

# DOCUMENTING PHONOLOGICAL CHANGE: A COMPARISON OF TWO JAPANESE PHONEMIC SPLITS\*

*Kathleen Currie Hall  
University of British Columbia*

## 1. Introduction

It is well known that pairs of sounds that were at one point allophonic can become contrastive and vice versa (e.g., Hock 1991). There is not, however, a clear means of determining how far such changes have progressed or which factors are affecting them. This paper illustrates one means of approaching this problem, applying a probabilistic metric for measuring phonological relationships (Hall 2009, 2012) to two pairs of sounds that are undergoing change in Japanese, showing a differential rate of change across the pairs and non-uniform effects of phonological context.

## 2. Phonological Relationships

### 2.1 Basic Definitions

A standard view of phonological relationships is that there are basically two kinds of relations that can hold between pairs of sounds (features, segments, etc.) in a language: *contrast* and *allophony* (e.g., Goldsmith 1995, Steriade 2007, Drescher 2011). Contrast is primarily characterized by the sounds' belonging to different categories, being used to distinguish between lexical categories, and being in at least partially overlapping phonological environments. Allophony, on the other hand, is characterized by the sounds' belonging to the same category and being in entirely complementary distribution. The latter characteristic for each relationship is one of the key ways in which the relationships are in fact determined: two sounds, *a* and *b*, are said to be contrastive if there is at least one phonological environment in which it is impossible to predict which of the two will occur, while they are allophonic if in every phonological environment, it is possible to predict which will occur. While there are other criteria (e.g., the presence or absence of alternations or the degree of phonetic similarity; see discussion in Hall 2013), the notion of predictability of distribution, applied in this way, is a major hallmark of the way that phonological relationships have been defined (e.g., Chao 1957, Jakobson 1990, Hall 2009). This criterion, then, will provide the basis for the documentation of changes in phonological relationships described in this paper. This choice is not intended to imply that other criteria would not be relevant in understanding the ways in which changes occur over time, but rather to provide a starting point for such explorations.

---

\* I am particularly grateful to Mary Beckman, Elizabeth Hume, Eric Fosler-Lussier, Daniel Currie Hall, and the audience at the 2013 CLA for discussion of aspects of this paper.

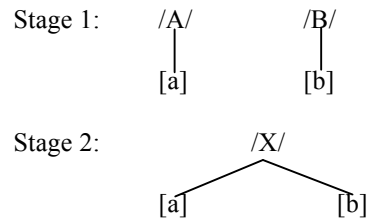
To preview the metric that will be discussed more fully in §2.3, we can reconceptualize the notion of predictability of distribution from a categorical split between “predictable” (i.e., allophonic) and “not predictable” (i.e., contrastive), into a continuum with perfect predictability on one end and perfect non-predictability on the other (cf. Goldsmith 1995, Hall 2009). Phonological relationships can be defined at any point along this continuum. One reason for adopting this continuous definition rather than a categorical one is to provide a means of documenting phonological changes, as described in the following section.

## 2.2 Changes in Progress

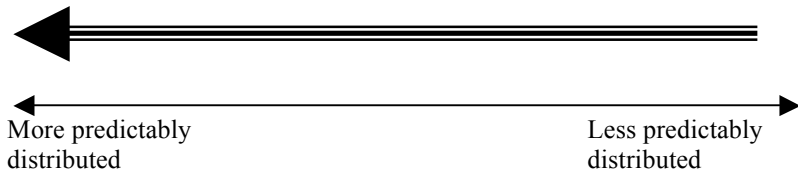
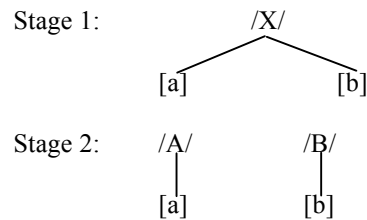
While the definitions given above in §2.1 can be applied to any snapshot of a phonological system, phonological relationships are not always stable over time. Pairs of segments can become more predictably distributed (i.e., merge into a single phonological category, with predictably distributed allophones) or less predictably distributed (i.e., split into separate phonological categories that have overlapping distributions) over time. Hock (1991), for example, describes a phonemic merger as a situation in which two unpredictably distributed segments (phonemes) merge into a single phoneme, either through the loss of one of the phonemes or through the introduction of predictable distribution of the two segments. He describes a phonemic split, on the other hand, as a situation in which two predictably distributed segments (allophones of a single phoneme) in a language split into unpredictably distributed segments (separate phonemes). Thus, mergers and splits can be illustrated as in (1), where mergers involve movement along a continuum from less to more predictable distribution, and splits involve movement in the opposite direction.

It is clearly not the case that language users abruptly shift from Stage 1 to Stage 2; there must be intermediate stages of predictability during the transition period from one stage to another. This is true at least insofar as one examines the surface predictability of distribution (where, indeed, it is probably quite rare to find perfect complementarity or perfect overlap of distributions). If, on the other hand, one considers the question from a traditional binary standpoint, such that all that matters is being allophonic (completely predictably distributed) or non-allophonic (unpredictably distributed in at least one case), then the adoption of even a single word that creates a contrast would restructure the entire system, if going from allophony to contrast (see Hall & Hall, 2013), and intermediate degrees of predictability are irrelevant. Similarly, one would not consider a “change” from contrast to allophony to have even occurred until all unpredictable conditioning environments had been lost. Such a binary split might indeed be useful for understanding the underlying structure of a system and for assigning contrastive features using the Contrastivist Hypothesis (Hall 2007, Dresher 2009). At the same time, it is certainly the case that the surface distributions may vary considerably during the timecourse of change and that studying those varying distributions may provide insight into exactly how such changes take place.

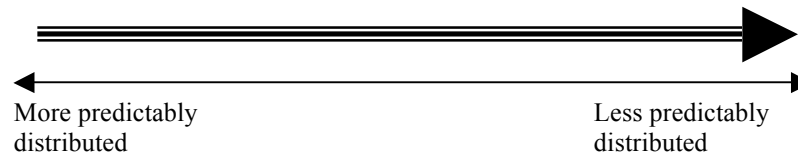
## (1) Example of phonemic merger (a) and phonemic split (b)

(a) *Phonemic Merger*

Movement is from less predictably distributed to more predictably distributed:

(b) *Phonemic Split*

Movement is from less predictably distributed to more predictably distributed:



## 2.3 A Metric

Hall (2009, 2012) provides a metric for quantifying surface predictability of distribution that is based on the information-theoretic concept of *entropy*, or uncertainty (e.g., Shannon & Weaver 1949). In this case, entropy provides a

measure of uncertainty in the choice between two sounds, *a* and *b*, in a particular phonological environment, and is quantified using the equation in (2).

$$(2) \text{ Environment-specific entropy: } H = - \sum p(X|e) \log_2 p(X|e)$$

In (2),  $p(X|e)$  is the probability of each element (*a* and *b*) in the given environment;  $X \in \{a, b\}$ , and the equation sums over these two probabilities. Because there are only two choices, entropy in this case will range between 0 (no uncertainty; analogous to allophony) and 1 (complete uncertainty; analogous to contrast).

To then calculate the overall, systemic relationship between *a* and *b* in a language (rather than simply their relationship in a single environment), the weighted average entropy is calculated, using the formula in (3).

$$(3) \text{ Systemic entropy: } \sum_e (H(e) * p(e|X))$$

In (3),  $H(e)$  is the entropy in each individual environment, calculated from (2), while  $p(e|X)$  is the probability of each environment occurring as compared to other environments, where the set of environments is limited to those where *a* or *b* can occur. This systemic measure again ranges from 0 to 1, and indicates the uncertainty of the choice between *a* and *b* on average across the entire system, allowing the behaviour of the two in more frequent environments to count more toward the overall measure.

### 3. Japanese

#### 3.1 Basic Japanese Sound Structures

The two pairs of sounds in Japanese that will be the main focus of subsequent discussion are [s] / [ɕ] and [t] / [tɕ]. Before describing the specific distribution of each of these pairs and how they are changing, a bit of background on Japanese phonetics and phonology more generally is warranted (see, e.g., McCawley 1968; Vance 1987; Tsujimura 1996; Akamatsu 1997, 2000; Labrone 2012). Only the facts that are relevant for an understanding of the distribution of the pairs of sounds will be provided; see the references above for a more comprehensive description of Japanese phonology.

The basic syllable-structure in Japanese is (C)V(N); a syllable consists of minimally a vowel, along with an optional onset and an optional coda; the only consonants allowed in coda position are nasals (and the first half of geminate consonants). There are no word-onset or word-coda consonant clusters; sequences of consonants occur only word-medially and are always homorganic—either a nasal plus homorganic obstruent or a geminate consonant.

Japanese has a five-vowel system: [i], [e], [a], [o], and [u]. Vowels can be either long or short: e.g., [to] ‘door’ vs. [too] ‘ten.’ The length of the vowel does not affect which consonants it can appear next to: if, for example, a consonant can appear before [i], then it can always also appear before [ii].

There is a common process of vowel devoicing in Japanese, by which a high vowel<sup>1</sup> is devoiced between two voiceless consonants (e.g., /kita/ ‘north’ is realized as [kᶱta]) or word-finally when unstressed and after a voiceless consonant (e.g., /muki/ ‘direction’ is realized as [mukᶱ]). Only voiceless segments can be adjacent to a voiceless vowel, but if a voiceless consonant can appear next to a voiced vowel, it can appear next to its voiceless counterpart.

There is, of course, much more to be said about the phonological structure of Japanese; however, the preceding remarks should suffice to allow a basic understanding of the distribution of particular consonant pairs in Japanese. Because all obstruent consonants appear in onset position whenever they occur (they may, of course, simultaneously appear in coda position if they are geminate), it is possible to focus exclusively on the *following* context when describing the distribution of consonants. Thus, only a two-segment window is needed to describe a sound’s environment: the consonant in question and the vowel following it.

### 3.2 Two Pairs Undergoing Change

The two pairs of interest here are [s] / [ç] and [t] / [tç], both pairs that involve an alveolar versus an alveopalatal place of articulation (e.g., Akamatsu 1997; Labrune 2012). Each pair has sometimes been considered to consist of allophones of a single phoneme in Japanese (e.g., Tsujimura 1996, Labrune 2012), largely because the members of the pairs have traditionally occurred in complementary distribution before front vowels. [s] and [t] do not occur before [i], while [ç] and [tç] do not occur before [e], at least in native Japanese words (with the possible exception of [tçe], an exclamation meaning roughly ‘ugh!’).

All four consonants, however, can surface before any of the back vowels, as in the (near) minimal pairs [soba] ‘buckwheat noodles’ vs. [çoba] ‘street market’ and [tobu] ‘to fly’ vs. [tçobo] ‘gamble.’ Thus, there is some evidence for their status as contrastive (at least on the surface), even within the native stratum, but if they are contrastive, that contrast is neutralized before front vowels.

The predictability of the distribution of [s] / [ç] and [t] / [tç] before front vowels is emphasized by the fact that each pair has alternations. For example, as shown in (4), the verb meaning ‘put out’ contains an [s] in the present, provisional, causative, and tentative forms, where it occurs before endings that start with [u], [e], [a], and [o], respectively. On the other hand, it contains [ç] in the past, participial, and conditional forms, where it occurs with endings that start with [i].

Similarly, [t] and [tç] alternate with each other. For example, the verb for ‘to wait’ contains a [t] when it appears before [a] in the negative form, [matanai], but contains [tç] when it appears before [i] in the polite present form, [maccimasu].

---

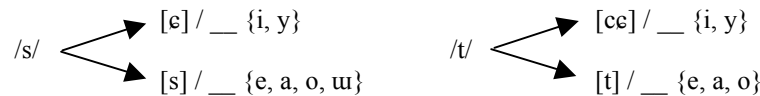
<sup>1</sup> To a certain extent, non-high vowels can also undergo devoicing, but it is less regular than high-vowel devoicing; see, e.g., Akamatsu (1997:36-40), Labrune (2012: 34).

- (3) Alternation between [s] and [ɕ] in the verb ‘put out’ (from McCawley 1968: 95)

<i>Form</i>	<i>Pronunciation</i>	<i>Vowel</i>	<i>Fricative</i>
present	[dasu]	[u]	[s]
provisional	[daseba]	[e]	[s]
causative	[dasareru]	[a]	[s]
tentative	[dasoo]	[o]	[s]
past	[dacita]	[i]	[ɕ]
participial	[dacite]	[i]	[ɕ]
conditional	[daeitara]	[i]	[ɕ]

Labrune (2012) analyzes these facts as in (4). There are two separate phonemes, /s/ and /t/; each has entirely predictable allophones in particular contexts. /s/ is realized as [ɕ] before both [i] and [y] (where [y] is a glide, not a high front rounded vowel), and as [s] elsewhere. Similarly, /t/ is realized as [cɕ] before both [i] and [y], and as [t] elsewhere (except before [u], where it has a third allophone, [t͡ɕ]). Thus, surface sequences of [ɕa], [ɕo], and [ɕu], or [cɕa], [cɕo], and [cɕu], are interpreted as being surface realizations of /syV/ and /tyV/, respectively, where V is any of {a, o, u}. The absence of surface [cɕ] and [cɕe] is attributed to a more general prohibition against \*[ye] sequences.

- (4) Analysis of allophonic, predictable distribution by Labrune (2012)



At the same time, the introduction of loanwords is eroding the predictable distribution of both of these pairs of sounds before front vowels; they are undergoing a phonological split. The alveolars can occur before [i] in words like [si] ‘letter C’ and [ti] ‘letter T,’ while the alveopalatals can occur before [e] in words like [cɕfu] ‘chef’ and [cɕekku] ‘cheque.’ Given their largely parallel phonetic nature and distribution, one might assume that the two pairs are progressing similarly in their changes. Labrune (2012), for example, treats both pairs (or rather, the [ɕ] & [cɕ] allophones of /s/ & /t/) equivalently. She almost exclusively refers to them together and claims that they are both “frequent, particularly in Sino-Japanese words and in recent loans, before all five vowels” (66), which suggests a relaxing of the \*[ye] constraint across the board. She gives, however, only one example in front of each vowel for each consonant (and in fact only provides examples before [a, o, u], though later provides one example with [cɕ] (p. 98)), and uses this as evidence that the consonants are both “frequent” and similar to each other. This leaves open the question, then, of

whether they are in fact both patterning the same way and consequently whether they should be treated the same way formally.<sup>2</sup>

#### 4. Applying the Metric to Japanese

##### 4.1 Corpora

Two corpora of Japanese were used to find the type and token frequency of each sound in question: the Nippon Telegraph & Telephone (NTT) lexicon and the Corpus of Spontaneous Japanese (CSJ). The NTT lexicon was used for all type frequencies (i.e., the number of words containing [si], [ei], [se], [ee], etc.); the CSJ was used for all token frequencies (i.e., how often those words actually occur in spoken Japanese). These frequencies were then used to calculate the probabilities of occurrence needed for the calculation of environment-specific and systemic entropy.

The NTT lexicon is a list of Japanese words based on the 3<sup>rd</sup> edition of the *Sanseido Shinmeikai Dictionary* (Kenbou et al., 1981; see Amano & Kondo 1999, 2000 for a description of the NTT lexicon). Only the phonetic transcriptions were used in the current analysis. Crucially, the distinctions among all the segments of interest are labelled, even when they are traditionally predictable. For example, both [s] and [ɕ] are transcribed; all tokens of [ɕ] that are predictable because they occur before [i] are transcribed as [sh], while all tokens of [ɕ] that are unpredictable are transcribed as [shy].

The CSJ is a collection of approximately 7,000,000 words recorded over 650 hours of “spontaneous” speech (the recordings involved planned topics if not planned word-for-word texts, though most texts were not designed specifically for inclusion in the CSJ). All of the speech is “standard” Japanese, similar to Tokyo Japanese, used by educated speakers in public situations; the speech was screened and all speakers with particular dialectal morphological and/or phonological markers were excluded. A description of the corpus is available online at: <http://www.kokken.go.jp/katsudo/seika/corpus/public/>; see also Maekawa, Koiso, Furui, & Isahara (2000), Furui, Maekawa, & Isahara (2000), Maekawa (2003, 2004). The CSJ “Core” contains about 500,000 phonetically transcribed words in 45 hours of speech, and it is this subset of the total that was used in the current analysis. No read speech was included. As in the NTT lexicon, distinctions among all of the segments in question are labelled, even when they are traditionally predictable. It is important to remember that the transcriptions in the CSJ are transcriptions of the actual acoustic signal, and not simply idealized phonetic transcriptions of the spoken text. Thus, the frequency counts from the CSJ accurately reflect the actual occurrences of the sequences in question and are not directly subject to, for example, a lexicographer’s bias toward a given pronunciation.

Two points should be particularly noted. First, all calculations are done on *surface* forms. As we have seen above, surface contrasts for each pair before back vowels can be analysed as stemming from underlying representations that

---

<sup>2</sup> Labrune does claim elsewhere, however, that [si] sequences are not allowed at all in loanwords, while [ti] sequences are in the most recent loans, indicating that there may in fact be differences between them.

do not involve the existence of separate phonemes /ɛ/ and /œ/ but rather just contrasts between /s/ and /sy/, or between /t/ and /ty/. If there is non-zero entropy for contrasts before [i] on the surface, however, then we might speculate that there has indeed been an actual phonemic split underlyingly as well. Second, it should be noted that results are calculated over all strata of the Japanese lexicon, without respect for which strata particular words might occur in; this follows Bloch (1950: 87), who deems it “unacceptable” to try to separate out different parts of the “necessarily single . . . network of total relationships among all the sounds that occur in the dialect.” Bloch, of course, is in the minority with respect to Japanese phonology, but subsequent authors have also pointed out the difficulty of actually relegating particular effects into one stratum or the other (e.g., Itô & Mester 1995; Labrune 2012). Here, the purpose is to examine the extent to which each pair of interest is undergoing phonological change; it is possible that the results would be cleaner were we to separate data in the Yamato (native) and Sino-Japanese strata from data in the *gairaigo* (recent loanword) stratum, with the expectation that the splits are really happening only in loans. This separation, however, should affect both pairs equally, and is expected to simply be reflected here in lower entropy numbers overall, given still large degrees of predictability in the non-*gairaigo* strata.

#### 4.2 Analysis

Slightly different search procedures were used to extract frequency data for the pairs of sounds in question from the two corpora, because of their different structures. For the NTT type frequencies, the raw corpus material consisted of a single long text file with phonetic transcriptions. These transcriptions indicate the mora boundaries within each word. A script was written in R (R Development Core Team, 2007) that separated out each transcription into its component morae and counted the number of occurrences of each mora within the corpus. This produces a frequency table of all morae in Japanese that occur in the NTT lexicon. These frequencies were used as type frequencies for each of the sequences of interest. For example, the mora [su] occurs 6222 times in the NTT lexicon. Note that the same mora can appear more than once in the same word—for example, in the word [su.su.mi] ‘progress’ the mora [su] appears twice. As a result, these two are counted separately as part of the 6222.

Thus, the type frequency of a sequence corresponds to the number of occurrences of that sequence in the Japanese lexicon, not strictly speaking the number of *words* that the sequence occurs in. This method of counting is preferable not only because it accurately represents the number of occurrences in the lexicon but also because it avoids the rather complicated issue of having to define a “word” in Japanese. It should be noted that the NTT lexicon also lists homophonous words separately. For example, there are six occurrences of the word [su.i.ta.i]. Jim Breen’s online dictionary of Japanese also lists six entries for this word, with six different meanings. Again, each instance of a mora across entries is counted separately; thus, the [su] from [su.i.ta.i] is counted six times in the current analysis.

For token frequencies, a different method was used because the CSJ corpus is much larger than the NTT lexicon, being a collection of actual spoken texts rather than a list of lexical entries. It is therefore not efficient to get the



frequency counts for all the morae. Instead, a list of all the possible CV sequences containing the consonants in question was developed. The corpus was then automatically searched for each occurrence of each sequence; the number of occurrences was counted and recorded. These counts were used as the token frequency measurements in the subsequent analysis.

In addition to the pairs [s]/[ɛ] and [t]/[cɛ], which are of particular interest, entropies for the pair [t]/[d] were calculated for comparison. This pair is entirely contrastive (i.e., the distributions are completely overlapping) in Japanese, and so the entropies will tend toward the higher end of the scale. Note, though, that even this pair doesn't always reach "full" contrast of  $H = 1$ , because there are slightly different type / token frequencies of words containing each sound.

### 4.3 Results

The results for each pair before [i] are shown in Figure 1; the results for each pair before [e] are shown in Figure 2. In each graph, the contrastive pair [t]/[d] is shown by the leftmost pair of columns, the pair [s]/[ɛ] in the middle, and the pair [t]/[cɛ] on the right. Within each pair of columns, the lefthand column represents the entropy based on type frequency calculations, and the righthand column represents the entropy based on token frequency calculations.

In both graphs, it is clear that [t]/[d] is in fact much more contrastive (i.e., unpredictably distributed) than either of the other pairs. Thus, while a split may be in progress for both pairs, it has not reached the stage of being analogous to other pairs of more "standardly" contrastive sounds. Interestingly, the type-based entropy for [s]/[ɛ] before [i] is in fact still 0, indicating that there were no words in the lexicon containing the [si] sequence, though a few words were pronounced with this sequence in the CSJ, leading to an entropy of 0.026. (It should be noted that these words were almost all loans, including "syllable," "sink," "CBS," "indexing," "proceedings," and "ICU"; many were in fact marked as being pronunciation errors, apparently because of the use of [si].) There is, of course, a striking difference between this pair and the pair [t]/[cɛ], which is much more advanced in its progress toward full contrast in this environment. The type-based entropy for the latter pair is 0.251, and the token-based entropy 0.363. Thus, changes in one pair do not in fact entail changes in the other.

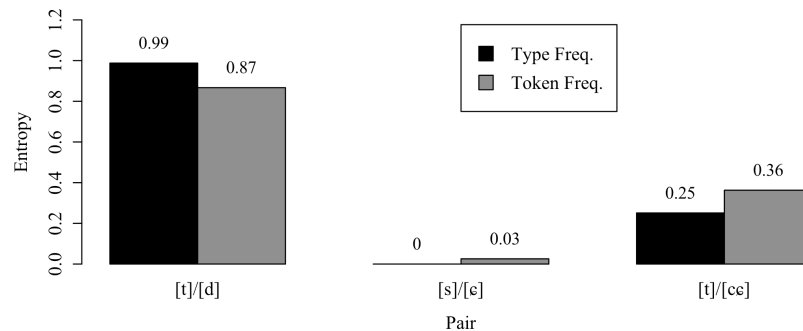
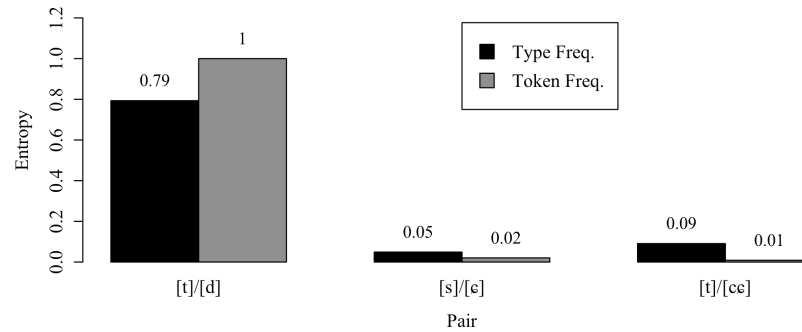


Figure 1: Entropies for each pair before [i]

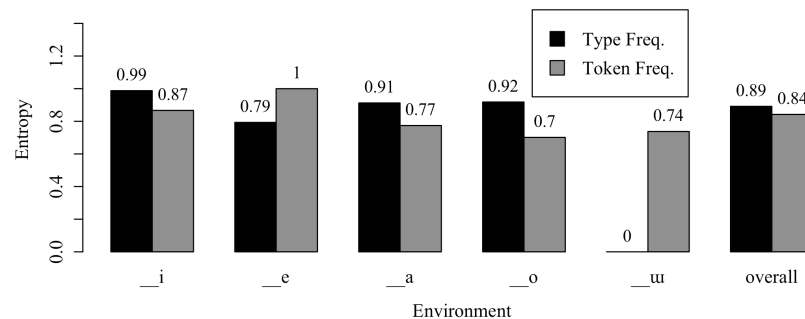
Looking at Figure 2, it is clear that before [e], the two splitting pairs are much more similar to each other, with neither pair having advanced much beyond 0 entropy. The entropy values for [s]/[ʃ] are 0.049 (type) and 0.020 (token), while those for [t]/[tʃ] are 0.091 (type) and 0.009 (token).



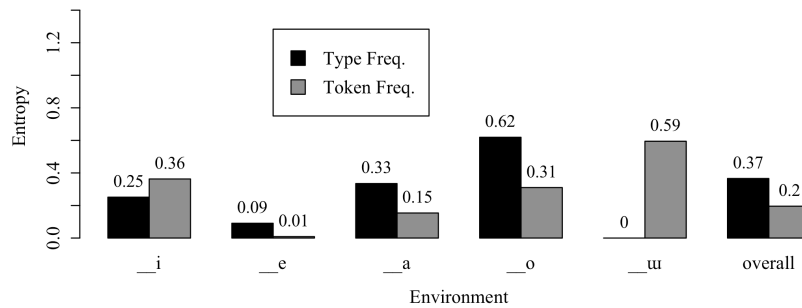
**Figure 2: Entropies for each pair before [e]**

We turn next to Figures 3-5, which show the entropy values (type and token) for each pair, respectively, across all five vocalic contexts and then the overall (systemic) weighted average entropy for the pair.

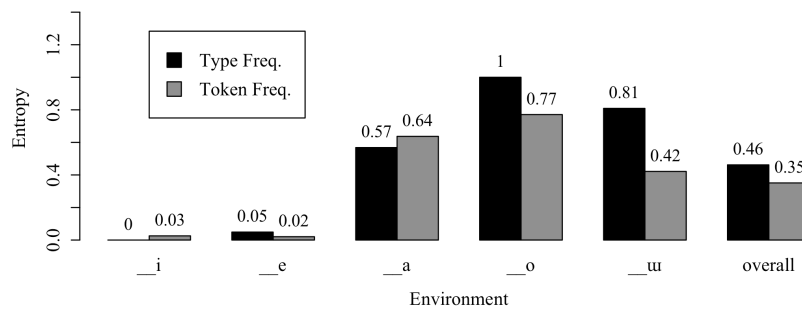
Figure 3 shows that the control pair, [t]/[d], is highly contrastive across almost all environments ( $H(e) \geq 0.7$ ). The one exception is before [u] in the NTT lexicon. Traditionally, both [t] and [d] are affricated in this context and so both “should” be non-occurring. There were, however, seven words in the lexicon containing the sequence [du], but none with the sequence [tu], making the entropy in this context quite low (0). In the CSJ, however, both [tu] and [du] sequences occurred, and so the entropy is again high. It is important to note, however, that the extremely low entropy in this context does not in fact have a significant impact on the overall entropy value for this pair, because this environment accounts for only 0.2% of occurrences of [t] and [d] in the lexicon.



**Figure 3: Entropies of [t]/[d] across environments**



**Figure 4: Entropies of [t]/[tʃ] across environments**



**Figure 5: Entropies of [s]/[ʃ] across environments**

Figure 4 shows the entropies of [t]/[tʃ] across environments, individually and overall. The calculations show that the entropy of [t] and [tʃ] before [i] is 0.251 (type-frequency based) or 0.363 (token-frequency based). Before [e], the entropies are 0.091 (type-frequency based) or 0.009 (token-frequency based). That is, the uncertainty of the choice between [t] and [tʃ] in these environments is greater than 0, as it would be if the two were still entirely predictable. Thus, these numbers reveal that the pair [t]/[tʃ] has split and become more contrastive in these environments.

Note, however, that the split before [i] is more advanced than the split before [e]. The fact that they are different is perhaps surprising given that a typical description of the predictable distribution is that it occurs “before front vowels.” Here, it seems that a breakdown of predictability in one specific environment ([\_i]) doesn’t in fact entail a breakdown before all environments fitting a more general description (“front vowels”).

Also interesting is the fact that the extent to which this pair is contrastive before back vowels is also quite variable, with [o] and [u] being the most unpredictable environments. (The caveat with [u] discussed above for the control pair [t]/[d] holds here, given that [t] again does not occur before this vowel in the NTT lexicon.) All of these environments have lower entropy than the control pair of [t]/[d], however, despite the fact that both pairs are “contrastive” in a categorical sense; this is because of the large discrepancy in the overall frequency of occurrence of [t] vs. [tʃ]. [t] is simply a much more frequent sound in Japanese—it occurs 2.6 times more often than [tʃ] in the NTT lexicon and 8.12 times more often than [tʃ] in the CSJ. In terms of raw

predictability, then, one is far more likely to be correct if one guesses that a [t] will occur in some environment than if one guesses [cɛ], even in the absence of phonological knowledge. Hence, even the environments that are fairly uncontroversially contrastive for this pair do show a lower degree of uncertainty than that of other contrastive pairs.

In terms of how far the split has progressed, the extent to which [t]/[cɛ] is contrastive before [i] is almost the same as the extent to which it is contrastive before [a]: 0.251 / 0.363 for type and token entropies before [i], vs. 0.335 / 0.154 before [a]. A threshold (other than >0) has not been determined for establishing contrast, but these numbers suggest that if we are to treat [t]/[cɛ] as contrastive before back vowels, we should treat it as such before [i] as well (though “contrastive” here may mean a contrast between /t/ and /ty/, not /t/ and /cɛ/).

Finally, consider the difference between the type and token frequency calculations. When type-based entropy is greater than token-based entropy, the split is more advanced in theory than in practice—that is, there may be a fair number of individual lexical items that show the split, but they simply are not used frequently in real speech. On the other hand, entropies that are higher when calculated on tokens than on types indicate that a contrast is quite strong in practice, in that it is actively used in words that are typically spoken, even if there aren’t many such words. The traditional reliance by linguists on one or even a handful of words to demonstrate distributional facts cannot reflect such subtleties, and reveals nothing about whether real speakers in fact know or use the words in question. In the case of [t]/[cɛ] in particular, it seems that the pair is *more* contrastive in practice before [i] than it is before [a] and [o]. This is partly, again, due to the overwhelmingly greater raw frequency of [t] as compared to [cɛ], which drives down the token-based entropy calculations before [a] and [o], but it is still the case that there’s not much evidence for making a distinction between the contrastive nature of this pair before [i] and before the back vowels.

Figure 5 shows the entropies of [s]/[ɛ] across environments. Taking the graph as a whole, we see a much different profile for this pair than we did for [t]/[cɛ]; [s]/[ɛ] still follows the traditional descriptions of its distribution. That is, it is strongly contrastive before back vowels and almost entirely predictable before front vowels. Despite reports similar to those for [t]/[cɛ] of loanwords disrupting this predictability, neither front vowel environment shows much evidence of the split’s actually taking hold: the entropy is maximally 0.049, for type frequency calculations before [e]. Thus, the splitting of one pair of sounds, [t]/[cɛ], doesn’t entail the splitting of another pair of sounds with a similar phonetic profile and a similar phonological distribution. At the same time, it should be noted that most of the entropies are *not* 0; from a purely categorical perspective, the fact that there are *any* sequences of [si] and [ɛe] would necessitate analysing this as a split having taken place. The quantification of the split in terms of entropy, however, clearly demonstrates that the split is hardly present and certainly has not taken hold in the same way as that of [t]/[cɛ].

## 5. Discussion and Conclusions

Recall that Labrune’s (2012) account of the standard distribution is that palatal [ɛ] and [cɛ] are derived from underlying sequences of /si/ or /syV/ and /ti/ or /tyV/. She then imposes an additional constraint against \*[ye] sequences. This kind of treatment is standard: it assumes that [ɛ] and [cɛ] are parallel in terms of analysis and distribution, and that the distribution is categorical in nature.

There are two possible interpretations of the data presented here. The first is categorical: if the entropy for a pair of sounds is greater than zero, then those sounds must be unpredictably distributed and hence contrastive in that environment. Under this interpretation, both [s] vs. [ɛ] and [t] vs. [tɕ] would be considered representative of underlying contrasts, at least if token frequencies as the language is used in the CSJ are considered. This underlying contrast might be either one of /s/ vs. /ɛ/ and /t/ vs. /tɕ/, or one of /s/ vs. /sy/ and /t/ vs. /ty/, as illustrated in (5); the bold values in (5b) indicate changes from (4). In the latter case, the basic analysis of the plain and palatalised consonants would still be one of allophony, but with a change in the conditioning environments. In either case, one could imagine a prohibition against palatals before [e], which would explain the relatively lower entropies before [e] as compared to before other vowels.

(5) Two possible interpretations of new contrasts

(a) Introduction of new phonemes

/s/ → [s] / \_\_ {e, a, o, u, i}      /t/ → [t] / \_\_ {e, a, o, i}

/ɛ/ → [ɛ] / \_\_ {a, o, u, i}      /tɕ/ → [tɕ] / \_\_ {a, o, i}

(b) Change in conditioning environments; keep allophonic analysis

/s/  $\begin{cases} \rightarrow [ɛ] / \_ \{y\} \\ \rightarrow [s] / \_ \{e, a, o, u, \mathbf{i}\} \end{cases}$       /t/  $\begin{cases} \rightarrow [tɕ] / \_ \{y\} \\ \rightarrow [t] / \_ \{e, a, o, \mathbf{i}\} \end{cases}$

The major interpretation of this data is that there is some cut-off other than non-zero that is needed before a “true” split has been achieved. For example, one might set an arbitrary cut-off of 0.05 or 0.10, claiming that entropies less than or equal to that number don’t signal new contrasts but rather non-integrated exceptions. Under this interpretation, /s/ has not changed at all; it surfaces as [ɛ] before [i] and [y], and as [s] elsewhere. /t/, on the other hand, would have undergone a change in its conditioning environments: [i] would now not condition [tɕ], but rather also be an environment for [t] to occur. That is, /t/ would be analysed as in (5b), while /s/ would still be analysed as in (4).

In either scenario, there is a remaining question: why is there a difference between /s/ and /t/? Despite apparent formal similarity, they are not parallel: either the split has proceeded further for /t/ than for /s/, or the conditioning environments have changed for /t/ but not /s/.

One possibility is suggested by Labrune’s description of the historical status of these two phones. She mentions (2012: 62) that /t/ in Ancient and possibly Middle Japanese was realized as [t] before all vowels, while /s/ was potentially realized as [ɛ] before all vowels and if not, at least before both [i] and [e] (2012: 66). Thus, [s] and [ɛ] have been linked, and perhaps predictably distributed, for much longer than [t] and [tɕ]. Given the other disturbances to the distribution of /t/, such as the prohibition against \*[tu], it’s possible that the formal analysis of /t/ given in (4) is simply wrong, and that instead, /t/ and /tɕ/ have always been contrastive since the latter entered the language. A specific constraint against \*[ti] (along, perhaps, with a more general prohibition against

palatals of any sort before [e]) might give the surface appearance of a distribution similar to that of /s/. But where /s/ is slowly moving individual vowels out of the conditioning environments for [ɛ], /t/ has simply been losing its co-occurrence constraints. While speculative, this hypothesis merits further investigation, particularly with the evidence given here of a decided difference between the apparent splitting of [ti]/[cɛi] compared to that of [si]/[ɛi].

Relatedly, Labrune's (2012) description of the data would lead one to expect a third scenario. Her claim is that /sye/ and /tye/ are entering the lexicon, giving rise to [cɛ] and [cɛɛ], but that [si] is not. Were this the case, we would expect a difference in the entropies of [s]/[ɛ] before [i], on the one hand, and [t]/[cɛ] and [s]/[ɛ] before [e] on the other. What is found instead, however, are very similar entropy values for all three cases; all three of these traditionally proscribed sequences do in fact occur, if sparsely, in the data. While the type-based calculations might support this scenario, assuming a categorical split between zero and non-zero entropy, the token-based calculations do not.

Finally, consider the status of a \*[ye] constraint or a constraint against palatals more generally before [e]. The non-zero entropy of the contrasts between [se]/[cɛ] and [te]/[cɛɛ] suggests that such a constraint is weakening, even if it is still present. If weakening were true overall, we would also expect to see new sequences of other consonants palatalised before [e], such as [kye] and [gye]. Labrune, however, gives no indication that such sequences are possible, and no such sequences are transcribed in either the NTT lexicon or the CSJ. Thus, perhaps we should revisit after all the possibility of separate /e/ and /cɛ/ phonemes, as in (5a), in combination with a \*[ye] constraint that pertains only to actual sequences of consonants followed by [y].

In sum, the results for Japanese suggest that [t]/[cɛ] are in fact essentially contrastive in all environments, though perhaps subject to a (weakening) \*[ye] constraint, while [s]/[ɛ] still essentially follow a pattern of predictable distribution or contrast neutralization (depending on one's treatment of their status before back vowels) before front vowels as a class. The results are not consistent with an analysis of these pairs as simply analogous to one another.

More generally, quantification of predictability of distribution as a measure of contrastiveness using entropy provides insight into the timecourse of phonological changes. Relying on a traditional handful of examples to illustrate that a split is happening does not capture the true state of affairs: are the examples in fact representative of the language as it is being used, or are they a set of rarified words that really have no discernible effect on the grammar of native speakers? It is quite difficult to tell in isolation. But, knowing the answers to such questions may prove crucial in our ability to accurately analyse the representation of sound systems and to understand how they change over time.

## References

- Akamatsu, Tsutomu. 1997. *Japanese phonetics: Theory and practice*. Munich; Newcastle: LINCOM EUROPA.
- Akamatsu, Tsutomu. 2000. *Japanese phonology: A functional approach*. Munich: LINCOM EUROPA.
- Amano, Shigeaki, and Tadahisa Kondo. 1999, 2000. *The properties of the Japanese lexicon*. Tokyo: Sanseido.
- Bloch, Bernard. 1950. Studies in colloquial Japanese IV: Phonemics. *Language*, 26(1), 86-125.

- Chao, Yuen-Ren. 1934/1957. The non-uniqueness of phonemic solutions of phonetic systems. In Martin Joos (Ed.), *Readings in linguistics I: The development of descriptive linguistics in America 1925–56* (4th ed., pp. 38–54). Chicago: University of Chicago Press.
- Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.
- Dresher, B. Elan. 2011. The phoneme. In M. van Oostendorp, C. J. Ewen, E. Hume & K. Rice (Eds.), *The Blackwell Companion to Phonology* (Vol. I, pp. 241–266). Oxford: Wiley-Blackwell.
- Furui, Sadaoki, Kikuo Maekawa, and Hitoshi Isahara. 2000. A Japanese national project on spontaneous speech corpus and processing technology. *Proceedings of ISCA ITRW ASR2000*, 244–248.
- Goldsmith, John. 1995. Phonological theory. In John A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 1–23). Cambridge, MA: Blackwell.
- Hall, Daniel Currie. 2007. *The role and representation of contrast in phonological theory*. Unpublished PhD dissertation, University of Toronto, Toronto.
- Hall, Daniel Currie, & Hall, Kathleen Currie. (2013). *Marginal contrasts and the Contrastivist Hypothesis*. Paper presented at the Linguistics Association of Great Britain, London.
- Hall, Kathleen Currie. 2009. *A probabilistic model of phonological relationships from contrast to allophony*. Unpublished PhD dissertation, The Ohio State University, Columbus, OH.
- Hall, Kathleen Currie. 2012. Phonological relationships: A probabilistic model. *McGill Working Papers in Linguistics*, 22(1).
- Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review*, 30(2), 215–275.
- Hock, Hans. 1991. *Principles of historical linguistics* (2nd ed.). Berlin; New York: Mouton de Gruyter.
- Itô, Junko, and R. Armin Mester. 1995. Japanese phonology. In John A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 817–838). Cambridge, MA: Blackwell.
- Jakobson, Roman. 1990. *On language*. Cambridge, MA: Harvard University Press.
- Kenbou, Hidetoshi, Haruhiko Kindaichi, Kyousuke Kindaichi, and Takeshi Shibata. 1981. *Sanseido Shinmeikai Dictionary*. Tokyo: Sanseido.
- Labrone, Laurence. (2012). *The phonology of Japanese*. Oxford: Oxford University Press.
- Maekawa, Kikuo. 2003. Corpus of spontaneous Japanese: Its design and evaluation. *Proceedings of ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR2003)*, 7–12.
- Maekawa, Kikuo. 2004. Design, compilation, and some preliminary analyses of the corpus of spontaneous Japanese. In Kikuo Maekawa and Kiyoko Yoneyama (Eds.), *Spontaneous speech: Data and analysis* (Vol. 3, pp. 87–108). Tokyo: The National Institute of Japanese Language.
- Maekawa, Kikuo, H. Koiso, Sadaoki Furui, and Hitoshi Isahara. 2000. Spontaneous speech corpus of Japanese. *Proceedings LREC2000*, 2, 947–952.
- McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague; Paris: Mouton.
- R Development Core Team. 2007. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Shannon, Claude E., and Warren Weaver. 1949. *The mathematical theory of communication*. Urbana-Champaign: University of Illinois Press.
- Steriade, Donca. 2007. Contrast. In Paul de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 139–157). Cambridge: Cambridge University Press.
- Tsujimura, Natsuko. 1996. *An introduction to Japanese linguistics*. Cambridge, MA: Blackwell.
- Vance, Timothy J. 1987. *An introduction to Japanese phonology*. Albany: State University of New York Press.