J. Linguistics **45** (2009), 111–138.
doi:10.1017/S0022226708005537 Printed in the United Kingdom

The role of psychoacoustic similarity in Japanese puns: A corpus study[1]

SHIGETO KAWAHARA

Rutgers University, the State University of New Jersey

KAZUKO SHINOHARA

Tokyo University of Agriculture and Technology

(Received 19 July 2007; revised 16 May 2008)

A growing body of recent work on the phonetics–phonology interface argues that many phonological patterns refer to psychoacoustic similarity – perceived similarity between sounds based on detailed acoustic information. In particular, two corresponding elements in phonology (e.g. inputs and outputs) are required to be as psychoacoustically similar as possible (Steriade 2001a, b, 2003; Fleischhacker 2005; Kawahara 2006; Zuraw 2007). Using a corpus of Japanese imperfect puns, this paper lends further support to this claim. Our corpus-based study shows that when Japanese speakers compose puns, they require two corresponding consonants to be as similar as possible, and the measure of similarity rests on psychoacoustic information. The result supports the hypothesis that speakers possess a rich knowledge of psychoacoustic similarity and deploy that knowledge in shaping verbal art patterns.

1. Introduction

A number of recent studies on the phonetics–phonology interface argue that many phonological patterns refer to PSYCHOACOUSTIC SIMILARITY – perceived similarity between sounds based on detailed acoustic information. Expanding on the general notion of faithfulness constraints in Optimality Theory (Prince & Smolensky 1993/2004, McCarthy & Prince 1995), several authors have proposed that speakers attempt to maximize the psychoacoustic similarity between inputs and outputs (Boersma 1998; Steriade 2001a, b, 2003; Côté 2004; Jun 2004; Fleischhacker 2005; Kawahara 2006;

[1] We are grateful to our informants for composing ample examples of Japanese puns for us. We would also like to thank Gary Baker, Michael Becker, John Kingston, Kazu Kurisu, Dan Mash, Lisa Shiozaki, Betsy Wang, and two anonymous *JL* referees for their comments on earlier drafts of this paper. Remaining errors are ours. An earlier version of this paper appeared as Kawahara & Shinohara (2007). Both the data and analyses presented in this version supersede those in Kawahara & Shinohara (2007).

Wilson 2006; Zuraw 2007). For example, several researchers have observed that nasals are more likely to undergo place assimilation than oral consonants (Cho 1990, Mohanan 1993, Boersma 1998, Jun 2004). Steriade (1994, 2001a), Boersma (1998), and Jun (2004) have proposed that this asymmetry arises because a place contrast is harder to perceive in nasals than in oral consonants, and therefore speakers can tolerate a change of articulation in nasal consonants but not in oral consonants: nasal pairs differing in [place] are 'psychoacoustically similar enough', whereas pairs of oral consonants differing in [place] are 'psychoacoustically too different' (see also Malécot 1956, Mohr & Wang 1968, Pols 1983, Hura, Lindblom & Diehl 1992, Kurowski & Blumstein 1993, Ohala & Ohala 1993, Davis & MacNeilage 2000 for the low perceptibility of [place] in nasals; see Kohler 1990, Hura et al. 1992, Huang 2001, Johnson 2003, Kawahara 2006 for the notion of 'perceptually tolerated articulatory simplification').

Another example of the role of psychoacoustic similarity in phonology comes from the observation that a voicing contrast is more likely to be neutralized than other manner contrasts. Steriade (2001b) points out that underlying voiced obstruents readily undergo devoicing in coda positions in many languages such as German, Polish, Russian and Turkish, but coda voiced obstruents never become nasalized in any languages. To explain this observation, drawing on previous psycholinguistic studies (Peters 1963, Greenberg & Jenkins 1964, Walden & Montgomery 1975, van den Broecke 1976), she proposes that a voicing contrast is less perceptible than other manner contrasts, and that speakers tolerate a change in voicing because the change is psychoacoustically the least salient. These proposals have led to – and have expanded on – the general idea of the P-map (for 'Perceptual'-map) (Steriade 2001a, b, 2003), which states that speakers know the perceptual distances between two phonological elements, and that based on this knowledge, they attempt to minimize the perceptual disparity between two corresponding elements in phonology (most commonly inputs and outputs).[2]

This paper contributes to the body of literature which argues for the importance of psychoacoustic similarity, but from a slightly different angle, namely from the perspective of verbal art. A sizable amount of work has shown that psychoacoustic similarity plays an important role in shaping phonological patterns, including verbal art patterns. Previous studies have revealed that speakers require psychoacoustic similarity between two

[2] The principle of minimization of perceptual disparity has been proposed in other correspondence relations: between base and derived words (Steriade 2000, Zuraw 2007), between base and reduplicant (Fleischhacker 2005), between source and loanwords (Steriade 2001a, Kang 2003, Kenstowicz 2003; cf. LaCharité & Paradis 2005), and between rhyming or alliterating elements (see below).

rhyming or alliterating segments, and this claim has been demonstrated in Middle English alliteration (Minkova 2003), English imperfect puns (Fleischhacker 2005), Japanese rap rhymes (Kawahara 2007), and Romanian half-rhymes (Steriade 2003). Building on this body of work, our corpus-based analysis demonstrates that Japanese speakers require two corresponding elements in puns to be as similar as possible, and the measure of similarity rests on psychoacoustic information. Our results show that both phonological patterns and verbal art patterns are governed by the same principle: the maximization of psychoacoustic similarity between two corresponding elements.

To make the case that speakers use psychoacoustic similarity in composing verbal art patterns, this paper analyzes *dajare*, which is a linguistic word game – more specifically, the creation of imperfect puns – in Japanese. In constructing *dajare*, speakers make a sentence out of two words or phrases that are identical or near-identical. An illustrative example is *huton-ga huttonda* 'the futon was blown away', which involves the juxtaposition of two similar words: *huton* and *huttonda*. An analogous English example would be something like 'Titanic in panic'. In many languages, similar sounds tend to form sound pairs in a range of verbal art patterns, including rhymes, alliterations, and imperfect puns (see Kawahara 2007 for a recent overview). The role of similarity in imperfect puns has been explored in English (Lagerquist 1980, Zwicky & Zwicky 1986, Sobkowiak 1991, Hempelmann 2003, Fleischhacker 2005) as well as in Japanese (Otake & Cutler 2001, Cutler & Otake 2002, Shinohara 2004). All of these studies agree that *some* kind of similarity plays a role in the formation of imperfect puns. We argue that it is specifically the PSYCHOACOUSTIC SIMILARITY that underlies imperfect pun parings.

Before closing this introductory section, a remark on our methodology is in order. In addition to lending further support to the role of psychoacoustic similarity in verbal art, our methodology is also innovative; that is, we undertake a statistical approach to the analysis of Japanese imperfect puns. The statistical approach allows us to uncover phonological patterns that could easily be obscured by non-phonological factors and therefore missed in an impressionistic survey. That is, speakers' focus on syntactic coherence and semantic humor may well obscure the purely phonetic and phonological aspects of two corresponding words (Hempelmann 2003). Here we draw on a large amount of data to verify whether discovered phonetic and phonological patterns are convincingly real.

The rest of this paper proceeds as follows. In section 2 we provide an overview of Japanese imperfect puns, and describe how we collected and analyzed the data. In section 3, as a preliminary, we statistically establish the role of similarity in the formation of Japanese imperfect puns. In section 4 we make the case that speakers deploy psychoacoustic similarity, rather than featural similarity, when they compose imperfect puns.

2. Method

2.1 *An overview of Japanese imperfect puns*

Puns are very common in Japanese, and some speakers create new puns on a daily basis (Mukai 1996, Takizawa 1996). In composing puns or *dajare*, speakers deliberately juxtapose a pair of 'similar-sounding' words or phrases – we argue in section 4 that this notion of 'similarity' is based on psychoacoustic information. The two corresponding parts can include identical sound sequences as in *buta-ga butareta* 'a pig was hit', an example of a PERFECT PUN; or the two corresponding words can include similar but non-identical sounds. In (1) we provide an example of an IMPERFECT PUN, which includes corresponding pairs of non-identical sounds (indicated in bold letters: [m–t], [h–b]). In this paper we focus on imperfect puns that include mismatched consonant pairs.[3]

(1) **m**an**h**attan-de su**t**an**b**attan
Manhattan-LOC stand-by
'I stood by in Manhattan.'

In (1), *manhattan* and *sutanbattan* are a pair of similar phrases that form an imperfect pun.[4] This paper uses the following notations. We mark CORRESPONDING DOMAINS in imperfect puns by underlining: we define corresponding domains as sequences of syllables in which corresponding vowels

[3] Other types of imperfect puns involve metathesis (e.g. *shi**mane**-no shi**nema*** 'Cinema in Shimane'), phrase boundary mismatches (e.g. *wakusee waa, kusee* 'Man, that planet stinks'), syllable intrusion (e.g. *bundoki bundo**tto**ki* 'Just take the protractor away from him'), etc. Analyzing these patterns is beyond the scope of this paper, but it is an interesting topic for future research.

[4] Speakers sometimes change the underlying shape of one form to achieve better resemblance to the corresponding form, as in (i) (Takizawa 1996, Shinohara 2004):

(i) Hokkaidoo-wa dekkai **d**oo
Hokkaido-TOP big PARTICLE
'Hokkaido is big.'

In this example, the sentence-final particle [doo] is pronounced as [zo] in canonical speech, but speakers pronounce /zo/ as [doo] here to make the word *dekkaizo* sound similar to *Hokkaidoo*. We also analyzed examples of puns like (i), but found that these types of puns are mainly created by mimicking non-standard speech styles. For example, for the change from /zo/ to [do] in (i), speakers are probably mimicking dialects of Japanese where the particle *zo* is pronounced as [do]. Another common pattern is palatalization, as in (ii):

(ii) Undooʒoo-o kaʃite kudasai. Un dooʒo.
playground-ACC lend please yes please.
'Please let us use your playground. Yes, please.'

The second [dooʒo] is usually pronounced as [doozo]; here /z/ is palatalized to mimic *undooʒoo*. This sort of palatalization is common in Japanese child speech and motherese (Mester & Itô 1989), and composers of puns may be mimicking these styles of speech. Our database thus excludes examples like (i) and (ii) which involve some change from underlying form to surface form.

are identical.[5] We use bold letters to indicate non-identical pairs of corresponding consonants, which is the focus of our analysis.

The pun in (1) is an imperfect pun in that it does not involve perfectly identical sequences of sounds. Though Japanese puns can involve non-identical pairs of consonants, previous studies have noted that imperfect pun pairs tend to involve 'similar' consonants (Otake & Cutler 2001, Cutler & Otake 2002, Shinohara 2004). Deferring detailed discussion of the precise definition of 'similarity', the importance of similarity in imperfect puns can be informally illustrated by examples like: *Ai**z**usan-no ai**s**u* 'ice cream from Aizu' ([z] and [s] are minimally different in voicing) and *okosa**m**a-o okosa-**n**aide* 'please do not wake up the child' ([m] and [n] are minimally different in place). Drawing on this observation, this paper statistically assesses what *kind* of similarity plays a role in imperfect puns, and demonstrates that speakers deploy their knowledge about psychoacoustic similarity based on detailed acoustic information. Our findings show that speakers possess a rich knowledge of psychoacoustic similarity and make use of that knowledge to form verbal art patterns (Steriade 2001b, 2003; Minkova 2003; Fleischhacker 2005; Kawahara 2007).

2.2 *Data collection*

This subsection describes how we compiled the database for this study. In order to investigate how similarity influences the formation of Japanese imperfect puns, we collected examples of imperfect puns from three websites.[6] We also elicited data from native Japanese speakers by asking them to create puns of their own 'out of the blue', from which we selected imperfect puns.[7] In total, we collected 2371 examples of Japanese imperfect puns.

Recall that we define 'corresponding domains' in imperfect puns as sequences of syllables in which corresponding vowels are identical. For

[5] In an example like *Ha**i**degaa-no zense-wa ha**e** dekka?* 'Was Heidegger a fly in his previous life?', the corresponding domain could be interpreted as coinciding with word boundaries; note that this example involves a pair of non-identical vowels ([i–e]). However, corresponding domains and word boundaries do not necessarily coincide with each other, as in *huton-ga huttonda* 'a futon was blown away'. In this study, we therefore define corresponding domains based on sequences of matching vowels rather than on word boundaries. Since examples like *Haidegaa-no zense-wa hae dekka?* were relatively rare in our corpus, our results below should not depend on the particular definition of the corresponding domain we have chosen.

[6] http://www.ipc-tokai.or.jp/~y-kamiya/Dajare/, http://www.geocities.co.jp/Milkyway-Vega/8361/umamoji3.html, http://www.dajarenavi.net/. Data were obtained from these websites in March 2007.

[7] The data collection was primarily done by the second author. She plays with *dajare* in her daily life, and contributed some examples to our database. However, she was unaware of the purpose of this project at the time of data collection.

example, in *okosa**m**a-o okosa**n**aide* 'please do not wake up the child', the corresponding domains include the first four syllables. In the domains defined as such, we counted pairs of corresponding consonants between the two words. We ignored identical pairs of consonants, because we are interested in similarity, not identity. Therefore, in the example *okosa**m**a-o okosa**n**aide*, for instance, we extracted just the pair [m–n].

A few remarks on our coding convention are in order. We counted only onset consonants, and ignored coda consonants because they assimilate in [place] to the following consonant in Japanese (Itô 1986). We also ignored singleton–geminate distinctions because our interest lies in segmental similarity. We coded the data based on surface forms, following a recent body of work which argues that similarity is based on surface forms rather than phonemic forms (Steriade 2001b; Kang 2003; Kenstowicz 2003; Minkova 2003; Kawahara 2006, 2007; Kawahara, Ono & Sudo 2006; though see Jakobson 1960; Kiparsky 1970, 1972, 1981; Malone 1987 for a different view). For instance, the onset of the syllable [ʃi] is arguably derived from an underlying /s/, but it was counted as [ʃ]. Finally, we excluded pairs which are underlyingly different but are neutralized at the surface because of independent phonological processes (even when such neutralization is optional). For example, the second consonant [dd] in *be**dd**o* 'bed' can optionally devoice in the presence of another voiced obstruent (Nishimura 2003, Kawahara 2006). Thus given an example like *be**dd**o-dai-wa be**tt**o itadakimasu* 'please pay for the bed separately', we did not consider the [dd–tt] pair as a pair of non-identical consonants, since the pun-composer could have pronounced /beddo/ as [betto] when s/he composed the pun. Similarly, many puns included examples of [y] corresponding with Ø in the context of [i _ a], as in *kanari**Ø**a-o kauno-wa kanari iya* 'I don't like having a canary'. We did not count these cases since /ia/ can surface as [iya] (Katayama 1998, Kawahara 2003).

2.3 *Data analysis: O/E ratios*

This paper investigates what kind of measure of similarity speakers use when they compose imperfect puns. For this purpose, we need a measure of combinability between two segments. We use O(bserved)/E(xpected) ratios as a measure of how well two elements combine. O/E ratios are ratios between how often a pair is actually observed (O-values) and how often it would be expected to occur if two elements were combined at random (E-values). Mathematically, the O-value of a sound [A] is how many times [A] occurs in the corpus, and the E-value of a pair [A–B] is $P(A) \times P(B) \times N$ (where $P(X)$ = the probability of the sound [X] occurring in the corpus; N = the total number of consonants). O/E ratios greater than 1 indicate that the given consonant pairs are combined as pun pairs more frequently than expected (overrepresentation), while O/E ratios less than 1 indicate that the given consonant pairs are combined less frequently than expected (underrepresentation) (see Trubetzkoy 1969, Frisch, Pierrehumbert & Broe 2004, Kawahara 2007 for the notion of relativized frequencies).

In our O/E analysis, we excluded consonants whose O-values are less than 20 because including them resulted in extraordinarily high O/E ratios: combining two rare consonants yields a very low E-value, so that even a single observed pair of that type would yield an artificially high O/E ratio (e.g. the O/E ratio of the [p^j–b^j] pair is 1121.7). For this reason, we excluded [ts], [dʒ], [y], [n^j], [r^j], and all non-coronal palatalized consonants. The exclusions left 535 pairs of consonants for the subsequent analysis.

Before completing the method section, a remark on our statistical analysis is in order. We did not test whether each individual O/E ratio was significantly larger than 1, for two reasons. First, applying the same test on multiple O/E ratios requires a drastic post-hoc α-level adjustment, and it is unlikely to find statistically significant overrepresentations – i.e. potential inflation of Type 2 errors (Klockars & Sax 1986, Myers & Well 2003: 84–86, 241–245). Second, since each O/E ratio is dependent on other O/E ratios, applying the same test on multiple O/E ratios violates the assumption of independence, which usually results in an increase in Type 1 errors (Seco, Menéndez de la Fuente & Escudero 2001). We thus instead applied a statistical test for general patterning of O/E ratios to test specific null hypotheses.

3. The correlation between similarity and combinability: a statistical preliminary

Table 1 shows the result of the O/E analysis, where shading indicates underrepresentation.

To statistically verify the correlation between combinability in puns on the one hand and similarity on the other, as a first approximation, we estimated the similarity of consonant pairs in terms of how many feature specifications they share (Klatt 1968, Fay & Cutler 1977, Shattuck-Hufnagel & Klatt 1977, Stemberger 1991, Bailey & Hahn 2005). We will later discuss the importance of psychoacoustic similarity in section 4 and argue that featural similarity ultimately fails to account for the detailed aspects of the pun pairing patterns. We nevertheless start with this simple version involving featural similarity in order to first statistically establish the correlation between similarity and combinability.

To estimate the numbers of shared feature specifications among the set of distinctive consonants in Japanese, we used eight features: [sonorant], [consonantal], [continuant], [nasal], [strident], [voice], [palatalized], and [place]. [Palatalized] is a secondary place feature which distinguishes plain consonants from palatalized consonants (*yoo-on*, in the traditional Japanese terminology). According to this featural similarity system, a [p–b] pair, for

	Ø	p	b	ɸ	m	w	t	tʃ	s	ʃ	z	d	j	n	r	k	g	h
Ø	0	.7	.24	.77	2.5	6.3	.64	.45	.36	.37	0	0	0	1.5	4.6	1.4	.46	3.7
p		0	8.5	5.6	0	0	1.1	0	1.1	0	0	0	0	0	.37	1.1	.24	.77
b			0	1.6	4.7	5.2	.30	0	0	0	.42	1.1	0	.35	0	.25	.65	2.9
ɸ				0	2.3	0	0	0	.81	.84	0	0	0	0	0	2.7	0	4.2
m					0	0	.64	0	.53	0	0	0	0	8.9	2.1	0	.34	1.7
w						0	0	0	0	0	0	1.2	2.1	1.5	0	.53	0	4.5
t							0	.57	.90	0	0	7.6	0	.62	.44	.87	.14	.94
tʃ								0	.95	7.9	0	0	13	0	.47	.23	0	.49
s									0	4.7	11	.2	0	1.6	.37	.73	0	.78
ʃ										0	0	0	6.8	.54	0	.38	0	.41
z											0	1.7	1.2	0	.62	0	0	0
d												0	.39	1.1	4.0	0	.39	.42
j													0	2.0	.72	0	.93	0
n														0	4.1	.25	1.7	1.1
r															0	.36	1.2	.38
k																0	8.0	1.7
g																	0	0
h																		0

Table 1

O/E ratio (shading indicates underrepresentation).

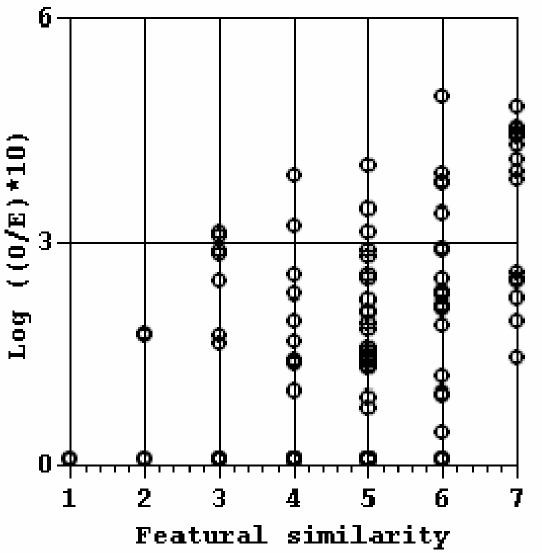

Figure 1
Correlation between the number of shared feature specifications and the likelihood of two consonants being paired in imperfect puns.

example, shares seven feature specifications, and hence is treated as highly similar. On the other hand, a pair like [ʃ–m] agrees only in [cons], and is thus considered as a highly non-similar pair.

We made the following assumptions when assigning feature specifications to Japanese sounds. Affricates are specified both as [+cont] and [−cont] (Sagey 1986, Lombardi 1990). We treated [h] as a voiceless fricative, not as an approximant (Lass 1976: 64–68, Jaeger & Ohala 1984, Sagey 1986, Parker 2002). We also assumed that sonorants are specified as [+voice] and thus agree with voiced obstruents in terms of [voice], although phonologically, voicing in sonorants behaves differently from voicing in obstruents in Japanese phonology (see section 4.3 for further discussion). The appendix gives the complete feature matrix used in this analysis.

The scatter plot in figure 1 illustrates the general correlation between similarity and combinability of consonant pairs. For each consonant pair, it plots the number of shared feature specifications on the x-axis and plots the natural log-transformed O/E ratios multiplied by 10 ($Log_e((O/E)*10)$) on the y-axis. We use log-transformed values on the y-axis because it allows us to fit all the data points in a reasonably small-sized graph. Log_e(zero) is undefined, so we replaced Log_e(zero) with zero. However, the log of O/E ratios smaller than 1 is negative, and hence would be incorrectly treated as a value smaller than zero (e.g. $0.5 > 0$, but $Log_e(0.5) < 0$). Therefore, we multiplied O/E ratios by 10 before log-transformation. The plot excludes all pairs in which one member is Ø (null), since it is difficult if not impossible to define Ø in terms of distinctive features; we discuss such pairs in section 4.4.

We observe in figure 1 an upward trend in which the combinability correlates with featural similarity: the more feature specifications a pair shares, the higher the corresponding O/E ratio is. To statistically assess the linear relationship in figure 1, we calculated a Spearman's rank correlation coefficient ρ, a non-parametric measure of linear correlation. We used a non-parametric measure because we cannot assume a bivariate normal distribution. The result is that ρ is .497 ($t(134) = 6.63$; $p < .001$), a statistically significant correlation.

Another measure of featural similarity can be calculated based on the number of shared natural classes divided by the number of shared and unshared natural classes $\left(\frac{\textit{shared natural classes}}{\textit{shared natural classses} + \textit{unshared natural classes}};\right.$ Frisch et al. 2004). Based on this formula, we calculated the similarity matrix among the Japanese consonants using a similarity calculator developed by Adam Albright (Albright n.d.). We then calculated the correlation between combinability and this similarity matrix, yielding a ρ value of .488 ($t(134) = 6.47$, $p < .001$), a value similar to the ρ value in the analysis above.

These two analyses establish statistically that a general correlation exists between combinability and some kind of similarity. Having established the statistical correlation, we now turn to some inadequacies of featural similarity in capturing imperfect pun patterns and argue for the importance of psychoacoustic similarity.

4. Beyond featural similarity: the role of psychoacoustic similarity

In section 3, we used distinctive features to estimate similarity. We argue in this section that, although featural similarity accounts for some broad patterns, it ultimately fails to capture some of the detailed aspects of pun pairings. We propose that Japanese speakers instead use psychoacoustic similarity. By psychoacoustic similarity, we mean perceived similarity between sounds based on acoustic details.

One example that supports the role of psychoacoustic similarity is the [r–d] pair, whose O/E ratio is 3.99. According to the distinctive feature system, this pair agrees in six features (all but [son] and [cont]); but the average O/E ratio of other consonant pairs agreeing in six distinctive feature is 1.49. The 95% confidence interval of the O/E ratios of such pairs is 0.39 ~ 2.59, and 3.99 lies outside this interval. Thus the high O/E ratio of the [r–d] pair is unexpected from the standpoint of featural similarity, but makes sense from a psychoacoustic perspective. Japanese [r] is a flap which involves a ballistic – or instantaneous – constriction (Nakamura 2002): [r] and [d] are similar to each other in that they are both voiced consonants with relatively short closures (Price 1981, Steriade 2000).[8]

Below we present four more kinds of arguments that pun composers exploit psychoacoustic similarity rather than featural similarity. First, speakers are sensitive to the different perceptual salience of the same feature in different contexts (section 4.1.). Second, composers take into consideration the different perceptual salience of different features (section 4.2.). Third, composers are sensitive to similarity contributed by a phonologically inert feature (section 4.3.). Finally, pun-makers are willing to pair Ø with consonants that are psychoacoustically similar to Ø (section 4.4.).

4.1 *Sensitivity to context-dependent salience of the same feature*

The first piece of evidence that pun composers make use of psychoacoustic similarity is the fact that composers are sensitive to different degrees of salience of the same feature in different contexts. Here, we focus on the perceptual salience of [place]. In (2), we list the O/E ratios of minimal pairs of consonants differing in place (the Japanese moraic nasal [N] appears only in codas, and was therefore excluded from this study).

(2) *O/E ratios of minimal pairs differing in place*

[m–n]: 8.85	[b–d]: 1.09	[p–t]: 1.11
	[b–g]: 0.65	[p–k]: 1.08
	[d–g]: 0.39	[t–k]: 0.87

We observe that the O/E ratio of the [m–n] pair is higher than that of any other minimal pair in (2) (by a binomial test, $p < .05$).[9] Thus the data in (2) indicate that composers treat the [m–n] pair as more similar than any minimal pairs of oral consonants: they treat the place distinction in nasal consonants as less salient than in oral consonants.

Evidence from previous phonetic and psycholinguistic studies supports the lower perceptibility of [place] in nasals than in oral consonants. First, a similarity judgment experiment by Mohr and Wang (1968) shows that nasal

[8] Other examples of this kind are the pairs [tʃ–kʲ] and [dʒ–gʲ], which are psychoacoustically – but not featurally – similar (Blumstein 1986, Ohala 1989, Guion 1998, Wilson 2006). A long burst of velar consonants resembles affrication, and raising of F2 in [kʲ] and [gʲ] due to palatalization makes them sound similar to coronals. Recall, however, that we excluded [kʲ] and [gʲ] from our O/E analysis (section 2.3), because they are rare consonants. Yet our original O/E analysis, which did include rare consonants, showed that both the [tʃ–kʲ] pair and [dʒ–gʲ] are highly overrepresented (O/E = 18.3, 21.6, respectively). To the extent that these high O/E ratios are statistically reliable, these examples provide further support for the importance of psychoacoustic similarity.

[9] Pairs of nasal consonants with different [place] specifications also seem to be common in English rock lyrics (Zwicky 1976), English and German poetry (Maher 1969, 1972), English imperfect puns (Zwicky & Zwicky 1986), and Japanese rap lyrics (Kawahara 2007).

minimal pairs are considered perceptually more similar to each other than oral consonant minimal pairs. Second, Pols (1983) demonstrated with Dutch speakers that [place] in nasals is less accurately perceived than [place] in oral consonants in noisy environments.[10] Place cues are less salient in nasal consonants than in oral consonants for two reasons: (i) formant transitions into and out of the neighboring vowels are obscured by coarticulatory nasalization, and (ii) burst spectra play an important role in distinguishing different places of articulation (Jakobson, Fant & Halle 1952, Stevens & Blumstein 1978, Blumstein & Stevens 1979), but nasals have weak or no bursts (see Malécot 1956, Pols 1983, Hura et al. 1992, Kurowski & Blumstein 1993, Ohala & Ohala 1993, Steriade 1994, Boersma 1998, Davis & MacNeilage 2000, Steriade 2001a, Jun 2004 for further discussion of the low perceptibility of nasals' [place] and its phonological consequences). To summarize, in measuring similarity due to [place], Japanese speakers take into account the lower perceptibility of [place] in nasal consonants. The lower perceptibility of [place] in nasals has a psychoacoustic basis: the blurring of formant transitions caused by coarticulatory nasalization and weak bursts. The data in (2) therefore show that speakers use psychoacoustic similarity in composing puns.

By contrast, if speakers were using featural similarity, then the higher combinability of the nasal pair would remain unexplained. One could postulate that featural similarity is affected less when two nasal segments disagree in [place] than when two oral segments disagree in [place]. However, introducing such weighting is post-hoc: it simply restates the observation and has no explanatory power.

An anonymous reviewer points out that feature weighting would not be ad-hoc 'if it is derived from an independently motivated model of feature relationship'. However, no previously established models of distinctive features distinguish [place] in oral consonants and [place] in nasal consonants. Feature geometry is one well-elaborated theory of feature relationships, which postulates that features are organized in a hierarchical structure (Clements 1985, Sagey 1986, McCarthy 1988, Halle 1992, Clements & Hume 1995). However, no versions of feature geometry postulate a structural difference between [place] in nasal consonants and [place] in oral consonants.

Another elaboration of feature relationships is the theory of underspecification. This theory postulates that non-contrastive or unmarked specifications are underlyingly unspecified (Kiparsky 1982, Itô & Mester 1986, Archangeli 1988, Mester & Itô 1989, Paradis & Prunet 1991, Itô, Mester & Padgett 1995; see also Lahiri & Reetz 2002 for a model of word recognition using underspecification). Underspecification theory cannot explain the

[10] Some studies found that voiced obstruent pairs are more similar to each other than voiceless obstruent pairs are (Greenberg & Jenkins 1964, Mohr & Wang 1968), but the difference is not observed in (2) – possibly because the difference is too small to be reflected in our database.

difference between [place] in nasal consonant and [place] in oral consonants either, because [place] is distinctive both in nasal and in oral consonants, and therefore [place] must be specified in both of these environments.

As a last alternative analysis, it has been claimed that when two sounds alternate with each other in phonology, their phonological relation may enhance similarity between them (Zwicky 1976; Malone 1987, 1988a, b; Hume & Johnson 2003; Shinohara 2004; Kawahara 2007; cf. Steriade 2003, who doubts this claim). However, it is impossible to derive the high O/E ratio of the [m–n] pair from their phonological alternating status. In Japanese phonology, [place] in nasals and [place] in oral consonants do not behave differently. In onset position, neither nasal nor oral consonants change their place specifications. In codas, both nasal and oral consonants place-assimilate to the following consonant (Itô 1986), as exemplified in (3).

(3) (a) *Nasal place assimilation*

ho**N**	'book'
ho**m-b**ako	'book-box'
ho**n-d**ana	'book-shelf'
ho**ŋ-g**a	'book-NOM'

(b) *Place assimilation of oral consonants*

bu**t**(u)	'to hit'	hi**k**(u)	'to pull'
bu**p-p**anasu	'to fire a bullet'	hi**p-p**aru	'to pull off'
bu**t-t**oosu	'to permeate'	hi**t-t**oraeru	'to arrest'
bu**k-k**orosu	'to kill'	hi**k-k**aku	'to scratch'

To summarize, no independent mechanisms of distinctive features derive the difference between [place] in nasals and [place] in oral consonants.

4.2 *Sensitivity to different salience of different features*

The second argument that pun composers deploy psychoacoustic rather than featural similarity rests on the fact that they treat some features as perceptually more salient than other features. A large body of psycholinguistic work has demonstrated the non-equivalence of perceptibility of different features (Miller & Nicely 1955, Peters 1963, Singh & Black 1966, Mohr & Wang 1968, Ahmed & Agrawal 1969, Singh, Woods & Becker 1972, Wang & Bilger 1973, Walden & Montgomery 1975, van den Broecke 1976, Benkí 2003, Bailey & Hahn 2005). Therefore, if speakers are using psychoacoustic similarity, it is predicted that we would observe that different features contribute to pun pairing frequencies to different degrees. We now show that this prediction is correct.

Here we discuss the perceptibility of [voice]. Some scholars have proposed that among manner features that determine spectral continuity ([cont, nasal, voice]), the [voice] feature contributes only weakly to consonant distinctions. The weaker perceptibility of [voice] obtains support from previous psycholinguistic findings, such as Multi-Dimensional Scaling (Peters 1963, Walden & Montgomery 1975; see also Singh et al. 1972), a similarity judgment task (Greenberg & Jenkins 1964, van den Broecke 1976, Bailey & Hahn 2005: 352), and an identification experiment in noise (Wang & Bilger 1973; see also Singh & Black 1966).[11] See also Kenstowicz (2003) and Steriade (2001b), who argue that [voice] is phonologically more likely to neutralize than other manner features because [voice] is less perceptible.

Given the relatively low perceptibility of [voice], if pun composers are sensitive to psychoacoustic similarity, we expect that pairs of consonants that are minimally different in [voice] should combine more frequently than minimal pairs that disagree in other manner features. To compare the effect of [cont], [nasal], and [voice], we compare the O/E ratios for the minimal pairs defined by these features, as in (4).

(4)	[cont]	[nasal]	[voice]
	[p–Φ]: 5.58	[b–m]: 4.68	[p–b]: 8.51
	[t–s]: 0.90	[d–n]: 1.12	[t–d]: 7.64
	[d–z]: 1.68		[k–g]: 8.03
			[s–z]: 11.3
			[ʃ–ʒ]: 6.81

A non-parametric Mann-Whitney test reveals that the O/E ratios of the minimal pairs that differ in [voice] are significantly larger than those of the minimal pairs that differ in [cont] or [nasal] (*Wilcoxon* $W=15$, $z=2.61$, $p<.01$).[12]

Japanese pun-makers treat minimal pairs that differ only in [voice] as very similar, which indicates that composers know that a disagreement in [voice] does not disrupt perceptual similarity as much as other manner features would. In other words, Japanese speakers have an awareness of the varying perceptual salience of different features.[13]

[11] A voicing contrast is more robust than other features in white noise (Miller & Nicely 1955). However, its robustness is probably due to the fact that voicing contrasts involve durational cues (such as VOT, preceding vowel duration, closure duration, and closure voicing duration; Lisker 1986), and white noise does not cover such durational cues. In fact, in signal-dependent noise, the difference in perceptibility between voicing and other manner features is highly attenuated, if not completely obliterated (Benkí 2003).

[12] Previous studies have observed that voicing disagreement is common in Japanese imperfect puns (Otake & Cutler 2001, Shinohara 2004); our study provides statistical support for this observation.

[13] The perceived similarity between minimal pairs of consonants differing in [voice] may be enhanced in Japanese because they are distinguished in Japanese orthography only by a diacritic, i.e. they are orthographically similar (see Itô, Kitagawa & Mester 1996 for the influence of orthography on verbal art). However, [voice]'s weak contribution to perceptual similarity has been demonstrated in languages other than Japanese (see the references cited above), showing that orthographic similarity is not the only factor. Moreover, voicing disagreement is also more common than nasality disagreement in other languages, including English imperfect puns (Lagerquist 1980, Zwicky & Zwicky 1986), slant rhyme in the

A theory that relies on featural similarity cannot explain why a [voice] difference does not reduce pun pairing likelihood as much as differences in other manner features do. To salvage the featural similarity view, one could augment the theory by postulating that [voice] contributes less to featural similarity; but this analysis is ad-hoc. This augmentation simply restates the observation and misses the insight that [voice]'s lesser contribution to the combinability of pun pairs has its basis in its lower perceptibility. Relatedly, Frisch et al. (2004: 204) consider augmenting their model of similarity by introducing weighting. However, weighting does not automatically follow from their model, but is rather 'perceptually and cognitively plausible' (p. 204) – Frisch et al. too, assume that weighting has a psychoacoustic basis.[14]

There are no independent mechanisms in feature theories that derive the observation that [voice] contributes to featural similarity less than other manner features. One might attempt to derive the low contribution of [voice] from its hierarchical position in feature geometry. McCarthy (1988) places [voice] under a Laryngeal node and places [nasal] and [cont] directly under a root node. One could argue that the fact that [voice] is dominated by a higher node – other than a root node – may make the contribution of [voice] smaller than that of [nasal] and [cont] (Golston 1994, cited in Holtman 1996: 252). However, it is not clear why being dominated by some higher node should result in a lower contribution to featural similarity. Nor, to our knowledge, have the proponents of feature geometry proposed to derive different degrees of contribution to featural similarity from hierarchical organization. Moreover, in other models of feature geometry, [nasal] and [cont] are dominated by some higher node as well; both of these features are dominated by a Manner node in Clements (1985); [nasal] is dominated by a Soft Palate and Supralaryngeal node in Sagey (1986) and Halle (1992) (although Sagey notes that the evidence for the Supralaryngeal node is not so strong); [cont] is dominated by an Oral Cavity node in Clements and Hume (1995) and by a Place node in Padgett (1991). Underspecification is of no help either, because all the features in (4) are contrastive in Japanese phonology.

Nor can [voice]'s weak contribution to featural similarity be derived from an alternating status of minimal pairs that differ in voicing. We do find a change in [voice] in Japanese phonology, namely Rendaku, which voices the initial consonant of the second member of compounds (Itô & Mester

poetry of Robert Pinsky (Hanson 1999, cited in Steriade 2001b), and many half rhyme patterns (English: Zwicky 1976, Romanian: Steriade 2003, and Slavic: Eekman 1974). This observation therefore shows that even without orthographic similarity, speakers know that a [voice] difference contributes to perceptual similarity less than other features.

[14] Frisch et al.'s model alone does not predict that a minimal pair of sounds differing in voicing would be most similar to each other. For example, the [p–b] pair has a similarity value of 0.56 while the [p–Φ] pair has a value of 0.67; similarly, the [t–d] pair has a value of 0.54 while the [z–d] pair has a value of 0.57.

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

Table 2
Probabilities of sonorants being paired with voiced obstruents and voiceless obstruents.

1986), as in (5a) (see also Nishimura 2003, Kawahara 2006 for another phonological alternation that changes consonantal voicing). However, [Φ] and [p] also alternate allophonically with each other in Japanese phonology, where [Φ] appears before [u] but [p] appears when geminated (Itô & Mester 1999), as in (5b). Moreover, voiced obstruents, when geminated, are realized as nasal–obstruent clusters, as in (5c) (Kuroda 1965; Itô & Mester 1986, 1999; Kawahara 2006).

(5) (a) *Rendaku*
tako 'octopus' oo-**d**ako 'big octopus'
kai 'shell' hotate-**g**ai 'scallop'
(b) *[ϕ] ~ [p] alternation*
Φuro 'bath' hito**p-p**uro 'one bath'
Φun 'minute' ro**p-p**un 'six minutes'
(c) *Coda nasalization*
/za**b**u+μ+ri/ → [za**mb**uri] 'splashing'
(cf. /ta**p**u+μ+ri/→ [ta**pp**uri] 'a lot')
/o+μ+**d**eru/→ [o**nd**eru] 'get out'
(cf. /o+μ+**k**akeru/→o**kk**akeru 'chase')
('μ' represents a floating mora that causes gemination)

Since Japanese phonology shows evidence for changes in all of [voice], [cont] and [nasal], it is difficult to derive the asymmetry in (4) from alternating statuses in phonology.

4.3 *Sensitivity to similarity contributed by voicing in sonorants*

The third piece of evidence that speakers resort to psychoacoustic similarity is that speakers are sensitive to similarity contributed by a phonologically inert feature – voicing in sonorants. Table 2 shows how often Japanese speakers combine sonorants with voiced obstruents and voiceless obstruents in imperfect puns in our database.

Out of 346 tokens of voiced obstruents, 63 of them are paired with sonorants, whereas out of 497 tokens of voiceless obstruents, only 30 are paired with sonorants. In other words, the probability of voiced obstruents corresponding with sonorants (0.18; s.e. (standard error) = 0.02) is higher

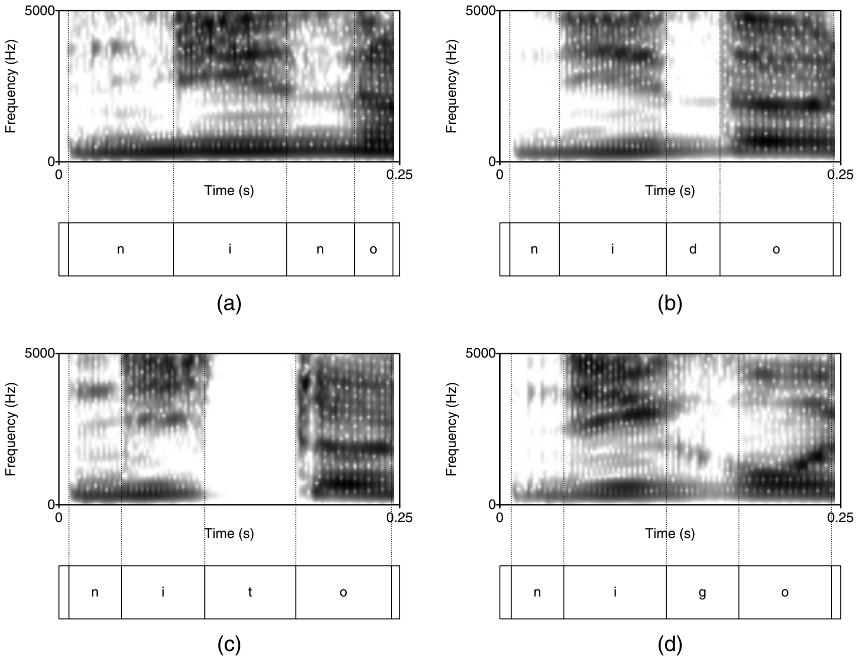

Figure 2
Illustrative spectrograms of Japanese [n, d, t, g] in the environment of [ni _ o]. Read by a female native speaker of Japanese. Pictures generated by Praat (Boersma & Weenink 2007).

than the probability of voiceless obstruents corresponding with sonorants (0.06; s.e. = 0.01). The difference between these ratios is statistically significant (by approximation to a Gaussian distribution, $z = 5.22$, $p < .001$). The pattern in table 2 suggests that pun composers treat sonorants as being more similar to voiced obstruents than to voiceless obstruents (see Stemberger 1991, Walker 2003, Côté 2004, Frisch et al. 2004, Rose & Walker 2004, Coetzee & Pater 2005 for evidence from other languages that obstruent voicing promotes similarity with sonorants; see Kawahara 2007 for the same pattern in Japanese rhymes).

The thesis that speakers use psychoacoustic similarity when forming puns correctly predicts the effect of sonorant voicing on their similarity with voiced obstruents. First, low frequency energy is present during the consonantal constriction in both sonorants and voiced obstruents (figure 2a, b), but not in voiceless obstruents (figure 2c). Second, Japanese voiced stops, especially [g], are often lenited intervocalically, resulting in formant continuity (figure 2d), as with sonorants. Voiceless stops do not spirantize, resulting in complete formant discontinuity, and thus differ from sonorants. For these reasons, voiced obstruents are acoustically more similar to sonorants than voiceless obstruents are.

By contrast, if speakers were deploying featural rather than psychoacoustic similarity, the pattern in table 2 would not be predicted, given the behavior of [+voice] in Japanese sonorants. Phonologically, voicing in Japanese sonorants behaves differently from voicing in obstruents. A well-known phonological restriction in Japanese requires that there be no more than one 'voiced segment' within a stem; but only voiced obstruents, not voiced sonorants, count as 'voiced segments'. Consequently, previous studies have proposed that [+voice] in sonorants is underspecified (Itô & Mester 1986, Itô et al. 1995), or that sonorants do not bear the [voice] feature at all (Mester & Itô 1989), or that sonorants and obstruents bear different [voice] features (Rice 1993). Regardless of how we featurally differentiate voicing in sonorants and voicing in obstruents, sonorants and voiced obstruents do not share the same phonological feature for voicing. Therefore, if speakers were deploying phonological featural similarity, the pattern in table 2 would not be predicted. The featural similarity view – augmented with underspecification or structural differences between sonorant voicing and obstruent voicing – in fact makes the incorrect prediction that voicing in sonorants and voicing in obstruents should not interact with each other.

4.4 *Sensitivity to consonants' similarity to Ø*

As a final piece of evidence for the importance of psychoacoustic similarity, we discuss consonants' similarity to Ø: consonants that correspond with Ø are those that are psychoacoustically similar to Ø. In some imperfect puns, consonants in one phrase do not have a corresponding consonant in the other phrase, as in ***h****ayamatte* ***Ø****ayamatta* 'I apologized without thinking too much' and ***Ø****akagaeru-ga* ***w****akagaeru* 'A brown frog will become young'. In (6) we list the set of consonants whose O/E ratios with Ø are larger than 1 – these are consonants which are paired with Ø more frequently than expected.[15]

(6) [w]: 6.25, [r]: 4.59, [h]: 3.72, [m]: 2.54, [n]: 1.49, [k]: 1.39

The consonants listed in (6) are in fact those that are likely to be psychoacoustically similar to Ø, which supports the claim that speakers use psychoacoustic information when they create imperfect puns.

First, since [w] is a glide, the transition between [w] and the following vowel is blurry, which can make the presence of [w] hard to detect.[16] Myers

[15] Shinohara (2004) also found that [r] and [h] are most likely to correspond with Ø in Japanese imperfect puns, although this observation was based on O-values rather than O/E ratios. See also Lagerquist (1980) for a similar pattern in English (cf. Zwicky & Zwicky 1986).

[16] Since the other palatal glide [y] was very rare in our data (O = 18), it did not enter the O/E analysis.

and Hanssen (2005) demonstrate that given a sequence of two vocoids, listeners misattribute the transitional portion to the second vocoid, effectively lengthening the percept of the second vocoid and shortening the percept of the first vocoid.[17] Thus, due to [w]'s blurry boundaries and consequent misparsing, the presence of [w] is perceptually hard to detect.

Next, Japanese [r] is a flap, which involves a brief and ballistic constriction in Japanese (Nakamura 2002). The shortness of the constriction makes [r] sound similar to Ø. Third, the propensity of [h] to correspond with Ø accords with the observation that [h] lacks a superlaryngeal constriction and hence its spectral property assimilates to that of neighboring vowels (Keating 1988, Pierrehumbert & Talkin 1992), making the presence of [h] difficult to detect; [h] is a perceptually weak consonant (see Mielke 2003 for the perceptibility of [h] in various phonetic contexts).

Fourth, for [m] and [n], we speculate that the edges of these consonants with flanking vowels are blurry due to coarticulatory nasalization, causing them to be interpreted as belonging to the neighboring vowels. Downing (2005) argues that the transitions between vowels and nasals can be misparsed due to this blurriness, and that the misparsing effectively lengthens the percept of the vowel. As a result of misparsing, the perceived duration of nasals may become shortened.

Finally, [k] extensively coarticulates with adjacent vowels in terms of tongue backness (Sussman, McCaffrey & Matthews 1991, Keating & Lahiri 1993, Keating 1996). As a result of this extensive coarticulation, [k] can fade into its environment and become perceptually similar to Ø (de Lacy & Kingston 2006).[18]

To summarize, speakers pair Ø with segments that are psychoacoustically similar to Ø – especially those that fade into their environments – but not with consonants whose presence is highly perceptible. Stridents, for example, are not likely to correspond with Ø because of their salient long duration and great intensity of the noise spectra (Wright 1996, Steriade 2001a). Additionally, coronal stops coarticulate least with surrounding vowels (Sussman et al. 1991, de Lacy & Kingston 2006) and hence they stand out perceptually from their environments. As a result, they are unlikely to correspond with Ø.

By contrast, phonological similarity does not offer a straightforward explanation for the set of consonants in (6). The list of consonants in (6) does not alternate systematically with Ø in Japanese phonology. We observe that

[17] The perceived shortening of the first vocoid may be exaggerated by durational contrast, whereby the percept of one interval gets shortened next to long intervals (Kluender, Diehl & Wright 1988, Diehl & Walsh 1989). However, we also acknowledge that durational contrast has not been replicated by some later studies (Fowler 1992, van Dommelen 1999).

[18] Given this explanation, one may wonder why [g] does not behave like [k]. We conjecture that since Japanese [g] is spirantized intervocalically (Kawahara 2006; see figure 2d above), it stands out from its environment with its frication noise.

(6) includes sonorant consonants, with the exception of [h] and [k];[19] but as Kirchner (1998: 17–18) notes, there is no sense in which sonorous consonants are similar to Ø phonologically. We should in fact not consider Ø as being sonorous: while languages prefer to have sonorous segments in syllable nuclei (Dell & Elmedlaoui 1985, Prince & Smolensky 1993/2004), no languages prefer to have Ø nuclei. Sonority thus fails to explain why speakers prefer to pair Ø with the set of consonants in (6).

Rather, the list of consonants in (6) includes sonorants because their phonetic properties make them akin to Ø: their edges with flanking vowels are blurry, fading into their environments. This view is supported by the fact that [k] and [h], which like sonorants blend into their environment, are treated as similar to Ø.

As another alternative, one could try to characterize Ø in terms of distinctive features, making use of the theory of underspecification. Recall that this theory postulates that non-contrastive or unmarked feature specifications are underlyingly unspecified (Kiparsky 1982, Itô & Mester 1986, Archangeli 1988, Mester & Itô 1989, Paradis & Prunet 1991, Itô et al. 1995). Based on this view, we could postulate that the segment which has the sparsest underlying specification is closest to Ø. This analysis, however, does not predict that the list of consonants in (6) would be close to Ø. For example, nasal consonants are marked while oral consonants are not, and sonorants are marked while obstruents are not (Chomsky & Halle 1968); therefore, [−nasal] and [−son] should be underspecified. As a result, oral obstruents are predicted to be closest to Ø because they lack underlying specifications for [±nasal] and [±son]. However, the list in (6) includes nasal and sonorant consonants. Moreover, no version of underspecification postulates that [k] is more sparsely specified than [t]: in fact, the proponents of Radical Underspecification argue for the opposite (Archangeli 1988, Paradis & Prunet 1991, Lahiri & Reetz 2002).[20] For these reasons, the theory of underspecification does not account for why the consonants in (6) are treated as close to Ø in Japanese imperfect puns.

4.5 *Summary: psychoacoustic similarity vs. featural similarity*

Japanese speakers perceive similarity between sounds based on acoustic information, and use that knowledge of psychoacoustic similarity to compose imperfect puns. Featural similarity fails to capture the bases of similarity

[19] Some proposals treat [h] as a (voiceless) sonorant (Chomsky & Halle 1968), but the [−son] status of [h] is supported by debuccalization processes (/s/ → [h]: Lass 1976, Sagey 1986) as well as by a psycholinguistic experiment (Jaeger & Ohala 1984) and an acoustic experiment (Parker 2002).

[20] Spring (1994) argues that in Axininca Campa, a velar glide is underlyingly underspecified, while [t] is not. However, Spring does not argue for general velar underspecification, and hence even under this theory [k] is no more underspecified than [t].

judgments that speakers make in composing pun pairs. We have presented four kinds of arguments: (i) context-dependent salience of the same feature, (ii) relative salience of different features, (iii) similarity contributed by a phonologically inert feature, and (iv) similarity between consonants and Ø. We have also argued throughout that these shortcomings of the featural similarity model cannot be salvaged by existing elaboration of feature theories, such as feature geometry or underspecification theory. Recall also that there are pairs like the [r–d] pair, whose high degree of over-representation cannot be explained in terms of featural similarity (see also footnote 8 for other potential examples of this kind).

Rather, Japanese speakers use psychoacoustic similarity when they make imperfect puns. A future investigation should construct a complete psycho-acoustic similarity matrix for Japanese sounds through psycholinguistic experiments (e.g. identification experiments in noise and similarity judgment tasks) and apply it to a further analysis of the imperfect pun patterns. Since this task will require a completely new set of experiments, we leave it for future research.

5. Conclusion

5.1 *Parallels between verbal art and phonological patterns*

In closing this paper, we discuss non-trivial parallels which we found between the imperfect pun patterns and phonological patterns. For example, in Japanese imperfect puns, [m] and [n] are more likely to correspond with each other than [p] and [t] are, just as in other languages' phonology in which [m] and [n] alternate with each other (i.e. nasal place assimilation) while [p] and [t] do not (Cho 1990, Mohanan 1993, Jun 2004). In both verbal art and phonology, nasal pairs are more likely to correspond with each other than oral consonant pairs.

Similarly, in the pun pairing patterns, minimal pairs that differ in voicing are more likely to correspond with each other than other types of minimal pairs that differ in nasality and continuency. This patterning parallels the cross-linguistic observation that a voicing contrast is more easily neutralized than other manner features (Steriade 2001b, Kenstowicz 2003). In both cases, a voicing mismatch between two corresponding elements is tolerated, more so than a mismatch in other features.

Third, in section 4.3 we observed that voicing in sonorants enhances similarity with voiced obstruents, and this finds its parallel in linguistic similarity avoidance patterns. In some languages, pairs of sonorants and voiced obstruents are disfavored, as in Arabic (Frisch et al. 2004: 195) and Muna (Coetzee & Pater 2005). Finally, we demonstrated in section 4.4 that when consonants correspond with Ø, speakers tend to choose those consonants that fade into their environment. Again, this pattern finds a parallel in phonology: when speakers epenthesize a segment, they typically choose segments that are the least intrusive (usually glottal consonants for consonants and schwa or high vowels for vowels: Dupoux et al. 1999, Steriade 2001b, Kenstowicz 2003, Howe & Pulleyblank 2004, Walter 2004, Kwon 2005, though see de Lacy & Kingston 2006 for the observation that [k] is never epenthetic in natural languages, presumably for a markedness reason).

With these results we conclude that there exist non-trivial parallels between verbal art patterns and phonological patterns.[21] As we have argued throughout this paper, the parallels exist because speakers use psychoacoustic similarity in shaping both phonological and verbal art patterns. Given the parallels, moreover, we further conclude that it can be fruitful to investigate speakers' grammatical knowledge by examining para-linguistic patterns (Ohala 1986, Itô et al. 1996, Fabb 1997, Yip 1999, Minkova 2003, Steriade 2003, Kawahara 2007 as well as references cited above). Our strategy has been inspired by a growing interest in using verbal art as a source of information about our linguistic knowledge. We hope our studies have shown the fruitifulness of investigating phonological hypotheses using external evidence (Churma 1979).

5.2 *Overall summary*

The combinability of two consonants in imperfect puns in Japanese correlates positively with their perceptual similarity. The analysis of imperfect puns has shown that speakers possess a rich knowledge of psychoacoustic similarity and deploy it in composing imperfect puns, supporting the tenets of the P-map hypothesis (Steriade 2001a, b, 2003; Fleischhacker 2005; Kawahara 2006, 2007; Zuraw 2007). Featural similarity, on the other hand, fails to capture the bases of the similarity decisions that speakers evidently make.

[21] Recall from section 4 that the effects of psychoacoustic similarity cannot be derived from patterns in Japanese phonology. In other words, the above parallels exist between Japanese imperfect pun patterns and phonological patterns in other languages. Therefore, the parallels arise not because speakers compose imperfect puns based on their knowledge of Japanese phonology, but rather because both Japanese imperfect pun patterns and phonological patterns in other languages are grounded in the same principle – the principle of the minimization of perceptual disparities.

APPENDIX

Feature matrix used for the analysis in section 3

	son	cons	cont	nasal	voice	strid	place	palatalized
p	−	+	−	−	−	−	lab	−
b	−	+	−	−	+	−	lab	−
Φ	−	+	+	−	−	−	lab	−
m	+	+	−	+	+	−	lab	−
w	+	−	+	−	+	−	lab	−
t	−	+	−	−	−	−	cor	−
d	−	+	−	−	+	−	cor	−
s	−	+	+	−	−	+	cor	−
z	−	+	+	−	+	+	cor	−
ʃ	−	+	+	−	−	+	cor	+
ʒ	−	+	+	−	+	+	cor	+
ʧ	−	+	±	−	−	+	cor	+
n	+	+	−	+	+	−	cor	−
r	+	+	+	−	+	−	cor	−
k	−	+	−	−	−	−	dors	−
g	−	+	−	−	+	−	dors	−
h	−	+	+	−	−	−	phary	−

REFERENCES

Ahmed, Rais & S. S. Agrawal. 1969. Significant features in the perception of (Hindi) consonants. *Journal of the Acoustical Society of America* 45, 758–773.

Albright, Adam (n.d.). Segmental similarity calculator [software]. http://web.mit.edu/albright/www/.

Archangeli, Diana. 1988. Aspects of underspecification theory. *Phonology* 5, 183–208.

Bailey, Todd & Ulrike Hahn. 2005. Phoneme similarity and confusability. *Journal of Memory and Language* 52, 339–362.

Benkí, José. 2003. Analysis of English nonsense syllable recognition in noise. *Phonetica* 60, 129–157.

Blumstein, Sheila. 1986. On acoustic invariance in speech. In Joseph Perkell & Dennis Klatt (eds.), *Invariance and variability in speech processes*, 178–193. Hillsdale, NJ: Lawrence Erlbaum.

Blumstein, Sheila & Kenneth Stevens. 1979. Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America* 66, 1001–1017.

Boersma, Paul. 1998. *Functional phonology: Formalizing the interaction between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Boersma, Paul & David Weenink. 2007. Praat: Doing phonetics by computer (version 5.0.34). http://www.praat.org.

Cho, Young-mee Yu. 1990. *Parameters of consonantal assimilation*. Ph.D. dissertation, Stanford University.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Churma, Don. 1979. *Arguments from external evidence in phonology*. Ph.D. dissertation, Ohio State University.

Clements, G. N. 1985. The geometry of phonological features. *Phonology Yearbook* 2, 225–252.
Clements, G. N. & Elizabeth Hume. 1995. The internal organization of speech sounds. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 245–306. Malden, MA & Oxford: Blackwell.
Coetzee, Andries & Joe Pater. 2005. Lexically gradient phonotactics in Muna and Optimality Theory. Ms., University of Massachusetts, Amherst.
Côté, Marie-Hélène. 2004. Syntagmatic distinctness in consonant deletion. *Phonology* 21, 1–41.
Cutler, Anne & Takashi Otake. 2002. Rhythmic categories in spoken-word recognition. *Journal of Memory and Language* 46, 296–322.
Davis, Barbara & Peter MacNeilage. 2000. An embodiment perspective on the acquisition of speech perception. *Phonetica* 57, 229–241.
de Lacy, Paul & John Kingston. 2006. Synchronic explanation. Ms., Rutgers University & University of Massachusetts, Amherst.
Dell, François & Mohamed Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7, 105–130.
Diehl, Randy & Margaret Walsh. 1989. An auditory basis for the stimulus-length effect in the perception of stops and glides. *Journal of the Acoustical Society of America* 85, 2154–2164.
Downing, Laura. 2005. On the ambiguous segmental status of nasal in homorganic NC sequences. In Marc van Oostendorp & Jeroen M. van der Weijer (eds.), *The internal organization of phonological segments*, 183–216. Berlin: Mouton de Gruyter.
Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25, 1568–1578.
Eekman, Thomas. 1974. *The realm of rime: A study of rime in the poetry of the Slavs*. Amsterdam: Adolf M. Hakkert.
Fabb, Nigel. 1997. *Linguistics and literature: Language in the verbal arts of the world*. Oxford: Blackwell.
Fay, David & Anne Cutler. 1977. Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry* 8, 505–520.
Fleischhacker, Heidi. 2005. *Similarity in phonology: Evidence from reduplication and loan adaptation*. Ph.D. dissertation, UCLA.
Fowler, Carol. 1992. Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics* 20, 143–165.
Frisch, Stephan, Janet B. Pierrehumbert & Michael Broe. 2004. Similarity avoidance and the OCP. *Natural Language & Linguistic Theory* 22, 179–228.
Golston, Chris. 1994. The phonology of rhyme. Presented at 1994 LSA Meeting.
Greenberg, Joseph & James Jenkins. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20, 157–177.
Guion, Susan. 1998. The role of perception in the sound change of velar palatalization. *Phonetica* 55, 18–52.
Halle, Morris. 1992. Phonological features. In William Bright (ed.), *International encyclopedia of linguistics*, vol. 3, 207–212. Oxford: Oxford University Press.
Hanson, Kristin. 1999. Formal variation in the rhymes of Robert Pinsky's *The Inferno* of Dante. Ms., University of California, Berkeley.
Hempelmann, Christian. 2003. *Paronomasic puns: Target recoverability towards automatic generation*. Ph.D. dissertation, Purdue University.
Holtman, Astrid. 1996. *A generative theory of rhyme: An optimality approach*. Ph.D. dissertation, Utrecht Institute of Linguistics.
Howe, Darin & Douglas Pulleyblank. 2004. Harmonic scales as faithfulness. *Canadian Journal of Linguistics* 49, 1–49.
Huang, Tsan. 2001. The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. In Elizabeth Hume & Keith Johnson (eds.), *Studies on the interplay of speech perception and phonology* (Ohio State University Working Papers in Linguistics 55), 23–42. Columbus, OH: OSU Working Papers in Linguistics.
Huffman, Marie K. & Rena A. Krakow (eds.). 1993. *Nasals, nasalization, and the velum*. New York: Academic Press.

Hume, Elizabeth & Keith Johnson. 2003. The impact of partial phonological contrast on speech perception. *15th International Congress of Phonetic Sciences* (*ICPhS*) *2003, Barcelona*, 2385–2388.

Hura, Susan, Björn Lindblom & Randy Diehl. 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech* 35, 59–72.

Itô, Junko. 1986. *Syllable theory in prosodic phonology*. Ph.D. dissertation, University of Massachusetts, Amherst.

Itô, Junko, Yoshihisa Kitagawa & Armin Mester. 1996. Prosodic faithfulness and correspondence: Evidence from a Japanese argot. *Journal of East Asian Linguistics* 5, 217–294.

Itô, Junko & Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17, 49–73.

Itô, Junko & Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 62–100. Oxford: Blackwell.

Itô, Junko, Armin Mester & Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26, 571–614.

Jaeger, Jeri & John Ohala. 1984. On the structure of phonetic categories. *Tenth Annual Meeting of the Berkeley Linguistics Society* (*BLS* 10), 15–26.

Jakobson, Roman. 1960. *Language in literature*. Cambridge, MA: Harvard University Press.

Jakobson, Roman, Gunnar Fant & Morris Halle. 1952. *Preliminaries to speech analysis* (Technical Report 13). Cambridge, MA: MIT.

Johnson, Keith. 2003. *Acoustic and auditory phonetics*, 2nd edn. Malden, MA & Oxford: Blackwell.

Jun, Jongho. 2004. Place assimilation. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*, 58–86. Cambridge: Cambridge University Press.

Kang, Yoonjung. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20, 219–273.

Katayama, Motoko. 1998. *Optimality Theory and Japanese loanword phonology*. Ph.D. dissertation, University of California, Santa Cruz.

Kawahara, Shigeto. 2003. On a certain kind of hiatus resolution in Japanese. *Phonological Studies* 6, 11–20.

Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of voicing in Japanese. *Language* 82, 536–574.

Kawahara, Shigeto. 2007. Half-rhymes in Japanese rap lyrics and knowledge of similarity. *Journal of East Asian Linguistics* 16, 113–144.

Kawahara, Shigeto, Hajime Ono & Kiyoshi Sudo. 2006. Consonant co-occurrence restrictions in Yamato Japanese. In Timothy Vance & Kimberly Jones (eds.), *Japanese/Korean Linguistics* 14, 27–38. Stanford, CA: CSLI Publications.

Kawahara, Shigeto & Kazuko Shinohara. 2007. The role of psychoacoustic similarity in Japanese imperfect puns. In Michael Becker (ed.), *University of Massachusetts Occasional Papers in Linguistics* 36, 111–133. Amherst, MA: Graduate Linguistic Student Association (GLSA).

Keating, Patricia. 1988. Underspecification in phonetics. *Phonology* 5, 275–292.

Keating, Patricia. 1996. The phonology–phonetics interface. *UCLA Working Papers in Phonetics* 92, 45–60.

Keating, Patricia & Aditi Lahiri. 1993. Fronted velars, palatalized velars, and palatals. *Phonetica* 50, 73–101.

Kenstowicz, Michael. 2003. Salience and similarity in loanword adaptation: A case study from Fijian. Ms., MIT.

Kiparsky, Paul. 1970. Metrics and morphophonemics in the *Kalevala*. In Donald Freeman (ed.), *Linguistics and literary style*, 165–181. Orlando, FL: Holt, Rinehart and Winston.

Kiparsky, Paul. 1972. Metrics and morphophonemics in the Rigveda. In Michael Brame (ed.), *Contributions to generative phonology*, 171–200. Austin, TX: University of Texas Press.

Kiparsky, Paul. 1981. The role of linguistics in a theory of poetry. In Donald Freeman (ed.), *Essays in modern stylistics*, 9–23. London: Methuen.

Kiparsky, Paul. 1982. Lexical phonology and morphology. In In-Seok Yang (ed.), *Linguistics in the morning calm*, 3–91. Seoul: Hanshin.

Kirchner, Robert. 1998. *An effort-based approach to consonant lenition*. Ph.D. dissertation, UCLA.

Klatt, Dennis. 1968. Structure of confusions in short-term memory between English consonants. *Journal of the Acoustical Society of America* 59, 1208–1221.
Klockars, Alan J. & Gilbert Sax. 1986. *Multiple comparisons*. London: Sage.
Kluender, Keith, Randy Diehl & Beverly Wright. 1988. Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics* 16, 153–169.
Kohler, Klaus. 1990. Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In William J. Hardcastle & Alain Marchal (eds.), *Speech production and speech modelling*, 69–92. Dordrecht: Kluwer.
Kuroda, S.-Y. 1965. *Generative grammatical studies in the Japanese language*. Ph.D. dissertation, MIT.
Kurowski, Kathleen & Sheila Blumstein. 1993. Acoustic properties for the perception of nasal consonants. In Huffman & Krakow (eds.), 197–224.
Kwon, Bo-Young. 2005. The patterns of vowel insertion in IL phonology: The P-map account. *Studies in Phonetics, Phonology and Morphology* 11, 21–49.
LaCharité, Darlene & Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36, 223–258.
Lagerquist, Linnes. 1980. Linguistic evidence from paronomasia. *Chicago Linguistic Society* (*CLS*) 16, 185–191.
Lahiri, Aditi & Henning Reetz. 2002. Underspecified recognition. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology* 7, 637–675. Berlin: Mouton de Gruyter.
Lass, Roger. 1976. *English phonology and phonological theory*. Cambridge: Cambridge University Press.
Lisker, Leigh. 1986. 'Voicing' in English: A catalog of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29, 3–11.
Lombardi, Linda. 1990. The nonlinear organization of the affricate. *Natural Language & Linguistic Theory* 8, 375–425.
Maher, Peter John. 1969. English-speakers' awareness of distinctive features. *Language Sciences* 5, 14.
Maher, Peter John. 1972. Distinctive feature rhyme in German folk versification. *Language Sciences* 19, 19–20.
Malécot, André. 1956. Acoustic cues for nasal consonants: An experimental study involving a tape-splicing technique. *Language* 32, 274–284.
Malone, Joseph. 1987. Muted euphony and consonant matching in Irish. *Germanic Linguistics* 27, 133–144.
Malone, Joseph. 1988a. On the global-phonologic nature of classical Irish alliteration. *Germanic Linguistics* 28, 93–103.
Malone, Joseph. 1988b. Underspecification theory and Turkish rhyme. *Phonology* 5, 293–297.
McCarthy, John. 1988. Feature geometry and dependency: A review. *Phonetica* 43, 84–108.
McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *University of Massachusetts Occasional Papers in Linguistics* 18, 249–384. Amherst, MA: Graduate Linguistic Student Association (GLSA).
Mester, Armin & Junko Itô. 1989. Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65, 258–293.
Mielke, Jeff. 2003. The interplay of speech perception and phonology: Experimental evidence from Turkish. *Phonetica* 60, 208–229.
Miller, George & Patricia Nicely. 1955. An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* 27, 338–352.
Minkova, Donca. 2003. *Alliteration and sound change in early English*. Cambridge: Cambridge University Press.
Mohanan, K. P. 1993. Fields of attraction in phonology. In John A. Goldsmith (ed.), *The last phonological rule: Reflections on constraints and derivations*, 61–116. Chicago: University of Chicago Press.
Mohr, B. & W. S. Wang. 1968. Perceptual distance and the specification of phonological features. *Phonetica* 18, 31–45.
Mukai, Yoshihito. 1996. *Kotoba asobi no jyugyoo zukuri* [Teaching with word games]. Tokyo: Meiji Tosho.

Myers, Jerome & Arnold Well. 2003. *Research design and statistical analysis*. Mahwah, NJ: Lawrence Erlbaum.

Myers, Scott & Benjamin Hanssen. 2005. The origin of vowel-length neutralization in vocoid sequences. *Phonology* 22, 317–344.

Nakamura, Mitsuhiro. 2002. The articulation of the Japanese /r/ and some implications for phonological acquisition. *Phonological Studies* 5, 55–62.

Nishimura, Kohei. 2003. Lyman's law in loanwords. Ms., University of Tokyo.

Ohala, John. 1986. Consumer's guide to evidence in phonology. *Phonology* 3, 3–26.

Ohala, John. 1989. Sound change is drawn from a pool of synchronic variation. In Leiv Egil Breivik & Ernst Håkon Jahr (eds.), *Language change: Contributions to the study of its causes*, 173–198. Berlin: Mouton de Gruyter.

Ohala, John & Manjari Ohala. 1993. The phonetics of nasal phonology: Theorems and data. In Huffman & Krakow (eds.), 225–249.

Otake, Takashi & Anne Cutler. 2001. Recognition of (almost) spoken words: Evidence from word play in Japanese. *European Conference on Speech Communication and Technology* 7 (EUROSPEECH 2001 Scandinavia), 464–468.

Padgett, Jaye. 1991. *Stricture in feature geometry*. Ph.D. dissertation, University of Massachusetts, Amherst.

Paradis, Carol & Jean-François, Prunet (eds.). 1991. *The special status of coronals: Internal and external evidence*. San Diego, CA: Academic Press.

Parker, Steve. 2002. *Quantifying the sonority hierarchy*. Ph.D. dissertation, University of Massachusetts, Amherst.

Peters, Robert. 1963. Dimensions of perception for consonants. *Journal of the Acoustical Society of America* 35, 1985–1989.

Pierrehumbert, Janet B. & David Talkin. 1992. Lenition of /h/ and glottal stop. In Gerard Docherty & Robert Ladd (eds.), *Papers in Laboratory Phonology II*, 90–117. Cambridge: Cambridge University Press.

Pols, Louis. 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.

Price, Patti. 1981. *A cross-linguistic study of flaps in Japanese and in American English*. Ph.D. dissertation, University of Pennsylvania.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA & Oxford: Blackwell.

Rice, Keren. 1993. A reexamination of the feature [sonorant]: The status of sonorant obstruents. *Language* 69, 308–344.

Rose, Sharon & Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80, 475–532.

Sagey, Elizabeth. 1986. *The representation of features and relations in nonlinear phonology*. Ph.D. dissertation, MIT.

Seco, G. V., I. A. Menéndez de la Fuente & J. R. Escudero. 2001. Pairwise multiple comparisons under violation of the independence assumption. *Quality and Quantity* 35, 61–76.

Shattuck-Hufnagel, Stefanie & Dennis Klatt. 1977. The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Behavior* 18, 41–55.

Shinohara, Shigeko. 2004. A note on the Japanese pun, *dajare*: Two sources of phonological similarity. Ms., Laboratoire de Psychologie Experimentale.

Singh, Rajendra, David Woods & Gordon Becker. 1972. Perceptual structure of 22 prevocalic English consonants. *Journal of the Acoustical Society of America* 52, 1698–1713.

Singh, Sadanand & John Black. 1966. Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *Journal of the Acoustical Society of America* 39, 372–387.

Sobkowiak, Włodzimierz. 1991. *Metaphonology of English paronomasic puns*. Frankfurt am Main: Lang.

Spring, Cari. 1994. Combinatorial specification of a consonantal system. In Phillip Spaelti (ed.), *Twelfth West Coast Conference on Formal Linguistics* (*WCCFL* 12), 133–152. Stanford, CA: CSLI Publications.

Stemberger, Joseph. 1991. Radical underspecification in language production. *Phonology* 8, 73–112.

Steriade, Donca. 1994. Positional neutralization and the expression of contrast. Ms., UCLA. [A revised, written-up version of 1993 NELS handout.]
Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Michael B. Broe & Janet B. Pierrehumbert (eds.), *Acquisition and the lexicon* (Papers in Laboratory Phonology V), 313–334. Cambridge: Cambridge University Press.
Steriade, Donca. 2001a. Directional asymmetries in place assimilation: A perceptual account. In Elizabeth Hume & Keith Johnson (eds.), *The role of speech perception in phonology*, 219–250. New York: Academic Press.
Steriade, Donca. 2001b. The phonology of perceptibility effect: The P-map and its consequences for constraint organization. Ms., UCLA.
Steriade, Donca. 2003. Knowledge of similarity and narrow lexical override. *29th Annual Meeting of Berkeley Linguistics Society* (*BLS* 29), 583–598.
Stevens, Kenneth & Sheila Blumstein. 1978. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64, 1358–1368.
Sussman, Harvey, Helen McCaffrey & Sandra Matthews. 1991. An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America* 90, 1309–1325.
Takizawa, Osamu. 1996. *Nihongo syuuzihyougen no kougakuteki kenkyuu* [Technological study of Japanese rhetorical expressions]. Tokyo: Tsuushin Sogo Kenkyuusho.
Trubetzkoy, Nikolai. 1969. *Principles of phonology*. Berkeley, CA: University of California Press.
van den Broecke, M. P. R. 1976. *Hierarchies and rank orders in distinctive features*. Amsterdam: Van Gorcum Assen.
van Dommelen, Wim. 1999. Auditory accounts of temporal factors in the perception of Norwegian disyllables and speech analogs. *Journal of Phonetics* 27, 107–123.
Walden, Brian & Allen Montgomery. 1975. Dimensions of consonant perception in normal and hearing-impaired listeners. *Journal of Speech and Hearing Research* 18, 445–455.
Walker, Rachel. 2003. Nasal and oral consonant similarity in speech errors: Exploring parallels with long-distance nasal agreement. Ms., University of Southern California.
Walter, Mary-Ann. 2004. Loan adaptation in Zazaki Kurdish. In Michael Kenstowicz (ed.), *Working Papers in Endangered and Less Familiar Languages* 6 (MIT Working Papers in Linguistics), 97–106.
Wang, Marilyn D. & Robert C. Bilger. 1973. Consonant confusions in noise: A study of perceptual features. *Journal of the Acoustical Society of America* 54, 1248–1266.
Wilson, Colin. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30, 945–980.
Wright, Richard. 1996. *Consonant clusters and cue preservation in Tsou*. Ph.D. dissertation, UCLA.
Yip, Moira. 1999. Reduplication as alliteration and rhyme. *Glot International* 4, 1–7.
Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog reduplication. *Language* 83, 277–316.
Zwicky, Arnold. 1976. This rock-and-roll has got to stop: Junior's head is hard as a rock. *Chicago Linguistic Society* (*CLS*) 12, 676–697.
Zwicky, Arnold & Elizabeth Zwicky. 1986. Imperfect puns, markedness, and phonological similarity: With fronds like these, who needs anemones? *Folia Linguistica* 20, 493–503.

Authors' addresses: (*Kawahara*)
Linguistics Department, Rutgers University, 18 Seminary Place,
New Brunswick, NJ 08901-1108, USA
kawahara@rci.rutgers.edu

(*Shinohara*)
Institute of Symbiotic Science and Technology,
Tokyo University of Agriculture and Technology, 2-24-16 Nakacho,
Koganei, Tokyo 184-8588, Japan
k-shino@cc.tuat.ac.jp