

A positional effect in sound symbolism: An experimental study

Shigeto Kawahara¹, Kazuko Shinohara², Yumi Uchimoto²

¹University of Georgia, ²Tokyo University of Agriculture and Technology

*Mental activity and underlying brain activity cannot be understood
outside of the context of bodily activities*
(Davis and MacNeilage, 2000: 230).

1. Introduction

It has been well-established in the psycholinguistic literature that word-initial syllables play a primary role in word recognition processes (Browman, 1978; Brown and MacNeill, 1966; Cole, 1973; Hawkins and Cutler, 1988; Marslen-Wilson, 1975; Mattys and Samuel, 2000; Nootboom, 1981). For example, when speakers recall only a part of word (“tip of the tongue phenomenon”), the most successful cue to recall the word in its entirety is the initial syllables (Browman, 1978; Freedman and Landauer, 1966). In a phoneme monitoring task, response times are generally faster for segments in word-initial syllables than for segments in non-word-initial syllables (Mattys and Samuel, 2000). These types of evidence show that initial syllables are important in word recognition; that is, they are *psycholinguistically prominent*. Reflecting initial syllables’ psycholinguistic importance, the cohort model of word recognition postulates that an incoming signal—or a spoken word—activates lexical entries starting with the same segment with the incoming signal. As speakers hear subsequent sounds, words which do not contain such sounds are dropped, until a unique entry is finally detected (Marslen-Wilson, 1980; Marslen-Wilson and Welsh, 1978).

The psycholinguistic prominence of word-initial syllables is also known to affect phonological patterns in at least three ways. First, initial segments can determine the phonological properties of the rest of the words in many harmony patterns (Beckman, 1998). In Sino-Japanese, for example, the backness of a second vowel in a morpheme following [k] matches the backness of the word-initial vowel preceding the [k] (*teki* ‘-like’ vs. *toku* ‘benefit’) (Tateishi, 1990). Second, initial segments can accommodate segments which non-initial segments cannot (Smith, 2002; Zoll, 1998). Again in Sino-Japanese, only initial syllables can license palatalized segments, fricatives, sonorants, and non-high vowels (Kawahara, Nishimura and Ono, 2002). Third, initial syllables are known to attract segments, especially those that are floating underlyingly (Beckman, 1998; Kawahara, 2006b; Zoll, 1998). A recently emerging suffix *-zu* in Japanese has a floating accent, and its accent is usually attracted to initial syllables (*tonneru* ‘tunnel’ → *to’nneru-zu*: Kawahara and Wolf,

2007). See Beckman (1998) and Smith (2002) for recent overviews and further exemplifications.

This paper demonstrates that the psycholinguistic prominence of word-initial syllables influences patterns of sound symbolism—the images evoked by particular sounds (Berlin, 2006; Hinton, Nichols and Ohala, 1994; Sapir, 1929 among many others; see below). Japanese speakers are known to associate voiced obstruents (or “*dakuon*”; i.e., [b, d, g, z]) with negative meanings such as dirtiness or sneakiness, especially in onomatopoeic words (Kubozono, 1999; Suzuki, 1962; Tamori, 1991: 57-58; see also Hamano, 1986). For example, *doro-doro* ‘murky/muddy’ has negative connotations compared to *toro-toro* ‘pulpy’: *doro-doro* and *toro-toro* both mean ‘thick liquid’, but *doro-doro* has a negative additional meaning ‘muddy-like’. Our experiments show that voicing in obstruents evokes negative images even in non-onomatopoeic words, and further that it evokes stronger negative images in word-initial syllables than in non-initial syllables. We thus conclude that the psycholinguistic prominence is an important factor in determining sound symbolism. Finally, we argue that the prominence of initial syllables exemplifies a paralinguistic external cognitive factor influencing linguistic patterns. This conclusion implies that linguistic patterns cannot totally be reduced to purely formal systems.

2. Experiment I

The first experiment was designed to verify the general hypothesis that voicing in obstruents evokes negative images in Japanese speakers’ minds.

2.1. Method

Obstruents include stops, affricates and fricatives, which are defined as a class of sounds in which oral occlusion is great enough to impede the oral airflow such that the intraoral airpressure rises and spontaneous voicing becomes impossible (Chomsky and Halle, 1968: 302). In Japanese, voiced obstruents are called *dakuon* ‘muddy sounds’, and their voiceless counterparts are called *seion* ‘transparent sounds’.

We compared twenty disyllabic nonce-words which contained voiced obstruents ([b, d, g, z]) in both syllables (e.g. *begi*) to twenty which contained only voiceless obstruents ([p, t, k, s]) (e.g. *peki*).¹ The Appendix lists the whole set of stimuli. All the stimuli consisted of light

¹ We did not test whether (spontaneous) voicing in sonorants would cause any negative meanings. Voicing in Japanese sonorants is phonologically inert in a sense that it does not count toward the violation of Lyman’s Law (Itô and Mester, 1986). It would therefore be

syllables. We avoided voiced affricates since their distribution is limited in Japanese phonology, and also avoided [h] because some studies consider [h] as a non-obstruent (i.e. a voiceless approximant: Chomsky and Halle, 1968). We excluded palatalized consonants from the stimuli also, because they are known to evoke images of “smallness or uncontrolledness” (Hamano, 1986; Mester and Itô, 1989). The actual auditory stimuli were created by combining a set of syllables, each separately pronounced by the third author (edited by Raku SpeechS²). In addition to forty test items, we added twenty filler items.

The stimuli were presented auditorily to the participants together with *roma-ji* orthography on an html web interface using a JAVA script. The order of the presentation was randomized. Thirty native speakers of Japanese participated in this study (fifteen females and fifteen males; ages: 18-24). They were asked to judge the “dirtiness” of the stimuli on a 1-4 scale (where 1 is considered dirtiest). Specifically, they were instructed to imagine an exotic language in which the stimulus words are used, and speculate the meanings of these words.

2.2. Results

Figure 1 illustrates the dirtiness ratings of the two conditions, one containing voiced obstruents, and the other containing voiceless obstruents. The Y-axis plots the dirtiness judgments where 1 is the dirtiest and 4 is the cleanest. The error bars represent 95% confidence intervals.

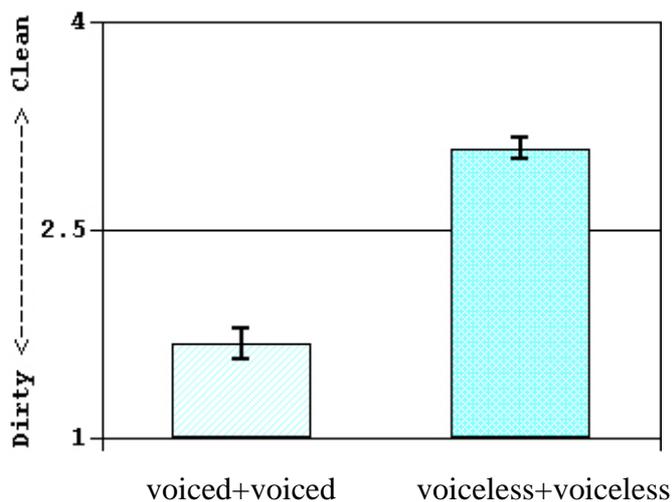


Figure 1: The effect of voicing on the dirtiness judgments.

As seen in Figure 1, the items that contained voiced obstruents (the first bar) were judged as

interesting to test whether the phonologically inert feature invokes some particular image.

² The software is downloadable at <http://frees.maxs.jp/index/modules/mydownloads>.

dirtier than the items that contained voiceless obstruents (the second bar). An ANOVA with *voicing difference* as a within-subject variable and *gender* as a between-subject variable reveals a significant main effect of the voicing difference ($F(1, 28)=329.26, p<.001$). Neither *gender* nor its interaction with *voicing difference* was significant ($F(1, 28)=2.21, p=.14$ and $F(1, 28)=1.84, p=.19$, respectively), which suggests that negative images associated with voiced obstruents were consistently observed in both genders. Overall, the results support the hypothesis that voiced obstruents evoke negative images in Japanese speakers' minds, regardless of the gender of the participants (Kubozono, 1999; Suzuki, 1962; Tamori, 1991).

3. Experiment II

Building on the results of Experiment I, the second experiment was designed to investigate whether the negative images of voiced obstruents identified in Experiment 1 are greater in initial syllables than in non-initial syllables.

3.1. Method

The stimuli were forty disyllabic nonce-words in Japanese. Half of them contained a voiced obstruent ([b, d, g, z]) in the initial syllable with a voiceless obstruent ([p, t, k, s]) in the second syllable (e.g. *bupe, zapo*), and the other half had a voiced obstruent in the second syllable and a voiceless obstruent in the initial syllable (e.g. *tgu, kuza*). See the Appendix for the list of stimuli.

Again thirty native speakers of Japanese participated in the experiment (fifteen females and fifteen males; ages: 19-24). No participants who participated in Experiment 1 participated in Experiment 2. The experimental procedure was identical to Experiment 1.

3.2. Results

Figure 2 illustrates the results of dirtiness rating in Experiment 2.

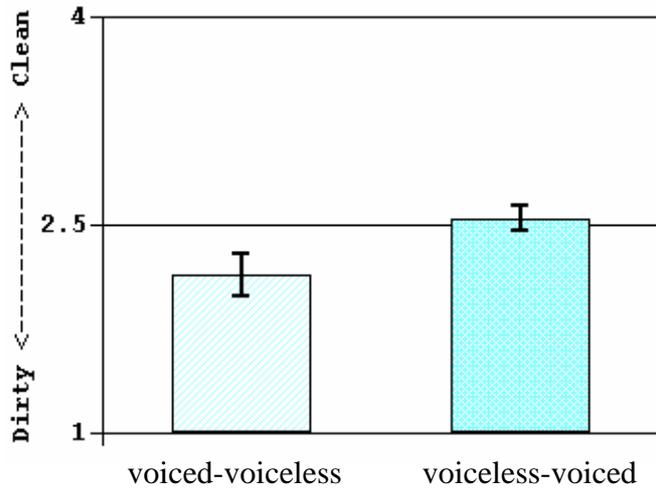


Figure 2: A positional effect of voicing on the dirtiness judgments in the two conditions.

The stimuli which contained voiced obstruents in initial syllables (the first bar) were considered as dirtier than stimuli containing voiced obstruents in second syllables (the second bar). An ANOVA reveals a statistically significant difference between the two conditions ($F(1, 28)=23.24, p <.001$); neither *the gender difference* nor its interaction with the condition reached significance ($F < 1, F(1, 28)=3.66, p=.07$). The significant difference between the two conditions suggests that the negative images evoked by voicing in obstruents were stronger in initial syllables than in non-initial syllables. The stronger effect of voicing in initial-syllables identified here accords well with the body of psycholinguistic literature that word-initial syllables have a privileged role in word recognition processes.

4. General discussion

4.1. On sound-meaning connections

In Experiment 1 we confirmed the correlation between negative images and voiced obstruents in Japanese. Building on that, Experiment 2 has identified the privileged role of initial syllables in sound symbolism. Our studies show that—contrary to the assertion that the sound-meaning relationship is arbitrary (Saussure, 1916)—there may be a consistent relationship between sounds and images that they evoke (Bentley and Varon, 1933; Hinton et al., 1994; Jakobson, 1978; Jespersen, 1922; Kamimura, 1965; Kelly, Leben and Cohen, 2003; Newman, 1933; Ohala, 1983a; Perfors, 2004; Sapir, 1929; Shinohara, Kawahara, Nakayama and Matsunaka, 2007; Ultan, 1978).

Given our results, we can reasonably ask ourselves where this negative image comes from. One possible hypothesis is that negative images arise from the aerodynamic difficulty of

maintaining voicing in obstruents (Ohala, 1983b).³ During a stop closure, intraoral air pressure rises quickly, making it difficult to maintain a drop in transglottal air pressure sufficient to produce voicing. For fricatives, the rise in intraoral air pressure required to create friction makes it difficult to sustain voicing. These aerodynamic difficulties may cause negative images in Japanese speakers' mind.

This hypothesis predicts that negative images associated with voiced obstruents should be universal, as the aerodynamic difficulty is also universal (it is a matter of a physical law—Boyle's Law).⁴ Our study in progress shows that English speakers do associate voiced obstruents with negative images just like Japanese speakers (Kawahara and Shinohara, in progress). We believe that this finding contributes one step toward the universality of sound symbolism, but needless to say, more cross-linguistic investigation is warranted to verify the universality of the correlation between obstruent voicing and negative images.

4.2. The role of psycholinguistic prominence in sound symbolism

Second, we have found an interesting parallel between phonology and sound symbolism. Our results support the importance of initial syllables in determining images associated with particular sounds—here, voicing in obstruents. This conclusion accords well with the common observation in phonology that initial syllables are strong positions (Beckman, 1998; Kawahara, 2006b; Kawahara and Wolf, 2007; Smith, 2002; Zoll, 1998). The primacy of initial syllables in sound symbolism and in phonology presumably has a single source; namely, their psycholinguistic prominence.

³ Alternatively, Kubozono (1999) hypothesizes that the negative image of voiced obstruents comes from low lexical frequencies of voiced obstruents in the Japanese lexicon. However, the log-transformed frequencies of voiced and voiceless consonants do not substantially differ in the Japanese lexicon (voiceless=19.4, voiced=18.3 (token); voiceless=12.6, voiced=11.7 (type), based on Amano and Kondo, 2000). Moreover, the correlation between dirtiness and low lexical frequencies is yet to be demonstrated cross-linguistically.

⁴ In fact, other cases of sound-meaning connections have been shown to hold across many languages and have articulatory grounding. For example, cross-linguistically it is commonly observed that the vowel height—or the size of oral cavity—correlates with the images of sizes: the higher the vowel, the smaller the oral cavity, and hence the smaller the image (Sapir, 1929; Shinohara et al., 2007; Ultan, 1978): the correlation between vowel height and size images is demonstrably universal. We suspect that this cross-linguistic robust correlation between size and the oral cavity holds because the connection has an articulatory—or bodily—grounding (Berlin, 2006; MacNeilage and Davis, 2001; Paget, 1930; cf. Archangeli and Pulleyblank, 1994).

4.3. Non-formal aspects of linguistic patterns

Finally, the psycholinguistic prominence of initial syllables instantiates non-formal aspects of linguistic systems. Thus, our findings support the fundamental tenet of Cognitive Linguistics that linguistic behaviors are not completely formal or computational, but are influenced by other cognitive factors to a non-negligible extent (Johnson, 1987; Lakoff and Johnson, 1980, 1999). This conclusion is also in concert with a thesis that phonological patterns cannot be understood if we ignore articulatory and perceptual bases of speech (Archangeli and Pulleyblank, 1994; Davis and MacNeilage, 2000; Donegan and Stampe, 1979; Kawahara, 2006a; Ohala, 1983b; Smith, 2002; Steriade, 2001; Zipf, 1949 among many others).

As discussed in §4.1, some images associated with sounds may have articulatory—or bodily—basis. For example, we proposed that the association of negative images with voiced obstruents may have its root in the aerodynamic challenge of articulating voiced obstruents. It has also been suggested that this aerodynamic challenge non-trivially affects phonological patterns (Kawahara, 2006a; Ohala, 1983b). Further studies of sound symbolism thus provide promising testing ground for how non-formal cognitive factors enter into—and interact with—linguistic systems, both in phonology and sound symbolism. Finally, we would like to suggest that cognitive linguists and (some) phonologists share the same philosophy—that linguistic patterns should be studied in the context of how they are spoken and how they are heard. We believe that this common interest has unfortunately remained unnoticed. In conclusion, then, both fields would benefit from each other by more extensive mutual communication.

References

- Amano, Shigeaki, and Kondo, Tadahisa (2000) *NTT database series: Lexical properties of Japanese, 2nd release*. Tokyo: Sanseido.
- Archangeli, Diana, and Pulleyblank, Douglas (1994) *Grounded Phonology*. Cambridge: MIT Press.
- Beckman, Jill (1998) *Positional Faithfulness*. Doctoral dissertation, University of Massachusetts, Amherst.
- Bentley, Madison, and Varon, Edith (1933) An accessory study of "phonetic symbolism". *American Journal of Psychology* 45: 76-85.
- Berlin, Brent (2006) The first congress of ethnozoological nomenclature. *Journal of Royal Anthropological Institution*: 23-44.
- Browman, Catherine (1978) Tip of the tongue and slip of the ear: Implications for language processing. *UCLA Working Papers in Phonetics* 42.

- Brown, Roger, and MacNeill, David (1966) The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5: 325-337.
- Chomsky, Noam, and Halle, Morris (1968) *The Sound Pattern of English*. New York: Harper & Row.
- Cole, Ronald (1973) Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics* 13: 153-156.
- Davis, Barbara, and MacNeilage, Peter (2000) An embodiment perspective on the acquisition of speech perception. *Phonetica* 57: 229-241.
- Donegan, Patricia, and Stampe, David (1979) The study of natural phonology. In *Current Approaches to Phonological Theory*, ed. D. A. Dinnsen, 126-173. Bloomington, IN: Indiana University Press.
- Freedman, Jonathan, and Landauer, Thomas (1966) Retrieval of long-term memory: "Tip-of-the-tongue" phenomenon. *Psychonomic Science* 4: 309-310.
- Hamano, Shoko (1986) *The Sound-Symbolic System of Japanese*. Stanford: CSLI Publications.
- Hawkins, John, and Cutler, Anne (1988) Psycholinguistic factors in morphological asymmetry. In *Explaining Language Universals.*, ed. J. A. Hawkins, 280-317. Oxford: Basil Blackwell.
- Hinton, Leane, Nichols, Johanna, and Ohala, John (1994) *Sound Symbolism*. Cambridge: Cambridge University Press.
- Itô, Junko, and Mester, Armin (1986) The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17: 49-73.
- Jakobson, Roman (1978) *Six Lectures on Sound and Meaning*. Cambridge: MIT Press.
- Jespersen, Otto (1922) Symbolic value of the vowel i. *Phonologica* 1.
- Johnson, Mark (1987) *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- Kamimura, Yukio (1965) Onsei no hyoushousei ni tsuite [On the symbolic aspects of sounds]. In *Gengo Seikatsu*, 66-70. Tokyo: Honami Shuppan.
- Kawahara, Shigeto (2006a) A faithfulness ranking projected from a perceptibility scale: The case of voicing in Japanese. *Language* 82: 536-574.
- Kawahara, Shigeto (2006b) Mimetic gemination in Japanese: A challenge for Evolutionary Phonology. *Theoretical Linguistics* 32: 411-424.
- Kawahara, Shigeto, Nishimura, Kohei, and Ono, Hajime (2002) Unveiling the unmarkedness of Sino-Japanese. In *Japanese/Korean Linguistics 12*, ed. W. McClure, 140-151. Stanford: CSLI.
- Kawahara, Shigeto, and Shinohara, Kazuko (in progress) The role of articulatory gestures in the cognition of voiced obstruents.
- Kawahara, Shigeto, and Wolf, Matthew (2007) The root-initial-accenting suffix in Japanese. Ms. University of Massachusetts, Amherst.
- Kelly, Barbara, Leben, William, and Cohen, Robert (2003) The meaning of consonants. Ms. Lexicon Branding INC.
- Kubozono, Haruo (1999) *Nihongo no Onsei: Gendai Gengogaku Nyuumon 2 [Japanese Phonetics:*

- An Introduction to Modern Linguistics*]. Tokyo: Iwanami.
- Lakoff, George, and Johnson, Mark (1980) *Metaphors We Live By*. Chicago: Chicago University Press.
- Lakoff, George, and Johnson, Mark (1999) *Philosophy in the Flesh*. New York: Basic Books.
- MacNeilage, Peter, and Davis, B. L. (2001) Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications. *Current Biology* 11: 696-700.
- Marslen-Wilson, William (1975) Sentence perception as an interactive parallel process. *Science* 189: 226-228.
- Marslen-Wilson, William (1980) Speech understanding as a psychological process. In *Spoken Language Generation and Understanding*, ed. J. C. Simon. Dordrecht: Reidel.
- Marslen-Wilson, William, and Welsh, Alan (1978) Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10: 29-63.
- Mattys, Sven, and Samuel, Arthur (2000) Implications of stress-pattern differences in spoken-word recognition. *Journal of Memory and Language* 42: 571-596.
- Mester, Armin, and Itô, Junko (1989) Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65: 258-293.
- Newman, Stanley (1933) Further experiments on phonetic symbolism. *American Journal of Psychology* 45: 53-75.
- Nooteboom, Sieb (1981) Lexical retrieval from fragments of spoken words: Beginnings vs. endings. *Journal of Phonetics* 9: 407-424.
- Ohala, John (1983a) Cross-language use of pitch: An ethnological view. *Phonetica* 40: 1-18.
- Ohala, John (1983b) The origin of sound patterns in vocal tract constraints. In *The Production of Speech*, ed. P. MacNeilage, 189-216. New York: Springer-Verlag.
- Paget, R. (1930) *Human Speech: Some Observations, Experiments, and Conclusions as to the Nature, Origin, Purpose, and Possible Improvement of Human Speech*. London: Routledge.
- Perfors, Amy (2004) What's in a name? A poster presented at CogSci 2004.
- Sapir, Edward (1929) A study in phonetic symbolism. *Journal of Experimental Psychology* 12: 225-239.
- Saussure, Ferdinand de (1916) *Cours de linguistique générale*, eds. C. Bally, A. Sechehaye and A. Riedlinger. Paris: Payot.
- Shinohara, Kazuko, Kawahara, Shigeto, Nakayama, Akira, and Matsunaka, Yoshihiro (2007) Onshoochoo to sintaisei [Sound symbolism and embodiment]. In *Kotoba to Ningen [Languages and Human]: Journal of Human Linguistic Circle* 6, ed. Y. Fujii, 1-12.
- Smith, Jennifer (2002) *Phonological Augmentation in Prominent Positions*. Doctoral dissertation, University of Massachusetts, Amherst.
- Steriade, Donca (2001) The phonology of perceptibility effect: The P-map and its consequences for constraint organization. Ms. UCLA.
- Suzuki, Takao (1962) Oninkookan to igibunka no kankei ni tsuite--iwayuru seidakuon dairitsu-o

chuushin toshite. *Gengo Kenkyu* 42.

Tamori, Ikuhiro (1991) *Nihongo Onomatope no Kenkyuu [A Research on Japanese Onomatopoeic Words]*. Kobe: Kobe Shoka Daigaku Keizai Kenkyuujo.

Tateishi, Koichi (1990) Phonology of Sino-Japanese morphemes. In *University of Massachusetts Occasional Papers in Linguistics 13*, eds. G. Lamontagne and A. Taub, 209-235. Amherst: GLSA.

Ulan, Russell (1978) Size-sound symbolism. In *Universals of Human Language II: Phonology*, ed. J. Greenberg, 525-568. Stanford: Stanford University Press.

Zipf, George Kingsley (1949) *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley Press.

Zoll, Cheryl (1998) Positional asymmetries and licensing. Ms. MIT.

Appendix: The list of stimuli

Experiment 1

[+voiced]-[+voiced]

badu
begi
bigo
bobe
buzo
dado
degi
dize
dozu
dubo
gazi
gedu
goda
guba
gugo
zabe
zebo
ziza
zoda
zuge

[-voiced]-[-voiced]

papo
peki
pite
posi
puto
tape
taso
teto
tiko
tusa
kapi
keso
kitu
kuke
kupa
saso
seka
seti
sipu
suko

Experiment 2

[+voiced]+[-voiced]

Basi
beka
bite
bopu
bupe
daso
dete
diki
dopa
dusu
gati
gese
geku
gike
gopi
zapo
zesa
zita
zoko
zuto

[-voiced]+[+voiced]

pazi
pegi
pide
poga
pudo
tadu
tebu
tigu
toze
tuzu
kade
kebo
kige
koda
kuza
sabe
sego
sizo
sodi
subi