonorants is inert: it does not trigger Lyman's Law:
 - ▶ Japanese does not allow two occurrences of voiced consonants within single stems. But only voiced obstruents count as voiced consonants.
 - ▶ The [+voice] feature on sonorants is underspecified (Itô & Mester, 1986), sonorants do not bear [voice] at all (Mester & Itô, 1989) or sonorants and obstruents bear different [voice] features (Rice, 1993).

Evidence 4: Similarity with ϕ .

- The list of consonants that often correspond with ϕ :
[w]: 6.25, [r]: 4.59, [h]: 3.72, [m]: 2.54, [n]: 1.49, [k]: 1.39.
- These are consonants that blend into their environments, and hence are perceptually close to ϕ .
 - ▶ [w] blends into surrounding vowels (Myers & Hansen, 2005).
 - ▶ [r] is very short (Nakamura, 2002).
 - ▶ [h] again blends into surrounding vowels (Keating, 1988).
 - ▶ [m] and [n] have blurry transitions (Downing, 2005).
 - ▶ [k] coarticulates with surrounding vowels (de Lacy & Kingston, 2006).

Summary of Part I

- Speakers minimize the perceptual disparities between two corresponding elements in puns.
- Here we find non-trivial parallels with purely phonological patterns.
 - ▶ Place assimilation asymmetry and the perceptibility of nasal and oral consonants.
 - ▶ Neutralizability of [voice] with respect to other features.
 - ▶ Speakers epenthesize perceptually least intrusive segments.

Part II: Vowel mismatch (Kawahara & Shinohara, 2008)

- Speakers can combine two words minimally different in one vowel.
- For example, *shibuya-wa shibiiya* and *haiidegaa-no zense-wa haedekka*.
- We collected 547 examples of such mismatched vowels, calculated the O/E ratios for each pair, and took their reciprocals as distance between the five vowels.

Vocalic distances

Distances of the five vowels are calculated as the reciprocals of O/E ratios.

Table: The O/E ratios of the five vowels.

	a	e	o	i	u
a	0	1.60	2.13	0.72	0.78
e		0	0.74	1.90	0.55
o			0	0.46	1.54
i				0	2.06
u					0

Vowel mismatch

Two dimensional representations of the distance matrix:

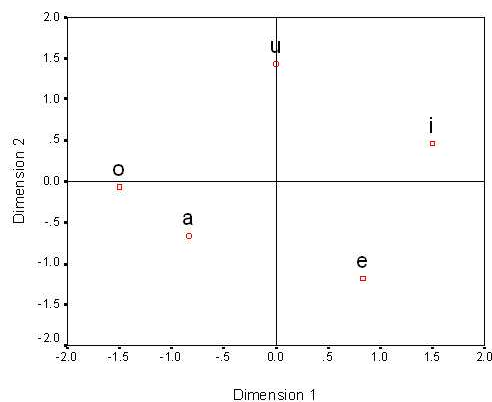


Figure: An MDS analysis of the vowel distance matrix computed based on combinability in puns.

Vowel mismatch

- Evidence for psychoacoustic similarity.
 - ① [a] is closer to [o] than [e].
 - ② [i] and [u] are closer to each other than [o] and [e].
- These patterns match with acoustic properties of Japanese vowels (Keating & Huffman, 1984; Nishi, Strange, Akahane-Yamada, Kubo, & Trent-Brown, 2008)
- The distance matrix statistically correlates with a distance matrix created based on euclidean distances of F1/F2 bark values of Hirahara and Kato (1992) ($r = .647, p = .043$).

Part III: Positional effects (Kawahara & Shinohara, to appear)

- We have seen that speakers minimize the differences between corresponding segments in imperfect puns, just as in phonology.
- Then it may be that positions of mismatches matter: in phonology, speakers avoid mismatches between inputs and outputs in phonetically and psycholinguistically prominent positions (Beckman, 1998; Steriade, 1994).

Experiment I: Psycholinguistic prominence

The first experiment tested whether speakers avoid mismatches in initial positions. Initial syllables play an important role in word recognition.

- Hearing initial portions of words help listeners to retrieve the whole words (Horowitz, Chilian, & Dunnigan, 1969; Horowitz, White, & Atwood, 1968).
- In “tip-of-the-tongue” phenomena, speakers can guess the first sound more accurately than non-initial sounds (A. Brown, 1991; R. Brown & MacNeill, 1966).
- Also, in tip-of-the-tongue situations, initial sounds help retrieve the whole word (Freedman & Landauer, 1966).

Psycholinguistic prominence cont'd

- Listeners are faster when detecting mispronunciations in non-initial positions (Cole & Jakimik, 1980; Cole, 1973)—once they hear initial syllables, they anticipate what's coming next.
- Sound symbolism—particular images associated with particular sounds—is stronger word-initially than non-word-initially (Bruch, 1986; Kawahara, Shinohara, & Uchimoto, 2008).

Phonological privilege of initial positions

Initial syllables exhibit a privileged status in phonology.

- In Sino-Japanese, while initial syllables can contain a variety of consonants, second syllables only allow [t] and [k] (Kawahara, Nishimura, & Ono, 2002; Tateishi, 1990).
- If there were an underlying form like /sasʊ/ (*as per* Richness of the Base), then speakers avoid changing the initial [s] but not the final [s] (perhaps to [satʊ]).
- In other words, speakers would avoid making changes in initial syllables.

Correspondence Theory

In terms of Correspondence Theory (McCarthy & Prince, 1995):

In phonology (input-output correspondence):

Input	/	s _i	a _j	s _k	u _l	/
Output	[s _i	a _j	t _k	u _l]

Likewise in pun formation (surface-to-surface correspondence):

Word 1	[s _i	a _j	s _k	u _l]
Word 2	[s _i	a _j	t _k	u _l]

Method 1

- The experiment was a wellformed judgement task.
- The stimuli were pairs of words that contain a pair of sounds that minimally differ in voicing ([t-d], [d-t], [s-z], [z-s], [k-g], [g-k]).
- To control for the distance between corresponding words, they were always separated by one-mora particle.
- Two conditions:
 - ▶ Initial mismatches (e.g. *sasetsu-ni zasetsu* 'I gave up turning left').
 - ▶ Internal mismatch (e.g. *hisashi-ni hizashi* 'Sunlight on the sun roof').

Method 2

- We asked two questions: how funny it is and how acceptable it is as a pun pair in a 1-4 scale.
- We included the first question, so that the participants would tease apart these questions.
- The questionnaire started with two sample questions, with one example which is clearly an example of a Japanese imperfect pun and one example which clearly is not.
- 37 speakers participated in this study, but we excluded eight of them because they did not consider the good example as a good pun or considered the bad example as a good pun.

Result

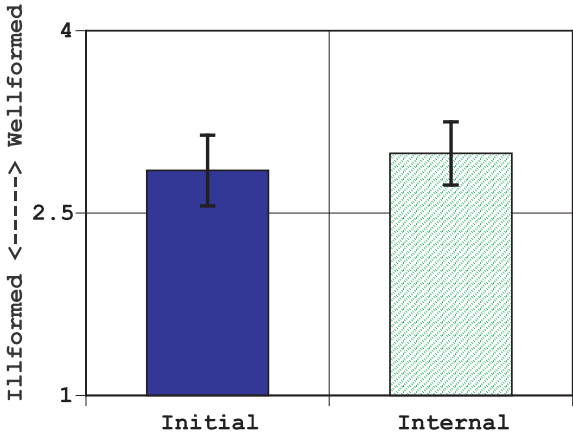


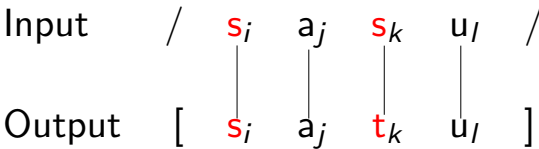
Figure: Wellformedness of puns with initial mismatches and those with internal mismatches. The error bars = 95% CIs.

Speakers judged mismatches in initial syllables less acceptable than those in non-initial syllables ($z = 2.59, p = .01$ by a non-parametric Wilcoxon signed-ranks test).

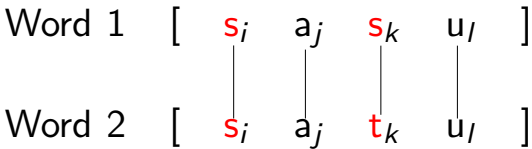
Discussion

Speakers avoid mismatches in a psycholinguistically prominent position, both in phonology and pun formation.

Phonology



Pun formation



The principle of positional faithfulness is observed both in puns and in phonology.

Experiment II: Introduction

The second experiment tested whether speakers avoid mismatches in long vowels.

- Long vowels are, by definition, phonetically long.
- Different long vowels are more different from each other than different short vowels (Steriade, 2003)—an [aa]-[ii] pair is more different than an [a]-[i] pair.
- A change in long vowels would be more perceptible also because speakers hyperarticulate long vowels more than short vowels. As a result, long vowels are more dispersed than short vowels (Hirata & Tsukada, 2003; Hisagi, Nishi, & Strange, 2008).

Phonological privilege of long vowels

- Hindi for example allows a surface nasality contrast in long vowels, but not in short vowels (Steriade, 1994).
- A hypothetical underlying /tããtã/ would map to [tããta].
- In phonology speakers avoid making changes—or neutralizing contrasts—more in long vowels than in short vowels.

Correspondence Theory again

In phonology (input-output correspondence):

Input / t_i **ãã**_j t_k **ã**_l /
Output [t_i **ãã**_j t_k **a**_l]

In pun formation (surface-to-surface correspondence):

Word 1 [t_i **ãã**_j t_k **ã**_l]
Word 2 [t_i **ãã**_j t_k **a**_l]

Method

- The design had two fully crossed factors: 10 vowel combinations ([a-i], [a-u], [a-e], [a-o], [i-u], [i-e], [i-o], [u-e], [u-o], [e-o]) × 2 lengths (short vs. long).
- An example of a crucial pair was: *jookuu-no jookaa* 'A joker in the sky' vs. *rippu-ga rippa* 'Lips are good'.
- Other details were identical to Experiment 1, except that we had four sample questions (two good examples and two bad examples).
- 26 speakers participated in the study.

Result

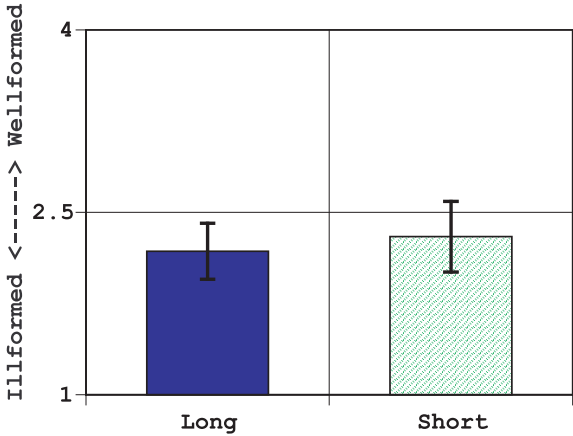


Figure: Wellformedness of puns with long vowel mismatches and short vowel mismatches.

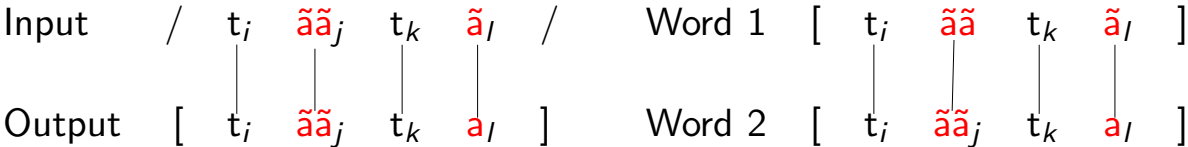
Speakers rated those with long mismatches as worse than short mismatches (2.93, $p < .001$).

Discussion

Japanese speakers avoid mismatches in long vowels more than mismatches in short vowels

Phonology

Pun formation



Mismatches in long vowels are perceptually salient because of their long duration (Steriade, 2003), and hence avoided by the participants.

Summary and discussion of Part III

- Speakers avoid disparities between corresponding elements more in prominent positions (initial syllables, long vowels) than in non-prominent positions.
- The same principle holds both in puns and in phonology.
- We again found non-trivial parallels between phonology and pun patterns.

General summary and discussion

- Japanese pun patterns have been understudied, but they have much to tell us about our linguistic knowledge of similarity.
- I hope our research stimulates further studies on Japanese puns.

Overall summary

- Speakers minimize perceptual disparities between corresponding segments in puns.
- In this sense, we find non-trivial parallels between pun pairing patterns and phonological patterns.
- We can **probe our knowledge of similarity through studying puns**.
- Pun patterns are understudied—they thus provide an untapped resource for future research.

Acknowledgements

Part I is based on Kawahara & Shinohara (2009). Part II is reported in more detail in Kawahara & Shinohara (2008). Part III is based on Kawahara & Shinohara (to appear) and Kawahara, Shinohara, & Yoshida (2008), which themselves are based on a BA thesis by Nobuhiro Yoshida. I would like to thank Donca Steriade, whose colloquium talk at UMass Amherst inspired me to take on this general project. I also owe to John Kingston, John McCarthy, Joe Pater, and anonymous reviewers of several journals for their comments. There is a summary website for this project on the author's website. The website contains many open issues for future research.

- Bailey, T., & Hahn, U. (2005). Phoneme similarity and confusability. *Journal of Memory and Language*, 52, 339-362.
- Beckman, J. (1998). *Positional faithfulness*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.
- Boersma, P. (1998). *Functional phonology: Formalizing the interaction between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Brown, A. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, 109(2), 204-223.
- Brown, R., & MacNeill, D. (1966). The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5(4), 325-337.
- Bruch, J. (1986). Expressive phonemes in Japanese. *Kansas Working Papers in Linguistics*, 11, 1-8.
- Cole, R. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics*, 13, 153-156.
- Cole, R., & Jakimik, J. (1980). How are syllables used to recognize words? *Journal of Acoustical Society of America*, 67(3), 965-970.
- Cutler, A., & Otake, T. (2002). Rhythmic categories in spoken-word recognition. *Journal of Memory and Language*, 46(2), 296-322.

- de Lacy, P., & Kingston, J. (2006). *Synchronic explanation*. (ms. University of Massachusetts and Rutgers University)
- Downing, L. (2005). On the ambiguous segmental status of nasal in homorganic nc sequences. In M. van Oostendorp & J. M. van der Weijer (Eds.), *The internal organization of phonological segments* (p. 183-216). Berlin: Mouton de Gruyter.
- Freedman, J., & Landauer, T. (1966). Retrieval of long-term memory: "tip-of-the-tongue" phenomenon. *Psychonomic Science*, 4(8), 309-310.
- Hirahara, T., & Kato, H. (1992). The effect of F0 on vowel identification. In Y. Tohkura, E. V. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech perception, production and linguistic structure* (p. 89-112). Tokyo: Ohmsha.
- Hirata, Y., & Tsukada, K. (2003). The effects of speaking rates and vowel length on formant movements in Japanese. In A. Agwuele, W. Warren, & S.-H. Park (Eds.), *Proceedings of the 2003 Texas Linguistic Society Conference* (p. 73-85). Somerville, MA: Cascadilla Press.
- Hisagi, M., Nishi, K., & Strange, W. (2008). Acoustic properties of

Japanese and English vowels: Effects of phonetic and prosodic context. In M. Endo-Hudson, S.-A. Jun, P. Sells, P. M. Clancy, S. Iwasaki, & S. Sung-Ock (Eds.), *Japanese/Korean linguistics 13*. Stanford: CSLI.

Horowitz, L., Chilian, P., & Dunnigan, K. (1969). Word fragments and their reintegrative powers. *Journal of Experimental Psychology*, 80(2), 392-394.

Horowitz, L., White, M., & Atwood, D. (1968). Word fragments as aids to recall: The organization of a word. *Journal of Experimental Psychology*, 76(2), 219-226.

Itô, J., & Mester, A. (1986). The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry*, 17, 49-73.

Jun, J. (1995). *Perceptual and articulatory factors in place assimilation: An optimality theoretic approach*. Unpublished doctoral dissertation, University of California, Los Angeles.

Kawahara, S. (2006). A faithfulness ranking projected from a perceptibility scale: The case of voicing in Japanese. *Language*, 82(3), 536-574.

Kawahara, S. (2007). Half-rhymes in Japanese rap lyrics and knowledge of

similarity. *Journal of East Asian Linguistics*, 16(2), 113-144.

Kawahara, S. (2008). Phonetic naturalness and unnaturalness in Japanese loanword phonology. *Journal of East Asian Linguistics*ast Asian linguistics, 18(4).

Kawahara, S., Nishimura, K., & Ono, H. (2002). Unveiling the unmarkedness of Sino-Japanese. In W. McClure (Ed.), *Japanese/Korean linguistics 12* (p. 140-151). Stanford: CSLI.

Kawahara, S., & Shinohara, K. (2008). *Calculating vocalic similarity through puns*. ms (Rutgers University and Tokyo University of Agriculture and Technology).

Kawahara, S., & Shinohara, K. (2009). The role of psychoacoustic similarity in Japanese puns: A corpus study. *Journal of Linguistics*, 45(1), 111-138.

Kawahara, S., & Shinohara, K. (to appear). Phonetic and psycholinguistic prominences in pun formation: Experimental evidence for positional faithfulness. In W. McClure & M. den Dikken (Eds.), *Japanese/Korean linguistics 18*. Stanford: CSLI.

Kawahara, S., Shinohara, K., & Uchimoto, Y. (2008). A positional effect in sound symbolism: An experimental study. In *Proceedings of the*

- japan cognitive linguistics association 8* (p. 417-427). Tokyo: JCLA.
- Kawahara, S., Shinohara, K., & Yoshida, N. (2008). *Positional effects in Japanese imperfect puns*. (A talk presented at Language, Communication, and Cognition (Brighton University, August 4th, 2008))
- Keating, P. A. (1988). Underspecification in phonetics. *Phonology*, 5, 275-292.
- Keating, P. A., & Huffman, M. (1984). Vowel variation in Japanese. *Phonetica*, 41, 191-207.
- McCarthy, J., & Prince, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Walsh Dickey, & S. Urbanczyk (Eds.), *University of Massachusetts occasional papers in linguistics 18* (p. 249-384). Amherst: GLSA.
- Mester, A., & Itô, J. (1989). Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language*, 65, 258-93.
- Mohanan, K. P. (1993). Fields of attraction in phonology. In J. Goldsmith (Ed.), *The last phonological rule: Reflections on constraints and derivations* (p. 61-116). Chicago: University of Chicago Press.
- Mohr, B., & Wang, W. S. (1968). Perceptual distance and the

- specification of phonological features. *Phonetica*, 18, 31-45.
- Myers, S., & Hansen, B. (2005). The origin of vowel-length neutralization in vocoid sequences. *Phonology*, 22, 317-344.
- Nakamura, M. (2002). The articulation of the Japanese /r/ and some implications for phonological acquisition. *Phonological Studies*, 5, 55-62.
- Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., & Trent-Brown, S. A. (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *Journal of Acoustical Society of America*, 124(1), 576-588.
- Nishimura, K. (2003). *Lyman's Law in loanwords*. Unpublished master's thesis, Nagoya University.
- Pols, L. (1983). Three mode principle component analysis of confusion matrices, based on the identification of Dutch consonants, under various conditions of noise and reverberation. *Speech Communication*, 2, 275-293.
- Rice, K. (1993). A reexamination of the feature [sonorant]: The status of sonorant obstruents. *Language*, 69, 308-344.
- Shattuck-Hufnagel, S. (1986). The representation of phonological

- information during speech production planning: Evidence from vowel errors in spontaneous speech. *Phonology Yearbook*, 3, 117-149.
- Shinohara, S. (2004). *A note on the Japanese pun, dajare: Two sources of phonological similarity*. (ms. Laboratoire de Psychologie Experimentale)
- Steriade, D. (1994). *Positional neutralization and the expression of contrast*. (ms. University of California, Los Angeles)
- Steriade, D. (2001). *The phonology of perceptibility effect: The p-map and its consequences for constraint organization*. (ms. MIT)
- Steriade, D. (2003). Knowledge of similarity and narrow lexical override. In P. M. Nowak, C. Yoquelet, & D. Mortensen (Eds.), *Proceedings of Berkeley Linguistics Society 29* (p. 583-598). Berkeley: BLS.
- Tateishi, K. (1990). Phonology of Sino-Japanese morphemes. In G. Lamontagne & A. Taub (Eds.), *University of Massachusetts Occasional Papers in Linguistics 13* (p. 209-235). Amherst: GLSA.
- Zwicky, A., & Zwicky, E. (1986). Imperfect puns, markedness, and phonological similarity: With fronds like these, who needs anemones? *Folia Linguistica*, 20, 493-503.