Abstract

This paper returns to the subject of Japanese Mimetic Palatalization. While this subject has been addressed previously in Mester & Ito (1989), and more recently in Zoll (1997), this paper raises questions about the analyses presented in these papers, and introduces a new analysis based on perceptual and functional grounds to take its place. Evidence for this new result is taken from a mathematical analysis of the data and a study conducted using nonsense words to test the current approach. The solution is first presented in a mathematical framework before returning to an Optimality Theoretic analysis for comparison with Zoll.

[Note: this very was copied from my old website, where it was saved in HTML, and is missing some equations.]

1. Introduction

In Japanese there is a subclass of the lexicon called mimetics, or in Japanese, *gitaigo*. These words are repetitive roots of the form CVCV and are used primarily as adverbs as described in Hamano (1998). Examples of these words are given in Table 1. These mimetic forms and others in Japanese are iconic and sound-symbolic, and are quite prevalent due to the limited differentiation of Japanese verbs. These examples are taken from Tsujimura (1996).

Table 1. Examples of Japanese Mimetics

a. kata-kata	[katakata]	“homogeneous hitting sound”
b. pota-pota	[potapota]	“dripping”
c. poko-poko	[pokopoko]	“up-and-down movement”
d. noro-noro	[noɾonoɾo]	“slow movement”
e. zabu-zabu	[zabuzabu]	“splashing”

Many of these CVCV mimetic roots have derivative forms with palatalization; that is, with a palatal element like a /y/ attached to one of the consonants in the root. Because mimetics are sound-symbolic, this palatal element attaches additional meaning to the word related to childishness or excessiveness. Examples of the words in Table 1 modified by this palatal element are given in Table 2.

Table 2. Examples of Mimetics with Palatalization

a. katya-katya	[katʃakatʃa]	“clattering sound”
b. potya-potya	[potʃapotʃa]	“dripping in large quantities”

c. pyoko-pyoko [p^jokop^{j}oko] "jumping around imprudently"

d. nyoro-nyoro [ɲoroɲoro] "slow wriggly movement"

e. zyabu-zyabu [dʒabudʒabu] "splashing indiscriminately"

f. dosya-dosya [doʃadoʃa] "in large amounts"

The important thing to note from these examples is that the palatal element will sometimes be applied to the first consonant in the root, and sometimes to the second consonant in the root. This is the basis of the analysis of these roots in previous studies, and in the account presented here. There are some restrictions on where the palatalization may be placed. The palatal element may never be placed in front of a vowel which is considered palatal, i.e. the mid and high front vowels /e/ and /i/. In addition, whenever two non-coronal consonants appear in a root, the palatalization falls on the first consonant in the root (see example in Table 2. c). Whenever two coronal consonants appear in a root, the palatal element falls on the second consonant in the root, unless the second is /r/ (see example in Table 2. d, f). Whenever a coronal consonant and a non-coronal consonant appear together in a word, the coronal consonant receives palatalization in preference to the non-coronal one, regardless of its position in the root (see examples in Table 2. a, b, e); however, /r/ is also avoided in these cases, and the non-coronal consonant receives the palatalization. Most mimetic roots contain one coronal consonant and one non-coronal consonant.

In previous efforts to analyze this problem, the descriptions given above have been taken rather literally, positing some exceptional behaviour for /r/, and phonologizing the rightward and leftward tendencies in the two types of consonant combinations. This approach is flawed, I believe, in two ways. The first is that both analyses posit the behaviour of /r/ to be exceptional. From a scientific standpoint, rules that contain outright exceptions are never to be preferred. Mester & Ito attempt to incorporate the exceptional behaviour of /r/ into their phonology, although their account of its implications have been challenged in later papers, particularly that of Zoll. Zoll, on the other hand, leaves the behaviour of /r/ as a mere stipulation. The second flaw is in taking the description of the behaviour of roots with two coronal consonants too literally. Zoll adopts a constraint that actively acts in these cases to align the palatalization to the rightward consonant; however, if we look at the distribution of palatalization in the language, there is only a slight tendency toward a rightward alignment of palatalization, one that, as I will show later, is not statistically significant. In general, we would like our grammar to be as unburdened as possible by special rules; therefore, the positing of this alignment constraint is certainly less than optimal, and should only be seen as a last resort. I intend to show that we need not appeal to this kind of constraint in order to solve this problem.

2. Linguistics of Japanese Mimetics

In this section, I would like to review the previous analyses of Japanese mimetics as described by Mester & Ito (1989) and Zoll (1997). However, before I do that, I would like to begin with a brief introduction to linguistics, including a review of the phonology of Japanese.

2.1 Phonology of Japanese

Linguistics itself is the study of the structure of language, and what this tells us about the structure of the human brain. Linguistics is divided up into several different areas of concern, primarily phonetics (the study of the sounds used in language), phonology (the organization of these sounds), syntax (sentence structure), and semantics (meaning). While there are other divisions, these are the ones that will be referred to in this paper. Phonetics and phonology in particular are intimately related. Functional phonology is an approach to phonology that appeals to phonetic principles such as articulation and perception in order to explain the behaviour of sounds in phonological systems, more perhaps than in other theoretical approaches.

In the past, a rule-based approach to phonology, which often mirrored the historical development of a language, was employed to explain changes and regular alternations in phonological systems across related words. In the mid-1990's, however, a constraint-based approach to phonological systems was developed by Prince & Smolensky (1993), that has become the dominant approach to phonological studies in some parts of the linguistic community. Rather than ordered rules, Optimality Theory employs ordered constraints, evaluating various possible output forms against the constraint ranking until only the optimal one remains. In order to discuss either one of these theories, a little bit of background on the phonology of Japanese is needed.

Japanese has a simpler syllable structure than does English. While English syllables may have several consonants at both the beginning or the end of a syllable, Japanese syllable onsets may have only one consonant at a time. Japanese may also only have one coda consonant, and these are limited to stop consonants matching the onset of the next syllable in the middle of a word, or a nasal consonant. Japanese does not rely as heavily on the concept of the syllable in its phonology the way that English does, rather, it relies on the concept of the mora, which consist of an optional onset consonant plus the vowel it precedes, or a coda consonant, but never both.

The sounds that occur in the Japanese language are fewer in number than the sounds that occur in English. Japanese has only five vowels, though it may use long and short versions of all of them. These can be approximated by the English vowels in "bee" [biː], "bay" [beː], "bah" [baː], "bow" (and arrow) [boː], and "boo" [bu].

The Japanese consonant inventory is also smaller, but there are some sounds in Japanese that do not occur in English. Japanese has a series labial consonants similar to those in English: /p/ and /b/ are the stop consonants; /m/ is the labial nasal; and /f/ is the labial fricative, and it alternates with /h/ in Japanese, but it is not pronounced like and /f/, it's pronounced by blowing air through the lips as though blowing out a candle. The phonetic symbols that go with these consonants are [p], [b], [m] and [ɸ]. Because /h/ and /f/ alternate, they are treated as the same consonant in the phonology, so [h] also patterns with the labial consonants.

Japanese has a series of velar consonants: /k/ and /g/ are the stop consonants. The velar nasal is not contrastive in Japanese, as is the case in English. There are no velar fricatives in Japanese. As we mentioned above, /h/ patterns with the labials, and not with the velars as it does (historically) in English.

Japanese also has a series of coronal consonants: /t/ and /d/ are the stop consonants; /n/ is the coronal nasal; /s/ and /z/ are the coronal fricatives; /r/ is the coronal approximant, although phonetically it is a flap. Japanese coronal consonants become palatalized before the high front vowel /i/; /t/ becomes [tʃ] like "chew"; /d/ becomes [dʒ] like in "joy"; /s/ becomes [ʃ] like in "she"; /z/ becomes [dʒ]. These phonetic changes are sometimes represented in Romanizations of Japanese words. In addition, the coronal stops /t/ and /d/ become [ts] and [z] in front of the vowel [u].

Japanese has contrastive palatalization before vowels such as /a/, /o/ and /u/. We described palatalization in the introduction as the attachment of a /y/-like element to a consonant. This is a simplification. The palatalized coronal consonants mentioned above participate in the system of palatalization contrast. In addition, there are also the palatalized coronal nasal [ɲ], written in Romanization as "ny"; the palatalized coronal approximant [ʎ], written as "ry". There are also palatalized labials and velars: [p^j] like in English "pew", [b^j] like in "bugle", [m^j] like in "mew", [h^j] like in "hue", [k^j] like in "queue", [g^j]. These consonants are all written in Romanization as the consonant plus -y-. Phonologically, in Japanese, however, these are treated as a single segment, which have an additional featural element of [-anterior] added to its structure. Coronal consonants that receive palatalization undergo a change in place of articulation, from coronal to palatal, but non-coronal consonants are doubly articulated, having two separate places of articulation. The non-coronal palatal consonants thus form a natural class in Japanese in terms of the complexity of their articulation.

There are, finally, two glides in Japanese: /y/ the palatal glide, and /w/ the labial glide. These are pronounced similarly to their counterparts in English, but they are somewhat more limited in distribution. The palatal glide does not appear before the vowels /i/ or /e/, and the labial glide only occurs with the vowel /a/ in modern Japanese, with only a few exceptions in unassimilated borrowing from English. The labial glide does not undergo palatalization, and neither of the glides participate in the palatalization of mimetic words that we will consider, except by serving to block the process altogether.

Japanese phonological system has several separate layers of the language, each with slight variations in the general phonological patterns of the language. Hamano (1998) in her discussion of mimetics describes these layers. The Yamato stratum is the portion of the language that is historically native to Japanese. The Sino-Japanese stratum is the portion of the language based on borrowings into Japanese from Chinese; this portion of the language, while maintaining some of its own features, is well assimilated into the larger phonology of the language. The portion of the language that consists of recent borrowing, mainly from English is the foreign stratum and is less well assimilated into the more general phonological patterns. Finally, there is the mimetic stratum; this portion of the language is sound-symbolic and allows for some sound patterns that do not occur in the non-foreign strata of the language, but these variations from the standard are not the same as those in the foreign stratum. The subset of mimetics that is under discussion in this paper, the CVCV reduplicative mimetic adverbs with palatalization, does not exhibit any of the marked differences from the non-foreign strata.

2.2 History of the Current Analysis

Based on the description of the reduplicative CVCV mimetic adverbs with palatalization that was laid out in Hamano (1998), phonologists attempted to place the behaviour of the palatalization process in linguistic terms. As described in §1, palatalization may take place on either of the two consonants in the root. Hamano describes the pattern of palatalization, and I have summarized it below in Table 3.

Table 3. Rules for Palatalization

- If the root has one (non-r) coronal consonant, place the palatal feature on the coronal consonant.
- If the root has two non-coronal consonants, place the palatal feature on the initial consonant.
- If the root has two (non-r) coronal consonants, the palatal feature falls on the second consonant.
- If one of the consonants in the root is /r/, place the palatal feature on the other consonant.

These "rules" are based on 87 lexical items that are listed in the appendices at the back of Hamano's text. The rules are descriptive. One can see the exceptional behaviour of /r/ written into them. The coronal consonant /r/ occurs relatively frequently in the lexicon, always as the second consonant in the root, but it never receives palatalization, regardless of the consonant it is paired with. Other than this exceptional behaviour of the coronal approximant, coronals otherwise preferentially receive palatalization, regardless of their location in the root. Examples of this were given in Table 2. In addition to these rules, if the vowel next to the consonant that should receive the palatalization based on the above rules is either an /e/ or an /i/, then palatalization does not take place. The palatal element is not moved to the other consonant. A complete analysis would account for this behaviour as well, though we will not be considering it in depth here.

Two accounts of the process are described below. Both accounts take the description provided in Table 3. and attempt to produce the same lexical outputs. The rule-based account was presented by Mester & Ito (1989). The constraint-based account was presented by Zoll (1997). Each will be described in turn below. Both successfully predict the lexical words presented in Hamano (1998).

2.2.1 The Rule-Based Account of Mester & Ito

Mester & Ito (1989) generalized the description of Japanese mimetic palatalization given in Hamano by positing the following rule: The palatalization element is a suffix that begins at the right edge of a word and moves leftward into the word seeking a suitable consonant upon which to dock its [-anterior] feature. If the first consonant it encounters is not deemed suitable, the feature continues to moves leftward until it reaches the left edge of the word. The [-ant] feature docks to this consonant by default.

The explanation of how this rule works depends on feature geometry. Feature geometry argues that each segment in a phonological representation has a feature tree. Among these features are the features of place of articulation. Coronal consonants will have a feature [coronal] in their feature tree; labial consonants will have a feature [labial] in the branch for place of articulation, and so forth. The [-

ant] feature is preferentially docked to a tree with the [cor] feature. The combination of the two features produces a non-complex palatal segment. When the [-ant] segment attaches to a non-coronal place feature such as [lab], this produces a complex, doubly articulated consonant.

Mester & Ito considered the behaviour of /r/ as evidence that not all coronal consonants had feature trees that included the feature [cor] in their underlying phonological representation. Other coronal consonants are contrastive by place. The coronal approximant /r/ is not contrastive with respect to place of articulation. In the feature geometry, the node for place of articulation is empty in the representation of /r/, and the feature [cor] is filled in by default when the phonology passes to the phonetic implementation phase. Because of this, the [-ant] segment, when it encounters an /r/ in its passage leftward through the word, cannot dock to the /r/ because its feature tree does not possess the feature [cor]. Examples of the implementation of this rule employing feature geometry are given in Figure 1. below. The example of Japanese mimetic palatalization became the standard example for contrastive underspecification.

Figure 1. Palatalization in Feature Geometry

<table>
<tr><td>[doʃa]</td><td>[tʃoko]</td><td>[p^{j}oko]</td><td>[k^{j}oɾo]</td></tr>
<tr><td>dosa</td><td>toko</td><td>poko</td><td>koro</td></tr>
<tr><td>| |</td><td>| ·</td><td>· ·</td><td>· ·</td></tr>
<tr><td>[cor][cor]</td><td>[cor]</td><td>\</td><td>\</td></tr>
<tr><td>\</td><td>\</td><td>·</td><td>·</td></tr>
<tr><td>[-ant]</td><td>[-ant]</td><td>\</td><td>\</td></tr>
<tr><td></td><td></td><td>[-ant]</td><td>[-ant]</td></tr>
</table>

One can see the effects of contrastive underspecification in the /koro/ example in Figure 1. When the palatal feature [-ant] begins at the right edge of the word and begins moving leftward, the first consonant it encounters is /r/. In the feature geometry, /r/ is not specified for the place of articulation [coronal], so the [-anterior] segment does not stop there; rather, it continues moving to the left until it finds /k/. This consonant does not have the feature [coronal] specified either. It cannot continue to move leftward because there is no more word left, and it cannot go backward in a rule-based account, so the [-ant] attaches itself as a complex segment. As one can see in the /dosa/ example, the first consonant the [-ant] segment attaches to the first consonant it finds with a [cor] feature upon which to dock. A word such as /kata/ would behave identically to /dosa/ because the [-ant] segment would never get to consider the leftward consonant in the word.

2.2.2 The Constraint-Based Account of Zoll

The constraint-based theory Zoll presents her work in is Optimality Theory. Zoll (1997) reexamines the behaviour of palatalization in Japanese mimetics and attempts to describe the process in these new theoretic terms. Like Mester & Ito, Zoll treats the palatalization element as a morpheme with a constraint aligning it to the right edge of the word. Zoll's analysis employs the concept of conflicting directionality, the interaction of constraints that pulls an element such as stress or palatalization toward opposite ends of the word. Zoll's analysis, as was Mester & Ito's, part of a larger theoretical argument. She takes the mimetic palatalization data as evidence for conflicting directionality, and as an argument against the contrastive underspecification of Mester & Ito as described above.

To account for the alignment of the palatal element, Zoll employs two principle constraints. The first of these constraints is the ALIGN LEFT (COMPLEX SEGMENT, PROSODIC WORD). The full definition Zoll gives is given in Table 4. below. This constraint aligns complex segments, such as non-coronal palatals, to the left edge of a word. Her evidence for this constraint comes from several different sources, including Steriade (1995) and others. Complex segments and other marked features cross-linguistically are licensed preferentially in word-initial position.

The second of her constraints is ALIGN RIGHT ([-ANTERIOR] SEGMENT, PROSODIC WORD). This constraint's full definition is also given in Table 4. below. It says that the [-ant] segment that represents the aspect of childishness or excessiveness in palatalized mimetics is a suffix. She argues that this feature prefers to orient itself toward the right edge of the word. I have reformulated the constraint here to attract the palatalization in the last consonant, rather than the last segment. While this changes the purely suffixal quality of the alignment constraint, it does allow the constraint to conform to the McCarthy (2002) paper against gradience in constraint violation and still be effective. The opposition of these two constraints provides a basis for conflicting directionality. Without gradience, the reformulation seems to be the only way to preserve the account of conflicting directionality, because another constraint would be forced to place the [-ant] feature on the consonant instead of the final vowel. The ranking for these constraints is also provided in Table 4.

A third OCP constraint is introduced limiting the appearance of the palatalization element in front of palatal vowels such as /e/ and /i/. Zoll states this constraint in a very limited fashion. She also pairs it with a low-ranked Parse Feature constraint to explain why a palatal element cannot appear in some mimetic adverbs. We will not consider this situation here.

Table 4. Optimality Theoretic Constraints

ALIGN LEFT (COMPLEX SEGMENT, PROSODIC WORD)

> $\forall$ Complex segments $\exists$ a prosodic word such that a complex segment coincides with the leftmost segment in the prosodic word

ALIGN RIGHT ([-ANTERIOR] SEGMENT, PROSODIC WORD)

> $\forall$ [-anterior] segment $\exists$ a prosodic word such that the [-anterior] segment coincides with the rightmost consonant in the word

Ranking: ALIGN L (COMPLEX SEGMENT, PWD) >> ALIGN R ([-ANT] SEGMENT, PWD)

The effects of the ALIGN L constraint can be seen in Tableau 1. below. This tableau inputs the root /toko/ and the [-ant] suffix. It shows how the ALIGN L constraints prevents the [-ant] segment from docking to the velar consonant /k/ and instead docks to the coronal consonant on the left edge of the word. Even though the ALIGN R constraint prefers to have the [-ant] feature on the right edge of the word, this is cannot be satisfied without violating the higher ranked ALIGN L constraint. The candidate in (b) best satisfied the constraint ranking, and so it becomes the surface form.

Tableau 1.

{toko, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	ALIGN R ([-ANT] SEGMENT, PWD)
a. tokʲo	*!	
b. ☞ tʃoko		*

The ALIGN R constraint successfully draws the palatalization feature to the right edge of the word when there is a (non-r) coronal consonant in the second position. The example in Tableau 2. shows an example with two coronal consonants. The example in Tableau 3. shows an example with an initial non-coronal consonant and a coronal consonant in the rightward position. As one can see, both examples illicit the same constraint violations and the same results in the surface form.

Tableau 2.

{dosa, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	ALIGN R ([-ANT] SEGMENT, PWD)
a. dʒosa		*!
b. ☞ doʃa		

Tableau 3.

{kata, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	ALIGN R ([-ANT] SEGMENT, PWD)
a. kʲata		*!
b. ☞ katʃa		

In the case where two non-coronal consonants appear in the word, the ALIGN L constraint once again draws the palatalization element toward the left edge of the word. This example is shown in Tableau 4. The violation of the ALIGN L constraint when the [-ant] segment docks to the rightmost consonant is fatal.

Tableau 4.

{poko, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	ALIGN R ([-ANT] SEGMENT, PWD)
a. pok^{j}o	*!	
b. ☞ p^{j}oko		*

The last part of Zoll's analysis that I will address here is her treatment of the coronal approximant /r/. Zoll makes the argument that palatalized /r/ is typologically rare and, based on this, makes the claim that palatalized /r/ must therefore be treated as a complex segment in the same class of behaviour as the palatalized non-coronal segments. Therefore, as shown in Tableau 5., the placement of the palatalization on the /r/ generates a violation of ALIGN L. The winning candidate is the one with the palatal feature docked to the initial consonant. Zoll argues that the behaviour of /r/ exemplified here undermines the contrastive underspecification account of Mester & Ito.

Tableau 5.

{koro, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	ALIGN R ([-ANT] SEGMENT, PWD)
a. koʎo	*!	
b. ☞ k^{j}oro		*

2.3 Questions to Be Answered

Both of the accounts described above have several common features. One important commonality of the two accounts is that they argue for the treatment of the palatalization feature to be a morpheme of its own, and in particular, a suffix. A problem with this analysis that has been highlighted both by Mester & Ito and Hamano is that the palatalization has the effect of adding an element of childishness or excessiveness to a root word; however, there is no fixed effect that the segment has on the root. It may mean any number of things according to Hamano (1998), including uncontrolledness, excessive energy, instability, unreliability, uncoordinated, diversity, noisiness, cheapness. Because of this lack of a single semantic definition, it is more likely the [-ant] feature is not an actual morpheme, but remains only a sound-symbolic feature. If the palatalization feature is not a morpheme, can we justify in either account the supposed leftward tendency of the feature that has been described as being suffixal?

The two accounts described above also have in common the exceptional behaviour of /r/. The Mester & Ito account uses a theoretical argument that attempts to account for this behaviour in broad, typological terms. The Zoll account specifically challenges the Mester & Ito account of the feature geometric treatment of /r/ and their proposed contrastive underspecification, and instead proposes to place the palatalized /r/ typologically with the non-coronal palatalized segments. She does this because of its cross-linguistic rarity. Other than the scarcity of examples, Zoll does not explain the phonological basis for the structure of /r/ that she proposes, nor does she give a phonetic argument to bolster her typological claims. Complex segments being defined as segments with more than one primary articulation, Zoll does not explain how the palatalized coronal /r/ would have such a double articulation and why it would be different than other coronal consonants. This seems to me to be the major flaw in the Zoll analysis.

Both accounts also begin from the description that Hamano provided in her account of the sound-symbolism of Japanese and attempted to create a framework in which her description of the phenomenon could be duplicated. Neither account attempts to look beyond the bare description of the lexical facts to provide a more general account of the results. Nor does either account attempt to explain the shape of the mimetic forms in the lexicon. A question that needs to be asked is: are the lexical forms given in Hamano restricted to any particular shape because of the phonological properties of Japanese or mimetics? Does the shape of the lexical items provide gaps in our vision of a more general property to mimetic palatalization? It has been noted by Hamano that /r/ never occurs initially in the class of mimetic roots under consideration here, while /r/ occurs quite frequently as the second consonant. Is this accidental, or is this part of the question that can be addressed in accounting for the behaviour of palatalized mimetic adverbs? These are some of the questions that I hope to answer in my account.

3. Mathematical Analysis

In this section, rather than just looking at the problem of mimetic palatalization from a purely linguistic standpoint, the problem will be examined with the help of mathematical techniques. Some of the techniques that will be employed are decision theory and game theory, and classical statistical analyses. We will also examine the results of a study conducted to test some of the claims made in previous accounts of palatalization.

A place to begin with this analysis is to look again at one of the questions posed in the previous section. One of the first questions about the previous analyses that I posed was whether the analysis of the [-ant] feature of the palatalization as a morpheme, and specifically a suffix, was really justified. There is a semantic argument against this that was presented above, but one can also make a statistical one. There are 87 CVCV mimetic adverbs in the lexicon that take palatalization that were listed in Hamano (1998). One of these, tokyo-tokyo, contains one coronal consonant and one non-coronal, but palatalizes the non-coronal consonant. In all other accounts of the palatalization process, this item is eliminated from consideration. It is possible that it is based on an analogy with the city Tokyo, but we will not consider it further. This leaves 86 lexical items to consider. Two of these have only one consonant, so we will ignore these. Of the remaining ones, 37 are palatalized on the first consonant, and 47 on the second. This gives us a ratio of 44% to 56%. We can generate a 95% confidence interval for secondary palatalization. The range for this interval to two significant digits would be: (45%, 67%). Since the confidence interval encompasses the 50% mark, the tendency toward secondary palatalization is not considered significant at the 5% level. In fact, these figures tell us that there is a chance that is slightly greater than one in four that the tendency toward palatalizing the second consonant in CVCV mimetic adverbs is an accidental phenomenon. Based on this, I would suggest then that it would be unwise to appeal to a rightward tendency in the palatalization unless there are no other possible alternatives. What then could be going on?

Coronals preferentially receive palatalization cross-linguistically, and this is well established, and for the data provided in Hamano for Japanese mimetics. We can perform a similar significance test for the coronal preference pattern. Of the 84 lexical items given in Hamano (1998), there are 67 that consist of a coronal and a non-coronal consonant in some order within the root. Of these, only the four that contain a non-coronal consonant and an /r/ do not palatalize the coronal consonant. This is a 94% coronal palatalization rate. This does not explain what is going on with /r/, but it does justify including a preference for palatalizing coronals in our analysis. This makes sense linguistically, because palatalized coronals are singly articulated and palatalized non-coronals would be doubly articulated.

There are seventeen lexical roots that consist of the same place of articulation: either two coronal consonants or two non-coronal consonants. I introduce these numbers in this section because I will be referring to them as I continue the discussion below.

3.1 Decision Theory

Decision Theory is a mathematical account of how thinkers—humans, machines, etc.—make decisions. Decision Theory is a blanket term that covers a large number of approaches to decision making. It can encompass game theory, decision trees, statistical reasoning, and so forth. In this paper we are going to consider decision trees and game theory first, and then we will consider a statistical analysis of the problem in §3.2.

3.1.1 The Decision Tree

Let's consider again the rules presented by Hamano (1998) for mimetic palatalization in reduplicative CVCV roots. We can try to analyze our data by creating a decision tree. Not all problems can be analyzed strictly in terms of a decision tree. When problems fail to yield to such a tree, and produce less than optimal results, this is called sequential incoherence.

Our rules tell us that coronals are generally preferred to non-coronals. They also tell us that a pair of non-coronals defaults to the left side, and a pair of coronals defaults to the right. However, /r/ must never receive palatalization. Let us consider a hypothetical mimetic root of the form xVyV, where x and y are the two consonants in our hypothetical root. We can form a decision tree that represents the rule-based account of Mester & Ito described above. This is given in Figure 2. The tree begins by asking if the second consonant in the root is a coronal. This agrees with the observation in Hamano that coronals are preferred to non-coronal consonants. If the answer to this question is "no", then the rule states we should palatalize the first consonant by default. If the answer to the first question is "yes", then the answered to another question is needed: Is the second consonant an /r/? If the answer is "yes", then the first consonant gets the palatalization. If the answer is "no", then the second consonant receives the palatalization.

The tree in Figure 2. is short, but it considers only the second consonant y in both questions. This is made possible by the shape of the lexical items: there are no /r/s in initial position. If this were not the case, the decision trees described here would be more complicated.

Figure 2. Mester & Ito Decision Tree

```
                xVyV
             Is y coronal?
               /   \
              /     \
             no     yes
            /         \
           /           \
Palatalize x.        Is y an /r/?
                        /     \
                       /       \
                      no       yes
                     /           \
          Palatalize y.       Palatalize x.
```

One can also propose alternative trees that do the same thing—correctly predict the output of the lexicon—but also remain more general. Let us also consider a decision tree in which our first decision about where to palatalize this root is based on the answer to the question: is there a difference in place of articulation between the two consonants? We ask this with the understanding that we have a two-way distinction, coronal and non-coronal—the distinction between labial consonants and velar consonants does not come into play. If we answer “no”, then we have two possible situations: either we have two non-coronal or two coronal consonants in the root, and another question will need to be asked.

The next question one would ask in this tree is: is the place of articulation of the two consonants coronal? This question also has a yes or no answer. If the answer is again "no", then the decision made would be to palatalize the leftmost or initial consonant in the root. If the answer is "yes", one must ask still another question: does the root have an /r/? Finally, now one can make decisions on both branches of the tree. If the answer is "no", palatalize the second consonant. If the answer is "yes", palatalize the non-/r/ consonant.

If we answer "yes", however, to the first question posed, then we have a root in which we have a coronal consonant and a non-coronal consonant, but what should our decision be here? One would like to say, based on the strong tendency toward palatalizing coronals over everything else mentioned above, palatalize the coronal. One cannot do that. Instead, another question has to be asked: is one of the consonants the coronal consonant /r/? If the answer is "yes", palatalize the other one. If the answer is "no", then palatalize the coronal. The full decision tree for this example is given below in Figure 3. While this tree does reproduce the lexical results, the tree asks the same question in two different places: is one of the consonants an /r/? Ideally, we would like to avoid this repetition.

Figure 3. Alternate Decision Tree #1

```
                                   xVyV

     Are x and y of different places of articulation (coronal or non-coronal)?

                                    /\

                                   /  \

                                  no   yes

                                 /        \

                Do we have two coronal       Is y an /r/?

                 consonants?                    /    \

                       /   \                  no      yes

                      /     \                /        \

                    no     yes        Palatalize     Palatalize

                   /          \        the coronal.       x.

           Palatalize x.        Is y an /r/?

                                  /  \
```

/ \

no yes

/ \

Palatalize y. Palatalize x.

Alternatively, let us consider a decision tree in which one makes the first decision based on the answer to the question: are there any coronal consonants in the root? If the answer to this question is "no", then the root contains only non-coronal consonants. The decision then is to place the palatalization marker on the leftmost consonant in the root.

If the answer to our initial question, however, is "yes", then we have to ask another question. We would like to make our decisions as general as possible in the beginning, and as we move down the tree, consider more specific questions. However, in order to avoid sequential incoherence, given the description in Hamano (1998), the next question in the tree has to be: is the second consonant in the root an /r/? If the answer is "yes", the decision is to palatalize the leftmost consonant.

If the answer is "no", then one has to ask another question. This time the question might be: are both consonants coronal? If the answer to this question is "yes", then the decision is to palatalize the second consonant. If the answer is "no", then the decision is to palatalize the coronal. This decision tree is given in Figure 4. This decision tree is shorter, but it asks a very specific question high up in the tree. We would like to be able to construct a tree that is both short, and avoids these questions about exceptionality.

Figure 4. Alternate Decision Tree #2.

xVyV

Is either x or y coronal?

/ \

/ \

no yes

/ \

/ \

Palatalize x. Is y an /r/?

/ \

no yes

/ \

/ \

Palatalize Palatalize x.

the coronal.

Let us try constructing a decision tree that is based on statistical probabilities. For example, if we return to Table 2., a simple decision tree can account for all of these items. This tree is given in Figure 5. We choose as our first decision to palatalize the coronal in the word. In half the examples in Table 2., we are finished making decisions. Otherwise, palatalize x. This is a short and general decision tree that avoids overly specific questions, about /r/ for instance, and doesn't refer to specific x or y segments frequently as the decision tree in Figure 3. does. This tree does, however, generate inconsistent predictions in seven lexical items, where non-coronals are paired with the coronal /r/. We can improve this slightly by appealing to salience (to be discussed later) instead of "palatalize x", but there are still four lexical items unaccounted for. An example of these exceptional roots is /koro/ which palatalizes to [k^joro], choosing the non-coronal over the coronal /r/ and generating an instance of sequential incoherence.

Figure 5. Decision Tree Based on Probabilities.

xVyV

Is only one of x or y a coronal?

/\

/ \

no yes

/ \

/ \

Palatalize x. Palatalize

the coronal.

All these trees represent rule-based accounts of the mimetic palatalization pattern. If we are to consider a rule-based account to be superior to some other account, such as the constraint-based account of Zoll, we need decision trees that are simpler still and more general. In order to do that, we need a utility function that takes into account all of the various factors involved. I have left this discussion for §3.1.3. A utility function based on probabilities alone generates a tree with sequential incoherence: it predicts at least seven lexical items incorrectly.

3.1.2 Game Theory

Game theory is a special case of decision theory involving two interacting players. The "game" under consideration here is the communication game. We can think of this game as involving one passive player and one active player. The active player is the speaker; the passive player is the listener. The game is won when the speaker and the listener successfully communicate. Boersma (1999) describes the elements of the grammar and the processing systems that are needed at the phonetics-phonology interface. This paper will only consider the simplified model represented by the speaker-listener interaction. We will first consider two possible strategies to make our game simpler and within the limited linguistic context of CVCV mimetic adverbs. The speaker can choose palatalization to convey the semantic content of childishness or excessiveness that the [-ant] element represents, or he can choose to allow only the context of the discussion imply the meaning. The listener is passive and cannot *choose* to hear or not to hear, but rather her "decisions" are based on environmental and contextual factors, such as noise level. We can see in the matrix representation of this game that either strategy may work, and that this is not a zero-sum game. It is the likelihood of success that differs between the two strategies. The expected probability of the contextual approach would be affected by how close the semantic content matches the general pattern.

Figure 6. Game Matrix for Mimetic Adverbs

Speaker

[-ant] context

Listener does/doesn't understand

The situation between a speaker and a listener is somewhat more complicated than the matrix given in Figure 6. The speaker may rely on a combination of strategies for success. This opens up the possibility of a random ideal strategy under uncertainty; though in a situation of perfect knowledge, such as within the lexicon, a simple non-random strategy would be expected. The speaker also has to choose which segment to palatalize. To represent this decision we can create a decision tree using matrices to represent the decisions at each step, as shown in Figure 7., or we may expand the initial game matrix to represent each decision in the tree in a single matrix as shown in Figure 8. We can expand this matrix further to represent our decisions for each possible situation, regardless of the method we employ in determining the likelihoods for each decision represented in the matrix, including Optimality Theory.

Figure 7. Decision Tree with Game Matrices

xVyV

[-ant]   context

/   \

/   \

[-ant]   context

x   y   \

…

Figure 8. Expanded Game Matrix

x[-ant] y[-ant] context

At issue in deciding which consonant is optimal for palatalization is affected not only by the ability of the listener to hear the palatalization, but the speaker also takes into consideration the articulatory effect involved in making the change. The speaker has to decide if the improvement in the odds of perception is worth the extra effort involved. The increased likelihood of communicating may be negligible. The speaker may have some knowledge of what the listener is expecting to hear, and this may also play a roll in his decision. Certainly this matters in cases of perfect knowledge, such as with lexical items. The listener, on the other hand, is affected not only by his own listening ability and expectations, but also by the environment around him. If the game is being played in an especially noisy environment, the chance of successfully communicating may be quite low but relying on perceptual salience will provide a better chance of success.

In §2.3 I described the range of semantic content that the [-ant] segment has in mimetic palatalization. I would argue that this makes it extremely difficult for a speaker to rely on context alone to convey the meaning behind the palatalization marker without actually employing it. The semantic content is difficult to recover without perceiving the [-ant] feature for two reasons. Not only because of the range of the semantic implied, but also because of the effect the palatalization marker has on the syntax of the word: none. Unlike many morphemes, the palatalization marker here does not induce any syntactic changes on the mimetic adverb. It does not become a new part of speech. It does not induce additional syntactic changes elsewhere in a sentence, nor does it necessarily move for emphasis. There is no additional method of recovering the semantic content if the marker is missed. As the [-ant] feature is the only feature of the marked vs. unmarked adverb that differs, the perceptual salience of the feature becomes an important consideration.

In order to employ the idea of perceptual salience in the game theoretic analysis here, one needs to have a way of measuring salience. Kawasaki (1982) refers to a principle that she calls *maximization of dissimilarity*. She employs this notion to describe natural acoustic limitations of sequences of sounds, but we can consider it here where salience, the difference between two utterances of different meaning, is also a consideration. She defines perceptual salience with the following formula:

Equation 1. Perceptual Salience

Alternatively, we can employ the analogous function:

Equation 2. Perceptual Salience (2)

The P_i function, one i for each segment being compared, represents the perceptual features in a phonological perspective or formants in acoustic waves in a phonetic perspective. Phonologically the segment is represented as a list of features that can be modeled with indicator variables, while phonetically there are acoustic properties and articulatory properties that can be represented as continuous functions. In Equation 2., the $P_{a,i}$ and $P_{b,i}$ represent the two segments being compared. In our case, it would be the palatal segment vs. the plain segment. Since the decision the speaker is trying to make involves choosing between two possible palatalization patterns, we would compare two different results of the equation. In an accurate grammar, the two approaches would produce the same judgments of salience.

In creating judgments of salience to be used in our game theoretic analysis, one must be constrained by the similarity of the plain and palatalized roots. The speaker does not want to insert any random feature, but rather he constrains his strategy by a measure of similarity that maintains the derivational features of palatalization. We can make use of the feature-contrast model introduced by Tversky (1977, & Gati 1978) that is given below in Equation 3.

Equation 3. Feature-Contrast Model.

$$\text{Similarity}(X,Y) = F[\theta\, f(X \cap Y) - \alpha\, f(X - Y) - \beta\, f(Y - X)]$$

Where F is an increasing function, θ, α, β are positive constants, f is a measure function of the features, X ∩ Y denotes the features shared by X and Y, X – Y denotes the features in X but not in Y, and Y-X denotes the features in Y but not in X.

Pierrehumbert (1993) introduced a similar model of similarity. She employed a normalizing factor, set F(x) = x, and set the weights of the constants, with θ =1, and α = β = 0. Her similarity model is given as Equation 4.

Equation 4. Normalized Similarity Model

$$\text{Similarity}(X,Y) = \frac{f(X \cap Y)}{f(X \cap Y) + f(X - Y) + f(Y - X)}$$

With this in mind, we have addressed the strong tendency in Japanese mimetics for palatalizing coronal consonants over non-coronal ones. This matches cross-linguistic data. We have seen that in a feature-geometry model this difference is represented by a change in the level of the complexity of the feature geometry. This model represents an abstraction of the articulatory reality of palatalized segments. Coronal consonants undergo a change in place of articulation and remain singly articulated, while non-coronal consonants become complex and doubly articulated. A difference in the behaviour of the two consonants can also be seen at the acoustic level, however. The change in place of articulation is much more perceptible than is the addition of a secondary articulation to the non-coronals. Perceptual prominence has been associated with the leftward tendency of complex segments in Steriade (1995). This strongly indicates that perceptual features are being taken into account in the placement of the palatalization feature. This raises questions again about the justification of the rightward tendency proposed for coronal-coronal consonant pairs. Perhaps, other perceptual features are also being considered in addition to place of articulation.

Let us now consider the seventeen remaining items that are coronal-coronal pairs or are non-coronal-non-coronal pairs. A listing of these items is given below in Table 5. One can see that there are a large number of coronal plus /r/ consonant pairs in these examples, but there are three that do not include /r/.

Table 5. CVCV Mimetic Roots with only Manner of Articulation Contrasts

Coronal-Coronal	Non-Coronal-Non-Coronal
a. tyari, tyara, tyoro, tyuru	g. pyoko

b. syari, syara, syuru

c. zyari, zyori, zyara

d. nyoro, nyuru

e. nitya, netya

f. dosya

h. hyoko

Let's consider the coronal-coronal pairs first. If we go back to the accounts of mimetic palatalization presented previously, the claim that there is an underlying rightward tendency rests solely on the last example in Table 5. If one adopts a stance that says /r/ is exceptional, twelve of the items in Table 5. above are not available for consideration because they tell us nothing. The phonological data in Japanese tells us that palatalization cannot occur with an /e/ vowel, and is non-contrastive in front of the /i/ vowel; therefore, if one excludes the items in Table 5. using the /e/ and /i/ vowels, the lexical items we can consider are reduced to just the one in Table 5.f. Basing an account of mimetic palatalization on a single lexical item seems to me to be seriously flawed, but let us begin in the same place as everyone else.

The root /dosa/ undergoes palatalization on the second consonant. Other than the difference in its position in the word, how else does the second consonant differ from the first? Perhaps one of these features can be generalized to the other examples. The /s/ differs from /d/ in that /s/ is a fricative, and it is voiceless; while /d/ is an oral stop, and it is voiced. Might either of these features tend to attract palatalization? While voicing is a feature of salience, the voicing between the plain and palatalized roots is not changing, so I will set aside the voicing issue in favour of the manner of articulation difference, because I find there is a clear distinction between the two that might be actively involved in attracting palatalization. Fricatives have a great deal more high frequency noise than do their oral stop counterparts. If one considers the effect that palatalization has on the acoustics of a consonant, modulation of high frequency noise is one of the principle means by which one distinguishes a non-palatal segment from a palatal one.

If we consider the other examples of coronal-coronal pairs we can make similar arguments. Suppose we neglect for a moment the vowels our lexical items in Table 5.e., oral stops when they become palatalized become affricated. Affricates, like fricatives, generate a large amount of high frequency noise. They are more acoustically distinct from their non-palatalized counterparts than are palatalized nasal stops. One can see from previous analyses of the data that in these examples, the vowels end up not playing a role in the placement of palatalization, and acoustic considerations can also produce similar results.

If we consider the case of coronals plus /r/, we can make a similar claim. All of these consonants should be singly articulated palatals. The fricatives and stops are clearly more acoustically distinct than are the /r/s. Nasal stops, because they are stops, would have a small release burst portion that can help carry some additional acoustic information about palatalization that the approximant would not be able to rely on. This suggests that there is a perceptual reason for palatalized /r/ being cross-linguistically rare. Rather than the segment being doubly articulated and complex, as Zoll (1997) has claimed, an alternative explanation is that the segment is rare because of its acoustic properties.

The coronal-coronal data given in Table 5. can allow us to construct a hierarchy of consonants in the order in which they prefer to be palatalized. The root /dosa/ shows us that /s/>>/d/. The roots /neta/ and /nita/ show that /t/>>/n/; the /noro/ and /nuru/ roots show us that /n/>>/r/. This preference hierarchy mimics the perceptual hierarchy associated with maximization of salience.

Suppose one looks then at the non-coronal-non-coronal roots. There are only two roots that contain two non-coronal consonants. One of them, /poko/ has no distinction in manner of articulation, so it seems clear that as Zoll suggested, there is a rightward tendency at work here. The second root, /hoko/, does not contradict the claim that manner of articulation and the associated acoustic properties might be at work here.

To explain the data entirely, we would also have to explain the twelve remaining roots that contain non-coronals and /r/. It has been observed, beginning with Hamano (1998) that /r/ behaves like a non-coronal for the purposes of mimetic palatalization. For the analysis to work, acoustic properties would also need to be able to explain the behaviour of /r/ here as well. Complex consonants are less acoustically distinct from plain consonants than are palatal consonants versus palatal ones. Based on the behaviour of /r/ in Japanese mimetic palatalization, one can surmise that /r/ is not more acoustically distinct than complex consonants. More to the point, palatal /r/ is *less* acoustically distinct than complex consonants. This can be tested phonetically. It is also supported cross-linguistically because palatalized /r/ is rarer than palatalized non-coronals. This lack of perceptual salience could explain why /r/ never occurs in initial position in CVCV mimetics, but this is a question I will return to in §4.

By employing perceptual salience as a strategy for the placement of the palatal feature in our hypothetical game, it seems possible that this single factor can be employed to predict the lexical results. In this section, the concept of salience has been discussed, and functions have been presented to represent the relationship between the two utterances (the plain and palatalized roots) under consideration. However, these functions only represent the differences or similarities between the two forms, and do not serve to quantify their usefulness. In order to put salience considerations into the game strategies, salience must be paired with its usefulness in leading toward successful communication with a given listener, or within a given language. In short, a speaker needs to construct a utility function for this game.

3.1.3 Utility Functions

One of the principle features of decision theory is the concept of the utility function. Utility functions are a way of representing judgments. One of the most common ways of thinking of utility functions is in the form of probabilities or bets. Using a person's judgment of the value of a certain event, and the probability of that event occurring, the expected value can be computed based on the size of the bet the person would be willing to take on the situation. Utility functions are not necessarily linear, and they are not comparable from one person to another; however, by comparing a utility function to a bet, one can have some way of making practical comparisons between functions. In this paper we will not be constructing money-based utilities, instead, we will judge utility functions based only on whether or not it accurately predicts the lexicon, and facilitates communication between the speaker and the listener.

One can take the factors that we discussed in the previous section and construct a utility function based on them. We discussed in §3.1.1 the idea of using a utility function based on the probabilities of various factors that are palatalized in Japanese mimetics. I mentioned that there would be a problem with sequential incoherence in a direct probabilistic approach shown in Figure 5. I will consider that in detail here for a moment. It was noted above that in roots with one coronal consonant, 94% of the time the coronal consonant was the one that was palatalized. We have seen however, that roots with /r/ are the exceptions to this rule, and the principle source of sequential incoherence with this method. A speaker than employed this probability directly as part of their utility function would need to make note of four lexical exceptions to this rule, and while lexical exceptions are not always unavoidable, one would like to make our utility functions as generalizable as possible.

Three features that have been discussed that contribute to the salience of plain versus palatalized segments are place of articulation, position in the word, and manner of articulation. Some combination of these features can be used to construct a utility function that can be tested against the lexicon for its usefulness in predicting the palatalization pattern. We can employ the salience and similarity equations given above to create one, but let us for the moment consider the nonlinear utility function presented below as Equation 5. In this utility function, f(x) represents the place of articulation of the consonant, g(x) the position within the root, and h(x) the manner of articulation. The function g(x) serves the same purpose as the ALIGN L constraint in the Optimality Theoretic analysis of Zoll. We can make the function even more general by taking into account factors like adjacency to a front vowel, but I will ignore that level of detail here.

Equation 5. Sample Utility Function

$$U(x) = (f(x) + g(x))\, h(x)$$

Where $f(x) = \{5, \text{ if } x \text{ is coronal}; 1, \text{ if } x \text{ is non-coronal}; 0, \text{ otherwise}\}$

$g(x) = \{0, \text{ if } x \text{ is in initial position}; -1, \text{ otherwise}\}$

$h(x) = \{3, \text{ if } x \text{ is a fricative}; 2, \text{ if } x \text{ is a stop}; 1, \text{ if } x \text{ is a nasal}; 0, \text{ otherwise}\}$

The utility function described above can be normalized by dividing by the maximum possible value of the function, which would be for an /s/ in initial position. The maximum for the unnormalized function is 15. Normalizing will facilitate the comparison of the function to a probabilistic analysis. The utility function described here is not particularly useful analyzed over a single consonant in the mimetic root. Rather, the speaker would compare U(x) and U(y) in the hypothetical root xVyV, and palatalize either x or y that generated the higher value for the utility function.

Let us test the usefulness of the utility function in Equation 5 by testing it on the lexical items. Table 6 gives the results from a sample of lexical tests using the normalized function. One can see that the utility function successfully predicts the palatalization pattern of the lexical items. The utility function succeeds in giving weight to all three factors that have been identified as influencing palatalization. It succeeds in attracting palatalization to all coronals except /r/. It likewise eliminates the palatalization of non-coronals in second position, and despite the advantages of initial position, manner of articulation successfully attracts palatalization into second position. By using U(x) described here, we can reduce our decision tree to one decision: Palatalize the consonant with the higher utility.

Table 6. Lexical Predictions Using U(x).

Lexical Root	U(x)	U(y)	Norm (x,y)	Predicted	Lexical Outcome
a. /kata/	(1+0)2=2	(5-1)2=8	(0.13, *0.53*)	[katʃa]	[katʃakatʃa]
b. /zabu/	(5+0)3=15	(1-1)2=0	(*1.0*, 0)	[dʒabu]	[dʒabudʒabu]
c. /koro/	(1+0)2=2	(5-1)0=0	(*0.13*, 0)	[kʲoɾo]	[kʲoɾokʲoɾo]
d. /huna/	(1+0)3=3	(5-1)1=4	(0.2, *0.27*)	[huɲa]	[huɲahuɲa]
e. /poko/	(1+0)2=2	(1-1)2=0	(*0.13*, 0)	[pʲoko]	[pʲokopʲoko]
f. /hoko/	(1+0)3=3	(1-1)2=0	(*0.2*, 0)	[hʲoko]	[hʲokohʲoko]
g. /noro/	(5+0)1=5	(5-1)0=0	(*0.33*, 0)	[ɲoɾo]	[ɲoɾoɲoɾo]
h. /neta/	(5+0)1=5	(5-1)2=8	(0.33, *0.53*)	[netʃa]	[netʃanetʃa]
i. /dosa/	(5+0)2=10	(5-1)3=12	(0.67, *0.8*)	[doʃa]	[doʃadoʃa]

For the lexical items given in Hamano (1998), the exact values of the functions f(x), g(x) and h(x) that make up the utility function U(x) can be modified slightly and still be capable of predicting the lexical outcomes. Variation is possible because not all of the forms that test the limits of the function exist in the lexicon. Forms where the initial consonant is an /r/ is an example; however, there are others. The non-coronal fricative never occurs in second position. This may be due to historical reasons, but because mimetics are sound-symbolic one may not wish to consider this as the only reason for this gap in the lexicon. There are also no pairings of the non-coronal nasal /m/ with another non-coronal, so that, like /r/, /m/ never attracts palatalization. In fact, the utility function given in Equation 5. imperfectly mimics the frequency of palatalization patterns in the lexicon. Coronals are palatalized most frequently, with fricatives palatalized slightly more frequently than stops. The consonants with the lowest utilities in U(x) are /m/ and /r/, and these never receive palatalization. Most of the lexical items tend to maximize the differences in the utility functions of the two consonants in the root.

More sophisticated versions of utility functions are already being used in linguistics. Optimality Theoretic constraint rankings essentially represent a utility function. A model of a strict interpretation of Optimality Theory is given below in Equation 6. Each constraint is modeled as an indicator variable I_{0j}. The ranking is represented as some permutation of the I_{0j}s and their constraint coefficients a_i. The ranking is achieved by the constraint given in Equation 7. Some alternatives to this model exist that allow constraints to be ordinal, or other minor modifications. Note that Optimality Theory is a linear utility function. When a constraint interaction is permitted (both constraints satisfied =1, otherwise 0), OT only allows for the interaction to be strictly higher in the ranking than either of the separate constraints. In this way, OT attempts to limit the power of interactions of this type. Whatever the form of the utility function, the real power of this approach is in predicting behaviour. This is something that I will look at more closely in §3.2.2.

Equation 6. Optimality Theoretic Utility Function

Equation 7. Ceofficients Constraint

3.2 Statistical Analysis

I have discussed briefly in previous sections some general statistics about Japanese mimetic palatalization. In §4.1 I will discuss a statistical analysis of the lexicon in more detail. I will also be presenting in §4.2 the results of a study conducted in 1998 that was first presented in McCall & Nagao (1999) that attempted to measure the predictive power of various accounts of mimetic palatalization.

3.2.1 The Lexicon

A summary of the statistics mentioned in previous sections is given in Table 7. below. Recall that we determined in §3 that the rightward tendency in mimetic palatalization that was relied upon in previous analyses was not statistically significant. We also saw that in Table 5., if we ignore the supposed exceptionality of /r/, the rightward tendency was not even common among roots with two coronal consonants. We did see that salience considerations, as determined by position in the word, place of articulation and manner of articulation were successful in predicting, through a utility function, the lexical outcomes. Let us consider these factors in our statistical analyses.

Table 7. Summary of Lexical Statistics

Total number of lexical items: 86

Number of items with one coronal consonant and one non-coronal consonant: 67

Percentage of coronal-non-coronal roots that palatalize the coronal: 94

Number of roots that palatalize the initial consonant: 37; the second consonant: 47

Percentage of roots that palatalize the initial consonant: 44; the second consonant: 56

Number of roots that have two non-coronal consonants: 2

Number of roots with two coronal consonants: 15

Number of roots with two non-/r/ coronal consonants: 3

Number of lexical roots with initial /r/: 0

There are a number of ways to represent the notions of place of articulation, manner of articulation and position in a lexical root. One can consider the place of articulation of the entire root (coronal first, coronal second or same place of articulation on whole word), likewise for manner of articulation. Or, one can consider variables for initial consonant place, initial consonant manner, and so forth. Other configurations are possible. Not all of these representations lead to useful results. In particular, while we would like to consider the whole word at once, features necessarily belong to particular consonants. Because of this, the analysis of the lexical items will include the following variables: x_1 = initial consonant place of articulation, x_2 = initial consonant manner of articulation (with four categories), x_3 = second consonant place of articulation, x_4 = second consonant manner of articulation (with four categories), and Y = position of palatalization (1 = initial position, 0 = second position). In §3.1.2, we discussed the justification for manner of articulation playing a role in the salience of the palatalization. In that discussion, it was noted that the salience of palatalization with various manners of articulation formed a scale of fricative, stop, nasal, approximant. I used this scale in my discussion of the utility function U(x) in §3.1.3. While a scale of salience does exist, the numerical values of it are not precisely known; therefore, this scale represents a categorical variable rather than a direct measure of salience. As a categorical variable, the variables x_2 and x_4 should be converted to three indicator variables apiece, z_1-z_3 (fricative, stop, nasal) and z_4-z_6 (fricative, stop, nasal). However, because the variables do represent a legitimate scale, I will consider the analysis using both formulations.

Equation 8. Initial Indicator Variable Model of the Lexicon

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 z_1 + \beta_3 z_2 + \beta_4 z_3 + \beta_5 x_3 + \beta_6 z_4 + \beta_7 z_5 + \beta_8 z_6 + E$$

If one runs the analysis of the lexical data using standard regression techniques, something interesting happens. The R^2 value of the resulting model is 1.0000. The coefficient β_0 is two. The coefficients β_1, β_2, β_3, and β_4 are equal to zero. The final model is given as Equation 7.

Equation 9. Final Indicator Variable Model of the Lexicon

$$Y = 2 - x_3 - z_4 - z_5 - z_6$$

Initially this model seems a bit unusual. The intercept is 2, but the outcome variable we desire is either 0 or 1, making this initially difficult to interpret. There is also perfect correlation. Something must be wrong. However, when we look carefully at the lexicon, these unusual features can be explained. The intercept is neutralized by the fact that at least one of the variables in the model is equal to one, and no more than two of the indicator variables can be equal to one. So, this model does give a result of either 1 or 0. Y is equal to one (palatalize the initial consonant) if the second consonant is the coronal /r/ (not a fricative, stop or nasal), or if it is a non-coronal consonant of any kind. Y is equal to zero (palatalize the second consonant) if the second consonant is both coronal and something other than /r/. This model mimics the analysis of Mester & Ito in their rule-based account. This model is possible because of the shape of the mimetic lexicon. The use of indicator variables here makes this model well-suited to an articulatory account of features because articulations either are fricative or not; they either are stops or they are not. This is a common feature of modern phonological approaches. Perceptual features are not necessarily categorical so perhaps this is not the best account of a perception-based theory. The true test of the power of this model is how well it predicts behaviour when the result is not known from experience, such as in the nonsense words discussed in §3.2.2 below.

Let us consider the categorical/ordinal variable for manner of articulation. The ordinal values are being considered here because they preserve the rank ordering for the perceptibility of the segments, though the actual values would have to be determined experimentally. The model using these variables is given in Equation 10. below. The R^2 for this model using multiple regression is 0.8803. The p-values for each of the coefficient estimates is significant at the $\alpha = 0.5$ level except for the coefficient for x_2 (the first consonant's manner of articulation) which is 0.951. I have left it in the model because I expect the initial consonant to attract palatalization (be positive) rather than repelling it (placing the palatalization on the second consonant). This justifies halving this p-value and admitting it into the model. Because our Y variable is an indicator variable itself, this model is a little more difficult to interpret using standard regression techniques. Unlike our previous model, this model does not generate 0-1 results, although despite the intercept, it does usually generate results between zero and one. Some sample results are given in Table 8. below. The model does produce results that give Y greater than 0.5 for initial palatalization, and a Y less than 0.5 for secondary palatalization. I include this model here because a similar model with come up again when we describe the results of the study described in McCall & Nagao (1999). The results of the model itself are at best suspicious.

Equation 10. Scale Variable Model for the Lexicon

$$Y = 1.15 + .24\ x_1 + .05\ x_2 - 0.62\ x_3 - 0.25\ x_4$$

Table 8. Sample Results of the Statistical Model in Equation 10.

a. katya-katya 1.15 + 0 + 0.1 – 0.62 – 0.5 = 0.13

b. kyoro-kyoro 1.15 + 0 + 0.1 – 0.62 – 0 = 0.63

c. pyoko-pyoko 1.15 + 0 + 0.1 – 0 – 0.5 = 0.75

d. nyoro-nyoro 1.15 + 0.24 + 0.05 – 0.62 – 0 = 0.82

e. zyabu-zyabu 1.15 + 0.24 + 0.15 – 0 – 0.5 = 1.04

f. dosya-dosya 1.15 + 0.24 + 0.1 – 0.62 - .75 = 0.12

I will not consider a logistic analysis of the lexical data at this point. I will return to a discussion of it when I consider the logistic analysis of the study data in §3.2.2. I will only say here that the logistic analysis of the lexicon gave poor results. Using indicator variables as in the model given in Equation 8., the algorithm converged to a similar model (with only the second consonant's variables having non-zero coefficients), but the p-values are extremely high, possibly for lack of frequency data. The algorithm would not generate a maximum likelihood model using the variables described for the model in Equation 10. I suggest that this is a result of the shape of the lexical items. We will return to this question in §4.1.

3.2.2 McCall & Nagao Nonsense Word Study

This study was conducted in order to test the predictive power of the existing analyses of Japanese mimetic palatalization. An accurate grammar will not only predict lexical results, but it will be generalizable to situations where the outcome is not strictly known. Speakers of English, for example, understand that complex patterns can be difficult to discern, and speakers do not always get them "right" on the first try. For example, there are various nominalizing suffixes in English and it is not always clear which one should be used when, but there in a general consensus about what most speakers of English think should be used in a particular case, and interactions with other speakers will weed out poor guesses. We expected similar results: the results would not be perfect, and there would be some noise in the data, but we did expect that most speakers would tend to follow a consensus of what was grammatically significant. The decision theoretic discussion before suggested that in such a case of decision-making under uncertainty we might see random strategies employed, not just across speakers, but even within a particular speaker's own results. We were not particularly interested in whether coronals were themselves significant compared to non-coronals. We were interested instead in testing the rightward tendency that has been claimed with roots containing pairs of coronals. Because of this, our data was heavily weighted toward creating nonsense words involving a pair of coronal consonants. Additional nonsense words were included to help mask this emphasis on coronality.

Our study was conducted using 80 native speakers of Japanese, almost entirely speakers of the Kansai dialect. Ten of our initial subjects did not provide complete information for the study, and their answers were eliminated from the final results, leaving 70 subjects. A total of 43 tokens were used, 30 of which were coronal-coronal consonant pairs. Subjects were asked to determine the optimal position for palatalization in the nonsense words provided. All the tokens were of the form CVCV. All the vowels were either /a/, /o/ or /u/ so that the effects of palatal vowels would not influence the outcomes. Our tokens included both combinations of consonants that appear in the lexicon and ten tokens that specifically do not occur. These ten tokens are of special interest. I have listed a sample of them in Table 9. below.

Table 9. Sample Nonsense Roots of Interest

Nonsense Root	% Receiving Initial Palatalization	95% Confidence Interval
a. sVdV (1)	72%	(60%, 83%)
b. sVnV (1)	81%	(72%, 91%)
c. zVtV	78%	(69%, 88%)
d. tVnV	68%	(56%, 79%)
e. rVsV	18%	(8%, 28%)

Table 9. shows that palatalization in coronal-coronal roots can be attracted to initial consonants. It also shows us that /r/ still resists palatalization in initial position. The feature that these nonsense roots have in common is that the consonant attracting palatalization more frequently is the consonant where the palatalization will be more salient. And we can see in the examples shown, that the confidence intervals do not contain the 50% mark, meaning that these tendencies are statistically significant and are unlikely to be random guesses. Not all of the coronal-coronal pairs had confidence intervals that did not contain the 50% mark at the 95% confidence level. Almost all of these cases, were cases where there was no place of articulation difference between the two consonants, such as nonsense roots of the form sVzV, or tVdV. Of the others, only two failed the significance tests at the 90% confidence level. Both of these were roots of the shape tVnV; however, a third root of this shape passed the significance test. If one considers the confidence interval taken over all three roots, the preference for the stop over the nasal is clear. In general, we see that as the two consonants are more similar the odds of initial palatalization approaches 1.

Additional support for the attraction of palatalization to more salient manners of articulation can be seen in the nonsense root from the study of the form pVhV. In the lexicon, roots with two non-coronal roots palatalize on the initial consonant only. However, 72% of our subjects preferred to palatalize this root on the second consonant. The confidence for this result is (61%, 84%). Clearly, manner of articulation can also attract palatalization away from initial position in non-coronal consonants.

Despite the presence of lexical items that clearly indicate the behaviour of /r/ in second position when paired with non-coronal consonants in initial position, speakers seemed especially unclear how to apply the lexical rules. This was also evident if /r/ was placed in the uncharacteristic initial position with a non-coronal consonant in second position. Other than one nonsense root of the form hVrV, the confidence intervals all contained the 50% mark. There are only four roots in the lexicon where the initial consonant is a non-coronal and the second is /r/. Two of these are of the form hVrV. Perceptual salience may be playing a role, but so also may be frequency in these examples.

In order to look at all of the data from the study at once, I considered a linear regression similar to that performed on the lexical data. It should be apparent that the regression that mimicked the Mester & Ito analysis fails to successfully predict the study results as reported here. However, the secondary analysis of the lexical data using the categorical variable proves to be more useful in predicting the study results. That model was given in Equation 10. Compare this to the model given in Equation 11. This model was obtained independently. While the coefficients are different—we are no longer trying to predict a 0-1 variable, but instead the probabilities of initial palatalization—the variables employed, and the signs of the coefficients and their relative strengths are similar. The coefficients are all significant at the $\alpha = 0.05$ level. The R^2 value of the model without the interaction term is 0.6251; with it, it is 0.6611. The coefficients are less similar without the interaction term, so I have not given that model here. The interesting feature of the salience model employing the scale variable is that it accurately predicts that there is likely to be more confusion when the difference in the salience characteristics of the two consonants is smaller.

Equation 11. Ordinal Regression Model for Study Results

$$Y = 1.09 + .20\, x_1 + .06\, x_2 - 0.67\, x_3 - 0.28\, x_4 + 0.16\, x_2\, x_4$$

We can consider a logistic analysis of the study data as well. We can create two logistic models, one with the indicator variables, and one with the ordinal variables. The indicator model differs from the indicator model generated for either of the two standard regression models. In the study, using the standard regression and just the indicator variables, we were unable to reproduce a model employing only the second consonant. We likewise have that result here. The model given in Equation 12. estimates the probability of palatalization of the initial consonant, and employs all the possible variables. The coefficients for the initial consonant are positive and naturally attract palatalization toward the initial consonant. The coefficients for the second consonant are negative, and attract palatalization away from the initial consonant. Fricatives (z_1 and z_4) attract more strongly than stops or nasals, forming the beginnings of a natural perceptual hierarchy. The difference in behaviour of stops and nasals between the first and second position may have to do with the tokens we employed in the study, or the distribution of frequencies in the lexicon. Further study would be needed to clarify this difference.

Equation 12. Indicator Logistic Model for the Study Results

The logistic model given in Equation 13. employs the ordinal version of the manner of articulation variables. This model maintains certain similar features to the model in Equation 12. These two models have similar deviance statistics suggesting that they model the data with similar accuracy. This provides more evidence that manner of articulation need not be treated as an indicator variable as the phonological approach would suggest, but that the ordinal approach, roughly matching a perceptual account, is as powerful and more general.

Equation 13. Ordinal Logistic Model for the Study Results

If we return to the utility function presented in Equation 5, §3.1.3, U(x) as described there successfully predicts the study results for more of the situations where there is a clear-cut difference in the salience of the consonants involved. The root of the form pVhV is not predicted to palatalize on the right side as the results of the study show. It may be possible to generate this behaviour in some version of this function, but I employed it for conceptual purposes only.

4. New Linguistic Results

In §3, I discussed the evidence supporting an alternative analysis for the behaviour of Japanese mimetic palatalization. We saw that rather than appealing to a tendency of the palatalization to be attracted to the rightmost consonant, we make a more general appeal to the manner of articulation—or more specifically, salience—to explain the lexical behaviour as well as the results of a study of nonsense words. In the §4.2 below, I will present an analysis employing these principles in Optimality Theoretic terms. First, I would like to discuss what other aspects of the mimetic lexicon can possibly be explained by appealing to the idea of salience, and make our game theoretic analysis easier to apply in practice.

4.1 Salience as a Restriction on the Lexicon

The exceptionality of /r/ in Japanese mimetic palatalization has been mentioned often in previous sections as a stumbling block in any analysis based solely on place of articulation. I have shown, however, that /r/ fits well into an analysis containing perceptual salience. The absence of /r/ in initial position in CVCV mimetic adverbs has also been noted. It can successfully be explained using an appeal to perceptual salience as a criterion. While Japanese does employ both [ɾ] and [ʎ] as contrastive features in the lexicon, subtle differences in meaning are not implied by the contrast anywhere else. Differences in meaning are readily apparent from the context and the [-ant] can be recovered. The semantic difference is more subtle and less readily retrievable from context. If salience is influenced by the three factors we've considered in this paper—place of articulation, manner of articulation, and position in the word—then potential areas where confusion might more likely would be avoided. Initial /r/ would be one of those potential areas, where the coronality and initiality would attract palatalization, but acoustic salience would disprefer palatalization. The small number of lexical items with non-coronals and /r/, and the pairing of the non-coronal fricative with /r/ in half of these cases, would also tend to maximize salience as much as possible in the overall lexical inventory.

The pairing generally of coronals with non-coronals maximizes both salience and ease of articulation as a general principle in the construction of mimetic adverbs. The absence of large number of roots with pairs of coronals has been noted elsewhere, as is the absence of large numbers of roots with pairs of non-coronals. Cases that we tested in our study, with both consonants of the same place of articulation and same manner of articulation number only one in the lexicon. Again, the lexicon appears to be minimizing areas where confusion may result.

Consider another specific segmental case, the case of /m/ in CVCV mimetic roots. Labial nasals acoustically are less salient compared to plain labial nasals and palatal nasals than are other manners of articulation with the same level of complexity. In Japanese mimetics, the labial nasal is always paired with coronals that receive the palatalization and it never receives palatalization itself. Perhaps, like /r/ it is insufficiently salient to carry the semantic content associated with the palatal element, and this is the method employed by the speakers to make the lexicon more transparent to palatal placement. Because the consonants /m/ is paired with are all coronal, however, there has never been any need to appeal to the exceptional behaviour of /m/ as has been done with /r/.

We might expect that a sound-symbolic system would be more flexible than other parts of the lexicon. We can rely even less on historical patterns to explain the gaps that we find. If these lexical gaps remain even across dialect patterns we can strengthen our argument that the lexical gaps are not accidental.

4.2 Optimality Theoretic Analysis with Salience

Let us reconsider the palatalization of mimetics in Optimality Theoretic terms. For the leftward tendency associated with initial prominence, we can maintain the ALIGN L constraint employed by Zoll (1997). Rather than appeal to the ALIGN R constraint, however, I wish to employ a constraint that appeals to perceptual salience. A constraint based on functional phonological principles seems to be in order. Alderete (2000) might suggest employing something like an anti-faithfulness constraint in order to force the palatalized root to be different than the plain root. The approach seems unnecessarily extreme to me. Faithfulness to the [-ant] segment may be sufficient. The maximization of perceptual salience is a long-standing principle in phonology and phonetics and we ought to be able to appeal to such a principle in Optimality Theory. Something simple like PERCEIVE is not specific enough, but does suggest a general direction. I propose instead a constraint maximizing the acoustic salience of the palatalization feature within a word. In Table 10. I give the constraints for the new analysis and their ranking, although its is inconsequential for this analysis.

Table 10. Optimality Theoretic Constraints

ALIGN LEFT (COMPLEX SEGMENT, PROSODIC WORD)

∀ Complex segments ∃ a prosodic word such that a complex segment coincides with the leftmost segment in the prosodic word

MAX SALIENCE ([-ANT])

Maximize the salience of the [-ant] feature while minimizing the articulatory effort

Ranking: ALIGN L (COMPLEX SEGMENT, PWD), MAX SALIENCE ([-ANT])

The MAX SALIENCE constraint proposed above is a non-linear constraint. As we can see from the results of Tableau 6., while we can separate out the ALIGN L constraint, which in itself is a perceptual prominence constraint as described in Steriade (1995) and Zoll (1997), we cannot separate the place of articulation and manner of articulation factors involved in salience. In other words, we cannot consider them in a linear ranking, nor does the conjoined constraint system help here.

Tableau 6.

{koro, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	MAX SALIENCE ([-ANT]) PLACE	MAX SALIENCE ([-ANT]) MANNER
a. ☹ koʎo			*
b. kʲoro		*!	

In Tableau 7., the ALIGN L constraint forces the palatalization onto the initial consonant. The MAX SALIENCE constraint is not violated because it is a constraint that compares the relative salience of the surviving candidates. The relative salience without regard to position in the word is the same for /p/ and /k/. In Tableau 8., the MAX SALIENCE constraint fatally violates the palatalization of the approximant. In Tableau 9., the constraint fatally violates the palatalization of a stop over a fricative. In Tableau 10., MAX SALIENCE is violated because of the place of articulation, but the ALIGN L constraint is also violated and either one would be fatal to the candidate in (a.).

Tableau 7.

{poko, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	MAX SALIENCE ([-ANT])
a. pokʲo	*!	
b. ☞ pʲoko		

Tableau 8.

{koro, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	MAX SALIENCE ([-ANT])
a. koʎo		*!
b. ☞ kʲoro		

Tableau 9.

{dosa, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	MAX SALIENCE ([-ANT])
a. dʒosa		*!
b. ☞ doʃa		

Tableau 10.

{toko, [-ant]}	ALIGN L (COMPLEX SEGMENT, PWD)	MAX SALIENCE ([-ANT])
a. tokʲo	*!	*
b. ☞ tʃoko		

These examples show that for the lexical items, the two constraints do not need to be ranked. However, if we consider the data from the study, a significant majority of the speakers appear to rank MAX SALIENCE over ALIGN L, as shown in Tableau 11. below. This ranking does not disturb any of the outcomes for the lexical items.

Tableau 11.

{pVhV, [-ant]}	MAX SALIENCE ([-ANT])	ALIGN L (COMPLEX SEGMENT, PWD)
a. pVhʲV		*
b. ☞ pʲVhV	*!	

5. Conclusions

The nonlinearity of the salience constraint is crucial to its understanding, and to its employment in solving mimetic palatalization. The interaction of articulation and perception is played out in this constraint through mimetic palatalization. This is a fundamental conflict in phonetics.

In this paper I showed that the previous analyses of Japanese mimetic palatalization cannot sufficiently explain either the behaviour of palatalization or the shape of the lexicon. I also showed that considerations of perceptual salience are important to understanding the behaviour not only in the lexicon, but in the behaviour of speakers applying grammatical rules to nonsense words. These factors were supported with a mathematical analysis employing aspects of Decision Theory and statistical analyses.

We asked questions in §2.3 that needed to be answered to resolve the problems in the previous linguistic analyses of Japanese mimetics. The question of the morphological character of the palatalization feature [-ant] is left unanswered. The question proves to be irrelevant to a final analysis. I was able to show that it is possible to incorporate the behaviour of /r/ into the lexical pattern without directly referring to it as an exceptional segment. Rather, it can be incorporated into a general account employing perceptibility. I was also able to avoid accounting for the pattern in Japanese mimetics without appealing to bare descriptive facts. I was able to employ previously accepted phonetic principles to explain not only the palatalization pattern, but also posit reasons for some of the gaps in the lexical data.

Further research can still be done to verify or clarify the results presented here. A follow-up study of the one described here may help to clarify some of the remaining gaps in the lexicon. It may be possible to construct a separate utility function for each speaker involved in such a study to measure the randomization pattern. We would expect that the accumulation of data would all converge to the same result that was seen here. This may reveal some of the flaws inherent in a study of nonsense words. A study of a different large dialect may also reveal whether the perceptual salience considerations are a dialectal feature or are present in any study of this data. Future research might also attempt to quantify the perceptual scale that justifies the categorical variable analysis presented in §3.2.2. The utility of ease of articulation vs. perceptibility should be measured, particularly with consonants that are nearby on the scale proposed here: such as between palatalized /n/ and /h/, between /t/ and /n/, and between /r/ and non-coronals.

This paper demonstrates that mathematical principles can be employed in the process of linguistic analysis. Further, it shows that mathematical descriptions of linguistic processes must live up not only to logical principles, but also linguistic ones. This paper also shows that many possible analyses may be accurate, and that the utility function employed in a language may not be completely general to all speakers of a language only that they should converge to the same result in order to make communication possible.

References:

Alderete, John. 2000. "Dominance Effects as Transderivational Anti-Faithfulness". [Rutgers Optimality Archive 407-0800 http://roa.rutgers.edu].

Berger, James O. 1985. *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer: New York.

Blackwell, David and M.A. Girshick. 1954. *Theory of Games and Statistical Decisions.* Dover Publications, Inc.: New York.

Boersma, Paul. 1998. *The Elements of Functional Phonology.* Doctoral dissertation. University of Amsterdam. [Rutgers Optimality Archive 173-0297 http://roa.rutgers.edu].

Boersma, Paul. 1999. "On the Need for a Separate Perception Grammar". [Rutgers Optimality Archive 358-1099 http://roa.rutgers.edu].

Chernoff, Herman and Lincoln E. Moses. 1959. *Elementary Decision Theory.* Dover Publications, Inc.: New York.

Cody, Ronald P. and Jeffrey K. Smith. 1997. *Applied Statistics and the SAS Programming Language,*4th ed. Prentice Hall: Upper Saddle River, NJ.

Dresher, Melvin. 1981. *The Mathematics of Games of Strategy: Theory and Applications.* Dover Publications, Inc.: New York.

Frisch, Stefan. 1997. *Similarity and Frequency in Phonology*. Doctoral dissertation. Northwestern University. [Rutgers Optimality Archive 198-0597 http://roa.rutgers.edu].

Hamano, Shoko. 1998. *The Sound-Symbolic System of Japanese.* CSLI Publications: Stanford, CA.

Hayes, Bruce P. 1996. "Phonetically Driven Phonology: The Role of Optimality Theory and Inductive Grounding". [Rutgers Optimality Archive 158 http://roa.rutgers.edu].

Hume, Elizabeth and Keith Johnson. 2000. "A Model of the Interplay of Speech Perception and Phonology." In Elizabeth Hume and Keith Johnson, eds., *The Role of Speech Perception in Phonology*. Academic Press: New York. [Rutgers Optimality Archive http://roa.rutgers.edu].

Johnson, Keith. 1997. *Acoustic & Auditory Phonetics.* Blackwell Publishers: Cambridge, MA.

Kadane, Joseph B., Mark J. Schervish and Teddy Seidenfeld, eds. 1999. *Rethinking the Foundations of Statistics.* Cambridge University Press: New York.

Kaplan, Mark. 1996. *Decision Theory as Philosophy.* Cambridge University Press: New York.

Kawasaki, Haruko. 1982. *An Acoustical Basis for Universal Constraints on Sound Sequences.* Doctoral thesis, University of California at Berkeley.

Kleinbaum, David G., Lawrence L. Kupper, Keith E. Muller and Azhar Nizam. 1998. *Applied Regression Analysis and Other Multivariable Methods*, 3rd ed. Duxbury Press: New York.

Kurisu, Kazutaka. 2002. *The Phonology of Morpheme Realization.* Doctoral dissertation. The University of California, Santa Cruz. [Rutgers Optimality Archive 490-0102 http://roa.rutgers.edu].

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the Worlds Languages.* Blackwell Publishers: Cambridge, MA.

McCall, Betsy & Kyoko Nagao 1999. "Aspects of Markedness in Japanese Mimetics". Presented at the 52nd Annual Kentucky Foreign Language Conference, April 22-24, 1999.

McCarthy, John. 2002. "Against Gradience". [Rutgers Optimality Archive 510-0302 http://roa.rutgers.edu].

Nagao, Kyoko & Betsy McCall 1999 "A Perception-Based Account of Mimetic Palatalization in Japanese". Presented at the International Congress of Phonetic Sciences satellite meeting, 'The Role of Perception in Phonology', July 30, 1999.

Mester, R.A. and J. Ito. 1989. "Feature predictability and underspecification: palatal prosody in Japanese mimetics." *Language* 65:258-93.

Myers, Raymond H. 1990. *Classical and Modern Regression with Applications*, 2nd ed. PWS-Kent Publishing Company: Boston.

Nagao, Kyoko & Betsy McCall 1999 "A Perception-Based Account of Mimetic Palatalization in Japanese". Presented at the International Congress of Phonetic Sciences satellite meeting, 'The Role of Perception in Phonology', July 30, 1999.

Pierrehumbert, J. (1993). "Dissimilarity in the Arabic Verbal Roots". *Proceedings of NELS 23.* GLSA Publications: Amherst, MA.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar.* MIT Press: Cambridge, MA.

Raiffa, Howard and Robert Schlaifer. 1960. *Applied Statistical Decision Theory.* John Wiley and Sons, Inc.: New York.

Resnik, Michael D. 1987. *Choices: An Introduction to Decision Theory.* University of Minnesota Press: Minneapolis.

Robert, Christian P. 2001. *The Bayesian Choice*, 2nd ed. Springer: New York.

Ross, Sheldon M. 1997. *Simulation*, 2nd ed. Harcourt Academic Press: New York.

Steriade, Donca. 1995. "Underspecification and Markedness". In John Goldsmith, ed., *The Handbook of Phonological Theory.* Blackwell: Cambridge. 114-174.

Tsujimura, Natsuko. 1996. *An introduction to Japanese linguistics.* Blackwell: Malden, MA.

Tversky, A. (1977). "Features of similarity". *Psychological Review.* 84:327-352.

Tversky, A. & I. Gati. (1978). "Studies of similarity." In E. Rosch and B. Lloyd (eds.) *Judgment Under Uncertainty: Heuristics and Biases.* Earlbaum: Hillsdale, NJ.

Zoll, Cheryl. 1997. "Conflicting directionality." *Phonology* 14:263-286. [Rutgers Optimality Archive 151-1096 http://roa.rutgers.edu].

---

[1] I would like to thank Ken de Jong, Natsuko Tsujimura, Barbara Margolius, John Holcomb, and participants at the 52nd Annual Kentucky Foreign Language Conference (1999), the Satellite Meeting of the International Congress of Phonetic Sciences (1999), and various members of the Linguistics Department at Indiana University for their helpful comments in the development of this paper. I would also like to acknowledge the special assistance of Kyoko Nagao, my coauthor on the two conference presentations, and with whom much of the linguistic analysis was begun.