Sound Change, Abstract Representations, and Simplicity

Wayne P. Lawrence
The University of Auckland

One criticism often leveled against generative/abstract phonology is that there is no psychological justification for the view that the brain/mind operates according to principles of ''simplicity'' and ''economy.'' For example, Jaeger (1986:74) writes that ''[n]early all studies of speech perception and production indicate that something close to surface forms exists in memory . . . and that words are stored with much redundancy.''

In this squib, I will present two historically and geographically unrelated examples of sporadic, irregular, historical changes in individual lexical items which strongly suggest that, in at least some speakers' mental representation of the prechange form, the words were being given a rather abstract representation not motivated by any language-specific details of the morphosyntax. Both examples are of words where adjacent syllables are (near-)identical: *kasusu* in one case, *sizimi* in the other. These change to *kakasu* and *hibimi*, respectively. These sound changes are amenable to a natural explanation only if they are considered to have taken place on some representation that explicitly notated them as instances of reduplication. What is noteworthy, however, is that these words are *not* historically due to any reduplication process, nor is there any synchronic reduplication process that appears to be relevant to them. In other words, the language speaker/learner has made the abstraction that these words contain a reduplicated syllable, with no more evidence for it than the actual pronunciation of the words.

1 Two Sound Changes

The Northern Okinawan dialect of the community of Yaka has the word *kakasu* '(edible) sea urchin', the Proto-Okinawan form of which is reconstructed as **gacucu*.[1] In Proto-Northern-Okinawan, both affricates (*c* = [ts]) in this form irregularly became *s*. The immediate predecessor of Yaka *kakasu* is thought to have been **kasusu*, and indeed this is the form attested in the dialect of the Kin community, located about seven kilometers east of Yaka. Assuming some kind of reduplicative function (informally represented here as a postposed *R*), the change **kasusu* → *kakasu* can be interpreted as **kasuR* → *kaRsu* (i.e., as a movement of the reduplicative function within the word), reflecting a probable confusion about the locus of reduplication on the part of the speaker who triggered this historical change.[2]

I would like to thank two anonymous reviewers for *Linguistic Inquiry* for very helpful comments. I am also grateful to the members of the University of Auckland informal formal-grammar discussion group for discussion at an early stage of this squib.

[1] The Okinawan dialects are a subgroup of the Ryukyuan language group. Proto-Ryukyuan is a sister language of the ancestor language of the mainland Japanese dialects.

[2] A Google search reveals examples like *Hononolu* (11 hits) and *Holululu* (8 hits) for *Honolulu*, *Kadudu* (for *Kakadu*) *National Park* (2 hits), *Jobaba* (for *Jojoba*) *oil* (2 hits), and *Tititica* for *Titicaca* (9 hits). As the anonymous

In this squib, I use *R* as a segmentally empty reduplicative function that is subject to specification of unit (mora, syllable, foot, etc.) and direction. I have arbitrarily selected reference to the left (i.e., to the preceding unit), but reference to the right would work as well: **kaRsu* → *Rkasu*. *R* differs from the function RED used in Optimality Theory analyses insofar as RED is restricted to being a morpheme (specifically, an affix) (McCarthy and Prince 1999:232), whereas there is no evidence to support morphemic status of the reduplicant in the examples given in this squib.

Standard Japanese *sizimi* 'corbicula (a small freshwater bivalve)'[3] is recorded in written form as *hibimi* on at least two separate occasions.[4] Confusion of *si* [ʃi] and *hi* [çi] is common in many dialects of Japanese, with *hi* tending to be pronounced as *si* in Tokyo Japanese (Vance 1987:22); so, in the Tokyo Japanese situation, *si* → *hi* would be an instance of hypercorrection. However, the change *zi* [$^{(d)}$ʒi] → *bi* in the second syllable is a very unnatural change that can be explained only if the second syllable is viewed as related to the word-initial syllable through reduplication. In Japanese, *b* is functionally the voiced equivalent of *h*, so *zi* stands in the same relationship to *si* as *bi* does to *hi*. Both the prechange and the postchange forms can be viewed as involving reduplication of the word-initial syllable, coupled with voicing (1).[5]

(1) s i R m i → h i R m i [ʃi$^{(d)}$ʒimi → çibimi]
[+voi] [+voi]

2 Discussion

The two examples given above are monomorphemic, both synchronically and historically as far back as their history can be traced. Word formation processes in these languages provide no evidence that reduplication is involved in the make-up of these words. In spite of this, sound changes indicate that the representation of these forms at the

reviewer who alerted me to such examples puts it, ''These seem to support the idea that the speaker knows there's reduplication somewhere but isn't sure where.'' Whereas these examples have possibly arisen in the process of committing words to writing, the Yaka dialect example differs in that the change has taken place without any influence of a writing system, Yaka (like most Ryukyuan dialects) being a purely spoken language with no orthography or written tradition.

[3] This is a common lexical item in oral Japanese, since sizimis are frequently used in making miso soup.

[4] One is reported in Gengoseikatu Editorial Board 1979:43, and a separate instance is reported on a number of Web sites (e.g., www.geocities.co.jp/HeartLand-Himawari/9254/top/k/k3._html).

[5] The voicing in *sizimi*, *hibimi* is *not* a case of the sequential voicing (*rendaku*) process that is observed in certain other reduplicative forms such as *hito-bito* 'people', *sju-zju* 'variety'. This is because sequential voicing applies only if the compound-final constituent is a free form, and *sizimi* is not a compound of /si-si/ or /si-simi/ where either *si* or *simi* is a free form.

time of the sound change involved the mapping of phonological information of one syllable onto another (reduplication). To the pre-sound-change speaker, the only external evidence for such a representation lay in the phonetic form itself.

In the framework of Optimality Theory, Prince and Smolensky (2004:225–230) propose a principle of Lexicon Optimization according to which, in the absence of alternations, learners construct lexical representations that are identical to the surface representations they hear.[6] Inkelas (1994:1) argues that the use of underspecification arises ''only when there are alternant surface forms all of which are predictable from context or grammatical defaults.'' Under this conception, the Yaka and Japanese underlying forms prior to sound changes would be /kakasu/ and /ʃiʒimi/, respectively, providing no base for the attested historical changes to occur. A proposed expansion of Lexicon Optimization, Pattern-Responsive Lexicon Optimization (Harrison and Kaun 2000), would allow pervasive phonological patterns, be they due to alternation, analogy, or predictable patterns of surface distribution, to yield underspecified input representations, but is there such evidence for a reduplicative structure (underspecified syllable) in Yaka and Japanese?

The Yaka dialect has yet to be described in any detail, but a glossary of the closely related Kin dialect (Okamura 1994) containing approximately 1,120 different vocabulary items (excluding compounds, conjugational variants, child language forms, and exclamations) contains only 3 words of the shapes $C_1V_2C_1V_2CV$ and $(C)VC_1V_2C_1V_2$ (of a total of 79 trimoraic trisyllabic words, or 3.8%), and only two words of the shape $C_1V_2C_1V_2$ (of a total of 86 bimoraic bisyllabic words, or 2.3%), where the voicing of the consonants has been ignored.[7] In Standard Japanese, a count of trimoraic trisyllabic noncompound words listed in Terakawa and Kusaka 1944 reveals that 6.9% (51 out of 741 such words) are of the shape $C_1V_2C_1V_2(C)V$ or $(C)VC_1V_2C_1V_2$ (again ignoring differences in consonant voicing). The same dictionary lists 627 non-Sino-Japanese noncompound nouns of the shape CVCV, and only 17 (2.7%) of these exhibit a reduplicative structure.[8] There is thus no pervasive phonological patterning in these languages that would motivate underspecification under Pattern-Responsive Lexicon Optimization.

[6] Prince and Smolensky (2004:30, 230) do permit the use of constraints of the *STRUC family to produce maximally simple lexical representations, but it is unclear what would motivate the ranking of the relevant *STRUC constraints above the relevant faithfulness constraints in the two cases discussed here. For arguments against *STRUC constraints, see Gouskova 2003.

[7] The forms found with potentially reduplicated syllables are *asasa* 'kind of cicada', *kasusu* 'sea urchin', *suzumi* 'drum', *nunu* 'cloth', *susu* 'hem'.

[8] There are no Sino-Japanese morphemes of the form $C_1V_2C_1V_2$, so including monomorphemic Sino-Japanese CVCV words would further reduce the percentage. There are also no verbs of the form $C_1V_2C_1V_2$.

Zuraw (2002a) proposes that ''there is a general drive, represented by the constraint REDUP, for all words to be construed as reduplicated'' (Zuraw 2002b). This ''aggressive reduplication'' strategy receives strong support from the sound changes discussed above.[9] However, one may ask *why* this general drive exists.

It is plausible that this ''aggressive reduplication'' is a reflex of a more general principle that values less ''complex'' underlying forms over more complex candidates. This would mean that $C_1V_2C_1V_2$ is deemed by the grammar to be structurally more complex than CVR. One way in which $C_1V_2C_1V_2$ is conceivably more complex than CVR is in the number of features in the underlying representation. The two sound changes discussed in this squib provide some support for such an evaluation metric, but further evidence will be required to establish and refine it.

3 Conclusion

The two sporadic sound changes discussed in this squib indicate that, in the speech of at least the speakers who instigated the changes, both the pre- and postchange forms were given an underspecified, reduplicative structure in their underlying representations. As this cannot be attributed to any language-specific detail of the languages concerned, we should seek a more general explanation for why the shortest possible derivation (i.e., underlying representation is identical to the surface representation) is not selected. These sound changes suggest that the language learner strives to reduce the structural ''complexity'' of the underlying form of words, adopting underspecified structure where this is recoverable.

References

Gengoseikatu Editorial Board, ed. 1979. Me: Kakikotoba memo-tyou (Eyes: Memo-book of written language). *Gengoseikatu* (Language Life) 336:43.

Gouskova, Maria. 2003. Deriving economy: Syncope in Optimality Theory. Doctoral dissertation, University of Massachusetts, Amherst.

Harrison, K. David, and Abigail Kaun. 2000. Pattern-Responsive Lexicon Optimization. In *NELS 30*, ed. by Ji-Yung Kim and Masako Hirotani, 327–340. Amherst: University of Massachusetts, GLSA.

[9] An anonymous reviewer points out that the Yaka example, **kasusu* > *kakasu*, bears on the question of whether construal of a word as reduplicated occurs during learning, because the resulting lexical representation would be simpler, or during generation, because the grammar enforces it (Zuraw 2002a: 424–430). The reviewer notes that [kakasu] appears to argue for the lexicalization interpretation, since under the generation account /kasusu/ would simply map to [kasusu], which does not violate REDUP constraints; there is no way for REDUP to predict the change to [kakasu].

Inkelas, Sharon. 1994. The consequences of optimization for underspecification. Rutgers Optimality Archive ROA-40. http://roa.rutgers.edu.

Jaeger, Jeri J. 1986. On the acquisition of abstract representations for English vowels. *Phonology Yearbook* 3:71–97.

McCarthy, John J., and Alan S. Prince. 1999. Faithfulness and identity in Prosodic Morphology. In *The prosody-morphology interface*, ed. by René Kager, Harry van der Hulst, and Wim Zonneveld, 218–309. Cambridge: Cambridge University Press.

Okamura, Toyo. 1994. *Cin-kutuba* (Kin dialect). Private edition.

Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, Mass.: Blackwell.

Terakawa, Kisio, and Kusaka Miyosi. 1944. *Hyouzyun nihongo hatuon daiziten* (Standard Japanese pronunciation dictionary). Kyoto: Taigadou.

Vance, Timothy J. 1987. *An introduction to Japanese phonology*. Albany, N.Y.: State University of New York Press.

Zuraw, Kie. 2002a. Aggressive reduplication. *Phonology* 19:395–439.

Zuraw, Kie. 2002b. Near- and pseudo-reduplication. UCLA Linguistics Talk abstract. http://www.linguistics.ucla.edu/colloquia/abstracts/2002/zuraw-2002-main.html.