Regularity vs. Productivity: The Puzzle of the Japanese Verb Paradigm

Kenneth Hanson

7/20/2022

1. Introduction

The allomorphy in Japanese verb stems and suffixes is extremely regular, and as such may be easily captured with a system of morpho-phonological rules (cf. Bloch 1946, McCawley 1968). But explaining why the system looks as it does is considerably more difficult. Many of the alternations appear to be phonologically motivated, yet are unattested elsewhere in the language. Nonetheless, several attempts have been made to derive the system from phonological principles (Davis and Tsujimura 1991; Ito and Mester 2004, 2015; Nasukawa 2005, 2010). On the other hand, poor performance in nonce verb inflection experiments led some researchers to question whether the system is productive at all (Vance 1987, 1991; Batchelder and Ohta 2000; Klafehn 2013). In this paper, I argue that the allomorphy seen in the Japanese verb paradigm cannot be explained entirely by phonology, nor as simple memorization in the lexicon. The system is morphological, and at least partially productive. In addition, I propose that an acquisition framework based on the Tolerance Principle (Yang 2016) in combination with Sequential Evaluation (de Chene 2016) provides a way to probe the productivity of the system, and conduct a simulation experiment to test the productivity of the r-epenthesis rule proposed by de Chene's (1985).

I begin by examining the claim by Ito and Mester (2004, 2015) that the allomorphy may be derived without any morphological information by selection of lexically specified allomorphs in an OT grammar. While their analysis correctly derives the alternations in a subset of suffixes, I show that it fails for the suffixes *-te*, *-ta*, *-tari*, and *-tara* (t-suffixes), resulting in both undergeneration and overgeneration. These along with several conceptual problems illustrate more generally why phonology cannot be the driving factor behind the synchronic allomorphy. Instead, I argue that the reason why the system appears this way is that it preserves the phonology of much earlier stages of the language. Specifically, the first paradigm enforces CV alternation, which was nearly surface-true in Old Japanese, and the t-suffix paradigm preserves a subset of the productive processes in Early Middle Japanese known as *onbin* (Frellesvig 2010).

Next, I address the question of whether the system is synchronically productive, and argue that it must be, at least partially. Even if speakers today are not making purely phonological generalizations, they are still able to make generalizations that include morphological context; morphologically conditioned phonology is a widely attested phenomenon (cf. Inkelas 2014, ch. 2). Despite the conclusions of many of the same researchers who conducted the nonce verb studies, their results in fact suggest that the CV alternation paradigm is productive for at least some classes of verb stems, specifically r-stems and the 'irregular' verb *suru*. Furthermore, de Chene (1985, 2016) shows that the emergence of novel suffix allomorphs across many Japanese dialects is elegantly explained by a rule of r-epenthesis extending its scope to additional suffixes.

Thus, the question is not *whether* speakers are making any generalizations, but *what* generalizations they are making and *why*. In order to probe these questions, I employ the Tolerance Principle (Yang 2016), which states that learners will adopt a generalization only if the number of exceptions within its scope falls below a certain threshold calculated according to type frequency in the input. But the Tolerance Principle has nothing to say regarding which generalizations learners consider in the first place. A partial solution can be found in de Chene's (2016) principle of Sequential Evaluation, which posits that speakers

first select the base form of two alternants, and only afterward do they attempt to form a rule to derive the non-base form. These two principles together provide a framework to predict morphological productivity.

I conclude with a case study demonstrating the application of this framework. I hypothesize that r-epenthesis is in fact only fleetingly productive according to the Tolerance Principle, which would better explain the piecemeal pattern of extension attested in the dialectal data. To test this hypothesis, I conduct a simulation experiment based on corpus data (cf. Kodner 2019). The results are highly sensitive to the set of word forms considered in the comparison class, but under reasonable assumptions show a developmental trajectory consistent with rare and temporary productivity.

The remainder of this paper is structured as follows. Section 2 consists of an overview of the Japanese verbal paradigm along with a traditional rule-based analysis. In Section 3, I present Ito and Mester's allomorph selection theory and show why their analysis fails to extend to the full paradigm. I also discuss conceptual problems with this and other theories which try to reduce all allomorphy to phonology. Section 4 broadens the discussion to evidence from diachrony, experiment, and dialectal variation. Here, I argue that the available evidence strongly suggests that speakers are indeed forming productive generalizations. In Section 5, I examine de Chene's two-step system of generalization, discuss how it can be cast in more explicitly acquisition-oriented terms using the Tolerance Principle, and present the results of a simulation experiment. Section 6 concludes, with an emphasis on the need for further research of this type.

2. Data overview

The alternations in the Japanese verb paradigm occur only at the boundary between a verbal stem and inflectional suffix, and involve processes not witnessed elsewhere in the grammar. The paradigm is naturally divided into two sub-paradigms: the first, which I call the *CV alternation paradigm*, covers most inflectional suffixes; the other, the *t-suffix paradigm*, includes *-te*, *-ta*, *-tari*, and *-tara*. There is an extensive literature analyzing this allomorphy. The modern linguistic analysis originates with Bloch (1946). Classic generative treatments include Kuroda (1965) and McCawley (1968). Davis and Tsujimura (1991) present an autosegmental account. The detailed differences between these accounts are not directly relevant. In this section I present the core data along with a simple rule-based analysis, which is similar to that given in Tsujimura (2013). All data is for the Tokyo dialect, though other dialects are broadly similar.

2.1. CV alternation paradigm

Most verbal suffixes in this group differ only in the presence or absence of an initial consonant or vowel. A few suffixes are irregular, i.e., not related in form. The choice of suffix allomorph depends on the phonological form of the stem: consonant-final stems (C-stems) select vowel-initial suffixes, and vowel-final stems (V-stems) select consonant-initial suffixes. A summary of the patterns is shown below. Although the focus is on the suffix immediately following the stem, the non-past suffix is added to create a fully inflected verb where applicable. Note that the passive, causative, and potential suffixes inflect as V-stem verbs, while the negative suffix is an adjective.

Verb Form	Suffix	Ex. C-Stem nom- ‘drink’	Ex. V-Stem tabe- ‘eat’	Alternation
non-past	-(r)u	nom-u	tabe-**r**u	∅ ↔ C
passive	-(r)are	nom-are-ru	tabe-**r**are-ru	
causative	-(s)ase	nom-ase-ru	tabe-**s**ase-ru	
conditional	-(r)eba	nom-eba	tabe-**r**eba	
volitional	-(y)oo	nom-oo	tabe-**y**oo	
negative	-(a)na	nom-**a**na-i	tabe-na-i	V ↔ ∅
infinitive	-i/∅	nom-**i**	tabe-∅	
potential	-e/rare	nom-**e**-ru	tabe-**rare**-ru	V ↔ CVCV
imperative	-e/ro	nom-**e**	tabe-**ro**	V ↔ CV

Table 1: Listing of all major CV alternation paradigm suffixes

There are, in fact, many other verbal suffixes. The vast majority of these are derivational suffixes, all of which display an alternation between an initial /i/ and zero. Here, they are considered to be invariant forms that attach to the infinitive suffix. Additionally, there are two suffixes, the polite suffix *-(i)mas-u* and the desiderative suffix *-(i)ta-i*, which could reasonably considered inflectional. It is worth considering whether these attach directly to the stem, and simply happen to have the same alternation as the infinitive. I will return to this issue in Section 5.

These complications aside, the allomorphy seen in this paradigm can be derived by simple morpho-phonological rules, as shown in the table below. The analysis which requires the fewest exceptions takes the long form of each suffix to be basic, and derives the short form with rules deleting the second of two sequential consonants or vowels. These rules are specific to the boundary between a verbal stem and suffix. A more complete analysis would have to account for the fact that these rules do not apply to the t-suffixes; see Davis and Tsujimura (1991) for one way of doing this. Alternative analyses involving insertion rather than deletion are also possible; de Chene (2016) argues for one of these alternatives. This issue will be taken up in Section 5. The key point to note for now is that all analyses of this sort must make reference to morphological context.

	Rule	Example
1a.	C → ∅ / C $_{Vb}$] [$_{Af}$ __	/nom+ru/ → nom-u (*nom-ru)
1b.	V → ∅ / V $_{Vb}$] [$_{Af}$ __	/tabe+anai/ → tabe-nai (*tabe-anai)

Table 2: Rule-based analysis of the CV alternation paradigm suffix allomorphy

The CV alternation pattern could be analyzed as being motivated by a pressure to maintain CV syllable structure. But Japanese tolerates vowel hiatus, and typically resolves consonant clusters in other ways. This presents a paradox, which Ito and Mester (2004) attempt to explain with their allomorph selection analysis, discussed in Section 3.

2.2. T-suffix paradigm

The t-suffixes (-te, -ta, -tari, and -tara) behave differently from the first paradigm but identically to each other. Both the suffix and stem may alternate, and all alternations are fully predictable by the final segment of the stem. The alternative stem allomorphs that co-occur with the t-suffixes are known traditionally as the *onbin* forms.

The full set of alternations can be summarized as follows. The stem may be replaced by a geminate consonant or a moraic nasal, which must match the following consonant in place of articulation. Other

stems undergo insertion of epenthetic /i/, or replacement of a velar consonant by /i/. Finally, initial /t/ in the suffix may be voiced. A summary is shown below. In all examples, I take the past tense suffix *-ta* as representative.

Class	Ex. Verb Stem		Past Form	Stem Process	Suffix Process
V	tabe-	‘eat’	tabe+ta → tabe-ta	n/a	n/a
t	tat-	‘drink’	tat+ta → tat-ta	gemination	n/a
w	kaw-	‘buy’	kaw+ta → kat-ta		
r	kaer-	‘go home’	kaer+ta → kaet-ta		
n	sin-	‘die’	sin+ta → sin-da	nasalization	voicing
m	nom-	‘drink’	nom+ta → non-da		
b	yob-	‘call’	yob+ta → yon-da		
s	kas-	‘lend’	kas+ta → kasi-ta	epenthesis	n/a
k	nak-	‘cry’	nak+ta → nai-ta	velar conversion	n/a
g	oyog-	‘swim’	oyog+ta → oyoi-da		voicing

Table 3: Listing of all stem classes for t-suffix paradigm

The processes seen here can also be derived using a handful of morpho-phonological rules. As before, a simplified analysis is given below.[1] For a more complete and general analysis, I refer the reader to Davis and Tsujimura (1991).

	Process	Rule	Relevant Alternation
2a.	Gemination	[-cons] → [+cons, -cont, -voice] / __ $_{Vb}$] [$_{Af}$ [-cont]	stem: {r, w} → t
2b.	Nasalization	[labial, -cont] → [+nasal] / __ $_{Vb}$] [$_{Af}$ [alveolar]	stem: b → n
2c.	Place Assimilation	[labial] → [alveolar] / __ $_{Vb}$] [$_{Af}$ [alveolar]	stem: w → t, {m, b} → n
2d	T-Suffix Voicing	[alveolar] → [+voice] / [+voice, -cont] $_{Vb}$] [$_{Af}$ __	suffix: t → d
2e.	/i/ Epenthesis	∅ → [i] / s __ $_{Vb}$] [$_{Af}$ [alveolar]	stem: s → si
2f.	Velar Conversion	[velar] → [i] / __ $_{Vb}$] [$_{Af}$ [alveolar]	stem: {k, g} → i

Table 4: Rule-based analysis of the t-suffix paradigm allomorphy

Gemination and nasalization occur elsewhere in the grammar, and can potentially be motivated by phonotactic constraints. But /i/-epenthesis and velar conversion are unattested elsewhere, making them difficult to motivate in this manner. It is precisely these alternations that are problematic in an OT-based analysis, and Ito and Mester’s account does not handle them correctly.

3. Phonological accounts, and why they fail

It is difficult to find analyses that attempt to explain the entire verbal paradigm using only phonological rules or constraints, without making use of morphological contexts or diacritics. In most cases, only a subset of the data is considered. Here, I take Ito and Mester’s (2004, 2015) allomorph selection account as representative, because it seems to me to be the most promising theory available. As an Optimality Theory-based analysis, it seeks to derive the pattern as an effect of independently motivated markedness constraints. Each alternating suffix is lexically specified with several underlying forms, which are selected from as part of the phonological derivation without any additional morphological information. If

1 Gemination (2a) and Nasalization (2b) must precede Place Assimilation (2c). T-Suffix Voicing (2d) must precede Velar Conversion (2f).

successful, this could be considered be a true phonological explanation.

Ito and Mester (2004) deal only with the CV alternation paradigm. Ito and Mester (2015) claim that their analysis extends to the t-suffixes, but do not demonstrate this. In fact, their analysis does *not* handle them successfully, and the ways in which it fails highlight just why the verbal allomorphy should not be considered to be phonologically motivated.

3.1. OT with allomorph selection in the CV alternation paradigm

The basic aim is to derive the verbal allomorphy using constraints on syllable structure. The key markedness constraints are ONSET, NOCODA, and CODACOND. The first two, ONSET and NOCODA, enforce CV syllable structure, and must be ranked low in modern Japanes, since vowel hiatus and certain consonant clusters are common. The third, CODACOND, disallows codas with independent place features, and is needed to derive the deletion pattern and must be highly ranked. In additional, we make use of the faithfulness constraints MAX-IO, DEP-IO, IDENT-IO, which disallow deleting, inserting, changing a segment, respectively.

A straightforward OT analysis would require duplicate constraints making reference to morphological context in order to derive deletion only where desired, parallel to the restrictions in the rule-based version. Ito and Mester (henceforth I&M) propose instead that we allow multiple URs to be specified for each lexical item, consider all mappings from UR to SR for all combinations, and let OT select the best mapping as usual. If we do this, we can derive the deletion pattern without introducing any *ad hoc* constraints.

A illustrative example is given below for the C-stem verb *nom-* ‘drink’. In these tableau, I provide both the UR and SR for each candidate. While I&M do not do this, it seems fairly clear that this is the intended interpretation. We see that in the present case, all SRs derived from the UR /nom+ru/ violate at least one constraint, while the candidate [nom-u], unchanged from UR /nom+u/, does not violate any. Any alternative SR derived from /nom+u/ would also necessarily violate at least one faithfulness constraint. Thus, [nomu] is correctly selected as the output form. We can correctly select the allomorphs of V-stems and irregular suffixes in exactly the same manner.

	/nom+{ru,u}/	CODACOND	MAX-IO	DEP-IO	IDENT-IO	ONSET	NOCODA
➔	nom-u → nom-u						
	nom-ru → nom-ru	*!					*
	nom-ru → nom-u		*!				
	nom-ru → no-ru		*!				
	nom-ru → nomi-ru			*!			
	nom-ru → nom-mu				*!		*

Figure 1: Allomorph selection analysis for C-stem verb with non-past suffix

I&M claim that this approach requires only mechanisms that are independently necessary: a theory of phonology and a lexicon capable of storing multiple URs. It is not obvious that allowing OT to evaluate all combinations of lexically specified URs is as innocuous as I&M claim, but for now, let us accept it. In this case, we have simplified the grammar by eliminating any morphological constraints or diacritics, making use only of general phonological principles. This was not entirely without cost, since we needed to enrich the lexicon. Is this a good enough reason to adopt the analysis? Perhaps, though as we will see shortly, the analysis is neither empirically adequate nor as simple as it seems.

3.2. Allomorph selection fails for the t-suffixes

Now, I extend I&M's analysis to the t-suffix paradigm, and show that it falls short on many levels: (1) it is unnecessary to derive some of the alternations, yet (2) it fails to derive all of those for which it is needed. Additionally, (3) it creates massive redundancy in the lexicon, and (4) it introduces proportional over-generation of output forms.

3.2.1. No allomorph selection needed for some stems

First, for the t/w/r stems (and possibly n/m/b-stems), we can derive all alternations using a small number of additional phonological constraints. To the extent that this can be justified, classic OT without allomorph selection is sufficient. An example for a w-stem verb is given below. We introduce an additional markedness constraint *DD, which bans voiced geminates, these being unattested in the native Japanese lexical substratum (Ito and Mester 1995). With just this one addition, [kat-ta] is correctly selected as the output from among the phonotactically licit options, under the reasonable assumption that MAX-IO outranks IDENT-IO.

	/kaw+ta/	*DD	CODACOND	MAX-IO	IDENT-IO	NOCODA
	a. kaw-ta		*!			*
→	b. kat-ta				*	*
	c. kaw-wa	*!			*	*
	d. ka-ta			*!		
	e. ka-wa			*!		

Figure 2: Single UR analysis for w-stem verb with past suffix

A more nuanced analysis might appeal to autosegmental representations in order to motivate gemination as the superior repair strategy (cf. Davis and Tsujimura 1991), in which case the winning candidate would not actually violate IDENT-IO. I&M also seem to implicitly assume this. For present purposes, the simple segment-based representations will suffice.

As for the remaining alternations of this sort, t- and r- stems also involve gemination, and are derived in the same manner as w-stems. N-, m-, and b- stems require additional constraints, which can nonetheless be phonologically motivated, though perhaps not as convincingly, in which case we would resort to allomorph selection. We consider this case next.

3.2.2. Allomorph selection and lexical redundancy

The alternations in the s-, k-, and g-stems are not amenable to the approach used to derive the gemination pattern, because the processes of /i/-epenthesis and velar conversion are not attested elsewhere in the grammar. To do so would require increasingly *ad hoc* constraints, which we would like to avoid. Instead, we make use of allomorph selection for these three stem classes. For example, we must posit allomorphs with and without epenthetic /i/ for every s-stem in the language. It is not clear whether this is plausible, since we would be claiming lexical specification of a huge number of allomorphs whose relations are completely predictable. This is in stark contrast to the CV alternation paradigm, where we needed to specify only a small number of allomorphs, some of which were irregular anyway.

That being said, we can correctly derive the past tense of s- and k-stem verbs under the allomorph selection analysis. An example tableau for s-stems follows. As in the early example with the non-past suffix, all SRs derived from the UR /kas+ta/ violate at least one constraint, while the fully faithful SR [kasi-ta] violates no constraints, and is therefore the winner.

	/{kas,kasi}+ta/	CODACOND	IDENT-IO	NOCODA
➔	a. kasi-ta → kasi-ta			
	b. kas-ta → kas-ta	*!		*
	c. kas-ta → kat-ta		*!	*
	d. kas-ta → kas-sa		*!	*

Figure 3: Allomorph selection analysis for s-stem verb with past suffix

At this point, the allomorph selection analysis appears to be faring reasonably well where it has been needed. The fact that we have introduced massive redundancy into the lexicon should give us pause, but we are at least able to capture the facts. But we will now turn to cases where the analysis does fail on empirical grounds.

3.2.3. The opacity problem

Allomorph selection cannot correctly derive t-suffix allomorphy in the g-stems, which requires opaque voicing in the suffix. Instead, it predicts that the suffix should remain unvoiced, as in the s- and k- stems. An example tableau is given below. As before, all SRs deriving from UR /oyog+ta/ violate at least one constraint. However, the target form /oyoi-da/ also violates IDENT-IO due to voicing of the suffix. As always, the output form [oyoi-ta] violates no constraints at all, which this time is not the correct result.

	/{oyog,oyoi}+ta/	CODACOND	IDENT-IO	NOCODA
➔	a. oyoi-ta → oyoi-ta			
☹	b. oyoi-ta → oyoi-da		*!	
	c. oyog-ta → oyog-ta	*!		*
	d. oyog-ta → oyoi-da		**!	
	e. oyog-ta → oyot-ta		*!	*

Figure 4: Allomorph selection analysis for g-stem verb with past suffix

The problem is that we need to select the vowel-final allomorph of a g-stem verb when combining with a t-suffix, but we also need the information in the consonant-final allomorph in order to derive voicing in the suffix. In a ruled-based derivation: /oyog+ta/ → oyog-da → [oyoi-da]. It is not surprising that I&M's analysis should falter here, since opacity is a common source of difficulty for bi-level OT analyses in general.

In fact, there is another case of opacity not yet discussed, and also not treated by I&M, which is w-deletion, as in /kaw+ru/ → kaw-u → [ka-u]. Hall et al. (2018) use this fact to argue for allomorph selection in combination with Harmonic Serialism, i.e., neither alone is adequate. For the purposes of the present discussion, any variant of OT which drops strict parallelism, such as Stratal OT, would work just as well (Kawahara 2017). We just need to be able to select the vowel-initial suffix before w-deletion applies, as in the rule-based derivation.

But this solution is not available for the g-stem voicing case under the allomorph selection analysis. The reason we are positing lexical specification of the *onbin* stem in the first place is that it cannot be derived from the main stem in a principled manner. We could tell a story in which the [+voice] feature on the stem consonant needs to be preserved, resulting in its transmission to the suffix consonant. But this would be a completely *post hoc* explanation. In any case, the most severe problem is yet to come.

3.2.4. The over-generation problem

Once we introduce multiple URs for the lexical entry of a given morpheme, the allomorph selection

analysis predicts that it should be available in any appropriate phonological environment. For the suffix allomorphs of the CV alternation paradigm, this produced the correct result, such that vowel initial suffixes were selected for C-stems, and vice versa. Now that we have introduced lexically specified allomorphs of verbal stems, the allomorph selection analysis predicts that they should be available with any suffix whose form is phonologically compatible. But the allomorphs are in fact only available with the t-suffixes. The result is massive overgeneration.

An example tableau for an s-stem with the present tense suffix is given below. Two output forms are predicted to be fully unmarked and fully faithful, the target form [kas-u] as well as the incorrect form [kasi-ru], leaving no way to choose between the two. There are, of course, technical fixes that one could add to salvage the analysis, such as stipulating which allomorph is to preferred when both yield perfectly unmarked output forms. But to do so would completely defeat the purpose of the analysis, which is to explain rather than stipulate.

	/{kas,kasi}+{u,ru}/	CODACOND	FAITH	ONSET	NOCODA
→	a. kas-u → kas-u				
	b. kas-ru → kas-ru	*!			*
	c. kasi-u → kasi-u			*!	
→	d. kasi-ru → kasi-ru				

Figure 5: Allomorph selection with multiple stem and suffix allomorphs

3.3. Discussion

At first glance, it appears as if I&M's theory derives much with very little, but when the full verbal paradigm including the t-suffixes is examined, the situation turns out to be very much the opposite. The portions of the t-suffix paradigm that appears most amenable to a phonological explanation (the gemination and nasalization cases) do not require allomorph selection at all. On the other hand, those that might have benefited from allomorph selection demand lexical specification of a huge number of stem allomorphs whose relations are completely predictable, and the resulting empirical predictions are extremely poor. Some of these problems can probably be fixed by adding additional technical assumptions, but this would not be desirable, as it would defeat the entire purpose of the analysis. Instead, the conclusion we should draw is that *the alternations in the t-suffixes are not optimizing for phonological markedness*.

Furthermore, any complexity that we managed to eliminate is really being shifted to the lexicon, not eliminated entirely. This same criticism applies to other phonologically-oriented theories such as Davis and Tsujimura (1991) and Nasukawa (2005, 2010), which also enrich the underlying forms of the relevant morphemes in order to account for the alternations. If we were to reject I&M's claim that allomorph selection does not require any additional grammatical machinery, then their argument becomes even weaker. In fact, from a computational perspective there seem to be two additional operations being added to the OT pipeline (Graf, p.c.). First, we must take the Cartesian product of the sets of allomorphs for each morpheme in the input string, producing a set of strings to be evaluated individually. Second, and perhaps more concerning, we must take the union of all resulting tableau in order to select the output form.

In the remainder of this paper, I argue instead for a view of the Japanese verb paradigm as being primarily morphological within the synchronic grammar, and that this view allows us to better understand and explain the patterns it exhibits.

4. Diachronic, experimental, and dialectal evidence

From this point forward, I return to the rule-based analysis as a baseline description of the facts, and again attempt to explain them, beginning with an examination of the history of the verbal paradigm. We will see that the patterns are a remnant of phonological processes that were active in previous stages of the language, but no longer are. This is is the basic reason for the fundamental paradoxical quality of the present day allomorphy, which appears to be phonologically driven, but in light of the previous discussion cannot be. From there, we will consider whether any of the patterns can be considered active in the synchronic grammar, drawing upon evidence from nonce verb experiments and dialectal variation, and see that there is good reason to believe that the system includes productive morphology.[2]

4.1. Diachrony

The relevant aspects about the history of Japanese are well studied, and as far as I can tell, uncontroversial. The facts cited in this section are from Frellesvig (2010), unless otherwise noted.

The CV alternation paradigm dates to Old Japanese, attested in 8th century, with limited written records going back somewhat earlier. At this point, Japanese syllable structure was strictly (C)V, with no coda consonants.[3] Most of the suffixes in this paradigm also date to this time period, including the non-past, passive, causative, conditional, infinitive, and imperative. The long form *-rare* of the potential suffix (homophonous with the passive) was also present, but the short form *-e* was a later innovation. The negative suffix *-(a)na-i* in modern Tokyo Japanese also appears later, taking the place of the older *-(a)n-u*. Crucially, the t-suffixes are entirely absent except for *-te*, which at this stage attaches to the infinitive suffix. This is the ultimate origin of the vowel /i/ which appears in the *-s/k/g* stems today.

There were three regular inflectional classes in this period. The first two correspond to the regular C-stems and V-stems today, and featured the same alternations. Suffixes showing the regular deletion pattern with C-stems and V-stems today are essentially unchanged from then. The third inflection class, the *nidan* or ‘bigrade’ class, merged into the V-stems in most dialects early in the modern period (1600–present). So we see that the deletion pattern in the CV alternation paradigm dates to a time when they could be plausibly motivated by surface syllable structure.

During Early Middle Japanese (800–1200), the language gained new syllable types due to extensive borrowing from Chinese, including syllables with a geminate consonant and a moraic nasal in coda position. Consequently, CV alternation, which was dominant on the surface in Old Japanese, was no longer surface true, and the alternations in the CV alternation paradigm became specific to the verbal stem-suffix boundary.[4] During the same time period, there were synchronic phonological processes which optionally mutated certain sound sequences. These processes are known as *onbin* in traditional Japanese grammar, and included formation of geminates and nasal codas as well as the deletion of intervocalic velar consonants. The t-suffixes paradigm of today originated in this period, in which they induced *onbin* alternations on their stems.[5] All modern t-suffixes derive from *-te* combined with subsequent auxiliaries, and this is why they display exactly these processes. While the alternations subsequently ceased to be productive, they were preserved on verbal stems appearing with t-suffixes. Different dialects of Japanese preserved different *onbin* stems; for example, *katta* ‘bought’ in modern Tokyo Japanese is *koota* in Osaka.

2 From this point on, I use the term “rule” to refer to any productive generalization. Nothing here hinges on having a rule-based analysis, and I remain agnostic regarding the correct choice of theoretical machinery.

3 Vowel initial words were possible. Vowel hiatus was less common, but seems to have existed.

4 This assumes that said alternations are properly analyzed as phonological rules before 800, that is, they were not already morphologized. What is certain is that this happened by the end of Early Middle Japanese.

5 Why only *-te* began to attach directly to the verbal stem is not clear, since the desiderative suffix *-tai* also existed then, and also attached to the infinitive.

In summary, both the CV alternation paradigm and t-suffix paradigm probably were productive phonological alternations in the periods in which they developed. As the phonology changed, the processes ceased to be productive *within the phonology*, but were preserved in the morphology in a limited context. By the end of the Early Middle period, the system already looked much as it does today. Interestingly, there has been comparatively little change to the core features of Japanese phonology since then, a point which Frellesvig notes explicitly (Frellesvig 2010, p. 10). Because of this, there has been relatively little disruption to the system, whether to trigger new alternations or to level old ones.

4.2. Nonce verb studies

Having established the reason why the Japanese verbal allomorphy appears deceptively phonological in nature, I turn to the issue of its synchronic productivity. Given the extreme regularity of the system, it would be surprising if speakers do not form at least some productive rules. In this section I review the results of several experiments designed to test native speakers' ability to inflect nonce verbs. We will see that despite the conclusions of the researchers who conducted these experiments, there is good reason to believe that the system is at least partially productive.

Among nonce verb experiments on Japanese speakers, Vance's (1987, 1991) adult forced choice experiments are the best known. In these experiments, participants are prompted first presented with sentences containing a nonce verb in two inflected forms (non-past and infinitive), which should be enough to establish the stem. They are then asked to select a form for the conditional, volitional, negative, and past forms. Three answer choices are provided, only one of which is appropriate for the intended stem class. Test items include an r-stem (*muru*), an m-stem (*homu*), a k-stem (*hoku*), and a p-stem (*kapu*), the latter of which is unattested. Also tested are the verbs *riru* and *meru*, presented as V-stems (*ri-ru*, *me-ru*), but which also could serve as r-stems (*rir-u*, *mer-u*). He gives participants one of two sets of answer choices, only one of which includes a choice appropriate for an r-stem.

Vance's results show mixed performance, with "analogically correct" responses ranging from about 85% to under 40% depending on the test case. Performance is the best for the r-stem *muru*, middling for m- and k-stems, and worst for the p-stem. The V-stems are also in the middle range, and do much worse if an r-stem answer choice is provided. Vance notes that the results are "consistent with the claim that even morphologically regular Japanese verb forms are stored in the lexicon...however, any number of other interpretations are also possible" (Vance 1991, p156). As far as I can tell, these results are in fact highly suggestive that r-stems are productive, even if no other class is.

Batchelder and Ohta (2000) conducted an elicited response experiment with similar results. They assume from the outset that most C-stem allomorphy is not productive and test only r-stems, which unlike other C-stems do occasionally admit new members. They also test nonce verbs ending in *suru* 'to do', usually considered irregular, on the grounds that nearly all novel verbs are formed in this manner. Finally, they test verbs ending in -iru/-eru presented as V-stems, C-stems, or *suru*-stems. They prompt only for negative and past forms, counterbalanced across participants. "Correct" responses are about 90% for unambiguous r-stems and 85% for *suru*-stems, and about 75% for -iru/eru stems presented as C-stems. Interestingly, they found a meager 52% for -iru/-eru stems presented as V-stems, with the vast majority of errors in the direction of regularization towards r-stems.

Batchelder and Ohta conclude that "the Japanese mental lexicon does not lean heavily on rule-like representations of the various verb paradigms" and "neither does it provide for analogic processes of creation of novel inflected forms based on various existing irregular verbs, like the English possible *bling, blang, blung"*. These conclusions do not seem to me to follow from their results. If anything, the results show that Japanese speakers *do* have productive rules for r-stems and perhaps also verbs formed with *suru*. On the flip side, the results suggest that V-stems are *not* productive. This would explain why

participants regularized in the direction of r-stems. It simply does not follow that because performance is not perfect, speakers are not forming any generalizations.

Neither of the above studies tested children. Klafehn (2013), however, tested both adults and children. Like Vance (1991), he includes nonce verbs with stems ending in r/k/m/i/e and tests several inflected forms for each. Like Batchelder and Ohta (2000), the experiment is elicited response rather than forced choice. Klafehn's results are the most negative of the group. Adult performance averages 36%, ranging from 51% to 0% depending on the stem, and some treated the C-stems as part of an r-stem. For example, participants would inflect *homu* as if the non-past form was *homuru* (with stem *homur-*). Children aged 8-11 performed similarly to adults, with average performance of 26%, and children aged 5-6 produced almost no target forms, with an average of 8%. Klafehn notes that they tended to produce inflected forms using *suru* or *yaru* (which also means 'to do').

Like Batchelder and Ohta, Klafehn concludes that "native speakers of Japanese do not produce the inflected forms of novel verbs in a way that is consistent with a rule-based approach to verbal inflection". However, his results are quite different from theirs, and he does not discuss their experiment, or even cite it. On the other hand, he draws attention to the phenomenon of regularization in the direction of r-stems and *suru*, results which echo the previous studies.

It seems to me that while more research is needed, the evidence from nonce verb experiments suggests that inflection of r-stems and verbs formed with *suru* is productive, while inflection of V-stems and other C-stems is not. If this characterization of the facts is correct, then the puzzle may be restated as follows: if the system is as regular as it appears, then why is it only partly productive?[6]

4.3. Dialectal variation

The last source of evidence for the productive morphological character of the verb paradigm I will consider is from variation among Japanese dialects. De Chene (1985, 2016) argues based on such data that the correct analysis of the verbal paradigm takes the suffix allomorphs which attach to C-stems as basic, with V-stem suffix allomorphs beginning with /r/ derived by a process of r-epenthesis. Others suffixes in the CV alternation paradigm, then, are irregular, whether they involve insertion, deletion, or suppletion. This contrasts with the analysis presented in Section 2, which takes the long form of each non-suppletive suffix to be the basic form, with the short form derived via deletion.

The primary evidence for this analysis is that novel forms of suffixes across dialects overwhelmingly show replacement of the V-stem suffix allomorph by the corresponding the C-stem allomorph with an initial added *r*. In essence, the idea is that speakers occasionally overextend the r-epenthesis rule, leading to morphological change. Crucially, it is not the case that V-stems are simply being reanalyzed as r-stems, as only a subset of suffixes are affected in each dialect. In the Tokyo dialect, for example, only the potential suffix *-e/rare* is affected, being regularized to *-e/re*. This phenomenon is known popularly in Japanese as *ra-nuki* (ra-dropping), though as de Chene argues, this is not the correct characterization of the process. Furthermore, the t-suffixes are completely unaffected by this process; if they were being turned into r-stems then we would see evidence of the *t* becoming a geminate *tt*.[7]

De Chene's analysis is provocative in that it rejects the simplest analysis in favor of one that requires a greater number of exceptions. De Chene (2016) proposes that this is because speakers select the base form first, and only afterward do they attempt to form a rule to derive the non-base form, a principle

6 De Chene (1982) also presents data from a nonce verb inflection experiment. This study is different from those discussed here in that it tests a specific hypothesis about the coinage of new verbs such as *demor-u* 'to hold a demonstration'. I refer the interested reader to de Chene (1985), which reanalyzes the same data and rejects this hypothesis in favor of r-epenthesis.

7 The only place this occurs is in southern Kyushu, where historical V-stems behave as r-stems in all contexts. The implication is the extension of r-epenthesis has gone to completion, leading to reanalysis of the stem. See de Chene (1985) for details.

which he refers to as S*equential Evaluation*. At each step, the choice of base form or rule is chosen to privilege the alternant or alternation with the highest type frequency; de Chene refers to this principle as *Generalized Type Frequency*. The specific claim is that since C-stems vastly outnumber V-stems by type frequency, speakers will always select the C-stem allomorph of each suffix as basic; then, the V-stem suffix allomorphs that show an additional *r* are the most frequent, and so a regular rule will be used to derive them. I will have more say about this proposal in Section 5. For now, it suffices to note that the dialectal data is explained by a theory where speakers may posit, and occasionally overextend a rule.[8]

4.4. Interim summary

Taking into account the full range of evidence discussed so far, there is good reason to believe that Japanese speakers are making productive generalizations about the allomorphy in the verbal paradigm. This is true even if the generalizations are not purely phonological, but morphologically conditioned rules which preserve the effects of historical phonological processes. It also may be the case that the rules that speakers posit do not match the most straightforward linguistic analysis. The question, then, is how to proceed from here. I wish to suggest that an acquisition-focused perspective could shed light on the problem. This is the topic of the following section.

5. Testing the productivity of r-epenthesis with the Tolerance Principle

In this section, I discuss several conceptual problems with de Chene's explanation for why r-epenthesis should be productive in Japanese, and suggest how it may be improved by integrating the Tolerance Principle (Yang 2016). The result is a predictive model of language acquisition which can be tested against corpus data. While it is difficult to gather large enough child corpora, it is also possible to simulate such a corpus using existing adult corpora (Kodner 2019). I conduct such an experiment in order to test the hypothesis that r-epenthesis is only fleetingly productive in modern Japanese.

5.1. Sequential Evaluation and the Tolerance Principle

The Tolerance Principle states that language learners will adopt a rule (=generalization) only if the number of exceptions falls below a certain threshold θ, calculated by type frequency. If there are N candidate items and *e* exceptions then *e* must be no greater than θ, which is N divided by the natural logarithm of N:

(1) $e \leq \theta_N \stackrel{\text{def}}{=} N / \ln(N)$

Thus, we can predict learner behavior by estimating the number of items they know which fall under the proposed rule. For example, if some verbal inflection has a "regular" form and one or more "irregular" forms, we predict that learners will form a productive generalization only if the number of irregular forms they know is much less than the number of regular forms. Ideally this is done using child language corpora such as those in the CHILDES database (MacWhinney 2000), but Kodner (2019) has shown that the distribution of high token frequency items in adult corpora is similar enough that they can be used as a reasonable substitute. This is the approach I pursue here.

What the Tolerance Principle cannot predict is what generalizations learners entertain to begin with. In other words, it does not provide the hypothesis space. This is a problem in for the Japanese verb paradigm, which allows for a great many logically possible analyses. I believe that de Chene's (2016) principle of Sequential Evaluation provides a possible solution. The claim is that speakers do not deal

8 It is worth noting that Ito and Mester (2004) also attempt to explain the regularization of the potential suffix in Tokyo Japanese, but not the regularization of other suffixes in other dialects. De Chene (2016) argues against their approach on the grounds that it involves comparison of the forms of a single suffix, while the correct generalization can only be made by looking at the entire system. On the latter approach can handle the variation across dialects.

with all elements of a morphological generalization simultaneously, but in two discrete steps. First, learners select a base form from the for each alternating morpheme. Then, and only then, they attempt to form a rule to generate the non-base form; any non-base forms that cannot be derived by the chosen rule are listed as irregular forms. This, of course, presupposes that learners have already correctly segmented the relevant word forms. So we really have a three step process:

(2) Steps of Sequential Evaluation (revised version)

1. Segmentation
2. Selection of base form
3. Rule formation

It seems likely that the process may not always be this clean, and that, for example, learners end up revising their segmentation decisions in a way that contradicts the choices made for later steps. It is worth considering whether they will always revisit the later steps as well. For the purposes of the present discussion, I will assume that segmentation is not an issue here, though I will return to this issue in Section 5.3.

De Chene proposes that the decision at steps 2 and 3 is made according to the type frequency of the conditioning contexts. In the case of the CV alternation paradigm, the form of the suffix is controlled by whether the stem is a C-stem or a V-stem, and the former outnumber the latter by approximately 2:1, so learners will select the C-stem alternant as the base form.[9] When they attempt to form a rule to derive the non-base forms, which are the V-stem alternants, they will notice that 3 out of 9 (following the table in Section 2) can be derived by adding /r/. No better generalization can be made, so a rule of r-epenthesis will be adopted.

Because the same type-based calculation is made separately for each step, for each case of alternation, de Chene refers to this principle as Generalized Type Frequency. As far as I can tell, the assumption seems to be that a simple majority (or plurality) suffices. This is obviously in direct conflict with the Tolerance Principle, which sets a much stricter threshold. I propose that the Tolerance Principle should apply at step 3, which directly corresponds the cases where it has been previously applied. In contrast, it is not clear that it should apply at steps 1 and 2, which are not so much generalizations as simple choices among a finite number of alternatives. Accordingly, I adopt de Chene's assumption for base form selection.

5.2. Is r-epenthesis actually productive?

This leads us to the following problem. De Chene takes the scope of comparison for the r-epenthesis rule to be the set of CV alternation paradigm suffixes, which, from the result of step 2, all begin with an initial vowel. But the Tolerance Principle breaks down at small values of N, and there are (depending on one's analysis) only about 9 suffixes. If this value of N is taken to be sufficient (see Schuler 2017 for evidence that it is), then the Tolerance Principle predicts that r-epenthesis will *never* be productive, because the threshold for an N of 9 is only 4, but there are 6 exceptions. On the other hand, if the initial *r* in the V-stem imperative suffix *-ro* is taken to fall under the scope of the r-epenthesis rule (this is in fact what is proposed in de Chene 1985), and the V-stem potential suffix *-rare* is counted separately, then the number of exceptions is only 4, and the rule predicted to *always* be productive.

This is possibly a good result, but it is worth considering an alternative. Assuming that r-epenthesis is a real phenomenon—and the dialectal data speaks strongly in its support—then it is somewhat of a mystery why overgeneralization does not occur more frequently. De Chene (1985) suggests based on an earlier

9 There is no notable semantic difference between C-stems and V-stems, so there should be no correlation between stem class and its frequency of occurrence with different suffixes.

experiment (de Chene 1982) that adults occasionally do this, but I am not aware of any other reports of this behavior, particularly in child speech. Furthermore, it is left unexplained why some dialects have extended the rule to wildly varying degrees, affecting only the potential suffix in Tokyo but going to completion in southern Kyushu.

Suppose, for the moment, that r-epenthesis is in fact only sporadically productive in most dialects of Japanese, past and present. This could be the case if, for example, children adopt the rule before they have robustly learned all suffixes. Even if the rule becomes unproductive later in acquisition (i.e., a U-shaped development trajectory), there is a small chance that they will produce enough overgeneralized forms during this period that this could become input to other children, occasionally leading to establishment of the innovative suffixes (cf. Kodner 2020). This is the hypothesis I wish to pursue.

I propose one further change to de Chene's analysis, which is that the scope of the rule is not the set of CV alternation paradigm suffixes, but rather the set of *inflected verbs* hosting such suffixes. More precisely, the scope is the set of each pair of verbal stem and the immediately following suffix. In addition to the issue of small values of N mentioned above, there is in fact good reason to believe that this is correct, because the actual r-epenthesis rule refers to to *both* the stem and suffix in its context.[10]

(3) $\varnothing \rightarrow$ r / V $_{\text{Vb}}]\ [_{\text{Af}}$ __ V

In this case, it is easy imagine a situation in which r-epenthesis is usually not productive, but where variation in the input could cause it to occasionally reach the threshold for productivity according to the Tolerance Principle. Testing this hypothesis is the topic of the following section.

5.3. A simulation experiment

First, I briefly outline the procedure for simulating the predicted productivity of r-epenthesis using the Tolerance Principle. I run the same experiment using data from two corpora, one written and one spoken. I also discuss some issues with the content and annotation of these corpora, as well as the issue of which suffixes ought to be included in the comparison class. Finally, I present the results of both experiments.

5.3.1. Methods

For this experiment, I followed the procedure described in Kodner (2019). First, a word frequency list is generated from two sets of corpus data. Since the scope of the r-epenthesis rule includes only V-stem verbs and CV alternation paradigm suffixes, only these are considered. Thus, the "words" in this case are pairs of stem and first following suffix. Next, I created simulated input for 100 learners by sampling 400 stem-suffix pairs, weighted by token frequency. Learners are assumed to learn items in an order that correlates strongly with their token frequency so items are sorted by descending frequency in order to approximate their order of acquisition. In effect, the number of items learned serves a proxy for the learner's age. Finally, in intervals of size 25, I calculated the number of learned items which fit the rule and number of exceptions in order to compare them against the value of θ according to the Tolerance Principle. All programming was done using Python and Pandas.

5.3.2. Data sources

I used data from the Balanced Corpus of Contemporary Written Japanese (BCCWJ) (Maekawa et al 2014) and the Corpus of Spontaneous Japanese (CSJ) (Maekawa et al 2004), a spoken corpus. Each was accessed using the Chunagon (Oka et al 2020) search tool, which allows users to query for "short words" (individual roots and affixes, roughly) according to criteria including morphological annotation. I used only the "core" section of each corpus, with a total of 1 million words and 500,000 words, respectively.[11] I

10 See de Chene (2016), Section 2.4.

11 The core subsets of the corpora have higher accuracy of annotations. Additionally, the full corpora are large enough that

analyzed each data set separately rather than attempting to combine them, since any differences in results could be informative.

The annotation scheme used in these corpora is based on traditional Japanese grammar, which treats the verbal morphology differently from the modern linguistic analysis. As a consequence, it is not possible to search directly for the a verb followed by any member of the set of target suffixes. Instead, a custom query is required for each stem-suffix pair. Additionally, there are some irregularities in the annotation scheme. For example, the C-stem potential suffix is not annotated directly, but treated as part of the stem. After downloading the query results, I consolidated and corrected the data using a Python script.

Each corpus has its drawbacks. Written Japanese differs substantially in style from spoken Japanese, especially informal registers. However, the relevant aspects of verbal morphology should be the same, i.e., there are still 9 major CV alternation paradigm suffixes with the same allomorphs and stem classes. The CSJ, on the other hand seems to contain a large proportion of polite verb forms, probably far higher than in child-directed speech. This could distort the data, since the polite suffix *-masu* appears in a middling position in the verb, after the causative, passive, and potential, but before other suffixes.

5.3.3. A problem with the infinitive suffix

Recall from Section 2 that it is unclear whether all suffixes displaying an alternation between /i/ and zero should be analyzed as infinitive suffix followed by an invariant auxiliary. These include the polite suffix *-masu* and the desiderative suffix *-tai*, both of which are plausibly inflectional and extremely frequent. If the infinitive suffix is taken to intervene, this leads to a huge prevalence of infinitive forms in the data, due to the number of suffixes that may follow. Simulations run under this assumption result in a proportion of exceptions to r-epenthesis well in excess of 50% while the Tolerance Principle requires exceptions to be below 20% even for modest values of N, making it impossible for the rule to be productive.

Also recall the discussion from Section 5.1 regarding the uncertainty as to whether we can assume that children have necessarily segmented all word forms completely. Thus, even if the analysis that includes the intervening infinitive suffix is correct, it is plausible that young children will not yet have made this decision. In consideration of these issues, I have decided to run simulations that include the polite and desiderative suffixes, but *not* the infinitive itself. Ideally, bare infinitives and other (lexical) suffixes would be coded separately, but due to practical difficulties in designing the correct queries I have decided to omit them for now.

5.3.4. Results

First, I present data on the number of types of stem-suffix pairs returned by the corpus queries. A total of 4474 types occurred in the BCCWJ, and 1973 in the CSJ. Mirroring de Chene's (2016) statistics, C-stem verbs outnumbered V-stem verbs by a factor of about 2:1. The exact ratio varies somewhat according to the suffix in question, but is always roughly in this range.

conj. class	caus	cond	desid	imp	neg	npst	pass	pol	pot	vol	total
c-stem	169	150	140	77	290	1129	462	307	280	156	3160
v-stem	33	65	77	26	150	500	220	155	n/a	88	1314
total	202	215	217	103	440	1629	682	462	280	244	4474

Table 5: Number of types of stem-suffix pairs (BCCWJ)

Chunagon truncates the results of the relevant queries. While the truncation appears to be consistent, it is unclear how the decision is made, and in any case distorts the frequency information, which is critical to this type of study.

conj. class	caus	cond	desid	imp	neg	npst	pass	pol	pot	vol	total
c-stem	33	49	60	24	155	451	179	207	138	55	1351
v-stem	15	27	35	9	78	197	110	122	n/a	29	622
total	48	76	95	33	233	648	289	329	138	84	1973

Table 6: Number of types of stem-suffix pairs (CSJ)

Note that there is no data for V-stems with a potential suffix. This is because the corpus annotation does not distinguish the two functions of the suffix *-rare*, which is formally ambiguous between potential and passive. There is no way to separate these mechanically, so I coded all as passive. This distorts the data in two ways. First, it *decreases* the number of total types, since a verb stem occurring with both suffixes is counted the same as one that occurs with only one or the other. Second, it *increases* the frequency of passive stem-suffix pairs, which causes them to be "learned" earlier in the simulation. At present, I have no workaround. The effect of this is that the results are skewed in favor of r-epenthesis being productive.

Now, I present the simulation results for each corpus. In the graphs below, colored lines represent the percentage of stem-suffix pairs compatible with r-epenthesis for each simulated learner. The black dotted line indicates the percentage required for productivity by the Tolerance Principle. For the BCCWJ, there is a period in which the learners come extremely close to productivity. One can imagine how, with many more learners, or sampling from a slightly different input distribution, some would cross the threshold.

For the CSJ, the percentage of stem-suffix pairs is well below the threshold set by the Tolerance Principle throughout the simulation. An inspection of the most frequent pairs shows that this is driven largely by excess of polite forms, realizing the concern mentioned earlier. This is shown in the table below.

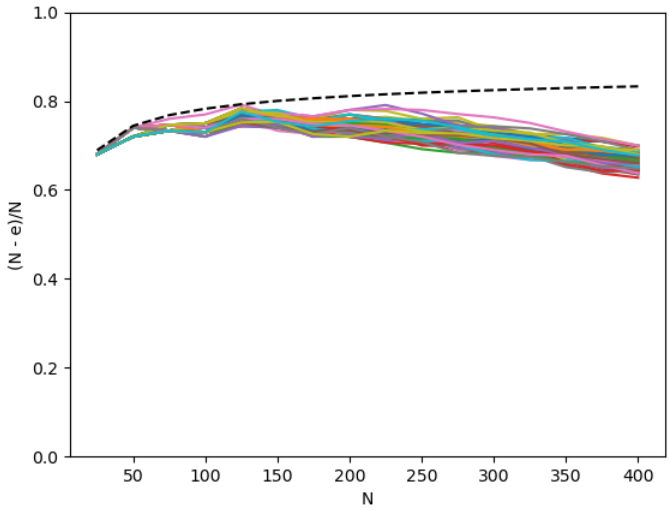

Figure 6: Percentage of learned forms meeting the context for r-epenthesis (BCCWJ)

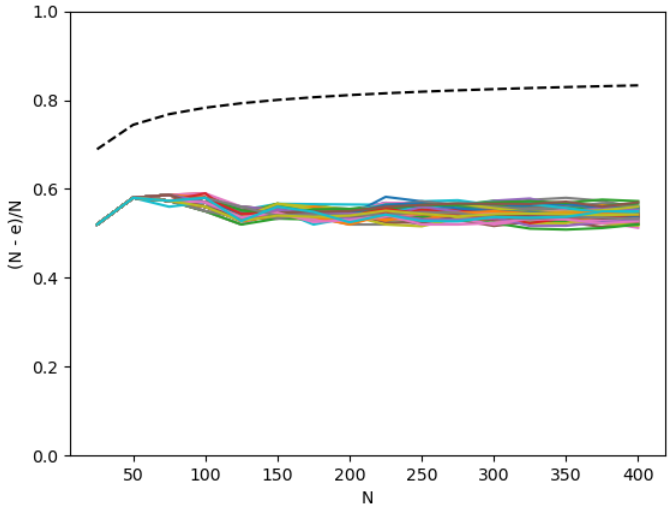

Figure 7: Percentage of learned forms meeting the context for r-epenthesis (CSJ)

Suffix	BCCWJ	CSJ
non-past	171	118
conditional	15	16
passive	70	62
causative	8	6
volitional	17	12
negative	45	55
polite	35	84
desiderative	12	19
imperative	2	3

Table 7: Comparison of top 400 stem-suffix pairs (types) by corpus

These results above provide preliminary support for the hypothesis developed in this section. While the results are highly sensitive to the choice of corpus and to analytic choices regarding segmentation, under reasonable assumptions we do find the evidence in favor of rare, temporary productivity of r-epenthesis in based on data from the BCCWJ. It seems plausible that a spoken corpus with an appropriate register might also shown the same results.

6. Conclusion

In summary, morphology is a necessary part of the solution to the puzzle of the Japanese verb paradigm. Analyses such as Ito and Mester's allomorph selection that attempt to derive all alternations within a phonological theory are doomed to failure because the patterns are not phonologically motivated. While the allomorphy likely was phonological in nature in past stages of the language, this does not mean that it must be so today. Speakers may form morphological generalizations conditioned by a variety of factors. It just so happens that, because of the history of the language, the relevant factors controlling the allomorphy today include both phonological and morphological context.

On the other hand, the notion that speakers make no productive generalizations at is directly contradicted by the available evidence. The experimental evidence paints a complex picture, which suggests that some but perhaps not all stem classes are productive, and the dialectal data suggests that speakers are selecting the based forms of suffixes in a way that contradicts the "simplest" analysis. While many details remain to be worked out, this state of affairs is wholly consistent with the view that the system includes productive morpho-phonological alternations.

An explicitly acquisition-oriented approach could be the key to further progress. The principle of Sequential Evaluation and the Tolerance Principle together make it possible to predict what generalizations speakers will make, and the simulation experiment presented here provides support to de Chene's r-epenthesis analysis. Similar simulations could be designed to predict the productivity of other aspects of the system, such as the stem alternations in the t-suffix paradigm.

Simulations can take us only so far, however. Further experimental work is sorely needed. Previous nonce verb experiments were focused on the question of *whether* speakers could infect novel verbs or not, but it should now be clear that this is the wrong question. Some high priority questions include the following. How do speakers treat nonce verbs with stems types not tested in any of the studies discussed: *w*, *t*, *n*, *b*, and *g*? How does the task type, i.e. forced choice vs. free response, affect the results? Do children ever show evidence of over-extending a rule of r-epenthesis? With the right focus, there is much more that could be learned learn about gap between regularity and productivity by examining the Japanese verb paradigm.

References

Batchelder, Eleanor Olds and Atsuko Ohta (2000). "Rule vs. rote in Japanese verb inflection". In: LACUS Forum (Fullerton, CA). Vol. 26, pp. 55–66.

Bloch, Bernard (1946). "Studies in colloquial Japanese I: Inflection". In: Journal of the American Oriental Society 66.2, pp. 97–109.

de Chene, Brent (1982). "The segmentation of Japanese verbs: Experimental evidence". In: Journal of Japanese Linguistics 8.1-2, pp. 29–62.

— (1985). "-(r) epenthesis and the Japanese verb". In: Journal of Japanese Linguistics 10.1-2, pp. 170–207.

— (2016). "Description and explanation in morphophonology: the case of Japanese verb inflection". In: Journal of East Asian Linguistics 25.1, pp. 37–80.

Davis, Stuart and Natsuko Tsujimura (1991). "An autosegmental account of Japanese verbal conjugation". In: Journal of Japanese linguistics 13.1, pp. 117–144.

Frellesvig, Bjarke (2010). A history of the Japanese language. Cambridge University Press.

Graf, Thomas (n.d.). Personal communication.

Hall, Erin, Peter Jurgec, and Shigeto Kawahara (2018). "Opaque allomorph selection in Japanese and Harmonic Serialism: a reply to Kurisu 2012". In: Linguistic inquiry 49.3, pp. 599–610.

Inkelas, Sharon (2014). The interplay of morphology and phonology. Vol. 8. Oxford University Press.

Ito, Junko and Armin Mester (2004). "Morphological contrast and merger: Ranuki in Japanese". In: Journal of Japanese Linguistics 20.1, pp. 1–18.

— (2015). "Word formation and phonological processes". In: Handbook of Japanese phonetics and phonology. De Gruyter Mouton, pp. 363–396.

Kawahara, Shigeto (2017). Opacity in Japanese. Ms.

Klafehn, Terry (2013). "Myth of the wug test: Japanese speakers can't pass it and English-speaking children can't either". In: Proceedings of the Annual Meeting of the Berkeley Linguistics Society. Vol. 37.

Kodner, Jordan (2019). “Estimating child linguistic experience from historical corpora”. In: Glossa: a journal of general linguistics 4.1.
— (2020). “Language acquisition in the past”. PhD thesis. University of Pennsylvania.
Maekawa, Kikuo, Hideaki Kikuchi, and Wataru Tsukahara (2004). “Corpus of Spontaneous Japanese: design, annotation and XML representation”. In: Reproduction 16.16, pp. 5–5.
Maekawa, Kikuo, Makoto Yamazaki, et al. (2014). “Balanced corpus of contemporary written Japanese”. In: Language resources and evaluation 48.2, pp. 345–371.
McCawley, James D. (1968). The phonological component of a grammar of Japanese. Mouton.
Nasukawa, Kuniya (2005). A Unified Approach to Nasality and Voicing. De Gruyter Mouton.
— (2010). “No consonant-final stems in Japanese verb morphology”. In: Lingua 120.10, pp. 2336–2352.
Oka, Teruaki et al. (May 2020). “KOTONOHA: A Corpus Concordance System for Skewer-Searching NINJAL Corpora”. English. In: Proceedings of the 12th Language Resources and Evaluation Conference. Marseille, France: European Language Resources Association, pp. 7077–7083. isbn: 979-10-95546-34-4. url: https://aclanthology.org/2020.lrec-1.875.
Schuler, Kathryn D. (2017). “The acquisition of productive rules in child and adult language learners”. PhD thesis. Georgetown University.
Tsujimura, Natsuko (2013). An introduction to Japanese linguistics. 3rd. John Wiley & Sons.
Vance, Timothy J. (1987). An introduction to Japanese phonology. SUNY Press.
— (1991). “A new experimental study of Japanese verb morphology”. In: Journal of Japanese linguistics 13.1, pp. 145–166.
Yang, Charles (2016). The price of linguistic productivity: How children learn to break the rules of language. MIT press.