

## Abstract

The suffixal alternations of Japanese verbal inflection have been analyzed in at least four distinct ways in the literature. In this paper, working in the context of a general model of the inferential relation between synchrony and diachrony in inflectional morphophonology, I compute the predictions for potential change for three analyses of those alternations and show that only one set of predictions is consistent with the ongoing changes evident in a nationwide survey of inflection. I conclude that the analysis generating the correct predictions is the unique descriptively adequate analysis of the system of alternations in question. With regard to the explanatory principles governing the choice of that analysis from the set of observationally adequate alternatives, I show that the Japanese case counterexemplifies a wide range of proposals that have been made about the operation of morphophonological analysis and change and propose that the choice of both base forms and rules is due to a principle of Generalized Type Frequency. Among the general themes of the paper are the grammatical reality of language-specific phonological rules, the relevance of analogical processes to synchronic description and explanation, and the limited role played in morphophonological analysis by global considerations of predictability.

**Keywords** Japanese, verb inflection, descriptive adequacy, type frequency, analogy

“[The] rules [of morphophonology] represent the result of many historical processes of reanalysis and generalization. Their original nucleus can be a sound change, a difference in morphemic composition, or simply a pattern perceived at one point by speakers in a paradigm composed of unrelated morphemes, much as an imperfect gestalt is seen in a random arrangement of pebbles. Once these rules become established, however, they serve as the tracks along which ‘analogical change’ proceeds. Forms not conforming to them are now ‘exceptional’ and over long periods of time get reshaped so as to become regular. This process of regularization ... is what ‘analogical change’ really refers to.”

Paul Kiparsky (1972:280)

## 1 Introduction

### 1.1 Description and Explanation in Phonology: A Research Program

Since the introduction of the concepts of descriptive and explanatory adequacy in Chomsky 1964b (see also Chomsky 1964a), the terms “description” and “explanation” have defined the twin projects of generative linguistic inquiry, the goal of description being the characterization of the steady-state endpoint of the language acquisition process and that of explanation being the characterization of the starting point of that process, the initial state or innate endowment. These two goals apply to all areas of linguistic competence and are independent of whether the transition from the initial state to the steady state is taken to be mediated by an evaluation procedure (Chomsky 1957:51), by parameter setting (Chomsky 1981:4), by constraint (re)ranking (McCarthy 2002:208), or by some other mechanism.

In the area of morphophonology, the existence of multiple observationally adequate analyses for many data sets (exemplified below) made it evident quite early that descriptive adequacy could not be attained merely on the basis of analyzing patterns of distribution and alternation. At the same time, the existence of cases in which speakers seemed to have arrived at analyses strikingly different from those predicted by the standard assumptions of phonologists (see e.g. Hale 1973) made it clear that explanatory adequacy could not be attained merely by adopting a priori a particular definition of simplicity—for example, the feature-counting evaluation metric of Chomsky and Halle 1968, taken to apply to the lexicon (see pp. 381–382) as well as to rule schemata. Two representative quotations illustrating these realizations, the first focusing on the problem of description, the second on the problem of explanation, are given below.

“Generative grammarians have ... claimed that a description of the phonological structure of a language is

simultaneously a characterization of the linguistic knowledge of native speakers .... But in order for this implication to be valid, we must be able to corroborate it by corpus-external evidence ....” (Kenstowicz and Kisseberth 1979:153–154)

“... [S]howing that introducing the alternation condition can lead to more complex analyses cannot by itself refute the alternation condition, since the point at issue is precisely whether simplicity is the correct evaluation measure. .... To avoid begging the question in investigations of this problem we must look for *external evidence* as to the correctness or incorrectness of specific analyses which are required or forbidden by the constraints at issue.” (Kiparsky 1971/1982:59–60 (*italics in original*))

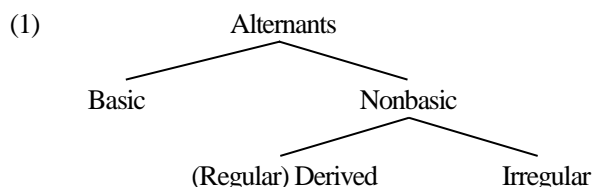
Both of the above quotations imply a research program, set forth explicitly in Kenstowicz and Kisseberth 1977:3 (and anticipated in slightly different terms in Andersen 1969:829), that would seek to determine the relevant explanatory principles through examination of a critical mass of cases for which the descriptively adequate analysis is known from external evidence. The principles thus discovered could then be used to predict the descriptively adequate analysis in cases for which no such evidence is available. The present paper, using evidence from morphophonological change, aims to advance that research program by showing that substantive results are obtainable within it. We begin in section 1.2 by developing a general model of the inferential relation between synchrony and diachrony in inflectional morphophonology, following the dictum of Kiparsky (1978/1982:217) “that structure can determine change, with the corollary that change can therefore be diagnostic of structure.” Section 2 deals with the suffixal alternations of Japanese verb inflection, which have been analyzed in at least four distinct ways in the literature. Based on the synchrony-diachrony relationship as explicated in 1.2, predictions for potential change are computed for three observationally adequate analyses of the Japanese alternations. It is then shown that only one set of predictions is consistent with the facts of change in progress, as documented in a nationwide survey of inflection; we conclude that the analysis that correctly predicts the ongoing changes (“Analysis A”) is the unique descriptively adequate analysis of the system of alternations under consideration.

Moving in section 3 to questions of explanation, we begin in 3.1 by considering a number of hypotheses about morphophonological analysis that have been entertained in the literature but are counterexemplified by speakers’ choice of Analysis A in the Japanese case or by the relevant system of alternations itself. With the space of analytic possibilities thus narrowed, section 3.2 argues that morphophonological analysis is a two-stage process involving choice of underlying representations (URs) and choice of a regular alternation in that order, rather than a one-stage process in which a combination of URs and a rule is chosen as a unit. A principle of Generalized Type Frequency is claimed to govern the selection of both URs and rules, and the choice of Analysis A is explained in terms of that principle. Section 4 briefly recapitulates the paper’s central conclusions.

## 1.2 The Relation between Synchrony and Diachrony in Morphophonology

We have suggested, following Kiparsky, that there is in principle a bidirectional inferential relationship between synchronic structure and diachronic change. But just how is it that synchronic structure limits the paths of possible change in morphophonology, so that change, conversely, can be diagnostic of structure? I will maintain that the link between synchrony and diachrony in this area follows from the typology of morpheme alternants that results from the assumption of basic forms and rules combined with a model of how the alternant types in question are stored, retrieved, and generated.

Two of the claims of classical generative phonology that distinguished it most sharply from the post-Bloomfieldian phonology that preceded it were that, in analyzing alternations, speakers postulate (a) basic or underlying forms for morphemes and (b) rules to derive nonbasic forms (see e.g. Halle 1962). Provisionally adopting these claims, and assuming further that basic forms in the cases of interest coincide with occurring morpheme alternants (see de Chene 2014a for discussion), postulation of basic forms divides the entire set of morpheme alternants for a given alternation into basic and nonbasic subsets. Further, the postulation of a rule governing that alternation divides nonbasic alternants into regular derived forms, which are predicted accurately by the rule postulated, and irregular forms, which are not. The result of these two binary divisions is the tripartition of alternants in (1).



The fundamental link between synchronic structure and diachronic change in morphophonology is, I claim, a consequence of this tripartition of alternants in that differential predictions concerning diachronic stability arise from the psychological interpretation of the three alternant types—in particular, from whether they must be stored in memory or can, on the other hand, be generated online. Before examining those questions, however, it is necessary to say something about the ways in which a morpheme alternant can be irregular.

I will take a single phonological matrix, in conjunction with a specification of meaning, to constitute the irreducible minimal form of a lexical entry and will count anything beyond that, from a diacritic feature to a suppletive alternant, as excess information, equating the latter concept with irregularity. At least four types of irregularity can be distinguished, depending first of all on whether the excess information is phonological or diacritic. In the most extreme case, that of suppletion, an irregular alternant (like a basic alternant) must be lexically listed in full; I will assume that, in contrast to a basic alternant, a suppletive alternant is listed with a specification of its environment. In the case of a morpholexical alternation, which shares with a suppletive alternation the property of being morpheme-specific but differs in that the alternants have phonological material in common, those characteristics of the morpholexical alternant that distinguish it from the basic alternant will have to be listed—again, I will assume, with a specification of environment. Thus, for example, the Japanese Causative suffix is *-sase-* after a vowel and *-ase-* after a consonant, where the *s*-zero alternation is morpheme-specific; if the latter alternant is basic, as I will claim, the lexical representation of the suffix will be  $\langle s \rangle ase \langle V \_ \rangle$ .<sup>1</sup>

Following Lakoff 1965 and Lightner 1968 (see also Andersen 1969 section 6.1), I will assume that an alternation that characterizes only a limited class of morphemes is to be captured by a “minor” rule that applies only to lexical items individually marked to undergo it. An irregular alternant that is the result of a minor rule will not need to be lexically listed; the excess information constituting the irregularity in question will take the form of a diacritic feature (“rule feature”) that triggers the minor rule. The excess information associated with failure to undergo a general (“major”) rule will be diacritic as well, consisting of an exception feature; in this case, the irregular alternant will typically be identical to the basic alternant, contrary to the prediction of the rule. In the context of such proposals, it is often suggested (see e.g. Kenstowicz and Kisseberth 1977:61,123) that there is a general tendency for speakers to eliminate diacritic information from lexical entries over time. I will assume that this tendency applies to all four types of irregularity we have just surveyed (rule feature, exception feature, suppletive alternant, morpholexical alternant)—that is, to any information in excess of a single phonological matrix, whether it be diacritic or phonological in content. Since an irregular alternant is one whose generation requires reference to such excess information, we expect irregular alternants, other things being equal, to be eliminated over time.<sup>2</sup>

Let us now return to the question of whether the three alternant types we identified above must be stored in memory or whether, on the other hand, they can be generated online, with an eye to the further question of whether they can be expected to display stability or instability across time. In considering this issue, I follow Marcus et al. (1992:vi,15–18) and especially Albright (2002:10–12) in taking the failure to retrieve irregular forms from memory as the central element in an account of the phenomenon of (over)regularization, whether in child language or in diachronic change. The basic intuition involved is that “morphological production involves competition between the retrieval of memorized forms and the creation of new ones by rule ... if existing ones are not learned, remembered, or accessed fast enough.” (Garrett 2008:128)

Basic alternants, to begin with, must be stored in memory—that is, lexically listed. As a consequence of

<sup>1</sup> In practice, shared segmental material (as opposed to subsegmental features) will normally be required to save a morpheme-specific alternation from suppletive status. For example, I will take the shared features [–high –low] (noted by a reviewer) of the two alternants *-e* and *-ro* of the Japanese Imperative suffix (see Table 3 below) to be coincidental rather than representing a significant generalization, with the result that the alternation is judged to be suppletive rather than morpholexical.

<sup>2</sup> As noted by a reviewer, the major factor capable of confounding this expectation is high token frequency.

Saussurean arbitrariness, if lexical retrieval should for any reason fail for a basic alternant, there will be no way to generate a substitute. There is thus no way for a basic alternant to show instability as the result of competition between listed and generated forms. Regular derived alternants, in contrast, may be stored in memory, but they need not be, because they may also be generated online by rule. The result of retrieval from memory and the result of online generation will be identical, though, so there is no way for a regular derived alternant to exhibit instability either.

Irregular alternants, finally, are similar to basic alternants in that they require lexical listing. More precisely, the excess information required by an irregular alternant, whether phonological or diacritic, must be noted in the lexical entry of the corresponding morpheme. We have hypothesized that this excess information tends to be eliminated from lexical entries over time. When lexical retrieval fails for the excess information required by an irregular alternant, however, the speaker is free to derive a substitute from the corresponding basic alternant by application of the relevant rule. This substitute, of course, will be an “overregularized” form, distinct from the result of retrieval from memory. As a result, if failure to retrieve the excess information required by an irregular alternant occurs with sufficient frequency, we will observe variation or fluctuation between that alternant and an innovative regular substitute, typically progressing over time to a state in which the innovative alternant has been generalized and the irregular alternant eliminated.

Of the three alternant types of (1), then, it is only the irregular alternants that are potentially unstable, and there is a single type of change that is predicted to occur if that instability is realized, namely the replacement of irregular alternants by regular derivatives. It is in this way that synchronic structure, in the form of basic forms and rules, restricts the pathways of change in morphophonology. Further, any hypothesis about synchronic structure, in the form of an analysis that induces the tripartition of forms seen in (1), makes falsifiable predictions about what changes in the system are possible, predictions which may then be confronted with the facts of change in progress. Conversely, change in progress can illuminate synchronic structure. In the change scenario sketched above, basic forms and regular derivatives will be stable, while irregular forms will be subject to replacement by regularly derived substitutes. Given a situation where, for a particular class of morphemes, alternants of type A are stable, but alternants of type C are subject to replacement by alternants of a type B that bears a constant relationship to type A, the indicated analysis will thus be that A is basic, C is irregular, and B is generated by rule from A. These inferential relations between synchrony and diachrony are summarized in Table 1 below.<sup>3</sup>

Postulated Analysis	Basic A Rule $A \rightarrow B$ Irregular C	Inferred Analysis
↓	Replacement of $A \sim C$ with $A \sim B$ (Extension of $A \sim B$ )	↑
Predicted Change		Observed Change

Table 1 Inferential Relations between Analysis and Change in Morphophonology

To this point, we have assumed the existence of all three of the alternant types of (1), designated A, B, and C in Table 1, and have seen that the diachronic prediction associated with the situation in which all three exist is extension of the regular alternation  $A \sim B$ . Several further analytical possibilities are naturally treated as degenerate cases of this state of affairs, cases in which one or more of the alternant types are lacking. If, to begin with, speakers postulate a basic (unmarked/default) vs. nonbasic (marked) distinction but no (major) rule to derive nonbasic forms, all nonbasic alternants will be irregular—that is, they will either be suppletive or morpholexical or will be the result of a minor rule. In this situation, then, only alternant types A and C will exist; the associated diachronic prediction will be leveling of the alternation  $A \sim C$  in favor of A, since this is what will result from the elimination of irregularity. If, on the other hand, all nonbasic alternants can be treated as regular, only alternant types A and B will exist, and, with no lexical irregularity, the alternation will be predicted to remain stable over time.

The two degenerate cases of (1) we have just seen have in common that the distinction between basic and nonbasic alternants is maintained, but the distinction between regular and irregular nonbasic alternants is neutralized. If, finally,

<sup>3</sup> In the generative tradition, early examples of such reasoning can be found in works such as Vennemann 1972, 1974a and Skousen 1975.

speakers decline to postulate even a basic-nonbasic distinction, all alternants will be listed, and all will be accorded the same status in the lexicon. In this case, there are two sub-possibilities, namely that all alternants are accompanied by environmental specification, and that none are (for examples of these two types of analysis as applied to Japanese verb inflection, see note 5 below). Let us refer to such analyses, which postulate neither rules nor (unique) basic forms, as “symmetric” treatments of alternations. Under a symmetric treatment of an alternation, while there will of course be lexical information for each morpheme in excess of a single phonological matrix, no alternant can be said to be irregular to the exclusion of any other. As in the case where all nonbasic alternants are regular, so that only alternant types A and B exist, I will take the absence of lexical irregularity in symmetric analyses of alternations to mean that such analyses predict diachronic stability for the alternations in question. More generally, I will assume that once a morphophonological analysis of an alternation is in force, the elimination of lexical irregularity is the only reason for that alternation to display instability.

We have seen that when we include degenerate cases, the classification of alternants in (1) can be interpreted as covering a total of four distinct analytic possibilities. Below, however, we will for the most part consider only analyses of the non-degenerate type, where all three alternant types of (1) exist.

We are now ready to consider such analyses in the context of Japanese verb inflection. Before we begin, however, there is one point concerning the nature of analogy that I believe bears clarification. It seems to me that considerable misunderstanding results from the notion that analogy is inherently or invariably sporadic or irregular, as suggested by a reviewer. Certainly a high percentage of the historical changes in alternations that are typically labeled “analogical” exhibit vacillation in one or both of two dimensions: leveling goes now in one direction, now in the other; or the alternation is extended in some instances but leveled in others (for an example of the latter situation, see Penny 2002:183). And if analogy is indeed inherently sporadic, it is but a short step from that fact to the conclusion that analogy has nothing to tell us about synchronic grammatical systems or the principles on which they are organized (Hale and Reiss 2008:242–243).

I take it to be the case, however, that there is a subtype of analogical change that, over long periods of time, is in fact more or less completely systematic. This is the type of change envisioned in the epigraphic quotation from Kiparsky 1972: the tracks for analogical change laid down by the rules of which Kiparsky speaks run in one direction only, and when the change processes they support have gone to completion, the resulting states of affairs are so regular that it is often unclear at first sight whether they are to be attributed to analogy or to sound change (Anttila 1989:98). Below, in addition to the Japanese case that we will examine in detail, we will refer briefly to instances of this type of change from Portuguese and Korean, and it is the existence of such examples that I will take to justify the stance that analogy does in fact have a good deal to tell us about synchronic grammar, both at the descriptive and at the explanatory level.

## 2 Description

After introducing the inflectional suffix alternations of the Japanese verb, the present section examines three observationally adequate analyses of those alternations, computing the predictions regarding potential instability and change for each and comparing the results with the facts of change in progress. Throughout, Japanese placenames (in isolation) and personal names are transcribed in a variant of Hepburn romanization (without macrons); all other Japanese words are transcribed in a variant of *kunreisiki* romanization, with length represented as gemination.

### 2.1 Introduction: The Alternating Suffixes of Japanese Verbal Inflection

The inflectional suffixes of the Japanese verb can be divided into two sets depending on the alternations they display and on whether or not they induce alternations in the stems they attach to. The members of the larger set of suffixes alternate between a vowel-initial form after consonant-final stems and a consonant-initial form (or zero) after vowel-final stems; in combination with such suffixes, stems (apart from *su-/si-/se-/s-* ‘do’ (and compounds thereof) and *ku-/ki-/ko-* ‘come’) are nonalternating. The members of the smaller set of suffixes (“*t*-suffixes”) are consonant-initial after all stems and undergo only a voicing alternation ( $t \sim d$ ), but they trigger a relatively complex set of stem alternations. Here we will be concerned only with the former set of suffixes, which will be seen to constitute a well-defined morphophonological subsystem. As an introduction to those suffixes and their alternations, Table 2

displays two representative examples, the first showing a consonant-zero alternation, the second a vowel-zero alternation (the stems are *mat-* ‘wait’ and *mi-* ‘see’; the Negative suffix is given for convenience in its Western Japanese form).

	C-Stem	V-stem
Indicative	mat- <u>u</u>	mi- <u>ru</u>
Negative	mat- <u>an</u>	mi- <u>n</u>

Table 2 Sample Alternations (Indicative, Negative)

Concerning the lexical representations of the suffixes of Table 2, there are at least the following possibilities: (a) C-stem alternants (*-u*, *-an*) are basic; (b) V-stem alternants (*-ru*, *-n*) are basic; (c) longer alternants (*-ru*, *-an*) are basic; (d) there are no unique basic forms; all alternants are lexically listed. Because the segment-zero alternations are not fully predictable given the shorter alternants, choice (c) is the one that maximizes phonological predictability, as we will see in more detail below.

For Tokyo and many other locations, the full set of suffixes that show V-initial forms after C-stems and C-initial forms after V-stems consists of the nine items of Table 3 (the Passive, Causative, and Potential suffixes create verb stems; the Eastern Japanese variant *-(a)na-* of the Negative suffix creates adjective stems).<sup>4</sup> There are also many locations that lack certain of the suffixes of Table 3 or display additional suffixes with parallel alternations.

	Suffix	C-stem	V-stem	Alternation
1	Indicative	mat-u	mi-ru	∅ ~ r
2	Provisional	mat-eba	mi-reba	∅ ~ r
3	Passive	mat-are-	mi-rare-	∅ ~ r
4	Hortative	mat-oo	mi-yoo	∅ ~ y
5	Causative	mat-ase-	mi-sase-	∅ ~ s
6	Infinitive	mat-i	mi-∅	i ~ ∅
7	Negative	mat-an(a-)	mi-n(a-)	a ~ ∅
8	Imperative	mat-e	mi-ro	e ~ ro
9	Potential	mat-e-	(mi-rare-)	(e ~ rare)

Table 3 Full Set of Alternating Suffixes (Tokyo and other locations)

As indicated in the last column of Table 3, suffixes 1–3 show the *r*-zero alternation that we saw in Table 2, suffixes 4–5 show alternations of other consonants with zero, suffixes 6–7 show vowel-zero alternations (including the *a*-zero alternation seen in Table 2), and suffix 8 shows an apparently suppletive alternation. The alternation of suffix 9 is distinct from the other alternations, I will claim, in resulting from syncretism: in conservative dialects, the Potential suffix has only the C-stem allomorph *-e-*, and the corresponding V-stem form *-rare-* results from the fact that *-(r)are-* is the default realization of a morphosyntactic feature that is shared by the categories Passive and Potential, as we will see in more detail below.

Published analyses of alternations 1–8 illustrate all four of the proposals for lexical representations that we noted above for the two suffixes of Table 2: C-stem alternants, V-stem alternants, and longer alternants as basic, and symmetrical listing of all alternants.<sup>5</sup> It is clear, then, even before considering the question of what phonological rule (if any) speakers postulate to relate basic and nonbasic alternants, that this is a system of alternations with multiple observationally adequate analyses. In section 2.2, we will examine three such analyses, involving respectively the

<sup>4</sup> A number of the suffixes of Table 3 are like the Negative in displaying regional variation; for example, the V-stem Imperative suffix is *-yo* or *-i* in most of Western Japan. Variation in the Negative suffix is recorded in this and subsequent tables because the Western Japanese variant *-an* will figure prominently in the story of change in progress to be recounted below.

<sup>5</sup> For C-stem alternants as basic, see de Chene 1985, 1987; for V-stem alternants as basic, see McCawley 1968: 93ff.; for longer alternants as basic, see e.g. Kuroda 1960 or Chew 1973; for symmetrical listing, see Bloch 1946/1970:24 (with environments specified) or Ito and Mester 2004 (with alternant choice determined by ranked constraints).

postulation as basic of C-stem alternants, V-stem alternants, and longer alternants.

## 2.2 Three Observationally Adequate Analyses and their Predictions

Below, the three analyses to be compared in this section will be introduced, each defined by a set of basic forms and a phonological rule. We will then compare the partition of suffix alternants into basic, regular derived, and irregular types induced by each of the three analyses and note, using one analysis as an example, how that partition determines the lexical representations of the alternating suffixes. Finally, we will examine the predictions of each analysis regarding what changes in the system are to be expected if regularization occurs, setting the stage for the confrontation of those predictions with the facts of ongoing change in section 2.3.

If the C-stem suffix alternants of Table 3 are basic, the rule that provides maximal predictability of nonbasic (i.e. V-stem) alternants will be one that inserts *r* intervocalically at verb stem boundary; such a rule will account correctly for the V-stem alternants of the first three categories of Table 3.<sup>6</sup> I will thus assume that the analysis that takes C-stem alternants as basic incorporates such an *r*-epenthesis rule; call this analysis “Analysis A”.<sup>7</sup>

### (2) Analysis A

- a. Basic Alternants: C-stem suffixes
- b. Rule:  $\emptyset \rightarrow r / V_{vb} [ \_ V ]$

In rule (2b), brackets mark morpheme boundary, and the label “Vb” indicates that the leftward morpheme bears the lexical category specification [Verb]. The precise set of environments in which the *r*-Epenthesis rule (2b) applies will be discussed in section 2.4.

If the V-stem alternants of Table 3 are basic, on the other hand, the rule that provides maximal predictability of nonbasic (i.e. C-stem) alternants will be one that deletes the second of two consonants at verb stem boundary; such a rule will account correctly for the C-stem alternants of the first five categories of Table 3. I will thus assume that the analysis that takes V-stem alternants as basic incorporates such a C-deletion rule; call this “Analysis B”.<sup>8</sup>

### (3) Analysis B

- a. Basic Alternants: V-stem suffixes
- b. Rule:  $C [+cont] \rightarrow \emptyset / C_{vb} [ \_ ]$

The consonant-deletion rule of Analysis B has been written as applying only to nonstops because the *t* of *t*-initial suffixes is unaffected, as illustrated by the examples /kar + ta/, /kaw + ta/ → [katta] of note 8. Alternatively, *t*-initial suffixes can be analyzed as added not to stems but to Infinitive forms, recapitulating their historical development (for a recent defense of this analysis, see Oshima 2014). In this case, the restriction to nonstops will be unnecessary, since clusters with second member *t* will result from a rule of syncope that can be ordered after Analysis B’s C-deletion rule, counterfeeding it.

If, finally, the longer suffix alternant is basic in each case, the rule that provides maximal predictability of nonbasic (i.e. shorter) alternants will be one that deletes both the second of two consonants and the second of two vowels at verb stem boundary; such a rule will account correctly for the shorter alternants of the first seven categories of Table 3. I

<sup>6</sup> The alternative of a rule deleting the second of two vowels at verb stem boundary would account for the alternations of only two suffixes, the Infinitive and Negative (Categories 6,7), and would create zero V-stem suffix allomorphs not only in the Infinitive, but in the Indicative and Imperative (as well as in the Potential, were that suffix to lose its restriction to C-stems).

<sup>7</sup> It has been proposed (Labrune 2014) that all Japanese *r* are ultimately epenthetic, arising from insertion intervocalically in Proto-Japanese. Two problems this thesis must address are (1) as a sound change, an “excrecent” segment is plausible only insofar as it can be explained as an articulatory consequence (as in *nr > ndr*) or an auditory interpretation (as in *ia > ija*) of the transition between adjacent segments; (2) many Japanese *r* are widely taken (Martin 1966, Whitman 1985, Whitman 2012) to be cognate with Korean *l*, with even scholars who are skeptical of a genetic relationship between the two languages admitting “reliable” cognate candidates showing this correspondence (Vovin 2010:238).

<sup>8</sup> While it is sometimes suggested (Hasegawa 1999:63, Miyagawa 1999:236) that deletion of the second of two successive consonants represents a general principle of Japanese phonology, the usual treatment of clusters is in fact assimilation of the first C to the second: /kar + ta/ ‘clip’ + Perfect → [katta], /kaw + ta/ ‘buy’ + Perfect → [katta], /but + sak-/ ‘beat’ + ‘tear’ → [bussak-] ‘tear violently’, /hik + sage-/ ‘draw’ + ‘hang’ → [hissag-] ‘carry’; /it + kai/ ‘one’ + ‘time’ → [ikkai], /it + sai/ ‘one’ + ‘year of age’ → [issai], /it + pai/ ‘one’ + ‘glass’ → [ippai].

will thus assume that the analysis that takes longer alternants as basic incorporates such a rule; call this “Analysis C” (I use the feature [syllabic] (Chomsky and Halle 1968: 354) to distinguish consonants and vowels in rule (4b)).

(4) Analysis C

- a. Basic Alternants: longer suffixes
- b. Rule:  $[\text{asyll}] \rightarrow \emptyset / [\text{asyll}]_{\text{vb}} [ \_ ]$

Under Analysis C, if *t*-initial suffixes are not to be treated as exceptions to deletion, they will have to be added to Infinitive forms rather than to stems, as suggested above in connection with Analysis B.

In introducing Analyses A, B, and C, we have already noted which suffix alternants will be basic and which will be regularly derived under each analysis; the irregular suffixes are the complement within the set of nonbasic suffixes of the regularly derived set. Table 4 below displays the tripartition of suffix alternants corresponding to each of the three analyses for categories 1–8 of Table 3, with basic suffixes unmarked in each case, regular suffixes underlined, and irregular suffixes shaded.

	Suffix	Analysis A		Analysis B		Analysis C	
		C-stem	V-stem	C-stem	V-stem	C-stem	V-stem
1	Indicative	mat-u	mi- <u>ru</u>	mat- <u>u</u>	mi-ru	mat- <u>u</u>	mi-ru
2	Provisional	mat-eba	mi- <u>reba</u>	mat- <u>eba</u>	mi-reba	mat- <u>eba</u>	mi-reba
3	Passive	mat-are-	mi- <u>rare-</u>	mat- <u>are-</u>	mi-rare-	mat- <u>are-</u>	mi-rare-
4	Hortative	mat-oo	mi-yoo	mat- <u>oo</u>	mi-yoo	mat- <u>oo</u>	mi-yoo
5	Causative	mat-ase-	mi-sase-	mat- <u>ase-</u>	mi-sase-	mat- <u>ase-</u>	mi-sase-
6	Infinitive	mat-i	mi- $\emptyset$	mat-i	mi- $\emptyset$	mat-i	mi- $\emptyset$
7	Negative	mat-an(a-)	mi-n(a-)	mat-an(a-)	mi-n(a-)	mat-an(a-)	mi-n( <u>a-</u> )
8	Imperative	mat-e	mi-ro	mat-e	mi-ro	mat-e	mi-ro

Table 4 Basic, Regular (\_\_\_), and Irregular (■) Alternants according to Analyses A, B, C

It is immediately clear from Table 4 that, as we have already observed, Analysis C is the most successful of the three analyses in terms of maximizing phonological predictability of suffix alternants, given that it treats seven of eight suffix alternations as regular. Analysis B, with five alternations treated as regular, is the next most successful analysis in this regard, and Analysis A, with only three alternations treated as regular, is the least successful.

The lexical representations of the eight suffixes of Table 4 under a given analysis can, to a first approximation, be read off the data of the table. Basic alternants will be lexically listed; irregular alternants (more precisely, the excess phonological or diacritic information that their generation requires) will be listed as well, along with a specification of environment. Concerning regular derived alternants, while we noted in section 1.2 that they may be stored in memory, I will assume here, following standard generative practice, that they are not lexically listed. I will further assume that insertion of inflectional elements into syntactic terminal nodes is governed by the principles (5) (Halle and Marantz 1993:123–124, Halle 1997:428).

- (5) a. Inflectional elements are inserted into terminal nodes based on feature matching.
- b. The features of an inflectional element must constitute a subset of the features of the node into which that element is inserted.
- c. An element matching in more features takes precedence over one matching in fewer.
- d. Among elements matching in the same features, an element with a more complex environment takes precedence over an element with a simpler environment.

Given the assumption that irregular (but not basic) alternants are listed with their environments, (5d) guarantees that the irregular and basic alternants corresponding to a single category will constitute a disjunctively ordered pair, with the basic alternant as the elsewhere case.



For the most part, the question of the morphosyntactic features realized by the inflectional suffixes of Tables 3–4 is orthogonal to the morphophonological issues we are concerned with here. An exception, however, arises from the syncretism that we have claimed obtains between the Passive and the Potential in the V-stem paradigm, for which we now need a concrete account. Syntactically, the Passive and the Potential are both unaccusative constructions: they lack an argument structure position for the transitive subject (Agent or Experiencer), with the result that it is the Theme argument that moves to the surface subject position and receives nominative case (although in the modern language a transitive case-marking pattern is also an option for the Potential). The two constructions were marked by the same suffix *-(r)are-* for all verbs from roughly the ninth through the fifteenth centuries (see e.g. Yamaguchi and Akimoto 2001:842-843,846-84); this situation was disturbed starting in the late fifteenth century by the gradual emergence of a dedicated Potential suffix *-e-* for C-stems, which took until about 1800 to become fully general (Aoki 2010:3–39).

I will assume, then, that the Passive and Potential share a morphosyntactic feature specification [+UA] (for “unaccusative”) and are differentiated in that the Potential is [+UA +Pot] while the Passive is [+UA –Pot]. As shown in (6) below, the suffix *-are-* (under Analysis A) or *-rare-* (under Analyses B and C) will be the default realization of the feature specification [+UA], preempted by *-e-* only when (a) [+UA] is accompanied by [+Pot] and (b) the environment of insertion is post-consonantal.

- (6) a. [+UA +Pot] ↔ e/C \_\_  
 b. [+UA] ↔ (r)are

We may note that the alternative account under which the Potential alternation *-e- ~ -rare-* is due (like, for example, the Imperative alternation *-e ~ -ro*) to lexically listed allomorphy would require claiming that there are two semantically parallel suffixes that coincidentally have the same form *-(r)are-*.

For the remaining seven suffixes of Table 3, I will assume for convenience a one-to-one correspondence between morphosyntactic features and inflectional categories. (7) below, incorporating (6), shows lexical representations for the nine suffixes of Table 3 under Analysis A; for reasons of space I leave calculation of parallel representations for Analyses B and C to the reader.

(7) Lexical Representations of the Suffixes of Table 3 under Analysis A

- |                 |   |  |
|-----------------|---|--|
| 1 [+Ind] ↔ u    | 4 [+Hort] ↔ <y>oo</V__>                                   | 7 [+Neg] ↔ >a< n(a)</V__>                          |
| 2 [+Prov] ↔ eba | 5 [+Caus] ↔ <s>ase</V__>                                  | 8 [+Imp] ↔ $\begin{cases} ro/V\_ \\ e \end{cases}$ |
| 3 [+UA] ↔ are   | 6 [+Inf] ↔ $\begin{cases} \emptyset/V\_ \\ i \end{cases}$ | 9 [+UA +Pot] ↔ e/C __                              |

Representations for categories 4 and 5 employ angled brackets, whereby the disjunction of XYZ and Y is represented as <X>Y<Z>. This allows us to write the material shared by the two alternants of the suffix, *oo* in the Hortative and *ase* in the Causative, only once, capturing the fact that the recurrence of that material is not coincidental. The possibility of using angled brackets for this purpose in categories 4 and 5, however, depends on the longer alternant occurring in the more complex environment; when the reverse is true, as in category 7, that convention is inapplicable. In order to allow in the latter case as well a representation of the suffix in which the material shared by the two alternants appears only once, I define the “reverse angled bracket” notation >X<Y <Z> (Y non-null) as abbreviating the disjunction of YZ and XY.

In section 1.2, we noted that our model of the retrieval and generation of basic, regular derived, and irregular morpheme alternants had the consequence that only the irregular alternants are potentially unstable, and that the only type of change predicted to occur if that instability is realized is replacement of irregular alternants by regularly derived substitutes. It is thus straightforward to compute the changes predicted by Analyses A, B, and C from the data of Table 4: the predicted replacement for each irregular (shaded) suffix alternant under a given analysis is the result of substituting the basic form of the suffix for the irregular alternant and applying the relevant rule.

To illustrate, consider the Imperative, a category for which all three analyses take one of the existing suffix alternants to be irregular. According to Analysis A, V-stem *-ro* is irregular; if retrieval fails for that irregular alternant, the basic alternant *-e* will be added to V-stems as well as C-stems, and the *r*-Epenthesis rule (2b) will apply to that alternant in the context of a V-final stem to produce an innovative suffix alternant *-re*. This derivation is displayed in (8).

- (8) Predicted innovative V-stem Imperative under Analysis A if retrieval of irregular *-ro* fails

Input [[mi<sub>vb</sub>][e]]  
 (2b) [[mi<sub>vb</sub>][re]]  
 Output [mire]

In the same way, C-stem Imperative *-e* is irregular according to Analyses B and C; if retrieval fails for that alternant, the basic alternant *-ro* (in much of Western Japan, *-yo* (note 4)) will be added to C-stems as well as V-stems, and the C-deletion rule (3b) of Analysis B or the more general deletion rule (4b) of Analysis C will apply to that alternant in the context of a C-final stem to produce an innovative suffix alternant *-o*. This derivation is displayed in (9).

- (9) Predicted innovative C-stem Imperative under Analyses B/C if retrieval of irregular *-e* fails

Input [[mat<sub>vb</sub>][ro]]  
 (3b)/(4b) [[mat<sub>vb</sub>][o]]  
 Output [mato]

More generally, for Analysis A, each regularly derived substitute for an irregular V-stem suffix alternant will consist of the corresponding C-stem alternant preceded by *r*; for Analysis B, each regularly derived substitute for an irregular C-stem suffix alternant will consist of the corresponding V-stem alternant minus its first consonant; and for Analysis C, each regularly derived substitute for an irregular short suffix alternant will consist of the corresponding longer alternant minus its first vowel or consonant. The predictions of the three analyses are summarized in Table 5, where the form given for the Negative suffix under Analysis B assumes that *n* (unlike *t*) is subject to deletion.

	Suffix	Analysis A		Analysis B		Analysis C	
		C-stem	V-stem	C-stem	V-stem	C-stem	V-stem
4	Hortative		mi-roo				
5	Causative		mi-rase-				
6	Infinitive		mi-ri	mat-			
7	Negative		mi-ran(a-)	mat-(a-)			
8	Imperative		mi-re	mat-o		mat-o	

Table 5 Predicted Changes in Case of Regularization

The three analyses, finally, each make one further prediction concerning what should be expected if Potential *-e-* is generalized to V-stems—that is, if the lexical condition that restricts that suffix to C-stems is lost. For Analysis A, that prediction is that we should see the emergence of a V-stem Potential alternant *-re-*. Analysis B, in contrast, includes no principle that would alter a sequence of vowels at verb stem boundary and thus predicts that in case of generalization, Potential *-e-* should appear unaltered after vowels as it does after consonants. Analysis C, finally, involves a rule deleting the second of two vowels at verb stem boundary and thus predicts that the V-stem Potential suffix alternant should be zero. In section 2.3, we confront these predictions and those of Table 5 with the facts of change in progress.

### 2.3 Change in Progress: Innovative *r*-Suffixes

With the publication, by the National Language Research Institute (known in English since October 2009 as the National Institute for Japanese Language and Linguistics (NINJAL)), of the six-volume, 350-map *Grammar Atlas of Japanese Dialects* (GAJ = Kokuritu Kokugo Kenkyuuzyo 1989–2006), data concerning verbal inflection has become

available for 807 locations in Japan, 41 from the Ryukyu dialect area and 766 from the Japanese dialect area proper; here we will consider only the latter.<sup>9</sup> Under the idealizing assumption that they have developed independently of each other, a set of related dialects like that recorded in the GAJ constitutes a kind of linguistic laboratory or natural experiment. More specifically, the GAJ in principle gives us 766 trials of an experiment in morphological and morphophonological development dating from the period of Proto-Japanese unity to the time of the survey, although there is evidence that the innovations with which we will be concerned date from no earlier than the year 1600 (see de Chene 2010a, section 4.1). The idealization of independent development, of course, is likely to be imperfectly realized when particular innovative forms appear in geographically contiguous areas, making transmission a possible factor in their spread.<sup>10</sup>

Data for the GAJ was gathered in interviews between local consultants (men most of whom were born before 1925 and were thus in their middle fifties and above at the time of the survey in 1979–1982) and survey workers, typically dialect specialists at local universities; each survey item asked consultants to provide a local equivalent for an expression presented in standard Japanese. The GAJ is thus fundamentally based on self-reporting and shares the potential drawbacks of that methodology (see Labov 2001:194ff. and references cited there), although the interview format did allow survey workers to suggest forms not volunteered by consultants. In many cases, furthermore, the GAJ does not add significantly to what was already known from the descriptive dialect literature (e.g. Iitoyo et al. 1982–1986 and references cited there) about the occurrence of particular forms or constructions in particular locations. Its great virtues, apart from the visual convenience of its maps, are that it is an exhaustive, uniform nationwide survey and that its results are quantifiable.

The verdict of the GAJ data is unambiguous. First, all predictions of Analysis A noted above are confirmed, both with regard to which suffix alternants should be stable and which unstable and with regard to the changes to be expected if instability is realized. Specifically, the five *r*-initial V-stem suffixes listed as predictions of Analysis A in Table 5 are widely attested,<sup>11</sup> as is the V-stem Potential suffix *-re-* that we noted is predicted by Analysis A in case of loss of the condition restricting Potential *-e-* to C-stems. Table 6 shows data from the GAJ on the frequency of those six innovative suffixes, identified in the leftmost column by category number, with one map chosen to illustrate each suffix.<sup>12</sup> The parenthesized items in the category 7 row refer to occurrences of *miyan* (19 from the Kii peninsula and 2 from southern Kyushu) that represent a regular phonological development of *miran*, as we will see in more detail below.

	Survey Item	Stem	Volume (Map)	Innovative Form	Innovative Locations	Reference Set	Innovative Percentage
4	ne-yoo	ne- ‘sleep’	3(108)	ne- <u>roo</u> etc.	141	505	28%
5	ko-sase-ru	ko- ‘come’	3(120)	ko- <u>rased</u> -ru etc.	267	747	36%
6	mi wa sinai	mi- ‘see’	4(161)	mi- <u>ry</u> -aa-sen etc.	109	504	22%
7	mi-na-i	mi- ‘see’	2 (74)	mi- <u>ran</u> (mi- <u>yan</u> )	122 (21)	738	17% (3%)
8	mi-ro	mi- ‘see’	2 (86)	mi- <u>re</u>	174	747	23%
9	kiru koto ga dekiru	ki- ‘wear’	4(176)	ki- <u>re</u> -ru	283	608	47%

Table 6 Frequency of Innovative *r*-Suffixes for Representative GAJ Survey Items

The percentages in the last column of Table 6, which give a rough idea of the strength of each innovative suffix, are the ratios of the figures in the two previous columns, where the reference set is the set of locations deemed potentially

<sup>9</sup> Ryukyu varieties, typically grouped into five languages (Amami, Okinawan, Miyako, Yaeyama, Yonaguni) but divided by Grimes (2000:649–650) into eleven, are by no means irrelevant to our topic, but their divergence from each other and from mainland Japanese puts them outside the scope of the present paper.

<sup>10</sup> It goes without saying that, as a reviewer notes, sociolinguistic research documenting in detail the course of ongoing change is also extremely important. Of the six “innovative *r*-suffixes” of Tables 6 and 7 below, such research is most plentiful for Potential *-re-* in Tokyo (see e.g. Matsuda 1993 and references cited there), but can be observed to some extent for other suffixes as well (see e.g. Jinnouchi 1996:56–57 on Negative *-ran* in Hakata (Fukuoka city)).

<sup>11</sup> The innovative Negative suffix occurs only in the Western Japanese form *-ran*; this is perhaps because speakers with irregular Eastern *-na-* identify it with the homophonous negative existential adjective, an identification that would be rendered impossible by innovative *\*-rana-*.

<sup>12</sup> Apart from irregular *su-* and *ku-* (section 2.1) and verbs that vary by region between C-stem and V-stem inflection, the V-stem survey items, by category, are (4) *ne-*, *oki-* ‘arise’, *ake-* ‘open (tr.)’; (5) *ake-*; (6) *mi-*; (7) *mi-*, *ne-*, *oki-*, *ake-*; (8) *mi-*, *oki-*, *ake-*; (9) *oki-*, *ki-* (4 items).

capable of showing the innovative suffix. The reference set was calculated, for each survey item in question, by subtracting from 766 (a) the number of nonresponding locations for the item in question and (b) either the number of locations that systematically lack the relevant basic alternant (i.e. the corresponding C-stem suffix) or, when that figure was unavailable due to a lack of relevant survey items, the number of locations whose responses to the survey item in question involved a construction or form distinct from that of the standard Japanese model.<sup>13</sup>

The figures of Table 6 are useful as a means of comparing the relative strength of the six innovative suffixes because in computing them, two factors that would otherwise have skewed the results were controlled for. The first of these, as already indicated, was variation by category in the number of locations that could in principle have displayed the relevant innovative suffix. The second is variation by category in the number of relevant GAJ survey items: for the Infinitive, whose many uses were not systematically investigated, there are only three relevant items (Maps 21 and 161–162), while for the Potential, there are seven (Maps 175–179 and 184–185). More revealing than the figures of Table 6 with regard to the overall strength of the evidence for Analysis A, however, is the percentage of the 766 GAJ locations that display one or more of the six innovative *r*-suffixes, showing that Analysis A is in force at that location. Even excluding Potential *-re-* (Category 9), the most common member of the set, over 60% of all locations display at least one innovative suffix, while, when the Potential is included, the percentage rises to over 80%. The number of locations and percentage of the whole represented thereby are presented in Table 7 for the six innovative *r*-suffixes individually and for the disjunctions of suffixes 4–8 and 4–9.<sup>14</sup>

	Suffix	Number	Percent
4	Hort <i>-roo</i>	159	20.8
5	Caus <i>-rase-</i>	273	35.6
6	Inf <i>-ri</i>	129	16.8
7	Neg <i>-ran</i>	142	18.5
8	Imp <i>-re</i>	205	26.8
9	Pot <i>-re-</i>	419	54.7
Any of 4–8		472	61.6
Any of 4–9		620	80.9

Table 7 Frequency of Innovative *r*-Suffixes over All GAJ Survey Items

We may note that the significantly higher frequency of Potential *-re-* compared to the other innovative *r*-suffixes plausibly reflects the distinct mechanism of change we have postulated for the Potential—specifically, loss of a contextual restriction as opposed to loss of irregular allomorphy (i.e. excess phonological or diacritic information).

The evidence supporting Analysis A, then, is extremely widespread. Further, there are a number of reasons why even the figures of Table 7 understate the extent of that evidence. One reason, already referred to, is the fact that the Infinitive form was not a focus of investigation for the GAJ; that form occurs in the survey items that do contain it more or less by accident. Systematic investigation would almost certainly have revealed many more instances of innovative Infinitive *-ri* than are recorded in Table 7, as can be inferred from examples of that suffix that show up in the GAJ data for survey items other than the three on which the Table 7 figure was based.

For example, Tokyo Japanese displays a number of weak imperative constructions that involve the Infinitive (see Martin 1975:964–965), and it is clear that the same is true of many regional varieties. Only in a few areas, however, is such a construction salient enough as a substitute for the Imperative to emerge as a response to an Imperative survey question; when it does, we sometimes see innovative Infinitive *-ri*, as in forms like *okirin* (~ *okin*) ‘get up!’ from location 656816 on the Atsumi Peninsula in Aichi Prefecture (for discussion of this construction, (o)+Infinitive+n(*sai*), in the region in question, see Yoshikawa and Yamaguchi 1972:152–154). Location 636625 in Yamaguchi Prefecture shows parallel forms. The Infinitive-based potential expressions of Kyushu also reveal innovative Infinitive *-ri* in a

<sup>13</sup> By category number, the two figures in question are as follows (NR = nonresponding, BL = base lacking, DC = distinct construction): Cat 4: NR 1, BL 260 (Map 109); Cat 5: NR 19, BL 0; Cat 6: NR 55, DC 207; Cat 7: NR 1, BL 27 (Map 80); Cat 8: NR 8, DC 11; Cat 9: NR 9, BL 150 (Maps 173–174, 181–183) [one location in both sets].

<sup>14</sup> The location-by-location tabulation of GAJ data that underlies the figures of Table 7 is available from the author on request.

number of cases; sample forms, based on *ki-* ‘wear, put on’ are *kiri-kiru* ( $\sim$  *ki-kiru*) (738398, Kumamoto Prefecture) and *kiri-ga-naru* (839421, Kagoshima Prefecture) ‘can put on’ (Map 175). None of these occurrences of *-ri* are reflected in Table 7. It is safe to say that comprehensive data on Infinitive-based constructions would raise the Table 7 figure for that suffix considerably.

A second reason that the figures of Table 7 understate the extent of the evidence for Analysis A is that there are areas in which, to a first approximation, intervocalic *r* is subject to deletion, so that when irregular V-stem suffixes are eliminated, the expected *r*-initial substitutes show up without their *r*. In particular, there are 22 locations, for the most part overlapping with the 21 recorded in Table 6 as displaying *miyan*, that show generalization to V-stems of C-stem Negative *-an* in accordance with the predictions of Analysis A, but no evidence of *r*-Epenthesis for that suffix, which appears as *-yan*.<sup>15</sup> The fact that *-ran* occurs in neighboring areas and that 20 of the 22 locations show evidence for a process converting intervocalic *r* to zero or *y* in the form of *r*-less variants of Passive/Potential *-rare-* and/or Provisional *-reba* makes it plausible to suppose that, as suggested above, *-yan* represents a regular phonological development of *-ran*. Two interpretations of this development seem possible: *r* may either change directly to *y* or delete completely, with *y* then representing phonologization of the transition between the front stem vowel and the vowel of the suffix.<sup>16</sup> Since none of the 22 locations in question show any innovative *r*-suffix from categories 4–8 (although twelve of them have Potential *-re-*), admitting them as evidence for Analysis A would raise the Table 7 figures not only for the Negative but also for Categories 4–8 by slightly under three percentage points.

A third and final reason that not all of the evidence for Analysis A appears in Table 7 is that the set of innovative *r*-suffixes observed across the range of Japanese dialects is by no means limited to the six that we have examined. Let us look briefly at several additional examples of *r*-suffixes whose historical origin and development are relatively clear.

Among the many honorific auxiliary verbs (morphologically speaking, inflectional stem-forming suffixes) displayed by 17th-century Kyoto/Osaka Japanese (“Kamigata-go”) is one that takes the form *-assyar-*  $\sim$  *-asyar-* after C-stems and *-sassyar-*  $\sim$  *-sasyar-* after V-stems (Matsumura 1971:972, Yamaguchi and Akimoto 2001:904). This suffix thus in principle shows the same *s*  $\sim$   $\emptyset$  alternation as the Causative suffix, and is in fact composed historically of Causative *-(s)ase-* plus the old Indicative form *-raru* of the Passive suffix (Matsumura 1971:264) reanalyzed as *-rar-u*, with the subsequent development *-(s)aserar-*  $>$  *-(s)asear-*  $>$  *-(s)asyaar-*  $>$  *-(s)assyar-*  $\sim$  *-(s)asyar-*. In several areas, however, as shown by the data for GAJ Map 291, the V-stem alternant *-sassyar-* has been or is being replaced by *-rassyar-*, parallel to the replacement of Causative *-sase-* with *-rase-*.<sup>17</sup> Similarly, the honorific auxiliary *-(s)ahar-* of the Shonai region of Yamagata Prefecture, identified with *-(s)asyar-* by the *Nihon Kokugo Daiziten* (NKD = Nihon Kokugo Daiziten Dainihan Hensyuu Iinkai/Syooagakukan Kokugo Ziten Hensyuubu 2000–2002), displays variation between conservative V-stem *-sahar-* and innovative V-stem *-rahar-* (Kokuritu Kokugo Kenkyuuzyo 1953:226).

An example where the original alternation is of the form V  $\sim$   $\emptyset$  rather than C  $\sim$   $\emptyset$  is provided by the Hortative suffix *-(a)zu* of Nagano, Yamanashi, and Shizuoka Prefectures (GAJ Maps 106–111), which goes back to Muromachi period *-(a)uzu* and Kamakura period *-(a)muzu* (Matsumura 1971:968,962); the historical development will have been *kak-amuzu*  $>$  *kakaizu*  $>$  *kakauzu*  $>$  *kakaazu*  $>$  *kakazu* ‘let’s write’ and *oki-muzu*  $>$  *okiizu*  $>$  *okiuzu*  $>$  *okiizu*  $>$  *okizu* ‘let’s get up’. Variation between conservative V-stem *-zu* and innovative V-stem *-razu* is displayed by all five of the candidate stems (Maps 106–108, 110–111) at GAJ location 567318, which otherwise lacks innovative *r*-suffixes, and parallel innovations, with variant forms of the suffix, are observed at 660401 (C-stem *-ada*, V-stem *-da*  $\sim$  *-rada*), 661368 (C-stem *-aza*, V-stem *-za*  $\sim$  *-raza*), and 662454 and 663432 (C-stem *-asu*, V-stem *-su*  $\sim$  *-rasu*).

In our final example, the “spontaneous activity” suffix *-(r)asar-*, best known as characteristic of Hokkaido and northern Tohoku (and observable over that range as a potential suffix in GAJ Map 181), the replacement of an *a*  $\sim$   $\emptyset$  alternation with *r*  $\sim$   $\emptyset$  has evidently gone to completion in many locations; Sasaki and Yamazaki (2006), for example, cite no V-stem form other than *-rasar-*. Examples in the dialect section of the entry for the auxiliary *saru* in the NKD, however, provide evidence that *-(r)asar-* descends from a suffix of the form *-(a)sar-* that is attested in

<sup>15</sup> Two points showing *miyan* do not belong to this set of 22 because they have *-ran* for a V-stem other than *mi-* (*okiran* at 750472, *neran* at 833350); three points of the 22 do not show *miyan*, their *-yan* being restricted to V-stems other than *mi-* (*neyan* at 750046 and 750391, *neyan* and *akeyan* at 658735). The 19 points that the two sets share are 654612, 655476, 655506, 656387, 656423, 656514, 656673, 657332, 657379, 657543, 657685, 658088, 658438, 659044, 659147, 659501, 750408, 751369, 924994.

<sup>16</sup> The former interpretation is suggested by the change *ru*  $>$  *i* (*arusi*  $>$  *aisi* ‘exists and’ (Mie Prefecture, Umegaki 1962:114), *kuruma*  $>$  *kuima* ‘vehicle’ (Kagoshima Prefecture, Kyuusyu Hoogen Gakkai 1991:236), where *r* appears to be the source of palatality in the output vowel.

<sup>17</sup> Points 469792, 562484, 563227; 653922, 654853, 655824, 662070; 722975, 732095, 737096.

multiple meanings starting in the 16th century. As in our other examples, in other words, the conservative V-stem alternant (here, *-sar-*) was irregular; its loss resulted in the basic alternant (*-asar-*) being added to V-stems as well as C-stems, triggering *r*-Epenthesis and the emergence of a regular derived V-stem alternant (*-rasar-*).

It is clear, then, that evidence for Analysis A is not only extremely widespread but is open-ended, in that it is not limited to the six innovative suffixes of Tables 6 and 7. In contrast, neither the GAJ nor the dialectological literature provides any evidence I am aware of for Analyses B or C.<sup>18</sup> This unanimity suggests that Analysis A is the unique descriptively adequate analysis of the system of alternations that we saw first in Table 3—the only analysis of that system, in other words, that is permitted by universal principles.

There is one final issue that should be addressed before we leave the topic of the innovative forms, typified by the examples of Table 6, that we have discussed in this section. While we have interpreted those forms as supporting an analysis of Japanese verbal suffix alternations according to which C-stem suffixes are basic and regular V-stem suffixes are derived by *r*-Epenthesis, other researchers have proposed quite different interpretations of them. Space limitations preclude a full discussion of those interpretations here (see de Chene 2010a, section 3 for more detail), but in closing this section, I will briefly take up two of them and attempt to identify their shortcomings.

One interpretation of the innovative forms we have seen, codified in the phrase by which they are referred to in the native dialectological literature (*itidan katuyoo no ra-gyoo godanka*), is that those forms represent assimilation of vowel-stem conjugation to the conjugational pattern of *r*-stems. The *r* of innovative forms like Imperative *mi-re* and Hortative *mi-roo*, however, belongs to the suffix rather than to the stem, as shown by the persistence of unambiguously vowel-stem forms in the paradigm of verbs showing the innovations in question. Interpreting those innovations as the result of influence from *r*-stem forms thus requires that we countenance mutual influence of word forms based on surface similarity without regard for internal morphological structure, as formalized, for example, in the unsegmented proportions that are common in the dialectological literature. Since such a theory of possible “analogical” influences is less restrictive than one requiring that morpheme boundaries be respected, it is in principle incumbent on those who would propose it to show that the sacrifice of restrictiveness is necessary to achieve empirical coverage.

In addition to this theoretical argument, however, there is in the present case concrete evidence disconfirming the hypothesis that innovative *r*-suffixes are due to influence from *r*-stems. This is because that hypothesis predicts that V-stems should assimilate to *r*-stem conjugation in forms involving *t*-suffixes (Gerundive *-te*, Perfective *-ta*, Conditional *-tara*, etc.) as well as in the inflectional categories of Table 3, coming to show in the pre-*t* environment the same obstruent-final allomorph as do the *r*-stems (/sir-ta/ [sitta] ‘knew’; contrast /mi-ta/ ‘saw’). In general, this prediction is not realized, and its failure to be is considered mysterious in the dialectological literature (see Kobayashi 2004:593–594). But on the interpretation offered here, under which innovative *r*-suffixes are due to *r*-Epenthesis, there is no mystery: the environment of that rule is intervocalic, so that no prediction is generated for forms with consonant-initial suffixes like *-ta*. (When obstruent-final pre-*t* allomorphs of historical V-stems (*mit-ta* ‘saw’) do appear, notably in Southern Kyushu (see e.g. Tojo et al. 1961 vol 4:279), they are a sign that, because of the proliferation of *r*-initial suffixes in the V-stem paradigm, V-stems have been reanalyzed as *r*-stems (de Chene 1985:179–180).)

A second common interpretation, or class of interpretations, of the innovative forms we have seen in this section seeks to explain them as direct transformations of the conservative forms they replace, much as if one were to seek to explain the 17th century replacement of *raught* /rɔ:t/ by *reached* /ri:tʃt/ as the past tense of English *reach* by postulating a change process /ɔ:/ → /i:tʃ/. Such an account is particularly well known for innovative Potential *-re-*, a suffix that has attracted more attention than the other innovative *r*-suffixes because it is characteristic of Tokyo Japanese. Since the 1990s, the ongoing replacement of conservative Potential *-rare-* by *-re-* has become known as *ra-nuki* (‘*ra*-removal’),<sup>19</sup> a name that embodies the claim that Potential *-re-* arises from *-rare-* by deletion of the syllable *ra*;

<sup>18</sup> Perhaps the clearest prediction of Analyses B and C, common to the two analyses, is for C-stem Imperatives in *-o*. The GAJ has no C-stem Imperative survey items, reflecting a paucity of interesting variation for this category nationwide, but unpublished results of the preparatory survey for the GAJ do show a handful of apparent C-stem Imperatives in *-o* from the vicinity of Kanazawa city, Ishikawa Prefecture. Such forms, however, recorded in the dialectological literature with *-oo* as well as *-o* (Tojo et al. 1961 vol. 3:100, Iwai 1959:100ff.), are in origin imperative uses of Hortatives (thanks to Takuichiro Onishi for clarification on this point).

<sup>19</sup> The word is absent from the first edition of *Daizirin* (Matsumura 1988) and the fourth edition of *Koozien* (Shinmura 1991), two standard dictionaries, but present in the second edition of the former (Matsumura 1995) and the fifth edition of the latter (Shinmura 1998). Its origin, then, can be dated quite precisely.

this claim has been taken up by, among others, Ito and Mester (2004) and Fukushima (2004). At the same time, however, a parallel “transformative” account of the innovative Negative *-ran* of Kyushu and elsewhere, which replaces conservative *-n*, will require a change process inserting the identical syllable *ra*. The Potential and Negative together thus illustrate in the starkest possible manner that there is no generalization to be obtained by direct comparison of conservative and innovative forms; each of the morpheme-specific change processes postulated will require an entirely ad hoc explanation (for examples of such, see Fukushima 2004:188 for the Potential and Kobayashi 2004:59 for the Negative). Further evidence against transformative accounts of innovative *r*-suffixes is provided in some cases by their geographical distribution. Thus, a number of researchers (see Martin 1975:960, Kobayashi 2004:588, Lawrence 2004:25–26) have suggested that innovative Imperative *-re* is to be related in some way to conservative Imperative *-ro*. Innovative *-re*, however, commonly occurs in areas where the conservative suffix it replaces is not *-ro*, but *-yo* or *-i*; de Chene (2010a, section 3.2) identifies more than 60 such locations among the GAJ survey points.

In sum, it seems fair to conclude that neither an interpretation of innovative vowel-stem forms that attributes them to influence from *r*-stem conjugation or an interpretation that attributes them to morpheme-specific change processes poses a serious challenge to the view of them that we have developed here, namely that they result from the feeding of the *r*-Epenthesis rule (2b) by loss of irregular V-stem suffix alternants.

## 2.4 The Environment of *r*-Epenthesis

To this point, we have characterized the environment of rule (2b) as “intervocally at verb stem boundary” ( $V_{vb}$ ] [ \_\_V). In fact, however, the rule applies only in a proper subset of the cases that meet this condition. In the present section, we refine our characterization of the environment of *r*-Epenthesis by discussing five minimal pairs for application of the rule, where in each case the first member of the pair constitutes an *r*-Epenthesis environment and the second is a minimally different formation in which *r*-Epenthesis fails to apply and hiatus either remains unresolved or is resolved through coalescence and monophthongization.

The first minimal pair contrasts the stem *mire-* of the innovative Potential verb *mireru* ‘can see’, where *r*-Epenthesis applies at the boundary between the stem *mi-* ‘see’ and the Potential suffix *-e-*, with the stem *mie-* of *mieru* ‘be visible’, where hiatus remains unresolved (in the examples below, a single underline marks a sequence of vowels that undergoes *r*-Epenthesis, and a double underline marks an apparently parallel sequence that does not).

- (10) a. mi-e- [mire-] ‘can see’  
 b. mi-e- [mie-] ‘be visible’

The crucial difference between the two stems of (10) is that while (10b) is listed in the lexicon, (10a), as an inflectional formation, is naturally analyzed as syntactically derived. For others of the suffixes of Table 3, dependence on syntactic operations is even clearer: the Passive and Causative suffixes are associated with changes in argument structure, and the Indicative and Infinitive forms are distinguished in part in that the former ends a main clause, the latter a coordinate clause. There is thus little question that *r*-Epenthesis is part of the post-syntactic phonology. In line with the fundamental assumption of Lexical Phonology, inherited from Chomsky and Halle 1968, that the internal structure of one level is invisible at later levels (“bracket erasure” (see e.g. Kaisse and Shaw 1985:11)), it is natural to assume that the internal structure of a lexical stem is erased on entry into the derivation. If so, the observed difference will follow: when *r*-Epenthesis applies, the stem (10b) of *mieru* will be effectively monomorphemic.

The second minimal pair contrasts the stem of *mireru* (= (10a)) and that of *ieru* ‘can say’, both syntactically derived.

- (11) a. mi-e- [mire-] ‘can see’  
 b. i-e- [ie-] ‘can say’

The stem of ‘say’ is /iw-/, as shown by its shape before *a*-initial suffixes (Passive *iw-are-*, Causative *iw-ase-*, Negative *iw-ana-*). *w* eventually deletes before any vowel other than *a*, but must clearly still be in place when *r*-Epenthesis applies. I will assume a word-level and a phrase-level stratum within the postsyntactic phonology, with *r*-Epenthesis belonging to the former and *w*-Deletion to the latter; in the traditional terminology of Lexical Phonology, *w*-Deletion is

a postlexical rule. The postlexical status of *w*-Deletion is attested to by its exceptionlessness (for discussion, see Vance 1987:25–26).

The third minimal pair contrasts the stems of *mirareru* ‘be seen’ and *mioeru* ‘finish looking’, both syntactically derived (for the syntactic status of the latter, see Kageyama 1999:301–303 and references cited there).

- (12) a. mi-are- [mirare-] ‘be seen’  
 b. mi-oe- [mioe-] ‘finish looking’

I will take pairs like (12) to show that the rightward morpheme in the environment of *r*-Epenthesis must be grammatical rather than lexical. Syntactically, the difference between *-are-* ‘be (passive)’ and *oe-* ‘finish’ is that the former is an auxiliary verb and the latter a main verb, like their English counterparts *be* and *finish*, respectively; morphologically, the difference is that the former is an affix and the latter a stem, so that while (12a) is an example of affixation, (12b) is an example of compounding. Abbreviating the positive value of the feature [ $\pm$ Grammatical] as “Af” (for *affix*), the *r*-Epenthesis rule (2b) can thus be rewritten as in (13).

- (13)  $\emptyset \rightarrow r / V_{\text{vb}} [Af \text{ \_\_\_ } V]$

In conjunction with the minimal pair of (10), the upshot of (12) is that *r*-Epenthesis applies only in inflectional formations: it is excluded from derivation (and from lexical compounds) by its postsyntactic status in conjunction with the convention of bracket erasure, and it is excluded from syntactic compounds by the requirement that the rightward morpheme be affixal.

The fourth minimal pair contrasts the innovative and conservative Imperative forms of *miru* ‘see, look’ recorded at GAJ locations 641131, 642049, and 642157 in the Izumo region of Shimane Prefecture (the diaeresis is the IPA centralization diacritic; for Tokyo and most other locations, the same point can be made with reference to the irregular Imperative form *ko-i* ‘come!’).

- (14) a. mī-e [mīre] ‘look!’ (innovative)  
 b. mī-i [mīi] ‘look!’ (conservative)

The fifth minimal pair, parallel to the fourth, contrasts the innovative and conservative Hortative forms of *okiru* ‘get up’ recorded at GAJ locations 728951 (Nagasaki Prefecture) and 734265 (Fukuoka Prefecture).

- (15) a. oki-oo [okiroo] ‘let’s get up’ (innovative)  
 b. oki-u [okyuu] ‘let’s get up’ (conservative)

Like C-stem suffixes, and unlike the V-stem suffixes of Table 3, the V-stem suffixes of (14b) and (15b), Imperative *-i* and Hortative *-u*, are vowel-initial. In contrast to C-stem suffixes, however, Imperative *-i* and Hortative *-u* do not trigger *r*-Epenthesis when they are added to vowel-final stems. As shown by the surface form of (15b) and by the widespread occurrence of monophthongized Imperative forms *akee/akii* < *ake-i* ‘open!’ and *kee/kii/koo* < *ko-i* ‘come!’ (GAJ maps 87 and 90, respectively), they do trigger a different hiatus reduction process, namely the coalescence of two adjacent syllables into one, which is a necessary condition for the observed monophthongizations. Moreover, these are exactly the same processes undergone by the homophonous adjectival suffixes Indicative *-i* and Polite Infinitive *-u*, as illustrated by examples like (16) and (17) (the monophthongizations of (16) are dependent on speech level and region; for evidence that the vowel sequences in question are tautosyllabic even when monophthongization does not occur, see Vance 1987:73–76).

- (16) a. taka-i [takee] ‘expensive’  
 b. hido-i [hidee] ‘outrageous’  
 (17) a. arigata-u [arigatoo] ‘grateful(ly)’  
 b. yorosi-u [yorosyuu] ‘good, well’



Imperative *-i* and Hortative *-u*, then, are marked to undergo an independently motivated minor rule of Syllable Coalescence, which creates a bimoraic nucleus from a sequence of vowels in hiatus, and that rule is ordered before *r*-Epenthesis, bleeding it. This entails, in turn, that the two vowels constituting the environment of *r*-Epenthesis must belong to separate syllables.

In this section, we have seen that *r*-Epenthesis is a word-level rule of the post-syntactic phonology and that it applies only in inflectional formations between vowels in hiatus and is bled by the distinct hiatus-reduction rule of Syllable Coalescence.<sup>20</sup> In concluding, it is important to note that none of these restrictions is peculiar to *r*-Epenthesis; the rule of Analysis C, which deletes the second of two vowels across verb stem boundary, would have to be restricted in exactly the same way.<sup>21</sup>

### 3 Explanation

We now need to tackle the question of why it is Analysis A that has been chosen by speakers—the question, that is, of why Analysis A is the descriptively adequate analysis of the system of alternations we have been considering. In approaching this issue, I will proceed in the following manner. In section 3.1, I narrow down the space of possibilities for an account of the choice of Analysis A by identifying a number of proposals concerning morphophonological analysis that are counterexemplified by that choice or by the relevant alternations themselves. Then, in section 3.2, I sketch and motivate a set of principles that entail the choice of Analysis A.

#### 3.1 Implications for an Explanatory Account

An explanatory account of the choice of Analysis A must of course make clear why that analysis was chosen over Analyses B and C. At the same time, it must account for certain properties that Analysis A shares with those two competing analyses, notably the distinction between basic and nonbasic forms and the distinction between regular and irregular forms within the latter set (cf. (1) above). An explanatory account, in other words, will have to specify how basic forms and rules are chosen, either separately or as a unit analysis. In doing so, it will also exclude “symmetrical” analyses that involve lexical listing of all suffix alternants either with (Bloch 1946) or without (Ito and Mester 2004) their environments.

Another type of analysis that will be excluded (as a reviewer reminds me) is one in which inflected forms are not segmented into morphemes, but treated as unanalyzable units (Klaffehn 2003; for similar suggestions, see Vance 1987:206–207, Vance 1991:156). That this is not a likely state of affairs can be inferred from the fact that a Japanese inflected form like *utaw-ase-rare-na-i* ‘will not be made to sing’ combines phonological material that (like its English counterpart) is widely understood to correspond to five syntactic heads (V, Caus, Pass, Neg, T). The claim that such a combination is memorized as a unit thus conflicts with the understanding that it is constructed in the syntax from five distinct lexical elements. More generally, the further a language lies along the continuum from isolating to polysynthetic, the closer the claim that words are not segmented approaches the claim that sentences are not either.<sup>22</sup>

Other implications for an explanatory account of the choice of Analysis A arise from the fact that analysis, and the corresponding system of alternations itself, are inconsistent with several proposals that are sometimes made concerning speakers’ choice of rules and representations for inflectional alternations. I identify six such proposals below, stating each as a numbered hypothesis before discussing it briefly. Before taking up those proposals, however, I point out one

<sup>20</sup> de Lacy (2006:81–82), citing Mester and Ito 1989 and Lombardi 1998 (both of which refer to de Chene 1985), includes Japanese *r*-Epenthesis on a list of cases for which “Epenthesis is ... forced by some general prosodic requirement...” As the above discussion makes clear, *r*-Epenthesis in fact has a far more restricted range of application than is consistent with de Lacy’s claim.

<sup>21</sup> Fukushima’s (2004:195 (note 4)) criticism of the *r*-Epenthesis rule of de Chene 1987 for incorporating restrictions of the type indicated is thus misguided: the restrictions on where hiatus is (dis)allowed are a fact of the data rather than of any particular analysis thereof.

<sup>22</sup> Klaffehn (2003:vii) presents two main reasons for concluding that inflected forms are atomic, namely that native speakers perform poorly on nonce word verb inflection tasks (“wug-testing”) and that he finds no default error pattern in a database (130 hours) of child conversation (three children aged 18–36 months). I submit that the innovative *r*-suffixes documented in section 2.3 above constitute the relevant “default error pattern” in Japanese verb inflection and that, given the evidence they provide regarding the analysis of the system that is in force, an analysis that presupposes word-internal segmentation, the wug-testing reported must be judged to have failed to tap into linguistic competence.

implication that the line of analysis developed here has, not for the theory of morphophonology, but for the theory of exponence—that is, the mechanism by which morphosyntactic features are realized phonologically (thanks to a reviewer for raising this issue).

From the outset (section 1.2), we have assumed a model of how speakers analyze alternations that is neutral regarding whether the alternation affects stems or affixes. In particular, we have assumed that, for either type of alternation, speakers postulate basic forms for alternating morphemes and (in the general case) rules to derive nonbasic forms. Further, we have equated the irregularity that typically results from these analytic choices with excess information in lexical entries, so that regularization—both leveling and extension—results from the elimination of lexical information. Our model of morphophonological analysis and change, then, presupposes that stems and affixes have lexical entries of the same form, and thus requires a “lexical” theory of exponence (Stump 2001:1), one in which both stems and inflectional affixes constitute listed morphemes of the traditional kind. By the same token, it is apparently inconsistent with an “inferential” theory such as that of Anderson 1992 or Stump 2001, in which inflectional affixes are the by-products of spellout rules that operate on the combination of a stem and the full set of morphosyntactic features realized by a particular inflected form. As a result, insofar as our account of the alternations of Table 3 and the changes they are undergoing, an account which assumes the model in question, is successful, it provides evidence for a lexical as opposed to an inferential theory.

Let us turn now to proposals concerning morphophonological analysis that appear to be counterexemplified or called into question by the case of Japanese verbal inflection. Hypotheses 1 and 2 below concern choice of rules, while 3 through 6 concern choice of basic or underlying representations.

1. Speakers analyze inflectional alternations by postulating rules deriving inflected forms from one another within the paradigm of a given stem or lexeme.

Hypothesis 1 has been proposed by Bybee (1985:49–50) and Albright (2002:ix). While applicable to the analysis of stem alternations, however, it cannot plausibly be extended to the analysis of affix alternations, which are cross-lexemic in nature. To see this, consider the following schematic example. Assume a lexicon of 1000 stems, each of which occurs in five inflectional categories. Each stem will then have a paradigm of five inflected forms. Likewise, each inflectional category will have a “paradigm” of 1000 inflected forms, the set of forms in which that category is instantiated. While it might be plausible to treat apparent stem alternations in terms of rules that directly relate to one another the five inflected forms of a stem’s paradigm, it will not be plausible to treat apparent affix alternations in terms of rules that directly relate to one another the 1000 inflected forms of an inflectional category’s paradigm. Since, in the absence of evidence that stem alternations and affix alternations are fundamentally different phenomena, we need a theory of morphophonological analysis that applies to both, Hypothesis 1 must be judged unsatisfactory.

2. Faced with a mutually competing set of alternations resulting from a particular choice of URs, speakers postulate a corresponding set of stochastic rules each of which applies with a likelihood proportional to the prevalence in the lexicon of the corresponding alternation (thanks to a reviewer for raising this issue with respect to the Japanese case).

In decision-theoretic terms, hypothesis 2 represents a “probability matching” theory of morphophonology; the “probability maximization” alternative proposed here (see (18) below) is that the alternation with the highest lexical frequency becomes the default or regular alternation and that other alternations, being irregular in the sense of requiring extra lexical specification, are subject to elimination over time in favor of the default. Concrete proposals incorporating a probability matching approach to morphophonology typically include a mechanism guaranteeing that the stochastic rules in question come into play only when a novel item has to be inflected; for the existing vocabulary, lexical knowledge is held to take precedence (Albright, Andrade, and Hayes 2000:6; Zuraw 2000:xiv, 9; Becker 2009:viii, 23). Nevertheless, ongoing change in the existing lexicon has also been analyzed in probability matching terms (Jun 2010), and it is worth asking what the implications of that approach would be for the Japanese case.

Given the choice of C-stem alternants as basic for suffixes 1–8 of Table 3 above, an alternation of zero with *r* applies to three suffixes, an alternation of a vowel with zero to two, and alternations of zero with *y* and *s* to one each. As a result, a probability matching theory of morphophonology suggests that to the extent that innovative forms are

observed, those displaying vowel deletion should occur two-thirds as often as those displaying epenthesis of *r*, and those displaying epenthesis of *y* and *s* should occur one-third as often. It also predicts, of course, that the *r*-initial V-stem suffixes 1–3 should show variation just as do the V-stem suffixes of categories 4–8. In fact, while innovative forms with *r*-Epenthesis are extremely numerous, as we have seen, I am not aware of any unambiguous example of extension of any the other three alternations. Correspondingly, the *r*-zero alternation of suffixes 1–3 is (modulo unrelated changes) stable. It thus seems clear that Hypothesis 2 makes incorrect predictions about the Japanese case (for further discussion of Hypothesis 2, see de Chene 2014b).

### 3. Speakers obey word-level phonotactics in choosing basic or underlying representations.

Hypothesis 3 has been widely cited and discussed since it was proposed as a tendency by Hale (1973:419–420). In Japanese, consonant-final verb stems would be inadmissible as phonological words, and their postulation thus violates Hypothesis 3 (for doubts about the existence of C-final stems in Japanese, see Vance 1987:199 and Nasukawa 2010). Our discussion of Japanese verbal inflection and the changes it is undergoing, however, has depended at every point on postulating consonant-final stems; Analysis A, in particular, in taking C-stem suffixes as basic, presupposes the existence of such stems. Speakers' choice of Analysis A thus clearly counterexamples Hypothesis 3. I suggest that that hypothesis is too strong; the actual generalization, which will cover cases of the type Hale discussed but correctly refrain from making predictions about the Japanese case, refers to the concept of isolation form, an inflected form that coincides with the stem on which it is based. That generalization is that, at least when the existence of an isolation form holds across an entire lexical class, isolation forms tend to be taken as basic (broadly parallel proposals go back at least as far as Vennemann 1974b:364).

### 4. Speakers choose URs for (nonautomatic) inflectional alternations on semantic criteria, taking alternants with more general or inclusive semantics as basic vis-a-vis alternants whose semantics is more highly specified.

Hypothesis 4 is arguably the import of Kurylowicz's (1945–49) well-known second law, which mandates basic status for alternants or categories that occur in environments where a semantic contrast is neutralized (e.g. for masculines vis-a-vis feminines if the former category has a gender-neutral use), given that occurrence in an environment of neutralization is *prima facie* evidence of inclusive semantics. It is also a major tenet of Natural Generative Phonology (Vennemann 1972:240, Hooper 1979:114ff.), and has recently been defended by Garrett (2008:139–142). There are in general no semantic distinctions, however, among morpheme alternants whose distribution is determined by purely phonological factors, as is typically the case for suffix alternations. For a semantic account of the Japanese alternations analyzed above to be possible, for example, there would have to be some semantic feature common to all of the thousands of C-stem verbs and none of the thousands of V-stems. Such cases show, then, that there can be no general semantic theory of UR choice.

### 5. Speakers choose an analysis so as to maximize the predictability of nonbasic morpheme alternants.

The goal of maximizing the predictability of nonbasic alternants is a plausible one, since in so doing, speakers are minimizing the amount of lexical irregularity involved in the analysis. It is also very close to what is entailed by the feature-counting evaluation metric of classical generative phonology, since minimizing the number of feature specifications that must be used in recording irregular allomorphs will in general minimize the total number of feature specifications required for the lexical representations of the morphemes in question. It is clear, however, that the choice of Analysis A cannot be explained on the hypothesis that speakers wish to predict nonbasic alternants: as we have already noted, that analysis treats as irregular five out of eight nonbasic suffix alternants, as opposed to the three of Analysis B and the one of Analysis C. Under that hypothesis, then, Analysis A will be the least highly valued of the three observationally adequate analyses that we examined in section 2.

### 6. Speakers choose an analysis so as to maximize the predictability of inflected forms.

An alternative to the hypothesis that speakers are attempting to maximize the predictability of nonbasic allomorphs is suggested by the work of Albright (2002:ix), who proposes a model of paradigm acquisition incorporating the principle that base forms are chosen so as to “[permit] accurate productive generation of as many forms of as many words as possible.” While the hypothesis that speakers attempt to maximize the predictability of inflected forms across the entire lexicon has considerable conceptual plausibility, it, too, is counterexemplified by speakers’ choice of Analysis A. This is because, when inflectional stem-forming suffixes are taken into account, the great majority of inflected verb forms, those based on lexical C-stems as well as those based on lexical V-stems, are V-stem forms—that is, have V-stem endings. As a result, Analyses B and C, under which basic forms of suffixes coincide with V-stem alternants, will do better at predicting inflected forms than will Analysis A. Confirming this claim will require a certain amount of calculation.

To begin with, I assume that the set of forms based on a given lexical stem—that stem’s paradigm—includes not just forms consisting of the lexical stem plus an ending (i.e. word-forming suffix), but forms in which one or more inflectional stem-forming suffixes are added to the stem before an ending is. Thus, the paradigm of the Latin verb stem *laudā-* ‘praise’ includes not just the (2nd person singular) Present (Active Indicative) form *laudā-s*, with a single inflectional suffix, but the corresponding Imperfect *laudā-bā-s*, with two, and the Pluperfect *laudā-v-erā-s*, with three. In the case of Japanese verb inflection, Passive *-(r)are-*, Causative *-(s)ase-*, and Potential *-e-* are all inflectional stem-forming suffixes. As a result, the paradigm of our sample C-stem *mat-* ‘wait’ will include not only the forms resulting from adding endings to that stem, but also the forms resulting from adding the same endings to Passive *mat-are-*, Causative *mat-ase-*, Potential *mat-e-*, and, among combinations of those stem-forming suffixes, at least the frequently used Causative-Passive *mat-ase-rare-*. Because our definition of predictability for suffix alternants and inflected forms below will depend on the tripartition of alternants in (1), however, and because the fact that the Potential suffix has only a single allomorph means that that classification does not apply to it, we will leave Potential forms out of the predictability calculation. For purposes of that calculation, then, the paradigm of *mat-* will consist of the forms shown in Table 8 (for simplicity, I take Negatives to count as single forms regardless of further inflection).

	Lexical	Passive	Causative	Causative-Passive
Indicative	mat-u	mat-are-ru	mat-ase-ru	mat-ase-rare-ru
Provisional	mat-eba	mat-are-reba	mat-ase-reba	mat-ase-rare-reba
Hortative	mat-oo	mat-are-yoo	mat-ase-yoo	mat-ase-rare-yoo
Infinitive	mat-i	mat-are-Ø	mat-ase-Ø	mat-ase-rare-Ø
Negative	mat-an(a-)	mat-are-n(a-)	mat-ase-n(a-)	mat-ase-rare-n(a-)
Imperative	mat-e	mat-are-ro	mat-ase-ro	mat-ase-rare-ro

Table 8 Paradigm of *mat-* ‘wait’

As may easily be verified, 18 of the 24 forms of Table 8 (all except those of the leftmost column) are V-stem forms. For a V-stem like *mi-* ‘see’, all 24 of the paradigm’s forms will be V-stem forms.

To calculate the relative proportions of C-stem and V-stem forms in the set of all verbal inflected forms, we need to know the relative lexical frequency of C-stems and V-stems. As is often observed, C-stems are roughly twice as common as V-stems in the lexicon of Japanese, and have been throughout the recorded history of the language (see Suzuki 1977:202–210, Keino 1972:7–8, and de Chene 1985:205 (note 10) for statistics for various periods). Searching the relevant designations for conjugational class (variants of *godan* “quinguirade” for C-stems and *itidan* “unigrade” for V-stems) in electronic versions of Japanese language dictionaries (and adding one to the V-stem total for the Potential suffix, which, unlike the Passive and Causative, does not have a dictionary entry) confirms this ratio; for the *Iwanami*, *Sinmeikai*, and *Meikyoo Kokugo Ziten* (Nishio, Iwabuchi, and Mizutani 2000; Yamada et al. 2005; Kitahara 2002), the C-stem figures are 67.4%, 67.3%, and 65.6%, respectively, for an average of 66.8%.<sup>23</sup>

<sup>23</sup> The relative lexical frequency of the eleven occurring stem-final segments is approximately as follows: *e* (29.6%), *i* (3.0%), *r* (22.1%), *s* (17.4%), *k* (8.4%), *m* (8.2%), *w* (6.6%), *t* (2.2%), *g* (1.6%), *b* (0.8%), *n* (0.02%). These figures are based on searches of Nishio, Iwabuchi, and Mizutani 2000 and exclude irregular verbs (see below). (The .2% discrepancy between the C-stem total of 67.3% these results yield and the value of 67.5% reported above is the result of (a) rounding and (b) a shortfall of two C-stems in the stem-specific search results.)

This figure, however, requires an adjustment for verbs that dictionaries designate as irregular (*ka-hen* “*k*-irregular”, *sa-hen* “*s*-irregular”) and which are therefore not included in the above statistics—more precisely, for the subset of such verbs that can be judged, on independent grounds, to have been lexicalized. *Kuru* ‘come’ is the only *k*-irregular verb, and its paradigm consists entirely of V-stem forms, irregularity being confined to alternation of the stem vowel. *S*-irregular verbs are *suru* ‘do’ and compounds with *suru* as second member, and these too, in principle, have paradigms consisting of V-stem forms (although the Passive and Causative stems *s-are-* and *s-ase-* of *suru* itself are formed on the stem *s-*). With regard to compounds with *suru* as second member, Kurisu (2010) provides criteria distinguishing those that have been lexicalized, typically ones whose first member is a single Sino-Japanese morpheme, from those that are arguably generated post-lexically. The list of lexicalized verbs in his appendix is closely matched by the verbs marked *s*-irregular in the *Iwanami Kokugo Ziten* (IKZ = Nishio, Iwabuchi, and Mizutani 2000), and it is the listings of that dictionary that I have taken as the basis for the following calculation.

For our purposes, the crucial fact about lexicalized compounds with second member *suru* is that their stem tends to be reanalyzed, so that their conjugation is variable in contemporary Japanese. For compounds of this type in which the *s* of *suru* voices, first of all, there is a tendency to be reanalyzed as *i*-stems (*kanzuru* > *kanziru* ‘feel’). Such verbs thus display V-stem conjugation in any case and should be counted as V-stems. There are 71 such verbs listed as *s*-irregular in the IKZ; of these, 40 are crosslisted as *i*-stems and thus will already have been taken account of in the relative lexical frequency calculation for that dictionary. It is therefore appropriate to add the difference between these two figures (plus one each for *kuru* and *suru*) to the V-stem count of 1391 for IKZ, giving 1424. For lexicalized compounds with second member *suru* in which the *s* of *suru* does not voice, on the other hand, there is a tendency, varying in strength by phonological subtype and inflectional category, to be reanalyzed as *s*-stems (*aisuru* > *aisu* ‘love’). Such verbs thus display a mixture of V-stem and C-stem conjugation and in that sense provide only equivocal evidence regarding the relative lexical frequency of V-stems and C-stems. I will therefore assume that they should be disregarded for the purposes of the relative lexical frequency calculation. Since 28 of these verbs are crosslisted as *s*-stems in IKZ, it is appropriate to subtract that figure from the dictionary’s C-stem count of 2883, giving 2855. Recalculating the percentage of C-stems in IKZ on the basis of the revised V-stem and C-stem counts for that dictionary gives the result that 66.7% of all verbs are C-stems, a negative adjustment of .8%. Applying this adjustment to the average C-stem percentage for the three dictionaries, finally, reduces the latter from 66.8% to 66.0%, and it is this last figure that we will use in the calculations below.

Returning to the paradigm of Table 8, the fact that C-stems and V-stems are distributed in the lexicon in the ratio 66:34 means that across the entire set of verbal inflected forms, 75% (18 of 24) of the forms of 66% of all verbs, plus all of the forms of the remaining 34%, will be V-stem forms. Since  $(.75 \times .66) + (1.00 \times .34) = .835$ , this means that roughly 84% of all verbal forms are V-stem forms, and only 16% are C-stem forms. This imbalance will have a substantial effect on the calculations below (in which, as here, I will compute ratios to three places and round to two at the end of the calculation).

Let us now move to the question of what proportion of verbal forms are predictable under Analyses A, B, and C. I will assume that a suffix alternant is predictable under a given analysis if it is either basic or regularly derived (i.e. if it is not irregular) under that analysis and that an inflected form is predictable under a given analysis if and only if all of its suffixes are. For example, under Analysis A, *mat-oo* ‘let’s wait’ is predictable, since *-oo* is basic (and thus predictable) under that analysis, but *mi-yoo* ‘let’s look’ is not, since *-yoo* is irregular. Under Analysis B, both forms are predictable, because *-yoo* is basic under that analysis and *-oo* is regularly derived.

Consider first Analysis A. Under that analysis, all six of the inflected forms of the lexical C-stem *mat-* will be predictable, but only two each of the forms of *mat-are-*, *mat-ase-*, and *mat-ase-rare-*, namely the Indicative in *-ru* and the Provisional in *-reba* (see Table 8 above). Thus for a C-stem verb, 12 of 24 forms will be predictable, or 50%. Similarly, just two of the forms (Indicative and Provisional) of the lexical V-stem *mi-* and the corresponding Passive stem *mi-rare-* will be predictable, and none of the forms of *mi-sase-* or *mi-sase-rare-*, given that all such forms will include one or more suffix alternants that are irregular under Analysis A. For a V-stem, then, only 4 of 24 forms (16.7%) will be predictable. In sum, Analysis A predicts 50% of the forms of 66% of all lexical verb stems, and 16.7% of the forms of the remaining 34%. Since  $(.50 \times .66) + (.167 \times .34) = .387$ , this means that it predicts 39% of all verbal inflected forms.

Now consider Analysis B. Under that analysis, three of the six inflected forms of *mat-* will be predictable, namely

the Indicative in *-u*, the Provisional in *-eba*, and the Hortative in *-oo*, as will all six inflected forms of *mat-are-*, *mat-ase-*, and *mat-ase-rare-*. For a C-stem verb, then, 21 of 24 forms will be predictable, or 87.5%. Further, all 24 inflected forms of *mi-* will be predictable. In sum, Analysis B predicts 87.5% of the forms of 66% of all verbs, and 100% of the forms of the remaining 34%. Since  $(.875 \times .66) + (1 \times .34) = .918$ , this means that it predicts 92% of all verbal inflected forms, well over twice the percentage predicted by Analysis A. Analysis C, finally, will predict everything except Imperative *mat-e*. It will predict, that is, 95.8% (23 of 24) of C-stem forms and 100% of V-stem forms—in sum,  $(.958 \times .66) + (1 \times .34) = .97$  of all verbal forms. Just as with the hypothesis that speakers aim to maximize the predictability of nonbasic morpheme alternants, then, the hypothesis that speakers aim to maximize the predictability of inflected forms results in Analysis A, the descriptively adequate analysis, being judged least highly valued of the three observationally adequate analyses we have considered.

### 3.2 Toward an Explanation

The observations of section 3.1 considerably reduce the range of possibilities for an explanatory account of Analysis A. At the same time, the fact that Analysis A loses badly to Analysis B and Analysis C both with regard to prediction of nonbasic morpheme alternants and with regard to prediction of inflected forms underlines the challenge we face in attempting to explain why speakers have adopted Analysis A. In this section I will argue that the key to resolving this problem is the recognition that morphophonological analysis proceeds in two steps, with choice of base forms and choice of a rule constituting separate analytic decisions that are taken in that order. I will call this procedure “sequential evaluation” and propose that both the choice of base forms and the choice of a rule are governed by a principle that privileges alternants and alternations with high lexical frequency (“Generalized Type Frequency”). Throughout, I assume both sequential evaluation and Generalized Type Frequency to be specifically morphophonological principles. In terms of the life cycle of alternations that has been familiar to phonologists since the work of Kruszewski (1881) and Baudouin de Courtenay (1895), in other words, these are principles that, to a first approximation, come into play at the point when an alternation ceases to be interpreted as an automatic response to a phonotactic constraint and becomes subject to reanalysis. (For discussion of the analytic differences between automatic and morphophonological alternations, see de Chene 2014a.)

#### 3.2.1 Background: Evaluation, Frequency

In this section, I introduce the concepts of evaluation (evaluation procedure) and frequency (type and token frequency) as background to the argumentation of section 3.2.2. Starting with the former concept, we may note that in attempting to explain speakers’ choice of Analysis A, we have a choice between two kinds of account. On one of them, that analysis, consisting of basic forms and a rule, will be chosen as a unit as a result of comparison with other potential analyses, notably Analyses B and C. On the other, the two components of Analysis A will be the result of separate analytic decisions, with the choice of basic forms logically prior to that of a rule. The principles underlying either kind of account will constitute an evaluation procedure in the general sense of Chomsky 1957:51, 1975:79 (where, however, the object of evaluation is taken to be a complete grammar). Let us say that an evaluation procedure of the first kind, under which an analysis is chosen as an indivisible whole, instantiates “unitary evaluation”, and that an evaluation procedure of the second kind, under which basic forms and rule(s) are chosen separately and in that order, instantiates “sequential evaluation”. The two evaluative criteria we considered at the end of section 3.1, according to which speakers choose the analysis that maximizes prediction of (a) nonbasic morpheme alternants or (b) inflected forms, are examples of unitary evaluation. As we have already seen, however, both those criteria fail badly at the task of predicting the choice of Analysis A. Below, we will see first that there is reason to believe that sequential evaluation is correct and then show that that assumption allows a natural account of the choice of Analysis A.

Let us move now to the concepts of type and token frequency. Narrowly defined, these apply to linguistic units such as segments or words; while the type frequency of a particular unit is its frequency across the lexicon, its token frequency is its frequency in a text or corpus. Thus the type frequency of the segment [ð] in English is low because it appears in few words, but its token frequency in any substantial corpus will be relatively high because the few words in which it occurs are common (Frisch 2011:2138).

Bybee (2001:10,119) significantly broadens the applicability of the notion of type frequency by defining it as the lexical frequency of a “pattern”, where the latter concept is general enough to cover items as diverse as phonotactic configurations (e.g. initial three-member consonant clusters), rhythmic configurations (e.g. antepenultimate stress), and affixes. In asking what factors determine whether speakers will generalize a pattern, Bybee emphasizes that when type frequency and token frequency diverge, it is patterns with high type frequency, and not those with high token frequency, that tend to be generalized (Bybee 1995/2007:173–175, Bybee 2001:118–121 and references cited there). Indeed, as she (1995/2007:174) notes, high token frequency has actually been claimed to have an inhibiting effect on pattern formation (Baayen and Lieber 1991). Similarly, Albright (2009:207) shows that taking token frequency into account has a mildly negative effect on the ability of two learning models to mirror speaker intuitions about the generalizability of the Spanish diphthongization alternation in verbs. Below, I will take the evidence cited by these authors as establishing a strong *prima facie* case for basing UR choice and rule choice in morphophonology on a suitable concept of type frequency.

Specifically, we will need a concept of type frequency that, as a first condition, allows us to compare the relative strength of the allomorphs of an alternating morpheme, either a stem or an affix, considered as candidates for that morpheme’s UR (as above, I use “morpheme” in the sense of “unit with a lexical entry”, so that the term applies to stems regardless of whether or not they are strictly speaking monomorphemic). To that end, we may take an inflected form to be composed of a stem inserted from the lexicon and a morphosyntactic feature complex developed by the syntax. The type frequency of a stem allomorph can then be defined as the percentage of eligible morphosyntactic feature complexes with which it combines, and the type frequency of an affix allomorph can be defined as the percentage of eligible stems with which it combines.

For example, verb stems in Portuguese (and Romance languages generally) have both stressed and unstressed alternants. Taking a conservative variety of the language as representative, there are roughly 60 morphosyntactic feature complexes with which a stem can co-occur, each corresponding to an inflected form (six person-number combinations each for nine subparadigms; two imperative and four nonfinite forms). Of these 60 forms, precisely nine are stem-stressed. For any verb stem, then, the type frequency of its stressed alternant will be 15%, and that of its unstressed alternant 85%. A parallel example for suffix alternations is furnished by Japanese, where, as we have already noted, the ratio of consonant-final to vowel-final verb stems in the lexicon is 66:34. As a result, for each verbal inflectional suffix, the type frequency of its C-stem alternant will be 66%, and that of its V-stem alternant 34%.

### 3.2.2 Sequential Evaluation and Generalized Type Frequency

Let us call a choice of underlying representations “contrastive” if it targets alternants that occur in an environment allowing a maximal number of contrastive values for alternating features or segments and “neutralizing” if it targets alternants that occur in an environment allowing only a reduced number of contrastive values. For example, if a range of stem-final consonants contrast prevocally but are uniformly reduced to zero word-finally, then taking prevocalic alternants as basic will constitute a contrastive choice of URs, while taking word-final alternants as basic will constitute a neutralizing choice. Given those definitions, the evidence for sequential evaluation is the fact that speakers not infrequently make a neutralizing choice of underlying representations in situations where a contrastive choice would have been possible. The reason this argues for sequential evaluation is that while a neutralizing choice of URs typically has a natural explanation in terms of criteria that are applicable to UR choice alone, the loss of phonological predictability that neutralized URs entail means that, regardless of what rule is postulated, the analysis as a whole is likely to fare badly on plausible global measures of evaluation such as prediction of nonbasic allomorphs or inflected forms. In principle, either leveling in favor of neutralized values or extension of an innovative rule taking neutralized values as input will tend to support the hypothesis that speakers have made a neutralizing choice of URs in a particular case. Here, I briefly summarize two examples, the first involving extension of an innovative rule, the second involving both extension and leveling.

The Western Romance seven-vowel system was subject to a neutralization rule reducing unstressed *ɛ* *ɔ* to *e* *o*. This resulted in alternations in verb stems (‘deny’, ‘ask’ in Table 9 below) depending on whether stress fell on the stem or on a suffix; there were also stems with nonalternating *e* *o* (‘fish’, ‘prune’).

Gloss	Unstressed	Original Stressed	Original Alternation	Portuguese Stressed	Portuguese Alternation
‘deny’	neg-	nɛg-	e ~ ε	nɛg-	e ~ ε
‘ask’	rog-	rɔg-	o ~ ɔ	rɔg-	o ~ ɔ
‘fish’	pesk-	<u>pesk</u> -	————	<u>pesk</u> -	e ~ ε
‘prune’	pod-	<u>pod</u> -	————	<u>pod</u> -	o ~ ɔ

Table 9 Extension of Lowering in Portuguese Verb Stems

In Portuguese, the alternation in question was extended to originally nonalternating stems (see Williams 1938:206ff.), as the underlined forms of Table 9 illustrate. This extension can be seen as the result of two analytic choices. First, in a neutralizing choice of URs, unstressed alternants were taken as basic, creating an arbitrary lexical distinction between stems whose *e o* alternated with *ε ɔ* under stress and stems whose *e o* was nonalternating. In this situation, the alternating pattern was taken as regular, resulting in a rule of Lowering under stress in verb stems (Harris 1974:75). Non-alternating stems were originally lexical exceptions to Lowering; as the exception feature was lost from their lexical entries, they came to undergo the rule, so that today, the only verb-stem mid vowels that fail to do so are those that are exempt for phonological reasons such as nasalization, hiatus with a suffixal vowel, or adjacency to a palatal consonant (Hensey 1972:290–291, Azevedo 2005:74).

As a second example, Korean obstruents undergo neutralization of laryngeal and manner contrasts in syllable codas, reducing *p p<sup>h</sup> p'* to *p*, *t t<sup>h</sup> t' s s' c c<sup>h</sup> c' h* to *t*, and *k k<sup>h</sup> k'* to *k* (we set aside parallel alternations resulting from cluster reduction). As a result, marked stem-final obstruents alternate with their neutralized counterparts before a consonant-initial suffix or clitic, appearing unchanged before a vowel as a result of resyllabification; there are also stems ending in nonalternating *p*, *k*, and (for verbs) *t*. In noun stems, these alternations are undergoing a mixture of leveling and extension. Stems ending in coronal obstruents other than *s* display a tendency for their alternation to be replaced by the *t ~ s* alternation (‘stick game’ in Table 10 below), while stems ending in marked labial or dorsal obstruents display a tendency for their alternations to be leveled in favor of the neutralized syllable-final form (‘knee’, ‘kitchen’).

Gloss	Syllable-final	Conservative Resyllabified	Conservative Alternation	Innovative Resyllabified	Innovative Alternation
‘stick game’	jut	juc <sup>h</sup>	t. ~ .c <sup>h</sup>	jus	t. ~ .s
‘knee’	murip	murip <sup>h</sup>	p. ~ .p <sup>h</sup>	murip	————
‘kitchen’	puək	puək <sup>h</sup>	k. ~ .k <sup>h</sup>	puək	————

Table 10 Leveling and Extension in Korean Noun Stems

As in the Portuguese case, these developments can be seen as the result of two analytic choices. First, in a neutralizing choice of URs, syllable-final alternants have been taken as basic, resulting in arbitrary lexical distinctions among multiple types of basic *p t k* depending on the alternant that is observed prevocally. In this situation, the *t ~ s* alternation has been taken as regular for coronal stem-finals, resulting in a rule taking *t* to *s* when it is both syllable-initial and final in a noun stem. Nonalternating *t*-stems were originally lexical exceptions to this rule, but have been almost completely regularized (Ito 2010:363). For labial and dorsal stem-finals, in contrast, the null alternation (i.e. nonalternating pattern) has been taken as regular. For all three points of articulation, stem allomorphs ending in marked obstruents other than *s* (e.g. the *juc<sup>h</sup>*, *murip<sup>h</sup>*, *puək<sup>h</sup>* of Table 10) are lexically listed along with the corresponding default forms (*jut*, *murip*, *puək*), but with a condition restricting them to the prevocalic environment. As these environmentally restricted irregular allomorphs are lost, the corresponding nouns assume regular behavior, stem-final *p k* ceasing to alternate and stem-final *t* coming to alternate prevocally with *s* (as chart (8) of Ito 2010 shows, there are also competing subtendencies among the coronals).

We claimed above that the phenomenon of a neutralizing choice of URs, which we have seen to be illustrated by the Portuguese and Korean reanalyses just sketched, supports the idea that, for morphophonological alternations, base form candidates and rule candidates are evaluated separately and in that order, a procedure we called “sequential evaluation”.



Our reason for taking a neutralizing choice of URs to argue for sequential evaluation was that while the loss of phonological predictability that neutralized URs entail makes it implausible that they are chosen as part of an attempt to maximize some desirable property of the analysis as a whole, a neutralizing choice of URs can typically be motivated in terms of properties of the URs themselves. We have not yet indicated, however, what properties of the neutralized URs in the Portuguese and Korean cases are responsible for their selection as underlying.

In the Portuguese case, type frequency, as defined above for stem alternants, provides a plausible account for why unstressed alternants were chosen as basic: as we have already seen, the type frequency of stressless alternants is 85% in a conservative variety of Portuguese, and it will not have been significantly different at the historical stage where the alternation was reanalyzed. In the Korean case, type frequency is a less promising candidate as an explanation for why neutralized alternants were taken as basic, since structural and most local case clitics are vowel-initial after consonant-final stems and thus create resyllabification environments (see e.g. Sohn 1999:213–214); a related problem is that there is no well-defined nominal paradigm in Korean to begin with. A primary desideratum of an explanation for the neutralizing choice of URs in the Korean case is that it distinguish between nominal and verbal inflection, since essentially the same set of alternations that is undergoing leveling and extension in noun stems is entirely stable in verb paradigms, implying that there is no irregularity in the verbal system of alternations and thus that a contrastive choice of URs is still in force. In this connection, Kenstowicz (1996, section 3.2; see also Kenstowicz and Sohn 2001, section 7.5) points out that while verb stems do not appear uninflected (rare exceptions occur in compounds), noun stems appear in isolation both as citation forms and as the result of deletion of structural case markers. Since isolation forms coincide with neutralized preconsonantal alternants, the neutralizing choice of URs in nouns can be taken to follow from the principle, articulated in section 3.1 above, that isolation forms tend to be taken as basic.

We have seen that the neutralizing choices of URs observed in the Portuguese and Korean reanalyses, while difficult to understand on the hypothesis of unitary evaluation, can be motivated on plausible assumptions about what speakers might be looking for in URs and thus on the competing hypothesis of sequential evaluation. The Portuguese and Korean reanalyses pose a second explanatory problem as well, however, namely that of how speakers choose a regular alternation from among the multiplicity of competing possibilities that result from a neutralizing choice of URs. The answer is particularly clear in the Korean case, which we will take as representative.

Above, we noted that taking neutralized syllable-final alternants of stem-final consonants as basic in Korean nouns results in lexical distinctions among multiple types of basic *p t k* depending on the alternant that is observed prevocally. Concretely, and again abstracting away from alternations involving cluster reduction, there are alternations of basic *p* with *p* and *p<sup>h</sup>*, of basic *k* with *k*, *k<sup>h</sup>*, and *k'*, and basic *t* with (originally) *t*, *t<sup>h</sup>*, *c*, *c<sup>h</sup>*, and *s*. We also noted that while alternations involving marked labials and dorsals are being leveled, indicating that the null alternation has been taken as regular, the *t ~ s* alternation is being extended, showing that it is the regular alternation for coronals. As shown by corpus searches (see Albright 2008, Jun 2010, and de Chene 2014b for details), it is the alternation with the highest lexical frequency that has been chosen as regular in each case.<sup>24</sup> For stems ending in labials and dorsals, the null alternation is overwhelmingly more common than the non-null alternations involving marked obstruents, and among coronal stems, those showing the *t ~ s* alternation constitute a majority and are roughly twice as common as stems showing the next most common alternation.

In the context of our suggestion that basic alternants in the Portuguese case are chosen on the basis of type frequency, determination of the regular alternation in the Korean case on the basis of lexical frequency suggests that, at a suitable level of abstraction, there is a single frequency-based principle that determines both what alternants are basic and what alternations are regular. In either case, this principle will choose from a well-defined candidate set the member that has the widest distribution with respect to a specified set of cooccurring elements. In the case of UR choice, the candidate set will be the set of all alternants of a given stem or affix; in the case of rule choice, it will be the maximal set of alternations (rule candidates) sharing a given input configuration or structural description—for example, the set of three alternations with input *k* in the Korean case. Calling this principle “Generalized Type Frequency”, let us state it as in (18).

<sup>24</sup> The fact that leveling in labial and dorsal stems is a consequence of the same quantitative criterion that dictates extension of the *t ~ s* alternation for coronals means that no condition mandating uniform paradigms (see e.g. the papers in Downing, Hall, and Raffelsiefen 2005) will be necessary to explain the leveling in question. See Garrett 2008 for the proposal that, consistent with the Korean case, leveling is merely a subtype of extension.

- (18) **Generalized Type Frequency:** Given the maximal set of alternants of a morpheme, considered as candidates for the base form of that morpheme, or the maximal set of alternations sharing an input configuration, considered as candidates for the regular alternation with that input configuration, define the type frequency of each candidate as the percentage of eligible cooccurring elements with which it occurs. The successful candidate is then the one with the highest type frequency.

The set of eligible cooccurring elements will vary depending on whether the object of evaluation is a base form candidate or a rule candidate and on whether the relevant alternation involves stems or affixes, as detailed in (19) (where (19a) and (19b) duplicate the definition of type frequency for stem and affix alternants that was adopted above).

- (19) Eligible cooccurring elements for base form candidates (alternants) and rule candidates (alternations)
- a. for an affix alternant, the set of lexical stems with which the affix occurs
  - b. for a stem alternant, the set of morphosyntactic feature complexes with which the stem occurs
  - c. for an affix alternation, the set of affixes to which the alternation could in principle apply
  - d. for a stem alternation, the set of stems to which the alternation could in principle apply

There are thus four subcases of Generalized Type Frequency (GTF), depending on the set of cooccurring elements. The choice of basic forms in the Portuguese case is the result of subcase (19b), and the choice of regular alternations in the Korean case is the result of subcase (19d).

Returning now to the Japanese case and the question of how to explain the choice of Analysis A, I suggest that in the context of sequential as opposed to unitary evaluation, that choice follows from the remaining two subcases of GTF, cases (19a) and (19c). We have already noted that for each Japanese verbal inflectional suffix, the type frequency of its C-stem alternant will be 66%, and that of its V-stem alternant 34%. According to GTF case (19a), this dictates the choice of C-stem suffixes as basic. Assuming that choice, the set of candidates for a regular alternation, given the input configuration of hiatus at verb stem boundary, can be read off the rightmost column of Table 3 above: epenthesis of *r* (3 cases out of 8 eligible suffixes, type frequency 37.5%),<sup>25</sup> epenthesis of *y* (1 case, 12.5%), epenthesis of *s* (1 case, 12.5%), and deletion of the second vowel (2 cases, 25%). GTF case (19c), then, predicts correctly that the regular alternation should be epenthesis of *r*. The choice of Analysis A can thus be seen to follow from the conjunction of the principle of sequential evaluation and the principle of Generalized Type Frequency.

I take it to speak in favor of the above account of the choice of Analysis A that that account is embedded in a general, if partial, theory of how speakers analyze morphophonological alternations. In closing, however, I would like to touch on two other factors that might be thought relevant to the choice of Analysis A and in particular to the choice of the *r*-zero alternation as regular. First, it is possible that ambiguity avoidance plays a role in the selection of epenthesis rather than deletion as a means of hiatus resolution at verb stem boundary, given that a general rule deleting a suffix-initial vowel after a vowel-final stem would result (as we saw in note 6) in zero allomorphs, for vowel-stem inflection, in three of the eight alternating suffixes of Table 3. More generally, the cross-linguistic prevalence of intervocalic stem-boundary epenthesis, particularly of coronal consonants (de Chene 1985:198–199), suggests the existence of a tendency for languages to require that the URs of inflectional elements be fully realized in surface forms (cf. Kenstowicz 2013).

Second, and in response to a reviewer's suggestion, I provide two token frequency statistics of potential interest, obtained from the Balanced Corpus of Contemporary Written Japanese ([http://www.ninjal.ac.jp/corpus\\_center/bccwj/](http://www.ninjal.ac.jp/corpus_center/bccwj/)) with the Chunagon search tool (<https://chunagon.ninjal.ac.jp/>); details of the search procedure are available on request. Of 8.48 million instantiations of suffixes 1–8 of Table 3, (a) 4.76 million (56.1%) are C-stem alternants and 3.72 million (43.9%) are V-stem alternants; (b) 5.32 million (62.7%) represent suffixes 1–3, those with a pre-existing *r*-zero

<sup>25</sup> While Eastern Japanese V-stem Imperative *-ro* begins with *r*, I omit the Imperative from the set of suffixes whose alternations support postulation of *r*-Epenthesis because there is no independent evidence for *-ro*'s segmentability; the fact that *-ro* is lacking in many localities that nevertheless display clear evidence for *r*-Epenthesis shows that in any case it is not necessary for postulation of that rule. I further assume that the restriction of the Potential suffix to C-stems (in conservative dialects), by precluding that suffix from having any V-stem alternant at all, excludes it from the set to which candidate alternations could in principle apply.

alternation. The reduced C-stem/V-stem differential of result (a) relative to the 66:34 type-frequency split reflects the fact that many of the most common verbs are V-stems; the high value seen in result (b) is due primarily to suffix 1 (4.39 million occurrences), which itself accounts for 51.8% of all forms. If these figures are representative, then, pretheoretically, token frequency would appear to fare worse than type frequency as a potential explanation for the choice of C-stem suffixes as basic, but better than type frequency as a potential explanation for the generalization of the *r*-zero alternation. While the consensus on the irrelevance of token frequency to pattern formation that we noted above would seem to preclude explanatory significance for this last observation, it should be kept in mind as we attempt to deepen our understanding of the principles governing morphophonological analysis.

## 4 Conclusion

The study reported above is based on two premises, both introduced in section 1. The first is that evidence from change in progress may reveal which of several observationally adequate analyses of a system of alternations has actually been adopted by speakers, solving the problem of descriptive adequacy for those alternations. The second is that cases in which the descriptively adequate analysis is known with some confidence from such “external” evidence can provide crucial data bearing on the principles speakers employ in analyzing morphophonological alternations and thus on the problem of explanatory adequacy for morphophonology.

It seems fair to say that both of these premises have been validated by our study. With regard to descriptive adequacy, we saw that change in progress unambiguously supports a specific analysis of the system of alternations we considered, that which we called “Analysis A”. With regard to explanatory adequacy, we first noted six well-known proposals that appear to be counterexemplified by the choice of Analysis A or by the relevant system of alternations itself. We then, referring to two other cases of morphophonological reanalysis, developed an account of the choice of Analysis A founded on two principles, the sequential evaluation of base form candidates and rule candidates, and a criterion of type frequency that applies to both the choice of base forms and the choice of rules. If our proposals are on the right track, any adequate theory of morphophonology will have to allow not only for phonological rules and underlying representations, but also for substantive principles, stated in terms of concepts like lexical frequency, according to which those rules and representations are chosen.

**Acknowledgements** I would like to thank audiences at the First International Spring Forum of the English Linguistic Society of Japan (Tokyo, April 2008) and the Sixth Workshop on Altaic Formal Linguistics (Nagoya, September 2009), where portions of the above material were presented. I would also like to express my appreciation to the National Institute for Japanese Language and Linguistics (NINJAL) and to Takuichiro Onishi in particular for allowing me to view unpublished results of the preparatory survey for the Grammar Atlas of Japanese Dialects and for helpful correspondence on related matters. I am also grateful to NINJAL for permission to use the Chunagon corpus search tool and to Kikuo Maekawa for guidance in this regard. For relevant discussion, I thank Masae Matsuki and Takayuki Ikezawa. Finally, I am greatly indebted to the editors of JEAL and to reviewers for two journals for comments that have led to many improvements and clarifications; remaining errors are my responsibility. Preliminary and abbreviated versions of parts of this material have appeared as de Chene 2009 and 2010b; a longer version was posted online as de Chene 2010a.

## References

- Albright, Adam. 2002. *The identification of bases in morphological paradigms*. Dissertation, UCLA.
- Albright, Adam. 2008. Explaining universal tendencies and language particulars in analogical change. In *Linguistic universals and language change*, ed. Jeff Good, 144–181. Oxford: Oxford University Press.
- Albright, Adam. 2009. Modeling analogy as probabilistic grammar. In *Analogy in grammar*, ed. James P. Blevins and Juliette Blevins, 185–213. Oxford: Oxford University Press.
- Albright, Adam, Argelia Andrade, and Bruce Hayes. 2000. Segmental environments of Spanish diphthongization. MS, UCLA. Published in Albright, Adam, and Taehong Cho, eds., 2001, *UCLA working papers in linguistics, number 7: Papers in phonology 5*, 117–151. Los Angeles: UCLA Department of Linguistics.
- Andersen, Henning. 1969. A study in diachronic morphophonemics: The Ukrainian prefixes. *Language* 45:807–830.
- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
- Anttila, Raimo. 1989. *Historical and comparative linguistics*, 2nd ed. Amsterdam: Benjamins.
- Aoki, Hirofumi. 2010. *Gokeisei kara mita Nihongobunpoosi [Japanese grammatical history viewed from word-formation]*. Tokyo: Hituzi

- Syoboo.
- Azevedo, Milton M. 2005. *Portuguese*. Cambridge: Cambridge University Press.
- Baayen, Harald, and Rochelle Lieber. 1991. Productivity and English derivation: a corpus-based study. *Linguistics* 29:801–843.
- Baudouin de Courtenay, Jan. 1895. *Versuch einer Theorie phonetischer Alternationen*. Strassburg: Trübner. English translation in Baudouin de Courtenay, Jan, 1972, *A Baudouin de Courtenay anthology*, edited and translated with an introduction by Edward Stankiewicz, 144–212. Bloomington: Indiana University Press.
- Becker, Michael. 2009. *Phonological trends in the lexicon: the role of constraints*. Dissertation, University of Massachusetts, Amherst.
- Bloch, Bernard. 1946. Studies in colloquial Japanese I. Inflection. *Journal of the American Oriental Society* 66:97–109. Reprinted in Bloch, Bernard, 1970, *Bernard Bloch on Japanese*, edited with an introduction by Roy Andrew Miller, 1–24. New Haven: Yale University Press.
- Bybee, Joan. 1985. *Morphology*. Amsterdam: Benjamins.
- Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10:425–455. Reprinted in Bybee, Joan, 2007, *Frequency of use and the organization of language*, 167–193. Oxford: Oxford University Press.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Chew, John J. 1973. *A transformational analysis of modern colloquial Japanese*. The Hague: Mouton.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1964a. *Current issues in linguistic theory*. The Hague: Mouton. Reprinted in Fodor, Jerry A., and Jerrold J. Katz, eds., 1964, *The structure of language*, 50–118. Englewood Cliffs, NJ: Prentice-Hall.
- Chomsky, Noam. 1964b. The logical basis of linguistic theory. In *Proceedings of the ninth international congress of linguists*, ed. Horace Lunt, 914–978. The Hague: Mouton.
- Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Plenum.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- de Chene, Brent. 1985. *r*-Epenthesis and the Japanese verb. *Papers in Japanese Linguistics* 10: 172–207.
- de Chene, Brent. 1987. Keitai on'inron ni okeru sinniteki zituzaisei: Nihongo no doosi no baai [Psychological reality in morphophonology: the case of the Japanese verb]. In *Kotoba kara mita kokoro [Mind viewed from language]*, ed. Yukio Otsu, 145–180. Tokyo: Tookyoo Daigaku Syuppankai.
- de Chene, Brent. 2009. Description and explanation in inflectional morphophonology: the case of the Japanese verb. *Waseda Daigaku Kyooikugakubu Gakuziyutu Kenkyuu Eigoeibungakuhen [Waseda University School of Education Academic Research English Language and Literature Edition]* 57:1–22.
- de Chene, Brent. 2010a. Description and explanation in inflectional morphophonology: the case of the Japanese verb. <http://ling.auf.net/lingbuzz/001055>
- de Chene, Brent. 2010b. Description and explanation in inflectional morphophonology: notes on the Japanese verb. In *Proceedings of the 6th workshop on Altaic formal linguistics (WAFL 6) (MIT working papers in linguistics 61)*, ed. Hiroki Maezawa and Azusa Yokogoshi, 123–138. Cambridge, MA: MITWPL.
- de Chene, Brent. 2014a. The duality of phonology: automatic and nonautomatic alternations at the explanatory level. MS, Waseda University. Presented at the Workshop on Nonautomatic Alternations in Phonology, 47th Meeting of the Societas Linguistica Europaea, Poznań.
- de Chene, Brent. 2014b. Probability matching versus probability maximization in morphophonology: the case of Korean noun inflection. *Theoretical and Applied Linguistics at Kobe Shoin* 17:1–13.
- de Lacy, Paul. 2006. *Markedness*. Cambridge: Cambridge University Press.
- Downing, Laura J., T. Alan Hall, and Renate Raffelsiefen, eds. 2005. *Paradigms in phonological theory*. Oxford: Oxford University Press.
- Frisch, Stefan A. 2011. Frequency effects. In *The Blackwell companion to phonology*, ed. Marc van Oostendorp et al., 2137–2163. Oxford: Blackwell.
- Fukushima, Kazuhiko. 2004. Conspiracy of form and function for optimization of language change. *Journal of East Asian Linguistics* 13:181–196.
- GAJ (*Grammar Atlas of Japanese Dialects*) = Kokuritu Kokugo Kenkyuuzyo 1989–2006
- Garrett, Andrew. 2008. Paradigmatic uniformity and markedness. In *Linguistic universals and language change*, ed. Jeff Good, 125–143. Oxford: Oxford University Press.
- Grimes, Barbara, ed. 2000. *Ethnologue*, 14th ed., vol. 2. Dallas, TX: SIL International.
- Hale, Kenneth. 1973. Deep-surface canonical disparities in relation to analysis and change: an Australian example. In *Diachronic, areal, and typological linguistics (Current trends in linguistics 11)*, ed. Thomas A. Sebeok, 401–458. The Hague: Mouton.

- Hale, Mark, and Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.
- Halle, Morris. 1962. Phonology in generative grammar. *Word* 18:54–72. Reprinted in Fodor, Jerry A., and Jerrold J. Katz, eds., 1964, *The structure of language*, 334–352. Englewood Cliffs, NJ: Prentice-Hall.
- Halle, Morris. 1997. Distributed morphology: impoverishment and fission. *PF: Papers at the interface (MIT working papers in linguistics* 30), ed. Benjamin Bruening, Yoonjung Kang, and Martha McGinnis, 425–449. Cambridge, MA: MITWPL.
- Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20*, ed. Kenneth Hale and Samuel J. Keyser, 111–176. Cambridge, MA: MIT Press.
- Harris, James W. 1974. Evidence from Portuguese for the “Elsewhere Condition” in phonology. *Linguistic Inquiry* 5:61–80.
- Hasegawa, Nobuko. 1999. *Seisei nihongogaku nyuumon [Introduction to the generative grammar of Japanese]*. Tokyo: Taisyuukan Syoten.
- Hensey, Fritz G. 1972. Portuguese vowel alternation. In *Generative studies in Romance languages*, ed. Jean Casagrande and Bohdan Saciuk, 285–292. Rowley, MA: Newbury House.
- Hooper, Joan B. 1979. Substantive principles in natural generative phonology. In *Current approaches to phonological theory*, ed. Daniel A. Dinnsen, 106–125. Bloomington: Indiana University Press.
- Itoyo, Kiichi, et al., eds. 1982–1986. *Kooza hoogengaku [Course in dialectology]*, 10 vols. Tokyo: Kokusyo Kankookai.
- IKZ (*Iwanami Kokugo Ziten*) = Nishio, Iwabuchi, and Mizutani 2000
- Ito, Chiyuki. 2010. Analogy and lexical restructuring in the development of nominal stem inflection from Middle to Contemporary Korean. *Journal of East Asian Linguistics* 19:357–383.
- Ito, Junko, and Armin Mester. 2004. Morphological contrast and merger: *ranuki* in Japanese. *Journal of Japanese Linguistics* 20:1–18. Also in the Rutgers Optimality Archive, ROA-711.
- Iwai, Ryusei. 1959. Isikawaken Kanazawasi Hikofoo itibantyo. [(The dialect of) Hikofoo ichibancho in Kanazawa, Ishikawa Prefecture]. In *Nihon hoogen no kizyututeki kenkyuu [Descriptive studies on Japanese dialects]*, 87–110. Tokyo: Kokuritu Kokugo Kenkyuuzyo.
- Jinnouchi, Masataka. 1996. *Hokubu kyusyu ni okeru hoogen singo kenkyuu [Research on innovative dialect forms in northern Kyushu]*. Fukuoka: Kyusyu Daigaku Syuppankai.
- Jun, Jongho. 2010. Stem-final obstruent variation in Korean. *Journal of East Asian Linguistics* 19:137–179.
- Kageyama, Taro. 1999. Word formation. In *The handbook of Japanese linguistics*, ed. Natsuko Tsujimura, 297–325. Oxford: Blackwell.
- Kaisse, Ellen M., and Patricia A. Shaw. 1985. On the theory of lexical phonology. *Phonology* 2:1–30.
- Keino, Masaji. 1972. *Doosi no kenkyuu [Studies on verbs]*. Tokyo: Kasama Syoin.
- Kenstowicz, Michael. 1996. Base-identity and uniform exponence: alternatives to cyclicity. In *Current trends in phonology: models and methods*, ed. Jacques Durand and Bernard Laks, 363–394. Paris X and Salford: University of Salford Publications.
- Kenstowicz, Michael. 2013. Realize morpheme in Kaqchikel. In *Studies in Kaqchikel grammar (MIT working papers on endangered and less familiar languages* 8), ed. Michael Kenstowicz, 67–80. Cambridge, MA: MITWPL.
- Kenstowicz, Michael, and Charles Kisseberth. 1977. *Topics in phonological theory*. New York: Academic Press.
- Kenstowicz, Michael, and Charles Kisseberth. 1979. *Generative phonology*. New York: Academic Press.
- Kenstowicz, Michael, and Hyang-Sook Sohn. 2001. Accentual adaptation in North Kyungsang Korean. In *Ken Hale. A life in language*, ed. Michael Kenstowicz, 239–270. Cambridge, MA: MIT Press.
- Kiparsky, Paul. 1971. Historical linguistics. Lecture, University of Maryland. Published in Kiparsky, Paul, 1982, *Explanation in phonology*, 57–80. Dordrecht: Foris.
- Kiparsky, Paul 1972. Review of *Geschichte der indogermanischen Verbalflexion*, by Calvert Watkins. *Foundations of Language* 9:277–286.
- Kiparsky, Paul. 1978. Analogical change as a problem for linguistic theory. Forum lecture, LSA Linguistic Institute, University of Illinois. Published in Kiparsky, Paul, 1982, *Explanation in phonology*, 217–236. Dordrecht: Foris.
- Kitahara, Yasuo, ed. 2002. *Meikyoo kokugo ziten [Meikyoo dictionary of the Japanese language]*. Tokyo: Taisyuukan Syoten.
- Klafehn, Terry. 2003. Emergent properties of Japanese verbal inflection. Dissertation, University of Hawai‘i.
- Kobayashi, Takashi. 2004. *Hoogengakuteki Nihongosi no hoofoo [A dialectological approach to the history of the Japanese language]*. Tokyo: Hituzi Syoboo.
- Kokuritu Kokugo Kenkyuuzyo [National Language Research Institute]. 1953. *Tiiki syakai no gengo seikatu. Turuoka ni okeru zittai tyoosa*. [The linguistic life of regional society. A survey of Tsuruoka]. Tokyo: Syuuei Syuppan.
- Kokuritu Kokugo Kenkyuuzyo [National Language Research Institute]. 1989–2006. *Hoogen bunpoo zenkoku tizu [Grammar atlas of Japanese dialects]*, 6 vols. Tokyo: Kokuritu Kokugo Kenkyuuzyo.
- Kruszewski, Mikolaj. 1881. *Über die lautabwechslung*. Kasan: Universitätsbuchdruckerei. English translation by Robert Austerlitz in Kruszewski, Mikolaj, 1995, *Writings in general linguistics*, edited with an introduction by Konrad Koerner, 5–34. Amsterdam: Benjamins.

- Kurisu, Kazutaka. 2010. Japanese light verb voicing as a connective morpheme. In *Proceedings of the 6th workshop on Altaic formal linguistics* (WAFL 6) (*MIT working papers in linguistics* 61), ed. Hiroki Maezawa and Azusa Yokogoshi, 213–227. Cambridge, MA: MITWPL.
- Kuroda, S.-Y. 1960. *Gengo no kizyutu* [*The description of language*]. Tokyo: Kenkyuusha.
- Kurylowicz, Jerzy. 1945–49. La nature des procès dits «analogiques». *Acta Linguistica* V:15–37. Reprinted in Hamp, Eric, et al., eds., 1966, *Readings in linguistics II*, 158–174. Chicago: University of Chicago Press.
- Kyuusyu Hooogen Gakkai [Kyushu Dialect Research Society]. 1991. *Kyuusyu hooogen no kisoteki kenkyuu* [*Basic research on Kyushu dialects*], revised edition. Tokyo: Kazama Syoboo.
- Labov, William. 2001. *Principles of linguistic change, vol. 2: Social factors*. Oxford: Blackwell.
- Labrone, Laurence. 2014. The phonology of Japanese /r/: a panchronic account. *Journal of East Asian Linguistics* 23:1–25.
- Lakoff, George. 1965. *On the nature of syntactic irregularity*. Cambridge, MA: Computation Laboratory of Harvard University.
- Lawrence, Wayne P. 2004. High vowels, glides, and Japanese phonology. *Gengo Kenkyuu* 125:1–30.
- Lightner, Theodore. 1968. On the use of minor rules in Russian phonology. *Journal of Linguistics* 4:69–72.
- Lombardi, Linda. 1998. Coronal epenthesis and unmarkedness. *University of Maryland working papers in linguistics* 5:156–175. Revised version published as Lombardi, Linda, 2002, Coronal epenthesis and markedness, *Phonology* 19:219–251.
- Marcus, Gary F., et al. 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 57.4. Chicago: University of Chicago Press.
- Martin, Samuel E. 1966. Lexical evidence relating Japanese to Korean. *Language* 42:185–251.
- Martin, Samuel E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.
- Matsuda, Kenjiro. 1993. Dissecting analogical leveling quantitatively: the case of the innovative potential suffix in Tokyo Japanese. *Language Variation and Change* 5:1–34.
- Matsumura, Akira, ed. 1971. *Nihon bunpoo daiziten* [*Great dictionary of Japanese grammar*]. Tokyo: Meizi Syoin.
- Matsumura, Akira, et al., eds. 1988. *Daizirin*. Tokyo: Sanseido.
- Matsumura, Akira, et al., eds. 1995. *Daizirin*, 2nd ed. Tokyo: Sanseido.
- McCarthy, John J. 2002. *A thematic guide to optimality theory*. Cambridge: Cambridge University Press.
- McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.
- Mester, Armin, and Junko Ito. 1989. Feature predictability and underspecification: palatal prosody in Japanese mimetics. *Language* 65:258–293.
- Miyagawa, Shigeru. 1999. Causatives. In *The handbook of Japanese linguistics*, ed. Natsuko Tsujimura, 236–268. Oxford: Blackwell.
- Nasukawa, Kuniya. 2010. No consonant-final stems in Japanese verb morphology. *Lingua* 120:2336–2352.
- Nihon Kokugo Daiziten Dainihan Hensyuu Inkai/Syoogakukan Kokugo Ziten Hensyuubu [Great Dictionary of the Japanese Language Second Edition Editorial Committee/Syoogakukan Japanese Dictionary Editorial Division], eds. 2000–2002. *Nihon kokugo daiziten* [Great dictionary of the Japanese language], 2nd ed. Tokyo: Syoogakukan.
- Nishio, Minoru, Etsutaro Iwabuchi, and Shizuo Mizutani, eds. 2000. *Iwanami kokugo ziten* [Iwanami dictionary of the Japanese language], 6th ed. Tokyo: Iwanami Syoten.
- NKD (*Nihon Kokugo Daiziten*) = Nihon Kokugo Daiziten Dainihan Hensyuu Inkai/Syoogakukan Kokugo Ziten Hensyuubu 2000–2002
- Oshima, David Y. 2014. On the morphological status of *-te*, *-ta*, and related forms in Japanese: evidence from accent placement. *Journal of East Asian Linguistics* 23:233–265.
- Penny, Ralph. 2002. *A history of the Spanish language*, 2nd. ed. Cambridge: Cambridge University Press.
- Sasaki, Kan, and Akie Yamazaki. 2006. Two types of detransitive constructions in the Hokkaido dialect of Japanese. In *Passivization and typology*, ed. Werner Abraham and Larisa Leisiö, 352–372. Amsterdam: Benjamins.
- Shinmura, Izuru, ed. 1991. *Koozien*, 4th ed. Tokyo: Iwanami Syoten.
- Shinmura, Izuru, ed. 1998. *Koozien*, 5th ed. Tokyo: Iwanami Syoten.
- Skousen, Royal. 1975. *Substantive evidence in phonology*. The Hague: Mouton.
- Sohn, Ho-Min. 1999. *The Korean language*. Cambridge: Cambridge University Press.
- Stump, Gregory T. 2001. *Inflectional morphology*. Cambridge: Cambridge University Press.
- Suzuki, Kazuhiko. 1977. *Bunpoo no uturikawari* [*Grammatical change*]. In *Nihongo no rekisi* (*Nihongo kooza dairokukan*) [*The history of Japanese* (*Course in Japanese vol. 6*)], ed. Atsuyoshi Sakakura, 197–242. Tokyo: Taisyukan Syoten.
- Tojo, Misao, et al., eds. 1961. *Hoogengaku kooza* [*Dialectology course*], 4 vols. Tokyo: Tookyoodoo.
- Umegaki, Minoru. 1962. Miekken hooogen [The dialects of Mie Prefecture]. In *Kinki hooogen no soogooteki kenkyuu* [*A comprehensive study of*

- the Kinki dialects*], ed. Minoru Umegaki, 91–157. Tokyo: Sanseido.
- Vance, Timothy J. 1987. *An introduction to Japanese phonology*. Albany, NY: State University of New York Press.
- Vance, Timothy J. 1991. A new experimental study of Japanese verb morphology. *Journal of Japanese Linguistics* 13:145–166.
- Vennemann, Theo. 1972. Rule inversion. *Lingua* 29:209–242.
- Vennemann, Theo. 1974a. Restructuring. *Lingua* 33:137–156.
- Vennemann, Theo. 1974b. Words and syllables in Natural Generative Grammar. In *Papers from the parasession on natural phonology*, ed. Anthony Bruck, Robert A. Fox, and Michael W. LaGaly, 346–374. Chicago: Chicago Linguistic Society.
- Vovin, Alexander. 2010. *Koreo-Japonica*. Honolulu: University of Hawai‘i Press.
- Whitman, John. 1985. *The phonological basis for the comparison of Japanese and Korean*. Dissertation, Harvard University.
- Whitman, John. 2012. The relationship between Japanese and Korean. In *The languages of Japan and Korea*, ed. Nicolas Tranter, 24–38. London: Routledge.
- Williams, Edwin B. 1938. *From Latin to Portuguese*. Philadelphia: University of Pennsylvania Press.
- Yamada, Tadao, et al., eds. 2005. *Sinmeikai kokugo ziten* [*New Meikai dictionary of the Japanese language*], 6th ed. Tokyo: Sanseido.
- Yamaguchi, Akiho, and Morihide Akimoto, eds. 2001. *Nihongo bunpoo daiziten* [*Great dictionary of Japanese grammar*]. Tokyo: Meizi Syoin.
- Yoshikawa, Toshiaki, and Yukihiro Yamaguchi. 1972. *Toyohasi tihoo no hoogen* [*The dialect of the Toyohashi region*]. Toyohashi: Toyohasi Bunka Kyokai.
- Zuraw, Kie. 2000. *Patterned exceptions in phonology*. Dissertation, UCLA.