The Productivity of Apophony in Japanese: An Experimental Approach

Mark Irwin
（Yamagata University）
Laurence Labrune
（Université Bordeaux Montaigne）

母音交替は，日本語の複合性を示す形態プロセスの１つである。イネ+カケ→イナカケのように，母音交替では，語末母音が複合語前部要素として現れる場合，その母音が交替するときがある。とはいえ，母音交替は必須ではなく，イネ+コケ→イネコケのように，稲は母音交替を示さずに現れる場合がある。母音交替の有無は，音韻上，形態上，語彙上等の法則に基づいて予測することはできないのである。母音交替が現代日本語でどれほど生産的なのかという問題は，筆者が知る限りこれまで徹底的に調査されたことはないので，この目的のために（臨時語複合の調査の形で）実験テストを設計した。その結果（1）若年話者が使用する現代日本語では，母音交替が利用可能な意味でまだ生産的なプロセスであり，（2）前部要素のモーラ数が標準的な２モーラの母音交替出現率が，前部要素のモーラ数が非標準的な１モーラより高い，という分析が得られた。

1. Introduction: apophony, rendaku and compoundhood

Two morphophonological processes which mark compoundhood are apparent in the modern Japanese language.[1] The first of these is apophony. Here, the final vowel of a lexeme may alter when it appears as the initial element (E1) in a compound. In the Japanese tradition, the phenomenon is known as *ten'on* 転音 'altered sound' or *boinkōtai* 母音交替 'vowel alternation'. For convenience, lexemes may be grouped into three Types by the quality of their final vowel: *e~a* (Type I), *i~o* (Type II) and *i~u* (Type III). Each Type thus contains two allomorphs of the same lexeme, which differ only in their final vowel. The second vowel in each of the three Types, *a~o~u*, occurs only when the lexeme is bound. Such allomorphs are termed +apo and are exemplified in (1). On the other hand, the first vowel in each of the three Types, *e~i*, may occur in both bound or unbound forms since apophony is not compulsory: *ine* 'rice plant', *ki* 'tree' and *tuki* 'moon, month' may, for example, appear bound without exhibiting apophony, as illustrated in (2). These allomorphs are termed -apo. Moreover, the appearance of the +apo or -apo E1 allomorph in a compound – the appearance of the pattern in (1) or (2) below – is not empirically predictable in any phonological, morphological or lexical fashion. Some compounds even display +-apo doublets: *am**a**.tubu* ~ *am**e**.tubu* 'raindrop'.

(1)	*in**e***	+	*kake*	→	*in**a**.kake*
	'rice plant'		'rack'		'rack for drying rice sheaves'
	*k**i***	+	*kage*	→	*k**o**.kage*
	'tree'		'shade'		'bower'
	*tuk**i***	+	*yo*	→	*tuk**u**.yo*
	'moon, month'		'night'		'moonlight'

[1] The additional suprasegmental phenomenon of accent will not be dealt with here. For thoroughgoing accounts of pitch accent in compounding, see McCawley (1968), Kubozono (1999, 2002, 2008, 2018) or Irwin & Zisk (2019).

(2)	*in**e***	+	*koki*	→	*in**e**.koki*
	‘rice plant’		‘threshing’		‘rice threshing’
	*k**i***	+	*me*	→	*k**i**.me*
	‘tree’		‘eye’		‘grain, texture’
	*tuk**i***	+	*nami*	→	*tuk**i**.nami*
	‘moon, month’		‘equivalent’		‘monthly, banal, trite’

The second, more pervasive, morphophonological process marking compoundhood is rendaku, or sequential voicing, an alternation whereby the initial voiceless obstruent of a non-initial element (E2) in a compound may be voiced (hereafter +ren). This is illustrated in (3) for the four voiceless obstruents of Japanese, *k~t~s~h*. A number of different factors are known to (quasi-)systematically dampen or block rendaku. These include Motoori-Lyman’s Law (Motoori 1822; Lyman 1894), the right branch condition (Otsu 1980), and dvandva (Martin 1952). Rendaku is also known to vary markedly across vocabulary strata (Vance 1996, 2007; Irwin 2011:150–153).[2] Despite these rules and constraints, the appearance of the +ren form in a compound is – as with apophony – not wholly predictable in any phonological, morphological or lexical way: rendaku sometimes occurs where it should not, fails to occur where it should, and explanations are not always available for its presence or absence. As with apophony, some compounds even display both +-ren doublets: *mizu.**t**ori ~ mizu.**d**ori* ‘waterfowl’.

(3)	*k ~ g:*	*ha*	+	***k**uki*	→	*ha.**g**uki*
		‘tooth’		‘stem’		‘gums’
	t ~ d:	*hito*	+	***t**uma*	→	*hito.**d**uma*
		‘person’		‘wife’		‘married woman’
	s ~ z:	*ao*	+	***s**ame*	→	*ao.**z**ame*
		‘blue’		‘shark’		‘mako shark’
	h ~ b:	*zu*	+	***h**osi*	→	*zu.**b**osi*
		‘figure, mark’		‘star’		‘bull’s eye’

The two phenomena share a number of properties: they are both non-compulsory morphophonological devices signalling compoundhood by means of an allomorphy manifested either immediately to the left (apophony), or immediately to the right (rendaku), of the element boundary. The behaviour of +apo and -apo allomorphs in apophony is mirrored by that of +ren and -ren allomorphs in rendaku. The behaviour of the marked and unmarked allomorphs of both phenomena is identical: while the unmarked allomorph may occur in either simplex words or compounds, the marked allomorph is restricted to compounds[3]. The foregoing is summarized in Table 1.

[2] For more detailed general discussion of rendaku in general, see Vance (2015), Vance & Irwin (2016) or Irwin & Zisk (2019).

[3] A few exceptions, such as *ba* ‘place’, *zama* ‘appearance’ (derogatory) or *gawa* ‘side’, exist for rendaku.

phenomenon		apophony	rendaku
compulsory		no	no
site		immediately left of element boundary	immediately right of element boundary
unmarked allomorph	shows	no vowel change (-apo)	-ren
	occurs in	simplex words and compounds	simplex words and compounds
marked allomorph	shows	vowel change (+apo)	+ren
	occurs in	compounds only	compounds only

Table 1: Summary of apophony and rendaku

In other respects, however, apophony and rendaku differ. Only a very small proportion of the total Japanese lexicon is subject to apophony: a tiny fossilized subset of 22 native Japanese nouns consisting of 2 moras or fewer. These are listed in (4): see Labrune & Irwin (forthcoming) for further detail.

(4) ***e~a*** *ame~ama* 'rain', *ame~ama* 'heaven', *ine~ina* 'rice plant', *ue~uwa* 'top', *ure~ura* 'end, tip', *kaze~kaza* 'wind', *kane~kana* 'metal', *koe~kowa* 'voice', *sake~saka* 'alcohol', *suge~suga* 'sedge', *take~taka* 'bamboo', *tane~tana* 'seed', *te~ta* 'hand', *tume~tuma* 'fingernail', *hune~huna* 'boat', *me~ma* 'eye', *mune~muna* '(human) chest', *mune~muna* 'roofbeam'

i~o *ki~ko* 'tree, wood', *hi-ho* 'fire'

i~u *kami~kamu* 'god', *tuki~tuku* 'moon'

On the other hand, almost all native Japanese lexemes beginning in a voiceless obstruent are subject to rendaku, with the phenomenon having even established itself, albeit to a limited extent, in the Sino-Japanese stratum. This difference in scale means that, in contrast to the mass of lexemes in which rendaku can manifest itself and create a tacit awareness of its productivity amongst non-linguist Japanese L1 speakers, or even L2 speakers, the same cannot be said for apophony whose tiny number of applicable nouns is a closed set.

This probably explains why apophony has been given so little attention in the linguistic literature. Even if apophony is far less conspicuous than rendaku in contemporary Japanese, it is still nevertheless a pervasive process which cannot be considered marginal: some apophonic compounds occur with extremely high token frequency in the language. The issue of how productive apophony is in the contemporary language constitutes a real issue, which has never been thoroughly investigated to our knowledge.

The issue of productivity, the different notions the term has in the morphological literature and the lack of consensus among scholars about its definition makes it necessary for us to clarify what we mean by 'productivity'. Schematically speaking, it is commonly accepted that productivity has three different aspects: profitability (among all the possible words that can be created by a given process, how many actually exist?); regularity (is the process predictable on the basis of properties attached to the base it applies to?); and availability (is the process still active for native speakers, i.e. are new lexical units still being created, however many there may be?). For additional detail, see Corbin (1987), Plag (2003), Baayen & Lieber (1991), Bauer (2001) and others. A dictionary-based corpus like the one used by Labrune & Irwin (forthcoming) is more suited to investigate the regularity issue. One would need a textual corpus to test the productivity of apophony as a morphological process regularly and actively involved in the production of new words in contemporary Japanese. The third aspect, availability, is one of the two issues that this study will test.

A by-product of our research methodology allows us, in addition, to address the question of redundancy. It is not difficult to find compounds where +apo and +ren co-occur, as in (5). Here, compoundhood is doubly marked on both sides of the element boundary. At what statistical level of frequency does such redundancy occur? Like the productivity of apophony, is its frequency governed by the phonology, lexical stratum or semantics of E1, E2 or the compound as a whole?

(5)

*am**e***	+	*to*	→	*am**a**-do*
'rain'		'door'		'shutter'
*k**i***	+	*tati*	→	*k**o**-dati*
'tree'		'stand'		'thicket'

With this background in mind, this study addresses two research questions:

1. To what extent is apophony still productive (in the sense available) in contemporary Japanese?
2. How do apophony and rendaku interact with each other, and to what extent is their combination or non-combination frequent and predictable?

To this end, we designed an experimental test which takes the form of a nonce compound survey.

2. Methodology: The Survey

To address the question of productivity of the apophonic process, it was decided that the nonce compounds in the survey should include examples with E1s which are known (after Labrune & Irwin forthcoming) to frequently show a +apo allomorph. To address the question of redundancy in compoundhood marking, i.e. the 'double marking' of both apophony and rendaku, it was decided that nonce compounds in the survey should include examples with E2s that appear regularly as +ren (see Irwin 2016). Further, it was decided that the nonce compound survey:

- should consist of two-element nonce compounds only
- that an E1 should belong to the overwhelmingly dominant apophony Type I only
- that a compound's potential meaning be readily intelligible
- that the orthography employed to write a nonce compound not have a potential alternative reading
- that the survey should take no longer than 5 minutes to complete
- in order to maximize native Japanese readings, that Sino-Japanese elements be avoided and all E2s be restricted to the native or foreign strata

Table 2 shows the 28 nonce compounds on which the survey respondents were questioned, listed in the randomized order they appeared in the questionnaire. Every E1 may undergo apophony and appears in (4) above. Each is used at least twice across a maximum of three unique nonce compounds. As shown in the third column, the majority (24/28) of E2s were drawn from the native stratum, the remaining 4 from the foreign stratum. Of the 24 native stratum E2s, 12 have +ren forms and 12 do not fulfil the criteria in (3) and never undergo rendaku. Each E2 appears in the nonce survey

once only. The fourth column shows the English gloss for a given nonce compound and is the semantic interpretation of the authors – others are possible. The rightmost column shows the lemmas of both elements: e.g. *HUNE* represents the lemma of both +apo *huna* and -apo *hune*, while *HATA* represents the lemma of both +ren *bata* and -ren *hata*.

	nonce compound	**E2 type**	**English gloss**	**lemmas**
1.	船旗	rendakuable	ship's flag	*HUNE+HATA*
2.	雨ダンス	foreign	rain dance	*AME+DAᴺSU*
3.	稲病	non-rendakuable	rice blight	*INE+YAMAI*
4.	金罠	non-rendakuable	metal trap	*KANE+WANA*
5.	声虫	non-rendakuable	noise bug	*KOE+MUSI*
6.	手薬	rendakuable	hand ointment	*TE+KUSURI*
7.	風山	non-rendakuable	Mount Wind	*KAZE+YAMA*
8.	酒腹	rendakuable	booze belly	*SAKE+HARA*
9.	爪エステ	foreign	fingernail care	*TUME+ESUTE*
10.	目種	rendakuable	mote	*ME+TANE*
11.	上所	rendakuable	place up on top	*UE+TOKORO*
12.	胸匂い	non-rendakuable	smell of the chest	*MUNE+NIOI*
13.	風マップ	foreign	wind map	*KAZE+MAᵠPU*
14.	酒豆	non-rendakuable	bean eaten with saké	*SAKE+MAME*
15.	船猫	non-rendakuable	ship's cat	*HUNE+NEKO*
16.	爪屋	non-rendakuable	fingernail salon	*TUME+YA*
17.	雨風呂	rendakuable	rain bath	*AME+HURO*
18.	上釘	non-rendakuable	topmost nail	*UE+KUGI*
19.	目脇	non-rendakuable	corner of the eye	*ME+WAKI*
20.	手肌	non-rendakuable	skin of the hand	*TE+HADA*
21.	稲蝉	rendakuable	rice cicada	*INE+SEMI*
22.	胸皮	rendakuable	skin of the chest	*MUNE+KAWA*
23.	声心	rendakuable	core of the voice	*KOE+KOKORO*
24.	風国	rendakuable	windy country	*KAZE+KUNI*
25.	爪花	rendakuable	flower painted on fingernail	*CUNE+HANA*
26.	雨蛇	non-rendakuable	rain snake	*AME+HEBI*
27.	金柱	rendakuable	metal column	*KANE+HASIRA*
28.	酒コーヒー	foreign	coffee with saké	*SAKE+KOOHII*

Table 2: Nonce compounds employed in the survey

The survey was administered in written form to a total of 616 Yamagata University first-year students over an 11-day period between 10–20 April, 2017. Respondents were seated in classes of 30-40 students and given 5 minutes to complete the survey reproduced in Appendix A. No instructions were given beyond reading aloud the introductory sentence. Of the 616 students, 56.7% were male and 43.2% female (0.2% no response). The vast majority of students, 94.6%, were 18- or 19-years-old, with 3.6% 20 or 21, 0.6% 22 or 23 and 1.0% over 24 years old.

3. Results and discussion

Table 3 shows the results of an analysis by apophony rate. Each question has four possible results: +apo, -apo, no answer and 'other' (on which more below). The top 5 questions for each result type are shaded. Overall, the proportion of answers which were +apo was 25.6%. This figure does suggest strongly that the phenomenon is not as unproductive and unconscious as initially hypothesized. Five nonce compounds showed apophony more than 50% of the time: *HUNE+HATA*, *UE+TOKORO*, *HUNE+NEKO*, *ME+HURO* and *INE+SEMI*. Conversely, five nonce compounds showed apophony less than 2% of the time, four of these having a monomoraic E1 (*TE* or *ME*). This will be discussed further below. 'Other' answers made up 7.7% of the total: not a particularly satisfactory result. Two nonce compounds in particular, *INE+YAMAI* and *KANE+WANA*, evinced a particularly egregious number of 'other' answers. The answers *ina.byoo* and *ine.byoo* (split approximately evenly between forms with and without apophony, in both cases containing a Sino-Japanese E2) were common for the former; *kiɴ.wana,* with a Sino-Japanese E1, was common for the latter. Another noteworthy feature is the important disparity between compounds with regards to apophony: from 0.6% to 60.1%. But the most striking result is probably that a non-negligible number of nonce compound (6) exhibit a +apo rate of over 50%, which strongly suggests that apophony is still available in contemporary Japanese as a marker of compoundhood.

		+apo	**-apo**	no answer	other	**total (*n*)**
1	*HUNE+HATA*	60.1%	34.6%	0.0%	5.4%	**100.0% (*616*)**
2	*AME+DAɴSU*	20.6%	76.3%	0.0%	3.1%	**100.0% (*616*)**
3	*INE+YAMAI*	18.0%	48.4%	0.3%	33.3%	**100.0% (*616*)**
4	*KANE+WANA*	24.5%	55.0%	0.0%	20.5%	**100.0% (*616*)**
5	*KOE+MUSI*	27.6%	69.5%	0.2%	2.8%	**100.0% (*616*)**
6	*TE+KUSURI*	1.1%	96.4%	0.2%	2.3%	**100.0% (*616*)**
7	*KAZE+YAMA*	28.7%	64.6%	0.0%	6.7%	**100.0% (*616*)**
8	*SAKE+HARA*	19.8%	68.5%	0.3%	11.4%	**100.0% (*616*)**
9	*TUME+ESUTE*	0.6%	96.8%	0.0%	2.6%	**100.0% (*616*)**
10	*ME+TANE*	1.1%	87.0%	0.3%	11.5%	**100.0% (*616*)**
11	*UE+TOKORO*	53.9%	39.3%	0.3%	6.5%	**100.0% (*616*)**
12	*MUNE+NIOI*	37.5%	54.2%	0.5%	7.8%	**100.0% (*616*)**
13	*KAZE+MAQPU*	9.4%	86.2%	0.2%	4.2%	**100.0% (*616*)**
14	*SAKE+MAME*	30.7%	62.8%	0.3%	6.2%	**100.0% (*616*)**
15	*HUNE+NEKO*	60.9%	36.0%	0.3%	2.8%	**100.0% (*616*)**
16	*TUME+YA*	5.2%	90.3%	0.2%	4.4%	**100.0% (*616*)**
17	*AME+HURO*	51.6%	35.9%	0.0%	12.5%	**100.0% (*616*)**
18	*UE+KUGI*	39.9%	54.4%	0.6%	5.0%	**100.0% (*616*)**
19	*ME+WAKI*	1.5%	92.7%	0.5%	5.4%	**100.0% (*616*)**
20	*TE+HADA*	0.0%	98.5%	0.0%	1.5%	**100.0% (*616*)**
21	*INE+SEMI*	52.4%	34.9%	1.1%	11.5%	**100.0% (*616*)**
22	*MUNE+KAWA*	34.1%	54.9%	0.6%	10.4%	**100.0% (*616*)**
23	*KOE+KOKORO*	20.1%	75.5%	0.2%	4.2%	**100.0% (*616*)**
24	*KAZE+KUNI*	27.1%	64.6%	0.5%	7.8%	**100.0% (*616*)**
25	*TUME+HANA*	5.4%	87.0%	0.3%	7.3%	**100.0% (*616*)**
26	*AME+HEBI*	48.1%	43.5%	0.3%	8.1%	**100.0% (*616*)**
27	*KANE+HASIRA*	31.3%	52.1%	0.3%	16.2%	**100.0% (*616*)**
28	*SAKE+KOOHII*	4.4%	94.2%	0.5%	1.0%	**100.0% (*616*)**
TOTAL		**25.6% (4,409)**	**66.2% (11,421)**	**0.3% (50)**	**7.9% (1,368)**	**100.0% (17,248)**

Table 3: Nonce compound apophony rates

Table 4 shows the data in Table 3 grouped by E1. Here, 'no answer' and 'other' data are ignored and E1 are sorted by average apophony rate (AAR), the most productive first. The AAR across all 15,830 datapoints was 27.9% No clear trends were evident and the accentual issue requires further investigation.

E1	accent	(CV) CV.	(CV) CV.	length	AAR	*n*
HUNE	HL	*n*	*u*	2	63.1%	*1,180*
UE	LH (H)	-	*u*	2	50.0%	*1,155*
INE	HL	*n*	*i*	2	45.8%	*947*
AME	HL	*m*	*a*	2	43.6%	*1,700*
MUNE	LH (L)	*n*	*u*	2	39.6%	*1,113*
KANE	HL	*n*	*a*	2	34.3%	*1,004*
KOE	HL	-	*o*	2	24.8%	*1,187*
KAZE	LH (H)	*z*	*a*	2	23.3%	*1,729*
SAKE	LH (H)	*k*	*a*	2	19.6%	*1,727*
TUME	LH (H)	*m*	*u*	2	3.9%	*1,757*
ME	H (L)	*m*	-	1	1.4%	*1,123*
TE	H (L)	*t*	-	1	0.6%	*1,208*
TOTAL					**27.9%**	***15,830***

Table 4: Average apophony rate analysis by initial element

It is interesting to compare the AAR-based ranking in Table 4 with that in Labrune & Irwin (forthcoming) for the same E1s. This study was based on dictionary data extracted from the reference dictionary *Kōjien* (Shinmura 2006), and can thus be taken as a reliable indicator of the type frequency of apophony within the Japanese lexicon in general.

E1	data reproduced from Table 4	rank	Labrune & Irwin data	rank
HUNE	63.1%	1	95.2%	1
UE	50.0%	2	88.8%	2
INE	45.8%	3	74.6%	4
AME	43.6%	4	79.4%	3
MUNE	39.6%	5	66.7%	6
KANE	34.3%	6	57.8%	9
KOE	24.8%	7	53.1%	10
KAZE	23.3%	8	60.0%	8
SAKE	19.6%	9	72.4%	5
TUME	3.9%	10	66.1%	7
ME	1.4%	11	12.0%	11
TE	0.6%	12	9.3%	12

Table 5:Comparison between experimental nonce word data (this paper) and dictionary data (Labrune & Irwin forthcoming), ranked by E1 apophony rate

Table 5 shows the overall ranking is not overly different and similar tendencies may be observed. Most strikingly, the two top (*HUNE, UE*) and two bottom (*ME, TE*) ranked E1s are identical in both data sets.[4] The main difference between the two rankings lies in the position of *SAKE* and *TUME* which have a higher rank in the dictionary data than

[4] E1s which do not occur in the experimental test have been ignored in calculating the ranking of the dictionary data.

in the nonce word data. The divergence here may be partly explainable by the fact that they were among the four E1s combined with a foreign E2 in the nonce word experiment, a factor shown to dampen the use of the +apo allomorph by Labrune & Irwin (forthcoming).

These results strongly suggest that apophony is not an erratic, fossilized phenomenon in the contemporary language. The convergence in ranking between experimentally tested nonce words and existing dictionary words shows that apophony is still an active morphological process in the Japanese language.

The second Interesting observation to be made is that the type frequency of +apo forms in the lexicon finds reflection in the nonce word +apo rate. The higher the number of +apo forms for a given lexeme in the general Japanese lexicon, the higher the number of +apo readings assigned by speakers in the experimental test. The proportion is consistently around 2/3 for the 6 top items. In contrast, token frequencies in the lexicon do not seem to exert any influence on speaker choice in the nonce word experiment. If they did, then we would expect nonce compounds with a *TE* E1 to exhibit the same AAR as nonce compounds with an *INE* E1, since *TE* and *INE* both possess the same number of +apo compounds (44) in the reference dictionary *Kōjien* according to Labrune & Irwin (forthcoming), despite their AAR being quite different (9.3% vs. 74.6%). The nonce word survey shows that it is the AAR rate which speakers have internalized and are sensitive to, rather than the rough number of +apo compounds for a given E1.

A third noteworthy feature of the results in Table 5 is that the AAR in the experimental test is always lower than that of the dictionary date. We interpret this as signifying that, although still available to speakers as a compounding process, apophony is nevertheless losing regularity (as defined above, and taken as one of the various productivity indices) in contemporary Japanese. But apophony is not dead.

Table 6 shows the data in Table 3 grouped by E2. Here too, 'no answer' and 'other' data are ignored. The top third of Table 5 makes it clear that there is no significant difference in AAR depending on whether the E2 was rendakuable or not. There was, however, as can be seen from the lower third of Table 6, a significant dampening of apophony when the E2 was from the foreign stratum. With E2 element length, indicated in the middle of Table 6, we see a statistically non-significant reduction in AAR the further away an E2 moves from the canonical length of 2 moras (other = 1 mora and 4 mora E2s combined).

E2	**AAR**	***n***
rendakuable	32.9%	*6,707*
non-rendakuable	29.5%	*6,730*
2-moras	33.0%	*9,597*
3-moras	23.6%	*5,038*
other	4.9%	*1,195*
native stratum	31.2%	*13,437*
foreign stratum	9.0%	*2,393*
TOTAL	***27.9%***	***15,830***

Table 6:Average apophony rate analysis by final element

To examine the question of redundancy in compoundhood – to what extent and why a compound might exhibit both +apo and +ren – we whittled the data down to those nonce compounds containing a +ren possible E2. These 6,707 datapoints have an AAR of 32.9% and an average rendaku rate of 59.3% (somewhat lower than we would expect: +ren across the native stratum as a whole is 70.8% according to Irwin (2016)). The figures in Table 7 show that while one marker of compoundhood only was the most common result (45.4%), redundancy (the appearance of both compound

markers) still occurred in just under a quarter (23.4%) of all datapoints. However, these data include compounds with a monomoraic E1 which, as we have seen in Table 4, have extremely low AARs. Further excluding all compounds with a monomoraic E1 yields the data in Table 8. These 5,563 datapoints have a higher AAR of 39.4%, though a somewhat lower average rendaku rate of 55.5% (again, somewhat lower than we would expect). By this analysis, redundancy levels rise to 28.0% of all datapoints.

E2		***n***
no marker (-apo AND -ren)	31.2%	*2,091*
one marker (+apo OR +ren)	45.4%	*3,047*
apophony only (+apo -ren)	*9.5%*	*637*
rendaku only (-apo, +ren)	*35.9%*	*2,410*
both markers (+apo AND +ren)	23.4%	*1,569*
TOTAL	**100%**	***6,707***

Table 7:Redundancy analysis (including monomoraic initial elements)

E2		***n***
no marker (-apo AND -ren)	33.1%	*1,841*
one marker (+apo OR +ren)	39.0%	*2,167*
apophony only (+apo -ren)	*11.5%*	*637*
rendaku only (-apo, +ren)	*27.5%*	*1,530*
both markers (+apo AND +ren)	28.0%	*1,555*
TOTAL	**100%**	***5,563***

Table 8:Redundancy analysis excluding monomoraic initial elements)

Table 9 shows the data in Table 8 organized by nonce compound: i.e. by the 10 nonce compounds appearing in the survey (Table 3) which had both a bimoraic E1 and a +ren possible E2. Each of these compounds have between 514 and 589 datapoints, for an average of 556. A large degree of variation is apparent between those which exhibited majority levels of redundancy (*AME+HURO* and *UE+TOKORO* at 58.8% and 57.0%, respectively) and those which exhibited hardly any (*TUME+HANA* at just 4.2%). A longer E2 appears to promote redundancy. A 3-mora E2 (n = 1,677) exhibits redundancy 37.9% of the time as against with a 2-mora E2 (n = 3,886) which exhibits redundancy 23.7% of the time. A Mann-Whitney U test on this data proved non-significant, however.

	% +apo AND+ren	**E2 length**	***n***
AME+HURO	58.8%	2	*539*
UE+TOKORO	57.0%	3	*574*
INE+SEMI	42.0%	2	*538*
KANE+HASIRA	36.6%	3	*514*
HUNE+HATA	22.0%	2	*583*
KOE+KOKORO	20.4%	3	*589*
SAKE+HARA	17.5%	2	*544*
MUNE+KAWA	13.0%	2	*548*
KAZE+KUNI	10.4%	2	*565*
TUME+HANA	4.2%	2	*569*
TOTAL	28.0%		***5,563***

Table 9: Redundancy analysis by final element length

Our final analysis, shown in Table 10, examines E1 accent pattern. When analysed by the accent pattern of E1, we find again that there is a tendency for those with an HL pattern (n = 2,763) to exhibit redundancy (35.4%), as against those with an LH pattern (n = 2,800; redundancy = 20.6%). While, as we have already seen, HL E1s exhibit a significantly higher level of redundancy than LH E1s, it is also clear that LH E1s exhibit a significantly higher rate of what might be termed 'compound marker immunity', i.e. both -apo and -ren. An E1 with a LH accent pattern is both -apo and -ren 43.7% of the time, as against a figure of 22.3% for an E1 with an HL accent pattern.

	all	**HL E1**	**LH E1**	***n***
no marker（-apo AND -ren）	33.1%	22.3%	43.7%	***1,841***
one marker（+apo OR +ren）	39.0%	42.2%	35.7%	***2,167***
apophony only（+apo -ren）	*11.5%*	12.6%	10.3%	***637***
rendaku only（-apo, +ren）	*27.5%*	29.6%	25.4%	***1,530***
both markers（+apo AND +ren）	28.0%	35.4%	20.6%	***1,555***
n		**2,763**	**2,800**	***5,563***

Table 10: Analysis by initial element accent pattern

4. Discussion & Conclusions

The foregoing analysis has elicited the following:

i. Apophony is still a productive process, in the sense that it is available, in contemporary Japanese as spoken by very young adult speakers (15,830 datapoints).
ii. An E1 of 2 moras in length is more likely to show apophony than an E1 of 1 mora (15,830 datapoints).
iii. Apophony was dampened when E2 was from the foreign stratum (15,830 datapoints).
iv. Limiting the discussion to 2-mora E1s, those with an HL accent pattern seem to be more likely to show apophony (13,499 datapoints), and to show redundancy (5,563 datapoints), than those with an LH(H) accent pattern.

Result (i) is a groundbreaking result as the productivity of apophony has never been tested before.
With result (ii), we note that rendaku shows a similar trend. This was demonstrated in Irwin & Lyddon (2016), who hypothesized that homophony avoidance was the most likely explanation. The finding is also in accordance with Labrune & Irwin (forthcoming), who label this as the PRINCIPLE OF INITIAL MORA INTEGRITY. That study found that while an E1 of 2 moras or longer showed an AAR of .617, a single-mora E1 evinced an AAR of only .137.

As for (iii), Labrune & Irwin (2020; forthcoming) once again came to a similar conclusion, although we must qualify the foregoing by pointing out that the number of datapoints, at only 8, was very small. Moreover, it has long been known that rendaku fails to be triggered at all when the second element is from the foreign stratum, so it should not be surprising to discover that apophony evinces the same tendency.

Finally, (iv) requires further and more detailed investigation. We have no idea as to why an HL accent pattern should favour +apo, and intend to pursue the matter further in a future nonce survey whose compounds will be selected based on accent patterns.

References

Arisaka, Hideyo. 1931. Kokugo ni arawareru isshu no boin kōtai ni tsuite. *Kokugo On'inshi no Kenkyū*. Tokyo: Sanseidō.

Arisaka, Hideyo. 1934. Boin kōtai no hōsoku ni tsuite. *Kokugo On'inshi no Kenkyū*. Tokyo: Sanseidō.

Baayen, Harald & Lieber, Rochelle. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29-5. 801-843.

Bauer, Laurie. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.

Corbin, Danielle. 1987. *Morphologie dérivationnelle et structuration du lexique* (2 vol). Tübingen: Max Niemeyer Verlag.

Frellesvig, Bjarke. 2010. *A history of the Japanese language*. Oxford University Press.

Hasegawa, Yoko (ed.). 2018. *The Cambridge handbook of Japanese linguistics*. Cambridge University Press.

Irwin, Mark. 2011. *Loanwords in Japanese*. John Benjamins.

Irwin, Mark. 2016. *The rendaku database*. In Perspectives on Rendaku: Sequential Voicing in Japanese Compounds, ed. Timothy Vance & Mark Irwin. Amsterdam: John Benjamins

Irwin, Mark & Lyddon, Paul. 2016. Rendaku and Homophony. *Phonological Studies* 19:11-18.

Irwin, Mark & Zisk, Matthew. 2019. *Japanese Linguistics*. Tokyo: Asakura Publishing.

Kubozono, Haruo. 1999. *Mora and syllable.* In Tsujimura, 1999, pp. 31–61.

Kubozono, Haruo. 2002. Prosodic Structure of Loanwords in Japanese: Syllable, Structure, Accent and Morphology. *Onsei Kenkyū* 6.1: 79-97.

Kubozono, Haruo. 2008. *Japanese accent*. In Miyagawa & Saito, 2008, pp. 165–191.

Kubozono, Haruo. 2018. *Pitch accent*. In Hasegawa, 2018, pp. 154–180.

Kusters, Harold H. H. A. 2004a. Vowel alternation in Modern Japanese and sonority. *Phonological Studies* 7. 25-32.

Labrune, Laurence. 2012. *The Phonology of Japanese*. Oxford: Oxford University Press.

Labrune, Laurence & Irwin, Mark. 2020. Apophony, prosodic size and initial mora integrity. *Phonological Studies* 23:3-10, Phonological Society of Japan.

Labrune, Laurence & Irwin, Mark. Forthcoming. *Japanese Apophonic Compounds*. Journal of Japanese Linguistics.

Lyman, Benjamin. 1894. Change from surd to sonant in Japanese compounds. *Oriental Club of Philadelphia*, 1–17.

Martin, Samuel. 1952. Morphophonemics of standard colloquial Japanese. *Language* 28:3 (Part 2, Supplement).

McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.

Miyagawa, Shigeru & Saito, Mamoru (eds.). 2008. *The Oxford handbook of Japanese linguistics*. Oxford University Press.

Motoori, Norinaga. 1822. *Kojiki den*. Edo: Maegawa Rokuzaemon; Osaka: Matsumura Kyūbei; Kyoto: Imai Kihei; Owari: Katano Tōshirō.

Otsu, Yukio. 1980. *Some aspects of rendaku in Japanese and related problems*. In Theoretical Issues in Japanese linguistics, ed. Yukio Otsu & Anne Farmer, pp. 207–227. MIT Press.

Plag, Ingo. 2003. *Morphological productivity. Structural constraints in English derivation.* Berlin / New York: Mouton de Gruyter.

Sakakura, Atsuyoshi. 1966. *Gokōsei no kenkyū*. Tokyo: Kadokawa.

Tsujimura, Natsuko. 1999. *The Handbook of Japanese Linguistics*. Blackwell, Oxford.

Vance, Timothy. 1996. Sequential voicing in Sino-Japanese. *Journal of the Association of Teachers of Japanese* 30: 22-43.

Vance, Timothy. 2007. Reduplication and the spread of rendaku into Sino-Japanese. *Proceedings of the 31st Annual Meeting*, 56-64. Ōsaka: Kansai Linguistic Society.

Vance, Timothy. 2015. Rendaku. In *Handbook of Japanese Phonetics and Phonology*, Haruo Kubozono (ed.), 397–441. Berlin: De Gruyter Mouton.

Vance, Timothy & Irwin, Mark. 2016. *Sequential voicing in Japanese*. John Benjamins.

Appendix: The Survey

下記のアンケートは日本語音韻論研究プロジェクトの一部です。これから記入しする答案・情報等を第三者に提供すること一切はありません。ご協力有難うございます。

次の四つの質問の中から，それぞれ一つだけ選んで丸をつけて下さい。

1. 年齢：　18−19才　20−21才　22−23才　24才以上
2. 性別：　男　女
3. 私は日本語母語話者　です。　ではありません。
4. 私はこのアンケートを　以前，記入したことがあります。
   初めて記入します。

左欄にある熟語は臨時語（その場限りに用いる既成の語ではない）ですが，意味が大体わかる方が多いのではないかと思います。もし，仮に，この単語を「訓読み」するとしたら，どう読むでしょうか。 辞書を使わず，他人と相談せずに，最も適当と思う読み方を右の空欄にひらがなで記入して下さい。

例　家虫　いえむし

1. 船旗 _____
2. 雨ダンス _____
3. 稲病 _____
4. 金罠 _____
5. 声虫 _____
6. 手薬 _____
7. 風山 _____
8. 酒腹 _____
9. 爪エステ _____
10. 目種 _____
11. 上所 _____
12. 胸匂い _____
13. 風マップ _____
14. 酒豆 _____
15. 船猫 _____
16. 爪屋 _____
17. 雨風呂 _____
18. 上釘 _____
19. 目脇 _____
20. 手肌 _____
21. 稲蝉 _____
22. 胸皮 _____
23. 声心 _____
24. 風国 _____
25. 爪花 _____
26. 雨蛇 _____
27. 金柱 _____
28. 酒コーヒー _____