

The phonology of Japanese /r/: a panchronic account

Laurence Labrune, Université Bordeaux 3 & CNRS (Erss, UMR 5610)

ABSTRACT

This study seeks to provide a principled account of the many particularities of /r/ in Japanese. The purpose is to understand how /r/ emerged and developed in Proto-Japanese, and how the conditions of its emergence can shed light on its present phonological behavior. We argue that /r/ primarily developed in Japanese as a default, epenthetic consonant in the intervocalic position, by virtue of an « Emergence of the Unmarked » mechanism, and that the conditions of its development bear on its present day characteristics.

We first provide a description of the phonetic, phonological and morpho-phonological characteristics of /r/ in Japanese, through examination of a large array of empirical evidence. The picture that emerges is that of an unmarked, phonologically empty segment. Taking into special consideration the fact that Old Japanese /r/ did not occur word-initially, that it stood in complementary distribution with zero, and that it exhibits an overwhelmingly high frequency, we propose an analysis whereby /r/ should be interpreted as the intervocalic default realization of a zero consonant at the stage of Proto-Japanese, before becoming a fully contrastive phoneme later on by virtue of a phonologization process. Second, we propose a formal account for the emergence of /r/ in Proto-Japanese within the framework of Optimality Theory. It is shown that the phonological content of /r/ is acquired due to the application of well-formedness constraints (ONSET, ALIGN, MAXIO, DEPIO) as well as that of two sets of unmarkedness constraints (FEATURAL AGREEMENT and HARMONY SCALE), which ensure that the null input is mapped to the least marked output in terms of phonological features. In the final part of the paper, the discussion is broadened to typological and cross-linguistic issues.

1. Introduction¹

/r/ in Japanese stands out as a segment exhibiting many idiosyncratic peculiarities, which make it unique in the phonological system of the language. This paper explores the nature and status of /r/ in Japanese from a panchronic perspective, and aims at providing a principled account of its numerous properties, including (among many others) initial prohibition, overwhelming frequency, and complementary distribution with zero. The main claim of this research is that the phonological characteristics of the rhotic in Modern Japanese provide evidence for its diachronic emergence. We will argue that /r/ primarily developed in Proto-Japanese as a default, epenthetic consonant in the intervocalic position, by virtue of an « Emergence of the Unmarked » mechanism, and that the conditions of its development bear on its present day characteristics.

The paper is organized as follow: section 2 reviews the empirical phonetic, phonological, and morpho-phonological data characterizing /r/ in Japanese at different historical states of the language, which lead to interpret it as unmarked. Section 3 attempts to provide a theoretical account of /r/'s special nature within the framework of Optimality Theory. It will expose the main thesis of this paper, namely, that /r/ originally developed in Proto-Japanese in the context of a zero onset surrounded by two vowels. A conclusion and a discussion on additional typological issues are offered in Section 4.

2. Main characteristics of /r/ in Japanese: an overview

We shall first review the set of characteristics of /r/ in Japanese from a panchronic perspective, which include the following: phonetic variability, overwhelming frequency, complementary distribution with zero, phonotactic restrictions, role in the morphology, instability, phonological inertia and transparency, as well as functional equivalence of *-rV* suffixes with reduplication. Except for section 2.1 - which deals with the phonetic aspects of /r/ - and although the data will sometimes come from the modern language, the discussion below will be mostly centered on Old Japanese (OJ, 8-14th century), and more specifically on the indigenous component of the language. However, most of the characteristics reported below are still valid for Modern Japanese (MJ). Following the Japanese traditional linguistic terminology, the term Yamato is used to refer to what can be considered the native component of the lexicon, including mimetic words, and excluding items borrowed from foreign languages such as Chinese, Indo-European

languages, etc., from the 3rd century A.D on.

2.1. Phonetic realizations

There is only one liquid phoneme in Japanese. Both historical and dialectal data indicate that there is no reason to doubt that it has been so since the times of the most ancient documents (7-8th century) which provide the earliest detailed information on the Japanese phonological system.

The voiced apico-alveolar tap [ɾ] is generally assumed to be the prototypical realization of the liquid consonant in contemporary Japanese. According to Matsuno (1971), [ɾ] should be considered the neutral realization of the rhotic in the language, because its articulation is central compared to other variants. However, /r/ displays a large number of social, geographical or combinatory variants. Outside of [ɾ], the following phonetic (social or regional) realizations are widely attested: [l], [ʎ], [r], [r:], [d], [ɽ], [ɺ].

The apico-alveolar lateral [l] is a common variant, frequent before palatalized vowels (*rya*, *ryu*, *ryo*) and in young women speech (Ohnishi 1987, Tsuzuki & Lee 1992).

Retroflex [ʎ] is also encountered under the same conditions. The short and long apical trills, [r] and [r:] are socially marked variants, characteristic of Tokyo popular male Japanese. The higher the number of trills, the more socially-marked the rhotic will be. The voiced alveolar stop [d] is a combinatory variant which is frequent word-initially in certain dialects, or in children speech. It can also occur word-internally. The retroflex [ɽ] might be encountered initially before /u/, or intervocally in sequences such as /ere/, /ara/, /uru/, /oro/ (Tsuzuki & Lee 1992). The fricative voiced lateral [ɺ] is a combinatory variant occurring before the high vowels /i/ and /u/. It is also the most common realization of /r/ in some Ryukyuan dialects. This wide range of phonetic realizations is undoubtedly relevant to the unmarked status of /r/.

Phonetically, /r/ is also the shortest of all Japanese consonants (Kurematsu, 1997).

In addition, note that whereas the phonetic quality of /r/ is frequently influenced by the surrounding vowels, /r/ itself does not seem to have any significant phonetic influence on the neighboring segments.

2.2. Distribution and frequency facts

Let us now examine what constitutes, in our view, one of the most important issue regarding /r/ in Japanese; the relation which, we claim, exists, between: i) /r/'s distributional gaps; ii) its overwhelming frequency in the language; iii) the fact that it is more likely to occur in the final syllable of words; iv) the fact that it stands in complementary distribution with the zero consonant (empty onset). We consider that these properties are inter correlated and proceed from a single cause. Therefore, any analysis of /r/ in Japanese should be able to account for this state of affairs.

2.2.1. Word-initial rhotic prohibition

/r/ does not occur at the beginning of Yamato *lexical* words². Note that it nevertheless occurs at the beginning of Yamato *grammatical* words (affixes, verbal auxiliaries, etc.) and at the beginning of *non-Yamato* lexical words (Sino-Japanese or Western borrowings).³

It is frequently assumed that in human languages, root morphemes exhibit more *marked* inventory of segments than do affixes and content morphemes (Beckman 1999:52, 183, McCarthy & Prince, 1994, 1995). The positional asymmetry exhibited by Japanese /r/ may then be taken as a clue in favor of its unmarkedness.

2.2.2. Frequency in the language

Although /r/ is absent at the beginning of Yamato lexical words, it is paradoxically the most frequent – or (depending on calculation method) the second most frequent – of all consonants in Yamato words (including mimetics), in OJ and in MJ alike (see Labrune, 1993, for a detailed presentation of various statistical data). For instance, according to Kokuritsu Kokugo Kenkyûjo (1984:25), /r/ is the most frequent of all consonants in the medial position of modern Yamato words, whereas it is the least frequent word initially.

2.2.3. Distribution within words

/r/'s are more likely to occur late in words. Kuginuki (1982) establishes that out of 614 words of the archaic language containing a rhotic, 543 (88,4%) occurrences of those /r/'s appear in the *last* syllable of the word. In other words, the closer to the end of the word, the higher the probability to find a /r/. For Kuginuki (1982), such a distributional pattern makes

sense if one supposes that /r/ emerged relatively late in the history of Japanese. The hypothesis is that /r/ was originally added to the phonemic inventory in order to increase the length of words, which were mostly one or two syllable long in pre-archaic Japanese. Japanese being a suffixing language, these newly added *r*-syllables are expectedly most frequent at the end of words.

2.2.4. Complementary distribution with zero (empty onset)

In OJ, /r/ stands in complementary distribution with the zero consonant.⁴ It is a well-known fact that sequences of two adjacent vowels do not occur in OJ.⁵ Empty onsets (noted as ' below) are licit only word-initially, that is to say precisely where /r/ is forbidden. Another fact deserving special consideration is that /'/ happens to be the most frequent type of onset in the word initial position: words beginning with an empty onset amount to about 25% of the lexicon (Kokuritsu Kokugo Kenkyuujo, 1984, Sakakura, 1966, Otsubô, 1989, Labrune, 1993). This just parallels the frequency and distribution of /r/ which, as seen above, is the most frequent, or one of the most frequent of all, in intervocalic position. The correlation pattern between zero and /r/ can thus be summarized as in the chart below:

(1) Pattern of distribution and frequency of /'/ and /r/ in Yamato Japanese (OJ and MJ).

	Distribution		Frequency
	#_	V_V	
/'/	yes	no	most frequent type of onset <i>word-initially</i>
/r/	no	yes	most frequent type of onset <i>word-internally</i>

There is thus an apparent paradox in the distribution of /r/: although it does not occur word-initially in Yamato lexical words, it is remarkably frequent in other positions, i.e. in the medial position of Yamato lexical words and in the initial position of Yamato grammatical words.

2.3. Co-occurrence restrictions

In OJ, /r/ never co-occurs with itself in a single word, that is, there can be only one /r/ per word. This OCP restriction is especially remarkable in verb stems. Whereas *-ru* is the most frequent verbal inflectional ending in OJ according to Yoshida (1976:87, 101), it is not attested after roots which already contain an /r/ (Kuginuki 1982). Thus, while *kaheru*, *inoru*, *tonakaru*, *musaboru* are well-formed and attested OJ verbs, forms with more than one /r/ such as **kiroru*, **aramaru* or **somoriru* are impossible, and are indeed unattested (but for one exception: *hiroru* “to spread, to widen”).⁶ The same type of co-occurrence restriction is also operative in OJ nouns and other parts of speech.

2.4. Instability

/r/ is the most instable of all Japanese consonants, both diachronically and synchronically. Throughout the history of Japanese, *rV* syllables have frequently undergone syncope (*de aru* > *da* ‘copula’, *karite* > *kate* “provisions”; see also Kishida, 1984, for additional examples) or unexpected paragoge (*kabu* > *kabura* “turnip”, *sippo* / *sippori* “tail” [dialectal] ; see also Labrone, 1998).

2.5. Phonological inertia

A number of facts suggest that the phonological behavior of /r/ is that of a phonologically inert, transparent consonant, lacking phonological content. First, /r/ fails to undergo palatalization, an important phonological process which occurs in the mimetic stratum. Second, it cannot be geminated. Third, unlike most other consonants, /r/ is never the cause of an assimilation process in Japanese. A fourth additional process which we will consider in this section is /r/’s phonological transparency in Ryukyuan dialects, allowing coalescence of two non-adjacent vowels.

2.5.1 Palatalization

According to Hamano (1998), /r/ cannot undergo palatalization in mimetic forms (the data are from MJ, but the phenomenon is also true of OJ).

In mimetics, palatalization is correlated with the semantic value of ‘childishness’ or ‘excessive energy’. It normally occurs on the rightmost coronal consonant of a bisyllabic root (2a). If the rightmost consonant is not a coronal, then the initial consonant will

undergo palatalization, whatever its place of articulation (2b).

The liquid /r/ behaves in an exceptional manner here because it cannot be palatalized when it occurs in the second syllable, so that *noronoro* (2c) does not yield **noryonoryo*, but *nyoronyoro*. Moreover, the presence of /r/ in the root does not block the palatalization of a non coronal in the first syllable, as we see in *kyorokyoro*. So /r/ actually behaves like a non coronal with regards to this palatalization process.

(2)

a.

kasakasa	kasyakasya	*kyasakyasa	‘dry objects scratching each other’
dosadosa	dosyadosya	*dyosadyosa	‘something falling heavily’

b.

pokopoko	pyokopyoko	*pokyopokyo	‘hopping around’
zabuzabu	zyabuzyabu	*zabyuzabyu	‘splashing’

c.

noronoro	nyoronyoro	*noryonoryo	“wiggly and curving movement”
korokoro	kyorokyoro	*koryokoryo	“looking around restlessly”

Note that palatalized /r/’s can be encountered in lexical strata other than mimetics, so that /ryV/ is not an impossible sequence in Japanese. However, the difference between mimetic palatalization and non-mimetic palatalization results from the fact that in mimetics, palatalization is a “feature-sized morpheme” (Mester & Itô, 1989) which can be productively attached to a root under the conditions stated above.

/r/’s resistance to palatalization in mimetics has led Mester and Itô (1989) to the claim that /r/ is underspecified for [coronal] and that it is actually the unmarked sonorant of the system.

2.5.2. Gemination

The second process to which /r/ is transparent in Japanese mimetics, but also in other strata of the lexicon, is consonant gemination as occurring in *-ri* suffixed forms, as reported by Mester & Itô (1989, citing a p. c. by Poser).

When the second consonant of a mimetic root is a voiceless consonant, suffixation of the

adverbial ending *-ri* causes the second consonant of the root to undergo total gemination in case it is a voiceless obstruent (3)a, or partial gemination (i.e. prenasalization) in case it is a voiced obstruent or a sonorant (3)b. However, in case the consonant in question is /r/, neither gemination nor prenasalization can normally occur (3)c.

(3)

a. C2 = [-voiced] -> gemination

yapa	yappari	“after all”
bata	battari	“with a bang”
kaki	kakkiri	“exactly”
goso	gossori	“entirely”

b. C2 = [+voiced] -> prenasalisation

gena	gennari	“to satiety”
simi	sinmiri	“intimately”
boya	bon’yari	“absentmindedly”
yawa	yanwari	“gently”
koga	kongari	“to be nicely brown”
mazi	manziri	“sleepless”

c. C2 = /r/ -> nothing happens

koro	korori	?*korrori	?*konrori	“without effort”
------	---------------	-----------	-----------	------------------

Fully geminated /r/ are also unattested outside of the mimetic lexicon in the Yamato and Sino-Japanese strata, and they are only marginally reported in Western borrowings and in some recent mimetic derivatives (Schourup & Tamori, 1992:137). As for the moraic nasal /N/ + /r/ sequence, it does occur in Sino-Japanese and in Western borrowings, and in a few mimetic forms. However, in such cases, it represents a recent development and is to be phonologically analyzed as a combination of two distinct segments rather than the result of a prenasalization process in the strict sense.

According to Mester and Itô (1989), total gemination of /r/ is impossible because it violates the Nasal Coda Condition, requiring all voiced sonorant codas to be nasal. Partial

gemination is impossible because /r/ is underspecified, and thus has no distinguishable parts available for separate linkage. Schourup & Tamori (1992) have criticized this analysis, arguing that the non-occurrence of palatalization with /r/ in mimetics is best explained by articulatory difficulty. However, one can object that articulatory factors alone cannot account for the many other properties of /r/ in Japanese.

2.5.3. Assimilations

Whereas /r/ is frequently a target for assimilation, it is itself never the source of an assimilation. Numerous examples such as *wakaranai* (MJ) / *wakannai* (MJ) “understand + Negation”, *sakari-na* (OJ) > *sakanna* (MJ) “flourishing”, or *tori-te* (OJ) > *totte* (MJ) “take and...”, where /rV/ is assimilated to a neighboring segment - resulting in total gemination -, can be reported, but there are no example where a neighboring segment is assimilated to /r/ (see also the remark at the end of section 2.1 above). Such phonological inertia, or neutrality, can be argued to follow from /r/’s emptiness at the phonological level. /r/ is never the source for assimilation because there is simply nothing to assimilate in its phonological structure.

2.5.4. Vowel coalescence across /r/

In Ryukyuan dialects, coalescence of vowels over an intervening /r/ or /h/ is reported by Wayne (2000), as in *guru-han* > *goro-hon* “fast” or *pataki-ha* > *patake-he* “to the field”. Wayne also mentions the fact that in the Nakizin dialect, mimetics containing the mid vowels /e/ and /o/ in a light syllable are only attested in roots with a medial /r/ followed by a heavy syllable containing /e/ and /o/, as in *gerengeren* or *zoroozoro*. The vowels /e/ and /o/ appearing in the initial light syllable can therefore be accounted for only by positing a spreading of the second syllable vowel features over /r/. Within a feature geometrical analysis, such facts suggest that /r/, like the laryngeal /h/, lacks a supra-laryngeal node, thus allowing for the vowel features to freely coalesce across it. We take this phenomena as an additional piece of evidence in favor of the structural emptiness and unmarkedness of /r/.

2.6. /r/ epenthesis

2.6.1. Role in the verbal morphology

/r/ also plays an important role in the morphology of Japanese. It is crucially involved in the verbal flexion, as the following data illustrate.

(4) Verbal flexion (MJ)

	C-final base	V-final base
	“write”	“see”
basic form	<i>kak-u</i>	<i>mi-r-u</i>
negative	<i>kak-anai</i>	<i>mi-nai</i>
hypothetic	<i>kak-eba</i>	<i>mi-r-eba</i>
nominal	<i>kak-i</i>	<i>mi</i>
polite	<i>kak-i-masu</i>	<i>mi-masu</i>
passive	<i>kak-areru</i>	<i>mi-r-areru</i>
imperative	<i>kak-e</i>	<i>mi-ro</i>

Observe that before vowel-initial endings (*-u*, *-anai*, *-eba*, etc.), no consonant surfaces in the case of a consonant-ending base, contrary to what happens after a vowel-ending base. According to de Chene (1985) and Mester & Itô (1989), consonant-stem suffixes are basic, and initial /r/ in vocalic-stem suffixes is epenthetic. What is significant here is that the surfacing consonant is precisely /r/. The verbal morphology of OJ provides further arguments for analyzing /r/ as an epenthetic consonant (Labrune, 1996).

Furthermore, /r/ is also extremely frequent at the beginning of several other nominal, adjectival or verbal suffixes such as *-ra* (OJ, MJ) ‘plural’, ‘directional’, *-ra / -ro* (OJ) ‘adverbial’, *-raka* (OJ, MJ) ‘adjectival ending’, *-raku* (OJ) ‘nominalizer’, *-ru* (OJ) ‘passive’, ‘potential’, *-ri* (OJ, MJ) ‘adverbial’, *-ri* (OJ) ‘aspectual auxiliary’, *-re* (OJ, MJ) ‘deictic’, *-reru* (OJ) ‘passive, potential, honorific’, *-ro* (OJ, MJ) ‘imperative’, *-ro* (OJ) ‘particle (of obscure function)’, etc.

2.6.2. First syllable reduplication in OJ

According to Unger (1976:32), there was a productive process in early Japanese whereby the first syllable of a verb root was doubled to form a new root with extended meaning, as

in *tuke-* “attach” / *tutuke-* “continue”, *yam-* “get sick” / *yayam-* “be troubled”. Unger adds to the list *wogam-* “bow” / *worogam-* “revere”, *utape* “appeal” / *urutape-* ‘id.’, *name-* “line up” / *narabe-* ‘id.’, and *i-* “mint, cast” / *ir-* “scorch”. He then argues that exact reduplication of an *r* beginning initial syllable should be reconstructed for these forms, i.e. **rogam-* / **rorogam-*, **rutape-* / **rurutape-*, **rabe-* / **rarabe-*, **ri-* / **riri*.

According to him, /r/ later changed to zero before /i/, to /w/ before /o/ and /u/, and to /n/ before /a/ word-initially. As a result of these rules, no OJ words begins with /r/.

We claim that an analysis exactly opposite to Unger's proves to be more persuasive. First, we propose that the pair *name-* / *narabe-* does not belong to the pattern under discussion, and be left apart (even in Unger's analysis, it is not clear why /r/ would change to /n/ before /a/ rather than to zero). Following Unger, we consider that *worogam-*, *urutape-* and *ir-* were derived through reduplication of the initial syllable, but we posit instead that the original roots were **ögam-*, **utape-*, and **ir(i)* (and not **rogam-*, **rutape-* and **riri*). Reduplication of the initial vowel in each form produced a sequence of two onsetless syllables (**öögam-*, **uutape-*, **iir-*). In order to break the hiatus, /r/ was inserted after the first vowel, yielding **örögam-*, **urutape-*, and **ir(i)-*. The /r/ appearing in the reduplicant is therefore best analyzed as an intrusive, epenthetic segment.

Note that what Unger transcribes as *worogam-* phonemically corresponds to /örögam-/ (with /ö/ denoting a round, centralized vowel). According to a tradition dating back to Fujiwara no Teika (13th century), /ö/ preceded by an empty onset has been transcribed with a letter belonging to the w- column in the Japanese *kana* system, i.e. *wo* in roman transcription, but there is absolutely no evidence that a labio-velar approximant ever occurred before the vowel in this position. Consequently, the root for “to bow” is phonologically **ögam-*, with an onsetless initial syllable, and not **wogam-* as posited by Unger.

2.7. Functional equivalence with reduplication

A final piece of morphological evidence in support of the unmarked status of Japanese /r/ can be invoked. As claimed by Labrune (1998), there exists in Japanese a morpho-lexical class of words ending with a *r* + vowel suffix, which share the same type of linguistic values as reduplicated words, as *baka-ra-sii* / *baka-baka-sii* “foolish”, *toro-ro* / *toro-toro* “yam”, *ware-ra* / *ware-ware* “we”, *aka-ra* (OJ) / *aka-aka* “reddish”, etc. (unless

otherwise specified, all examples are MJ). Examples involving partial reduplication, such as *simi-mi* (OJ) / *simi-zimi* / *simi-ra* (OJ) “fully” are also attested. It can be argued that such a functional correspondence between reduplication and *-rV* augmentation reflects some ancient morphological process which involved some kind of iconic lengthening by means of unmarked segmental material. Reduplication, which copies segmental material from the base, appears as an unmarked lexical means to lengthen a base, whereas *-rV* suffixation (where V is either a vowel copied from the base, either a default segment) can be analyzed as some unmarked segmental means to lengthen the base. The choice of /r/ to express such iconic lengthening, in parallel with reduplication, can be taken as evidence for the unmarked status of the rhotic in the phonological system of the language, and as a case of emergence of the unmarked. The status of /r/ in the cases discussed in Labrune (1998) is thus comparable to the status of other unmarked segments involved in partial reduplication or reduplication with fix segmentism, for example /ʔ/ in Makassarese as described by McCarthy & Prince (1994).

3. Accounting for the phonological nature of /r/ in Japanese

On the basis of the evidence reviewed in section 2, we will claim that /r/ is unmarked, and that it originally developed in Proto-Japanese as a default, epenthetic consonant in the intervocalic position.

3.1. /r/ as an unmarked consonant

In a well-known paper dealing with palatal prosody in Japanese mimetics in relation to feature predictability and underspecification, Mester and Itô (1989) claim that /r/ is the unmarked sonorant of Japanese, that it is underspecified for the feature Coronal, and that, in a feature-geometric representation, “the underspecified *r* is nothing but a root node with internal major class specifications”. While their analysis is based on modern Japanese examples, one can assume that it is also valid for OJ since the data they rely on comes mostly from Yamato Japanese, the native stratum of the language.

In this paper, we will go one step further and argue that /r/ is totally featureless at the abstract level, even for the major class features, because all the information contained in its phonological representation is predictable. The analysis presented below focusses on the diachronic development of /r/ at the Proto-Japanese stage, but it is important to keep

in mind that our basic assumption is that /r/'s present idiosyncratic properties trace back to the circumstances of its development.

The range of facts reviewed in section 2 provides rich and various evidence that Japanese /r/ is phonologically unmarked. Take, for instance, the variability of its phonetic realizations (2.1). This arguably results from the fact that some features are not fully specified at the phonological level, thus allowing for a non-distinctive, optional surface variation. Features such as [±continuant], [±lateral], [retroflex] are supposedly acquired at a later stage, depending on sociolinguistic or contextual factors. Notice that all the features characterizing /r/ at the surface level appear as relatively unmarked: [+sonorant], [-nasal] and [+voiced] are unmarked in an intervocalic environment, whereas [coronal] can be seen as the unmarked place of articulation (Paradis & Prunet, 1991).

High frequency (2.2.2), too, can be taken as a clue to the relative unmarkedness of a segment (Greenberg, 1966). Concerning /r/'s instability (2.4) as well as its phonological inertia, we assume that the range of phenomena involved provides direct evidence for /r/'s lack of intrinsic phonological substance.

Recall also that, as already noted, two exceptions to the dissimilation principle which forbids identical consonants to co-occur in a mimetic root involve precisely /r/, in *rerorero* and *rorirori* (see footnote 3). Such examples are more than anecdotal exceptions: they are significant because they constitute one more illustration of the fact that /r/ behaves as a transparent, invisible segment. The OCP normally applying in mimetics fails to detect the succession of the same consonant in the case of /r/, because /r/ is empty in phonological content.

The behavior of /r/ in the phonetics, phonology and morpho-phonology of Japanese is thus characteristic of what any theory of phonology, whether structuralist, generativist or OTist, would recognize as an unmarked, default segment.

3.2. /r/ as a default epenthetic consonant

Our interpretation of /r/'s idiosyncratic behavior is that diachronically, /r/ emerged as the result from the filling of an intervocalic empty position with unmarked features. In other words, /r/ can be viewed as the phonologization of an empty onset in the intervocalic context.

The most direct evidence for this claim has been presented in 2.6. and 2.7. above, where we saw that /r/ appears as the default epenthetic consonant in a number of derivational patterns. Remember that in Yamato Japanese, /r/ always occurs in the V_V context (2.2). In a procedural, serialist model of phonology, one would consider that the features characterizing /r/ have been acquired, first, through propagation of neighboring features by means of assimilation. /r/ acquires its voicedness, its non-nasality, and its sonorancy because it is surrounded by two voiced, non-nasal, sonorant segments (i. e. vowels). When neighboring segments cannot provide /r/ with adequate melodic specification, more specifically, when they cannot provide /r/ with a place of articulation (which is actually always the case, since /r/ only occurs in an intervocalic context, and vowel and consonant places of articulation are, we assume, substantially different), the featural content of /r/ results from general or universal unmarkedness requirements. So /r/ is a coronal, because coronal is the unmarked place of articulation for consonants.

3.3. /r/ as a consonant of (relatively) recent development

As mentioned in 2.2.3 above, Kuginuki (1982), in an attempt to account for the fact that /r/ is more likely to occur late in words, has proposed the idea that /r/ developed relatively late in the history of Japanese. Another argument supporting his hypothesis, already discussed in section 2.3 is the fact that only one /r/ can occur in an OJ word. According to Kuginuki, this shows that /r/ did not exist in the primitive language, and that it was later added as a suffix to increase the distinctive power of words⁷.

We will follow Kuginuki's basic insight that /r/ appeared relatively late in the history of the Japanese language. But our analysis goes one step beyond Kuginuki's, because we assume that there exists a connection between /r/'s late development, the fact that it stands in complementary distribution with /ʀ/ in OJ, and its unmarkedness.

Its late development explains why, in contradiction with what would be assumed to be characteristic of a marked segment, rather than that of an unmarked and transparent one, /r/ tends to occur only once in OJ words, and is subject to positional restrictions. This presumably follows from the conjunction of three independent facts: first, /r/ appeared late in the inventory; second, Japanese is a suffixing language; third, Japanese words rarely contain more than one or two affixes.

3.4. A constraint-based analysis

In its general outlines, the analysis presented so far has been somewhat intuitive and pre-theoretical. We will now consider how the basic scenario that /r/ was originally, at the Proto-Japanese stage, a segment empty in feature content, resulting from the default filling of an empty onset slot in the intervocalic position, can be captured within the framework of Optimality Theory (Prince & Smolensky, 1993).

We assume that the phonological features of the Japanese phonological system are the following: Voiced, Sonorant, Nasal, Continuant, Labial, Coronal, Dorsal. We also assume that place features such as [coronal], [labial], and [dorsal] are one-valued (privative), whereas laryngeal and class features such as [voiced], [sonorant], [continuant] and [nasal] are two-valued (binary), thus [+voiced] contrasts with [-voiced], and so on.

In order to account for the output forms containing a full consonant in place of zero at the stage of Proto-Japanese, we posit the following two sets of constraints: i) well-formedness constraints (which include faithfulness constraints as well as structural constraints); ii) featural unmarkedness constraints.

The first set includes the following four general and familiar constraints:

(5) ONSET: syllables should have an onset

(6) ALIGN(PrWd, L, stem, L): the left edge of the prosodic word coincides with the left edge of the stem

(7) MAXIO: input segments must have output correspondents (no deletion)

(8) DEPIO: output segments must have input correspondents (no epenthesis)

The relative ranking of these four constraints is provided in (9).

(9) ALIGN >> ONSET >> MAXIO >> DEPIO

These constraints, and the proposed hierarchy, account for the fact that an empty onset will be filled in when occurring word-internally, as the following tableau demonstrates:

(10)

Input: V V	ALIGN	ONSET	MAXIO	DEPIO
1. V		*	*!	
2. VV		**!		
3. CVCV	*!			**
☞ 4. VCV		*		*

Tableau (10) shows that candidate n° 4, i.e. VCV, is the optimal output. The relative position of ONSET, which outranks MAXIO and DEPIO in the hierarchy is crucially what determines the choice of VCV over candidate n° 2 VV.

Next, we have to account for the feature content of the epenthetic C in VCV (candidate 4). This will be done by invoking two subgroups of unmarkedness constraints: FEATURAL AGREEMENT and HARMONY SCALE.

What we call FEATURAL AGREEMENT refers to a syntagmatic, contextual type of constraints prohibiting consecutive adjacent segments which do not share the same value for a given feature, hence $*[\alpha \text{ Feature}] [-\alpha \text{ Feature}]$. This type of constraint basically encodes an assimilation process. FEATURAL AGREEMENT resembles what Itô & Mester (1998) call “Sequential Unmarkedness”.

In our analysis, FEATURAL AGREEMENT subsumes the following four individual constraints:

(11) FEATURAL AGREEMENT: $*[\alpha F] [-\alpha F]$

a. $*[\alpha \text{Sonorant}] [-\alpha \text{Sonorant}]$

b. $*[\alpha \text{Voiced}] [-\alpha \text{Voiced}]$

c. $*[\alpha \text{Nasal}] [-\alpha \text{Nasal}]$

d. $*[\alpha \text{Continuant}] [-\alpha \text{Continuant}]$

Since there is no evidence for the relative sub-hierarchy of these four FEATURAL AGREEMENT constraints, we shall leave them unranked with respect to one another.

The second subgroup of unmarkedness constraints which has to be posited follows the universal Harmony Scale proposed by Prince & Smolensky (1993: 180 *sq*). It results from the relative ranking of a set of Featural Unmarkedness Constraints, shown in (12):

- (12) HARMONY SCALE:
*PL/DORSAL, *PL/LABIAL >> *PL/CORONAL

HARMONY SCALE constraints have to be ranked below the FEATURAL AGREEMENT constraints, otherwise the emerging default consonant would be a segment with no supra-laryngeal place of articulation, such as /h/ or /ʔ/, rather than /r/.

The general ranking of the constraints is provided in (13):

- (13) Constraint hierarchy:
ALIGN >> ONSET >> MAXIO >> DEPIO >> FEATURAL AGREEMENT >> HARMONY SCALE

Notice that none of the constraints invoked here are language specific. They are all very general constraints, based on widely recognized, theory-independent phonological principles.

Tableau (14) below illustrates how /r/ is indeed the most harmonic output for zero, given the constraints and the hierarchy proposed above. In order to keep the tableau to reasonable size, we will adopt the following conventions. First, violations to FEATURAL AGREEMENT will be marked by indicating which feature(s) cause(s) the constraint-infringement in the FEATURAL AGREEMENT column. The presence of a feature in the column is thus equivalent to the presence of a *, and a blank indicates that no violation occurs. Second, we only considered the most serious competitors to /r/ as an epenthetic segment, given what is generally assumed to have constituted the phonemic system of OJ, and with the addition of a couple of other consonants such as /h/ and /ʔ/ which appear as potential rivals to /r/ as default epenthetic segments.⁸ For any given input consisting of a sequence of two identical or different vowels (we chose *aa* here in order to illustrate the case) a number of eighteen serious candidates can be considered:

(14)

Input : a a	ALIGN	ONSET	MAXIO	DEPIO	FEAT-AGREEMENT	HARMONY SCALE		
						*Dors	*Lab	*Cor
1. a		*	*!					
2. aa		** !						
3. rara	*!			**			**	
☞ 4. ara		*		*			*	
5. ada		*		*	Son! Cont		*	
6. ata		*		*	Son! Vcd Cont		*	
7. ana		*		*	Nas! Cont		*	
8. asa		*		*	Son! Vcd		*	
9. aza		*		*	Son!		*	
10. apa		*		*	Son! Vcd Cont		*	
11. aba		*		*	Son! Cont		*	
12. ama		*		*	Nas! Cont		*	
13. aka		*		*	Son! Vcd Cont	*		
14. aga		*		*	Son! Cont	*		
15. aha		*		*	Son! Vcd			
16. awa		*		*		*	*	
17. aya		*		*		*		
18. aʔa		*		*	Son! Vcd Cont			

Candidates n° 3, 2 and 1 are eliminated straight away because contrary to the rest of the candidates, they display, respectively, one fatal infraction to ALIGN (n°3), two fatal infractions to ONSET (n°2) and one fatal infraction to MAX (n°1). All other candidates, from number 4 to 18, score identically with respect to these three top constraints: they all incur one violation to ONSET and one to the fourth constraint, DEP. Looking at the next constraint, FEATURE AGREEMENT, we see that only three candidates satisfy the featural agreement requirements: n° 4 /ara/, 16 /awa/ and 17 /aya/. All three display a non-nasal sonorant as the epenthetic internal segment, that is, a segment which shares the features [+Sonorant], [+Voiced], [-Nasal], and [+Continuant] with the surrounding vowels.⁹

If we now consider the HARMONY SCALE constraints, we see that candidates n°4 /ara/ is the optimal output. This is because /r/ is a Coronal, that is to say the least marked place of articulation after Labial and Dorsal. /w/ being a labio-dorsal and /y/ a dorsal, their place of articulation is more marked than that of a coronal such as /r/, and candidates n° 16 and

17 /awa/ and /aya/ thus loose over /ara/.¹⁰

Notice that /h/ and /ʔ/ (n° 15 and 18) would be better candidates than /r/ insofar as they do not have a place of articulation. However, they violate the FEATURAL AGREEMENT constraints, and therefore loose over /r/.

Candidate n° 4 /ara/ thus appears as the optimal output for the input /aa/.

In the light of the analysis provided here, the process underlying the emergence of /r/ in Proto-Japanese can be seen as an instance of “The Emergence of The Unmarked”, in the sense of McCarthy & Prince (1994).

3.5. Lexicon optimization and /r/ phonologization

In our view, MJ /r/ is the result of a two-stage diachronic process. Stage 1 corresponds to a step where zero in the input is mapped to [r] in the output, according to the constraint hierarchy shown above. At stage 2, zero in the input has been lost and replaced by /r/, by virtue of Lexicon Optimization (Prince & Smolensky, 1993: 192, Kager, 1999:32). New learners, who never hear [aa], have come to identify /ara/ as the underlying form for surface [ara] because /ara/ incurs the fewest violations to faithfulness, causing /r/ to be added as a new segment in the phonological inventory of Proto-Japanese. In other words, what we have at this stage is a phonologization process.

Our claim is that the properties which are exhibited by MJ /r/ result from the conditions of its emergence in the language¹¹. That is, modern /r/ has retained something of its “ancestor” zero, as can be inferred from its status and behavior in present-day Japanese.

4. Conclusion and further issues

This study has been primarily an attempt to provide a principled account of the many particularities of /r/ in Japanese from a panchronic perspective. The purpose is to understand how /r/ emerged and developed in Proto-Japanese, and how the conditions of its emergence can shed light on its present phonological behavior.

First, a description of the phonetic, phonological and morpho-phonological characteristics of /r/ in Japanese has been presented, through examination of a large array of empirical evidence. The picture that emerges is that of an unmarked, phonologically

empty segment. Second, we have proposed a formal account for the emergence of /r/ in Proto-Japanese within the framework of Optimality Theory. The basic idea is that /r/ was originally a segment empty in feature content at the input level. We have shown that the phonological content of /r/ is acquired due to the application of well-formedness constraints such as ONSET and ALIGN, as well as two sets of unmarkedness constraints, FEATURAL AGREEMENT and HARMONY SCALE, which ensure that the empty input is mapped to the least marked output in terms of phonological features. /r/ should therefore be interpreted as the intervocalic default realization of a zero onset at the stage of Proto-Japanese, before becoming a fully contrastive phoneme later on by virtue of a process of phonologization.

This study also has important typological implications, and a number of issues deserve future study and analysis. Strikingly enough, some of the particularities which have been encountered with Japanese /r/ can be found in a number of other languages of the world which are not necessarily closely related. Initial prohibition and high frequency figure among the most recurring properties of rhotics cross-linguistically. We were able to collect a list of more than 60 languages which, like Japanese, ban the occurrence of a liquid, typically a *r*-sound, in the initial position of native independent words.¹² It is particularly intriguing to note that several proto-languages including Proto-Altaic, Proto-Indo-European, Proto-Austronesian, Proto-Dravidian or Proto-Australian lack /r/ in the initial. This might suggest that /r/ could have developed relatively late in some languages inventories, following a scenario closely resembling the one we posit for Japanese, but this is only a tentative hypothesis that deserves further attention in future research.

In parallel to distributional asymmetries, another characteristic which we observed in Japanese, namely /r/'s outstanding frequency, is also well attested across languages. In a rather significant number of languages, the rhotic is the most frequent (French, Portuguese, Korean, Susu, Unya, etc.), or among the most frequent (English, Spanish, Latin, Italian, German, Mandingue, etc.) of all consonants. Additional data concerning phonemic statistics in various languages should be gathered in the future. They will undoubtedly confirm *r*'s relatively high frequency in the languages in which it exists. Much work thus remains to be done regarding the peculiar nature of liquids from a

typological point of view. It is clear that it is not only the gaps in the distribution which should interest the linguist, but also the fact that a consonant, which cannot occur in a given position, is the most frequent of all or among the most frequent of all consonants in some other position. This type of paradoxical pattern seems to be quite rare outside of liquids.¹³ To our knowledge, it has not received much attention from linguists. However, such cross-linguistic evidence constitutes a strong argument to contend that distributional asymmetries and outstanding frequency should definitely be added to the set of properties characterizing the rhotics from a general linguistic point of view, although these two traits generally do not figure among the list of possible generalizations about /r/ (see for example Lindau, 1985, Wiese, 2001, or Walsh-Dickey, 1997).

The frequently observed complementary distribution between /r/ and zero is also an issue which should deserve more careful examination, since it seems to occur in many languages outside of Japanese. The cases of English and Groningen Dutch discussed in van Oostendorp (2001) and Giegerich (1999) are particularly interesting. Van Oostendorp (2001: 115) mentions the fact that in Groningen Dutch, “/r/ alternates between [r] and zero, because of its placelessness”. He also cites Humbert (1996), for whom “/r/ [...] has a relatively unspecified vowel-like structure which could be filled in by the environment”. Van Oostendorp (2001:114) also mentions Giegerich (1999), who writes that “in English non-rhotic dialects there is an empty segment /0/ which is realized as [ə] before consonant or pause, or as [r] before a vowel. For Giegerich (1999:196), in English, “[r] and zero are surface realizations, in complementary distribution, of the same melodic underlier”.

It is important to document these issues more thoroughly in the future and to consider the reasons why such similar patterns of rhotic distribution, frequency, and relationship with zero frequently occur across languages. An open question is whether an analysis similar to the one developed in this paper in order to account for the distributional asymmetries of /r/ in Japanese might be extended to other languages.

REFERENCES:

- Anderson, Gregory D. S. 1997. "Burushaski phonology". In *Phonologies of Asia and Africa*, vol. 2, Alan S. Kaye (ed). 1021-1041. Winona Lake (Indiana): Eisenbrauns.
- Austin, P. 1981. *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.
- Beckman, J. 1999. *Positional Faithfulness: An Optimality Theoretic Treatment of Phonological Asymmetries* [Outstanding Dissertations in Linguistics]. New York and London: Garland.
- Benveniste, E. 1939. «Répartition des consonnes et phonologie du mot ». *Travaux du Cercle Linguistique de Prague* 8 : 27-35.
- Bhat, D.N.S.1974. "The phonology of liquid consonants". *Working Papers on Language Universals* 16: 73-104.
- Bird, Steven. 1999. "Dschang syllable structure". In *The Syllable: Views and Facts*, H. van der Hulst & N. Ritter (eds), Mouton, pp. 447-476.
- Breen, G. 1981. *The Mayi Languages of the Queensland Gulf Country*. Canberra: Australian Institute of Aboriginal Studies.
- Bril, I. 2000. *Dictionnaire nêlêmwa-nixumwak français anglais* (Selaf 383). Louvain / Paris : Peeters.
- Canu, G. 1981. « Le mò:re ». In *Les langues dans le monde ancien et moderne, Afrique subsaharienne, pidgins, et créoles*, J. Perrot (ed) : 111-122. Paris : Editions du CNRS.
- Caprile, J-P. 1981. « Le Mbai-moissala ». In *Les langues dans le monde ancien et moderne, Afrique subsaharienne, pidgins, et créoles*, J. Perrot (ed) : 243-257. Paris : Editions du CNRS.
- Chafe, W. L. 1979. "Caddoan". In *The Languages of Native America. Historical and Comparative Assessment*, L. Campbell & M. Mithun (eds), 213-235. Austin: University of Texas Press.
- Chamorro, Claudine. 2003. *Parlons Purepecha*, Paris: L'Harmattan.
- Charachidze, G. 1989. "Ubykh". In *The Indigenous Languages of the Caucasus vol. 2*, B. G. Hewitt (ed), 359-459. Delmar / New York: Caravan Books.
- Clements, G. N. & Hume, E. 1995. « The internal organization of speech sounds », in *The Handbook of Phonological Theory*, J. Goldsmith (ed): 245-306. Cambridge: Cambridge University Press.

- Dayley, J. 1989. *Tümpisa (Panamint) Shoshone Grammar*. University of California Publications in Linguistics vol. 115, Berkeley: University of California Press.
- De Chene, B. 1985. « R epenthesis and the Japanese verb ». *Papers in Japanese Linguistics* 10: 170-207.
- Delafosse, M. 1955. *La langue mandingue et ses dialectes*, 2 vol. Paris : Paul Geuthner.
- Dixon, R.M.W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.
- Dixon, R.M.W. 1977. *A grammar of Yidiñ*. New York and London: Cambridge University Press.
- Dixon, R.M.W. 1980. *The Languages of Australia*. Cambridge: Cambridge University Press.
- Dugast, I. 1967. *Lexique de la langue tunén*. Paris : Klincksieck.
- Durand-Forest, J, et al. 2000. *Parlons Nahuatl*. Paris : l'Harmattan.
- Dwyer, D. J. 1989. "Mande". In *The Niger-Congo Languages*, J. Bendor-Samuel and R. Hartell (ed). Lanham: University Press of America.
- Edward, C, 2003, *Awara, Phonology*. University of North Dakota, MA Thesis.
- Emeneau, M. B. 1971. « Kodagu and Brahui developments of Proto-Dravidian *r ». *Indo Iranian Journal* 13: 176-198.
- Fleury, E. 1944. *Précis de Phonétique Grecque*. Paris: J. de Gigord.
- Foley, W. A. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.
- Fujioka, K. 1908. « Nihongo no chii ». *Kokugakuin Zasshi* 14.
- Giegerich, H. J. 1999. *Lexical Strata in English*. Cambridge: Cambridge University Press.
- Gnerre, M. 1999. *Profilo descrittivo e storico-comparativo di una lingua amazzonica: lo shuar (jivaro)*. Naples: Istituto Universitario Orientale.
- Gomez-Imbert, E. 2003. « Phonologie des langues tukano: quelques aspects intéressants ». Paper presented at the Erss Phonology Workshop, March 2003, Bordeaux.
- Greenberg, J. H. 1966. *Language Universals*. The Hague: Mouton.
- Hamano, S. 1998 (first pub. 1986). *The Sound Symbolic System of Japanese*. Tokyo:

- Kuroshio.
- Hayden, R. 1950. « The Relative Frequency of Phonemes in General American English ». *Word* 6-3: 217-223.
- Houis, M. 1963. *Etude descriptive de la langue susu* [Mémoires de l'Institut Français d'Afrique Noire n°67]. Dakar : IFAN.
- Hualde, J. I. 1991. *Basque Phonology*. London : Routledge.
- Hug, M. 1979. *La distribution des phonèmes en français / Die Phonemverteilung im Deutschen*. Geneva / Paris: Slatkine.
- Humbert, H. 1996. “De Kameleontische aard van de Groningse /r/ verklaard vanuit representatieve perspectief. *Taal en Tongval* 48 (2): 139-162.
- Itô, J and Mester A. 1998. *Markedness and Word Structure, OCP Effects in Japanese*. Available on ROA 255.
- Johanson, L. and Csato, E. A. 1998. *The Turkic Languages*. London: Routledge.
- Kager, R. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kishida, T. 1984. *Kokugo on'in henka no kenkyû*. Tokyo: Musashino Shoin.
- Köhler, O. 1981. « La langue kxoe. Matériel phonétique » In *Les langues dans le monde ancien et moderne, Afrique subsaharienne, pidgins, et créoles*, J. Perrot (ed). Paris : Editions du CNRS.
- Kokuritsu Kokugo Kenkyûjo [National Center for Research on the Japanese Language]. 1984. *Nihongo no Goi to Kyôiku*. Tokyo: Kokuritsu Kokugo Kenkyûjo Publications.
- Kuginuki, T. 1982. “Jôdai Nihongo ra-gyôon-kô”. *Toyama Daigaku Jinbungakubu Kiyô* 6 : 192-206.
- Kurematsu, Akira, 1997, “Intsu jôhō no jôhō kagaku he no tekiyō”. In *Nihongo Onsei (2), akusento, intonêshon, rizumu to pôzu*, Sugito Miyoko (ed), 303-318. Tokyo: Sanseidō.
- Kurylowicz, J. 1927. « Origine indo-européenne du redoublement attique ». *EOS* 30 : 206-210.
- Labrune, L. 1993. Le Statut Phonologique de /r/ en Japonais et en Coréen, Phd thesis, Paris : Université Paris 7.
- Labrune, L. 1996. « La formation des désinences verbales en japonais ancien ». *Cipango* 5 : 141-171.
- Labrune, L. 1998. « De l'iconicité en japonais : suffixes en r+ voyelle et reduplication ». *Cahiers de Linguistique Asie Orientale* 27(1) : 79-118.

- Lafitte, P. 1979 ? (1st edited 1944). *Grammaire basque (Navarro-labourdin littéraire)*, édition revue et corrigée. Donostia (San Sebastian): Elkar.
- Laroche, Emmanuel. 1980. *Glossaire de la langue hourrite*. Paris : Klincksieck.
- Lehmann, W. 1951. "The distribution of Proto-Indo-European /r/". *Language* 27: 13-17.
- Lindau, L. 1985. "The Story of /r/". In *Phonetic Linguistics, Essays in Honor of Peter Ladefoged*, V. Fromkin (ed), 157-168. Orlando: Academic Press.
- Lynch, J. 1983. "On the Kuman 'Liquids'". *Linguistic Society of Papua New Guinea* 14: 96-110.
- Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
- Matsuno, K. 1971. "Some allophones of Japanese /r/ and their stylistic values". *Onseigakkai Kaihō* 136: 5-8.
- Matteson, E. 1965. *The Piro (Arawakan) Language*. Berkeley & Los Angeles: University of California Press.
- McCarthy, J and Prince, A. 1994. "The emergence of the unmarked: Optimality in Prosodic Morphology". *Proceedings of the North East Linguistic Society*. Amherst: GLSA: 333-379. (reference is made to the French translation, which appeared in *Langages*, 125, 1997.)
- McCarthy, J. and Prince, A. 1995. "Faithfulness and reduplicative identity". In *University of Massachusetts Occasional Papers in Linguistics 18*, J. Beckman et al. (ed), 318-366. Amherst: GLSA.
- Mester, A. and Itô J. 1989. "Feature predictability and underspecification: palatal prosody in Japanese mimetics". *Language* 65: 258-293.
- Naden, T. 1989. « Gur ». In *The Niger-Congo Languages*, J. Bendor-Samuel and R. Hartell (ed), 141-168. Lanham: University Press of America.
- Nichols, Johanna. 1997. "Chechen phonology". In *Phonologies of Asia and Africa*, vol. 2, Alan S. Kaye (ed). 941-971. Winona Lake (Indiana): Eisenbrauns.
- Ohnishi, M. 1987. "Dôteki jinkō kōgai ni yoru nihongo shiin no kōsatsu". *Onsei Gakkai Kaihō* 186: 19-24.
- Ôtsubo, H. 1989. *Giseigo no kenkyū*. Tokyo: Meiji Shoin.
- Paradis, C. and Prunet J-F. (eds). 1991. *The special Status of Coronals: internal and external evidence*. San Diego: Academic Press.
- Paris, C. 1989. "Esquisse grammaticale du dialecte abzakh (tcherkesse oriental)". In *The*

- Indigenous Languages of the Caucasus* vol. 2, B. G. Hewitt (ed), 155-260. Delmar / New York: Caravan Books.
- Patz, E. 1991. "Djabugay". in *the Handbook of Australian Languages* 4, R. Dixon and B. Blake (eds): 245-348. Amsterdam: John Benjamins.
- Prince, A. and Smolensky, P. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University & University of Colorado: ms.
- Prost, A. 1950. *La langue Bisa* [Collection Etudes Voltaïques]. Ouagadougou : Centre Ifan.
- Pulleyblank, E. G. 1984. *Middle Chinese, A Study in Historical Phonology*. Vancouver: University of British Columbia Press.
- Rivierre, J-C. 1973. *Phonologie comparée des dialectes de l'extrême sud de la Nouvelle Calédonie*. Paris : Selaf.
- Rood, D. S. 1979. « Siouan ». In *The Languages of Native America*, L. Campbell & M. Mithun (eds), University of Texas Press.
- Ross, M. D. 1992. "The sounds of proto-Austronesian: an outsider's view of the Formosan Evidence". *Oceanic Linguistics* 31 (1): 23-64.
- Sakakura, A. 1966. *Gokôsei no Kenkyû*. Tokyo: Kadokawa.
- Schadeberg, T. C. 1989. "Kordofanian". In *The Niger-Congo Languages*, J. Bendor-Samuel and R. Hartell (ed), 67-80. Lanham: University Press of America.
- Schourup, L. and Tamori, I. 1992. "Japanese palatalization in relation to theories of restricted underspecification". *Gengo Kenkyû* 101:107-145.
- Schulze-Fürhoff, W. 1994. « Udi ». In *The Indigenous Languages of the Caucasus* vol. 4, R. Smeets (ed), 447-514. Delmar / New York: Caravan Books.
- Sommer, F. 1947. *Hethiter und Hethitisch*. Stuttgart: W. Kohlhammer.
- Steever, S. B. 1998. *The Dravidian Languages*. London and New York: Routledge.
- Stenzel, K. 2001. *An Overview of Wanano Phonology*. Ms.
- Tersis, N. 1981. "Le zarma". In *Les langues dans le monde ancien et moderne, Afrique subsaharienne, pidgins, et créoles*, J. Perrot (ed) : 89-99. Paris : Editions du CNRS.
- Trombetti, A. 1923. *Elementi di glottologia*. Bologna: Nicola Zanichelli.
- Tsuzuki, M. and Lee H-B. 1992. "A phonetic study of the Korean and Japanese lateral, flap and nasal". In *Proceedings of the 1992 Seoul International Conference on Linguistics*, the Linguistic Society of Korea (ed), 761-780. Seoul: Koryo university.

- Ueda, I. and Davis S. 2001. "The acquisition of Japanese 'r'". *Gengo Kenkyû* 119: 111-139.
- Unger, J. M.. 1976. *Studies in Early Japanese Morphophonemics*. Phd Dissertation. Yale University.
- van Hout, R. and van de Velde, H. 2001. "Patterns of /r/ variation", In *r-atics, Sociolinguistics, Phonetic and Phonological characteristics of /r/*, R. van Hout & H. van de Velde (eds), [Etudes et Travaux n° 4],1-9.
- van Oostendorp, M. 2001. "The phonology of postvocalic /r/ in Brabant Dutch and Limburg Dutch", In *r-atics, Sociolinguistics, Phonetic and Phonological characteristics of /r/*, R. van Hout & H. van de Velde (eds) [Etudes et Travaux 4], 113-122.
- Walsh Dickey, L. 1997. *The Phonology of Liquids*. Phd thesis. Amherst: University of Massachussets.
- Wayne, L. 2000. "Shiin wo koete okoru boin no yûgô – Ryûkyû hôgen ni okeru genshō wo chûshin ni". *Onsei Kenkyû (Journal of the Phonetic Society of Japan)* 4 (1): 55-60.
- Wiese, R. 2001. "The unity and variation of German /r/". In *r-atics, Sociolinguistics, Phonetic and Phonological characteristics of /r/*, R. van Hout & H. van de Velde (eds) [Etudes et Travaux 4], 11-26.
- Wilson, W. A. A. 1989. « Atlantic ». In *The Niger-Congo Languages*, J. Bendor-Samuel and R. Hartell (ed). Lanham: University Press of America.
- Wioland, F. 1985. *Les structures syllabiques du français*. Geneva / Paris : Slatkine.
- Yoshida, K. 1976. *Nihongo Gogengaku no Hôhō*. Tokyo: Taishûkan.
- Yu J-W. 1985. *Uri mal yeoksun sajeon*. Seoul: Jeongeumsa.
- Zipf, G. K. and Rogers, F. M. 1939. "Phonemes and variphones in four present-day romance languages and classical Latin from the point of view of dynamic philology". *Archives Néerlandaises de Phonétique Expérimentale* 15: 111-147.
- Zuraw, K. 2000. *Patterned Exceptions in Phonology*. Phd thesis. Los Angeles : UCLA.

¹ Throughout this paper, the Japanese examples have been transcribed according to the Kunrei system, except for the presentation of Japanese reference materials cited in the bibliography, where the Hepburn system has been adopted.

The symbol > denotes diachronic evolution, whereas / indicates synchronic alternation. Translation of the examples is given in double quotes (“translation”). Gloss appears in simple quotes (‘gloss’).

² The absence of /r/ at the beginning of Yamato words is sometimes taken as a proof that Japanese belongs to the Altaic family (Fujioka, 1908, and subsequent researchers).

However, while it is true that Altaic languages do not have r-starting words, they are not the only languages in the world which present such a characteristic. Indeed, languages which lack /r/ at the beginning of lexical words are quite widely spread (see endnote 12).

³ True exceptions to this principle are nearly inexistent. We found only *ryûryû to* “with speed and force”, *rerorero* “grumbling” or “incoherently”, *rorirori* “restless”, *rin* (*rinyin*, *riin*, *riiin*) ‘ringing sound’, *runrun* “in good spirit”. All are mimetics. *Rerorero* and *rorirori* are the only ones with ancient attestation (*rerorero* is cited by the *Iwanami Dictionary of Old and Middle Japanese [Iwanami Kogo Jiten]*, and *rorirori* appears in the *Japanese Portuguese Dictionary [Vocabolario da lingua de Iapam com a declaração em Portugues]* compiled by Jesuit Missionaries of the Society of Jesus in 1603-1604). It is particularly significant to observe that *rerorero* and *rorirori* are exceptional in two respects: first because they begin with /r/, second because they infringe a strong phonotactic constraint, which is that identical consonants cannot occur twice in a mimetic root (Hamano, 1998:41-42, Schourup & Tamori, 1992:120). We believe that the fact that exceptions to this second principle involve precisely /r/ and no other consonant can be interpreted as an additional proof that /r/ is phonologically transparent, which explain why, unlike other consonants, it can be repeated.

Other MJ Yamato words such as *rutubo* “melting pot”, *reko* “this”, *risu* “squirrel” could appear as exceptions, but they are not from a diachronic point of view: *rutubo* results from the loss of an original initial vowel *i* (< *irutubo*), *reko* from the inversion of *kore* (language play), and *risu* is probably a borrowing from Chinese.

⁴ The complementary distribution between the zero onset and [r] in Japanese is reminiscent of what occurs between English [ə] and [r]. The proposal that [ə] and [r] are surface reflexes of an underlying zero has actually been made for RP English by Giegerich (1999). A similar analysis has been proposed for several varieties of Dutch (van Oostendorp, 2001, Humbert, 1996 cited by van Oostendorp, 2001). All the segments involved here in Japanese, Dutch and English are arguably placeless at the phonetic or phonological level. One anonymous reviewer has suggested that one would also be prepared to argue that Modern English lacks empty onsets because empty onsets are in complementary distribution with the velar nasal. However, it seems to us that two types of different phonological objects are compared here: a position (the onset, which is empty) and a segment (the velar nasal) which happens to occur only in the coda position in English. But the comparison works only if two positions, onset and coda, or two segments, the null segment zero and the velar nasal, are compared. Since zero codas, i.e. codaless syllables, do occur in English, zero and [ŋ] cannot be considered to be in complementary distribution. Moreover, it seems difficult to maintain that English [ŋ] is placeless and unmarked at the phonetic or phonological level. Therefore, in our view, the comparison between Japanese and English holds between [ə] and [r] in the case of English, rather

than between the zero onset and [ŋ].

⁵ In MJ the zero consonant can appear word-medially, but this is due to a relatively recent evolution of the language. Modern V'V sequences result from borrowing or from the loss of a consonant, generally /h/ or /w/, as in *kahi* > *kai* “shell”, or *kowe* > *koe* “voice”.

⁶ This constraint is not totally preserved in MJ (cf. *oriru* “to get down”, *ireru* “to insert”), but even in the modern language, most verbs containing two occurrences of /r/ are compounds involving two stems, or are derived by adjunction of the suffix *-eru* which is an innovation of the pre-modern language.

⁷ Child language phenomena might provide additional clues in favor of the late development hypothesis. Since /r/ figures among the last acquired phonemes in the chronology of Japanese children first language acquisition, one can ask how these data are consistent with the rest of the properties displayed by Japanese /r/. If one subscribes to the idea that ontogenesis mirrors phylogenesis, then the acquisition facts provide additional support to the hypothesis that /r/ emerged late in the history of Japanese.

⁸ Note that the consonant system of the language has not substantially changed over the past 1500 years, except for the development of /h/, and the phonemic split of /s/, /z/, /t/ and /d/ which gave rise to /s/-/ʃ/, /z/-/ʒ/, /t/-/tʃ/, and /d/-/(d)ʒ/.

⁹ We suppose that /r/ is phonologically [+Continuant], because most of its phonetic realizations are, but this feature is actually non distinctive for /r/ in modern Japanese (remember that /r/'s realizations are either [+Continuant] or [-Continuant], see section 2.1).

¹⁰ We assume that Japanese /y/ (IPA j) is a true palatal and not a coronal, because it is neither apical nor laminal, contrary to coronals such as /t, d, s, z, r/. Following Clements & Hume (1995:286), we consider that palatality is characterized in terms of a [-Back] dorsal feature. Consequently, /y/ has to be characterized as a dorsal, [-Back] consonant. It is interesting to remark that candidate n° 17 /aya/, with intervocalic /y/, appears to be the second best candidate after /ara/. Empirical evidence confirms this relative harmony of /y/, as sporadic alternation between /r/ and /y/ is attested in OJ. It seems to occur especially at the beginning of certain suffixes, for instance in *-ra-ru* / *-ra-yu* (OJ) ‘passive auxiliary’, *-raka* / *-yaka* ‘adjectival suffix’ (*name-raka* “smooth”, vs. *nigi-yaka* “cheery”). The existence of such an alternation, which seems to be frequent in the environment of a high vowel, might result from the assimilation of some feature in the adjacent high vowel onto the intervocalic slot.

¹¹ An alternative account of the distributional asymmetries of /r/ could be considered: one could suppose that /r/ is present underlyingly, but that it is deleted word-initially, due to the action of a constraint banning liquids from the initial position. Such a constraint is actually posited by Ueda & Davis (2001), in an attempt to provide an account of the acquisition of /r/ in Japanese. However, their approach fails to explain why /r/ is overwhelmingly frequent in the Japanese language, especially at the beginning of non-lexical suffixes, whereas our analysis is able to relate /r/'s frequency with that of /'/. Ueda and Davis are also incapable of accounting for the fact that liquids are allowed at the beginning of foreign loans. Finally, if one considers that /r/ is simply deleted word initially, one can no longer account in a unified manner for its phonological inertia, and for all the other properties presented in section 2 which lead to interpret it as an unmarked segment. We therefore conclude that the analysis proposed here is superior because it provides a systematic, principled and empirically-adequate account of the numerous

properties exhibited by /r/ in Japanese.

¹² Such languages include: Korean, Mongol, Tungus, Turkish and a number of other Altaic languages, *Proto-Turkic* (Johanson & Csato, 1998:71), *Proto-Altaic*; *Ancient Chinese* (Pulleyblank 1984); Telugu (vishakhapatnam dialect, Bhat, 1974), Toda (p.c. Brian F. Head, 2002), *Proto-Dravidian* (Emeneau, 1971:183, Steever, 1998:16); Yukaghir (Benveniste, 1939); Votyak (Bhat, 1974), Ostyak (Bhat, 1974), Samoyed (Trombetti, 1923:347), Kotto (Trombetti, 1923:484); Basque (Lafitte 1944 :15, Hualde 1991:12); Ancient Armenian (Lehmann, 1951), Archaic Greek (Fleury, 1944 :30, Benveniste, 1939), Hittite (Benveniste, 1939 cited by Lehmann, 1951), Gascon, *Proto-Indo-European* (Kurylowicz, 1927, Lehmann 1951, Benveniste, 1939); Udi (South-East Caucasian, Schultze-Furhoff, 1994 :455), Abzakh (North-West Caucasian, Paris, 1989: 158), Ubykh (North-West Caucasian, Charachidze, 1989:363), Chechen (Nichols, 1997), Old Caucasian languages (Sommer, 1947, cited by Lehmann 1951); Burushaski (Anderson, 1997); Gur (Naden, 1989: 152), Bisa (Prost, 1950 :15), Susu (or Soso, Houis, 1963 :27), Efik (Bhat, 1974), Kordofanian (Schadeberg, 1989 :75), Mandingue (Delafosse, 1955:68), Malinke (Delafosse, 1955), Bambara (Delafosse, 1955:449), Dioula (Delafosse 1955:58), Fula (Wilson, 1989), Moore (Canu, 1981:113), Tunen (Dugast, 1967), Agau (Trombetti, 1923:347), Nuba (Trombetti, 1923:347), Dschang (Bird, 1999), *Proto-Mande* (Dwyer, 1989:56), *?Proto-Fula* (Wilson, 1989); Zarma (Songhai) (Tersis, 1981:90), Mbai-Moissala (Caprile, 1981:245) ; Kxoe (Köhler, 1989 :486), Hottentot (=Nama) (Benveniste, 1939); Afar (Cushitic) (Trombetti, 1923:348), Saho (Cushitic) (Trombetti, 1923:348), *?Camito-Semitic* (Trombetti, 1923:351), Chaha (Gurage) ; Hourrite (Laroche, 1980) ; Piro (Arawakan) (Matteson, 1965:29), Nahuatl (Uto-aztec, Durand-Forest et al., 2000 :33), Shuar (Jivaro, Gnerre, 1999, cited by Gomez-Imbert, 2003), Tukano (Elsa Gomez-Imbert, p.c. 2002), Tuyuka (Gomez-Imbert, 2003), Kubeo (Gomez-Imbert, 2003), Wanano (Stenzel, 2001, cited by Gomez-Imbert, 2003), Arikara (Caddoan, Chafe, 1979), *?Sioux* (Rood, 1979:237), Cree (Algonquian, Christian Bassac, p.c. 2003), Purepecha (Tarascan, Chamorro, 2003:43), Tümpisa / Panamint Shoshone (Uto-aztec, Dayley, 1989); Yidiñ (Dixon, 1980, cited in Walsh-Dickey 1997:143), Diyari (Austin, 1981:19), Dyirbal (Dixon, 1972:3), Djabugay (Patz, 1991, cited in Walsh-Dickey 1997:143), Guugu Yimidhirr (Dixon 1980), Mayi (Breen, 1981, cited in Walsh-Dickey 1997:143), Warrgamay (Dixon, 1980), and most Australian Languages (Dixon, 1972:3), *Proto-Australian*; Unya Dialect (New Caledonia, Rivierre 1973), Nêlêmwa and Nixumwak (New Caledonia, Austronesian family, Brill 2000 :29); Kunimaipa (Bhat, 1974), Sentani (Bhat, 1974), Kewa (Bhat, 1974), Kuman (Lynch 1983, cited in Walsh-Dickey 1997:143), Yimas (Foley, 1986: 56), Fore (Foley, 1986:55), Awara (Edward, 2003); Old Tagalog (Zuraw, 2000:29); *Proto-Austronesian* (Ross 1992). Reconstructed languages appear in italics.

¹³ Another segment presenting initial prohibition in a number of languages is the velar nasal /ŋ/ but, contrary to /r/, /ŋ/ is not a phoneme which ranks high in terms of frequency.