

# HOKKAIDO PUMPKINS AND HUAWEI PHONES: ANTI-HIATUS TENDENCIES IN SLOVENE\*

**Nina GOLOB**

University of Ljubljana, Slovenia  
nina.golob@ff.uni-lj.si

**Mateja PETROVČIČ**

University of Ljubljana, Slovenia  
Mateja.petrovcic@ff.uni-lj.si

## Abstract

In Slovene, sound shapes of Japanese and Chinese words undergoing the process of assimilation may be quite different from their origins, which in the end should not come as a surprise as it is so in many other languages. However, the fact that there are many cases with two slightly different Slovene pronunciations of the same Japanese or Chinese word, of which one is closer to the original sound shape than the other, calls for a more comprehensive investigation on the factors that contributed to such a result. This research pays attention to vowel sequences in Japanese and Chinese, and how they appear in Slovene as the so called visiting lexica. Though the two languages carry out vowel sequences in different ways, similarities can be found in the way Slovene detects and resolves their vowel hiatuses. Authors stress the importance of metrical aspects of the original sound shapes, which Latin script includes inconsistently or does not express at all. Phonological approach to vowel sequences solution to vowel hiatus and offers systematic guidelines on pronunciation of the Latin script. Besides, it may further bring some new insights into possible solution on conjugation and declination forms of Japanese and Chinese loanwords in Slovene.

**Keywords:** Japanese; Chinese; visiting lexicon in Slovene; hiatus-avoiding mechanisms

## Povzetek

V slovenščini je lahko glasovna podoba besed iz japonščine in kitajščine, ki so v procesu prevzemanja v jezik, precej drugačna od tiste v izvornem jeziku. To navsezadnje niti ni tako presenetljivo; spremenjeno glasovno podobo je moč opaziti tudi v drugih jezikih. Z vidika jezikoslovja je bolj zanimivo dejstvo, da v ciljnem jeziku, tj. slovenščini, obstajata dve nekoliko različni glasovni podobi iste izvirne besede, kar kliče k podrobnejši raziskavi o razlogih zanj. V tem delu obravnavamo samoglasniška zaporedja v t.i. gostujočem besedju, ki naj bi pri tem igrala pomembno vlogo. Japonščina in kitajščina sicer samoglasniška zaporedja realizirata na precej drugačen način, vendar je ob poznavanju izvirne glasovne podobe obravnava precej podobna. Avtorja poudarita pomen metrične strukture besed v

---

\* The work presented in this article has been supported by the ARRS research programme P6-0243 "Asian languages and cultures".



izvornem jeziku, ki je v latiničnem zapisu pogosto obravnavana nekonsistentno ali pa je celo neizražena. Fonološki pristop k samoglasniškim zaporedjem v slovenskih prevzetih besedah omogoča zadovoljivo reševanje samoglasniških hiatov in ponuja sistematične smernice o izgovarjavi takšnih besed. Takšne rešitve so lahko osnova nadaljnim diskusijam o možnih sklonskih in spregativenih oblikah japonskih in kitajskih besed v procesu prevzemanja.

**Ključne besede:** japonščina; kitajščina; gostujoče besedje v slovenščini; antihiatki mehanizmi

## 1 Introduction: a visiting lexicon

Vocabularies of world's languages are built of words of different origin, and phonologically that makes words form the so-called phonological lexical classes of those vocabularies (Ito & Mester, 1999). Besides the etymologically prime lexicon, most languages consist of one or more borrowed lexica, which have already implemented a certain degree of assimilation.

A language may at particular events import foreign words for a single use. Such events are relatively sudden compared to a natural process of borrowing, and generally influence the spoken form only—and even that to the extent of unavoidable adjustments only—, while the written form is kept foreign. We will call such words “a visiting lexicon”. It is characteristic for a visiting lexicon that its words either have an important status on the international stage and therefore keep their internationally recognized written forms for practical purposes, or else have little or no long-term significance and very low frequency of appearance in a particular language and therefore assimilation does not occur. The level of assimilation for a visiting lexicon is similar to that of an early-stage borrowed lexicon named “non-assimilated foreign” by Ito and Mester (1999).

Not all languages implement all three stages in borrowing words, and some do not accept or recognize a visiting lexicon. In Slovene, a visiting lexicon generally represents foreign proper names or foreign naming expressions though it must be noted that not all foreign proper names or foreign naming expressions necessarily belong to it.

### 1.1 Orthography of a visiting lexicon in Slovene

Not much is to be written about those foreign proper names and foreign naming expressions that originate from languages with Latin, Arabic, Cyrillic and Greek scripts. Some of such proper names have in the past been slovenized due to historical, grammatical or other reasons (Chicago → Čikago; Bruxelles/Brussel → Bruselj; Paris → Pariz; Kyōto → Kjoto; Beijing → Peking; etc.)<sup>1</sup>. Apart from those, original foreign orthographic form has been used in case of a Latin script, and a Latin transliteration

---

<sup>1</sup> For detailed guidelines see *Seznam tujih zemljepisnih imen v slovenskem jeziku (List of foreign geographical names in Slovene)*.

was conducted for proper names from Arabic, Cyrillic and Greek scripts (e.g. Београд → Beograd; Панчево → Pančevo; Δάφνη → Dafni; Σαλαμίνα > Salamina).<sup>2</sup>

On the other hand, orthographic rules for a visiting lexicon from languages with non-alphabetic writing systems, including Japanese and Chinese visiting lexicon have been treated much more ambiguously. Decades ago when more and more slovenization of such words was thought to be the right and the best way to handle them. Recently, such thinking is facing changes, and internationally recognized Latin scripts for Japanese and Chinese are gradually getting into use. Still, in the Slovenian Orthography they are not recognized as the official Latin scripts of Japanese and Chinese yet (see Section 1.2 for more details).

While recognizing official Latin scripts also for the languages of logographic and syllabic writing systems for Slovene visiting lexicon certainly systematizes and simplifies orthographic rules in Slovene, it on the other hand adds to the confusion concerning the readings. Different languages' orthographies offer different degrees of correspondence between spelling and pronunciation<sup>3</sup>, and this is also the case for Latin scripts that non-alphabetic languages utilize. Let us look at the following sentence that includes several city names from all over the world.

- (1) Medtem ko je glavni kandidat za izvedbo naslednjega prvenstva francoski Toulouse, so kandidaturo vložila tudi mesta Seville (Španija), Nyíregyháza (Madžarska), Kyzyl-Oi (Kirgistan), Hai Phong (Vietnam), Daejeon (Koreja) in Växjö (Švedska).

'While the main candidate to host the next championship is the French Toulouse, candidacy was also filed by Seville (Spain), Nyíregyháza (Hungary), Kyzyl-Oi (Kyrgyzia), Hai Phong (Vietnam), Daejeon (Korea) and Växjö (Sweden).'

The process of recognizing the words origin to read them accordingly can be very difficult or sometimes even impossible. In case a reader is not familiar with the sound shape of a visiting word, a mismatching between a written and spoken form is unavoidable. The following name George exemplifies how different pronunciations of the same written word can get among languages:

<sup>2</sup> Original orthographic form has recently been in use even with some historical names that have in the past been assimilated (ex. Galileo Galilej → Galileo Galilei). (Dobrovoljc, 2013)

<sup>3</sup> International Phonetic Alphabet (IPA) was devised as a standardized representation of the sounds of spoken language. With it the relationship between sign and sound is the same no matter the language.

- (2) /dʒ'ɔrdʒə/ (Italian)  
/ʒɔʁʒ/ (French)  
/ge:ɔrtə/ (German)  
/dʒɔrč/ (Indian language Karnataka)  
/dʒɔrdʒe/ (Romanian)  
/gʲɔrgi/ (Russian)  
/'dʒɔrdʒ/ (American English)  
/dʒɔ:dʒ/ (British English)  
/xe.'ɔr.xe/ (Spanish)

## 1.2 Official Latin script for Japanese and Chinese

About a century ago, the problem of how to write proper names from languages other than Slovene was first mentioned in the manual *Slovenian Orthography* (1920) by Breznik (Pogačnik, 2012). His brief note on the importance of keeping the written word “as it is” (Breznik, 1920, §1-1-5) was dedicated to foreign languages of Latin script, however, it is thought to be a breaking point in the awareness and inclusion of something foreign into Slovene, cf. what we now call a visiting lexicon. It took nearly another half a century to include the guidelines for writing proper names from languages of non-alphabetical writing systems. At the time it was suggested that such words be written phonetically (*Slovenian Orthography*, 1962), which corresponds to a fully assimilated form.

There are no linguistic grounds for categorizing words in Slovene to those of Latin and non-Latin script, further treating the latter as either alphabetical or not with the purpose to decide on whether a foreign word in the borrowing process would undergo slovenization or not. Categorization above all produces discussions on how to treat border cases such as the name of the Nobel prize winner Ishiguro Kazuo, an Englishman of a Japanese descent, whose name can be found slovenized due to his Japanese roots (Kazuo Išiguro) or written with its original Latin spelling (Ishiguro Kazuo) (Lengar Verovnik, 2017).

The tendency to keep a visiting lexicon from languages of non-alphabetical writing systems in their internationally recognized Latin script is gaining importance though the guidelines from 1962 (and their revised versions from 1991 and 2001) have not been abandoned fully yet.

Several possible reasons—leaving aside their justifications—could be thought of for the present situation; with the existence of more than one international Latin script for one language it is difficult to choose and adopt one representative; adopting different international scripts for all non-Latin non-alphabetical writing systems would be confusing for a Slovene reader; general Slovene readers are neither used to the sound shapes of Japanese and Chinese neither to their Latin scripts, and besides instructions on how to read such scripts would get long and complicated, etc.

Let us have a brief look at the official Latin script for both Japanese and Chinese, and their status in Slovene.

Romanization of Japanese dates centuries and resulted in the existence of several different styles, among which Hepburn Latin script was most widely known. To unify the styles the former Ministry of Education compiled and send out Kunrei Latin script, which was in 1954 recognized as the correct style of romanization. In the same Cabinet notice, Hepburn system was allowed to be used as well, and consequently there are now two Latin scripts for Japanese; internationally standardized Kunrei Latin script (ISO 3602), and the widely used Hepburn Latin script for writing names and geographical locations.<sup>4</sup> Though well-established, the two Latin script for Japanese have not been recognized by the Slovene Orthography (*Slovenski pravopis*) yet despite the professional proposals to do so (Mlakar & Ilc, 2009).

Similar is the situation with Chinese. Discussions about how to write Chinese words (and proper names names within that) in Slovene date back to 1974 when the sinology scholar Mitja Saje presented Hanyu Pinyin to Slovene professional public, and have been a subject of debate since then (Toporišič, 2011). The role of Hanyu Pinyin as the official Latin script for Standard Chinese has been repeatedly rejected by Slovene scholars till recently. The fact that Hanyu Pinyin was adopted by The International Organization for Standardization (ISO) as an international standard in 1982, followed by the United Nations in 1986, was not convincing enough. Majority of publications followed the standard prescriptive orthography defined in Slovene Orthography, ignoring arguments from the sinological perspective.

In 2009, an entire issue of the *Asian and African Studies* (Volume 8, Issue 2) scholarly journal was devoted to the problem of writing Japanese and Chinese words in Slovene. Its clear message that proper names should not be forced into slovenization argued against paragraph §180 in the Slovene Orthography stating that both languages had already had their own Latin scripts and should therefore not be forced into slovenization” any further.<sup>5</sup>

Introducing internationally recognized Japanese Hepburn romanization and Chinese Hanyu Pinyin into Slovene with no doubt brings about the need to equip Slovene readers with information on the optimal pronunciation of the two scripts. It is therefore important to come to the solution which thoroughly considers characteristics of Slovene, Japanese and Chinese phonological systems.

---

<sup>4</sup>Very recent movements try to unify the two (The Mainichi, 2017).

<sup>5</sup>For the exact proposals regarding Chinese, see Petrovčič 2009a, 2009b, 2009c.

## 2 Research question

With the process of borrowing words a language sometimes has to deal with new sound sequences. Among them are sequences of vowels which may express different types of interrelationship and consequently different syllable structures.

Previous studies have shown that speakers have a very clear intuitions about the number of syllables in a word (Blevins, 1996), and that such intuitions are used for perceiving syllables of a foreign language (Carlisle, 2001; Zárata-Sández, 2011). In addition, a visiting lexicon with its foreign orthographic forms from different languages may further complicate the situation for a reader or speaker, as it is known that orthography also influences syllabic segmentation of a foreign language (Detey & Nesouplous, 2008).

A gap between writing and pronunciation of vowel sequences is observed in many of the world's languages. Pay attention to the pronunciation of the vowel sequence "ie" in the following example.<sup>6</sup>

- (3) Pyeongchang 2018 Olympic Winter Games hosted many athletes, including:
- |   |        |
|---|--------|
| Francesco Friedrich (Germany, bobsleigh) <sup>7</sup> | /i:/   |
| Emilien Jacquelin (France, biathlon)                  | /i.ja/ |
| Yurie Watabe (Japan, freestyle skiing)                | /i.e/  |
| Xie Jiaxuan (China, speed skating)                    | /je/   |

Based on the above, the main interest of this study is first to make an overview of the possible vowel sequences and their Latin scripts for Japanese and Chinese respectively, and look at how they get shaped when entering into Slovene. Focusing on the so-called visiting lexicon, for which in Slovene it is suggested to keep the internationally recognized Latin script, the study discusses those vowel letter sequences that are less familiar to the Slovene reader or seem problematic.

## 3 Vowel sequences: diphthong and hiatus

Two vowels in a sequence can appear either as a diphthong or a vowel hiatus. A diphthong refers to two vowels of different qualities<sup>8</sup> that represent a tautosyllabic sequence, cf. belong to the same syllable. On the other hand, vowel sequence in hiatus is heterosyllabic with the two vowels occurring across syllable boundary.

---

<sup>6</sup> The example itself is fictitious, however information about athletes are real.

<sup>7</sup> <https://www.pronouncenames.com/search?name=friedrich> (name to be typed in the last part of URL)

<sup>8</sup> Catford (1977, p. 215) treats a diphthong as "a sequence of perceptually two different vowel sounds within one and the same syllable", and Ladefoged (1982, p. 171) defines it as "a single vowel with continuously changing qualities".

There are morphological and phonological criteria that are used for the definition of the vowel sequence type. The morphological criterion is that two vocalic elements must appear within a morpheme rather than across two morphemes in order to form a diphthong. The phonological criterion concerns the sonority of the two vowels; the preceding vowel must be at least as sonorous as the following one to form a diphthong, and at the same time, vowels in sequence belong to different syllables if the following vowel is more sonorous than the preceding one. Potential exceptions to this are cases where preceding vowel becomes a glide, e.g. [ia] → [ja], [ua] → [wa], and cases where the following vowel becomes a schwa, e.g. [ia] → [iə]. And for a hiatus, vowels may come to meet at word or morpheme boundaries, or may arise due to an intervocalic consonant deletion in the middle of phonological word (Picard, 2003).

Slovene contains diphthongs and hiatuses. Diphthongs are relatively common (summarized in Šuštaršič, Komar & Petek, 1999), while hiatus may appear on morphemic boundaries in derivatives and in loanwords<sup>9</sup> (Jurgec 2004a).

Though vowel hiatus is not a rare phenomenon—there are languages such as Portuguese, Japanese, Bantu languages, and Polynesian languages having no or very few diphthongs but they perform vowel sequences that appear in hiatus—, it is interesting that languages are generally quite intolerant of hiatus, it being “phonologically unstable” (Kenstowicz, 1994, p. 23) and “apparently uncomfortable” (Trask, 1996, p. 65). Such languages employ several different anti-hiatus or hiatus-avoiding mechanisms.

### 3.1 Hiatus resolution

Languages that do not permit vowel hiatus may employ any of several processes that eliminate it in cases where it would otherwise arise. One of the most common means of avoiding hiatuses, resolving potential hiatuses, or breaking up actual ones is the elision of one of the two vowels. Other common pattern cross-linguistically are diphthong formation or glide formation, vowel coalescence, and epenthesis of a default consonant, while capturing a ‘floating’ consonantal melody or the spread of some (consonantal) melody from one of the vowel positions flanking the empty onset position are also observed (for details see Casali, 1997, 2011).

### 3.2 Anti-hiatus tendencies in Slovene

Jurgec (2004b) has shown that hiatus in standard Slovene attract cross-linguistically common hiatus-avoiding tendencies, with the epenthesis of [j] and the diphthongization of Vɤ on the level of general phonological system, and glide formation and elision for individual cases. Phonetically, reductive processes such as

---

<sup>9</sup> Hiatuses are very rare though.

diphthongization, elision and coalescence, and dissimilation also occur. Let us look at the prevalent ones through examples.

- (4) **Epenthesis of [j]: i + V → ijV**  
biologija /bijologija/ 'biology'  
Vietnam /vijetnam/ 'Vietnam'
- (5) **Glide formation: i [-stress] + V → jV**  
bianko /bjanko/ 'blank'  
bonboniera /bonbonjera/ 'box of chocolates'
- (6) **Epenthesis or glide formation**  
riviera /rivijera/ or /rivjera/ 'riviera'  
premier /premier/ or /premje/ 'premier'
- (7) **Diphthongization: V + u [-stress] → Vu, rarely V + i [-stress] → Vi, V + o → Vo**  
nauk /nauk/ 'lesson'  
čoln /tʃoun/ 'boat'  
aids /ajts/ 'aids'
- (8) **Elision: V1 + V2 → V1, V1 + V2 → V2**  
laringealen /laringalen/ 'laryngeal'  
poizkus /poskus/ 'try'

#### 4 Hiatus in Japanese and Chinese: its appearance in the Latin scripts

Both Japanese and Chinese possess a variety of vowel sequences. In this section we will look at such possible vowel sequences, examine their syllabic constitution, and look into how they are represented by the official Latin scripts for the two languages respectively.

##### 4.1 Hiatus and Japanese

According to Lindau, Nolin and Svantesson (1990), standard (Tokyo) Japanese can be classified into one of two thirds of the world's languages that lack phonemic diphthongs. There is a vivid debate on the topic.

Japanese has five monophthong vowels that can be combined freely with one another to form a vowel sequence. Across morpheme or lexeme boundaries such vowel sequences always form a hiatus. Vowel sequences within a morpheme/lexeme are, on the other hand, highly limited (Takayama, 2003) but among those V + i sequence makes the potential for a diphthong (Kubozono, 2015). Many scholars disagree with the interpretation that Japanese has diphthongs (Labrune, 2012; and others). Kawahara (2003) researched on hiatus resolutions in Japanese. He reports that epenthesis of [j] or [w] is the representative hiatus-avoiding mechanism but reductive



processes such as coalescence<sup>10</sup> can also be found. Concerning the diphthong debates, phonetic diphthongization on tautomorphic V + i sequence could be thought of.

In any case, discrepancies in phonological and phonetic interpretation of vowel sequences do not concern official Latin scripts for Japanese. In the discussion on Japanese we will use the Hepburn Latin script as this script is proposed for proper names (for details on the official Latin scripts for Japanese see Section 1.2). It should be noted that spelling of vowels sequences is the same for Hepburn and Kunrei Latin script.

With Hepburn romanization<sup>11</sup>, all vowel sequences of different vowels are in fact transliterations from the *kana* syllabic orthography, and are treated as heterosyllabic. In other words, each monophthong has its own vowel letter and vowel sequences are represented by the corresponding vowel letter sequences, all being treated as hiatus.<sup>12</sup>

- (9) /sa.so.u/    sasou    'to invite'  
       /ka.ru.i/    karui    'light (for weight)'  
       /ma.e/      mae      'ahead'  
       /o.i/        oi        'nephew'  
       /o.bo.e.ru/ oboeru 'to remember'  
       \* The period (.) marks a syllable break.

Therefore, there is no need to distinguish between vowels sequences that occur within or across morpheme boundaries in compounds.

- (10) ha + isha /ha.i.ɕa/ haisha 'out of service vehicle'  
       hai + sha /ha.i.ɕa/ haisha 'dentist'  
       \* The plus (+) marks a morpheme boundary.

Perhaps a special note should be made for a sequence of vowels of the same quality<sup>13</sup>. Such a sequence is, as any other vowel sequence, generally referred to as hiatus. Except for the /ii/ vowel sequence which is always written with two successive vowel letters because the visual distinction between "i" and "i" is miniscule, sequences

<sup>10</sup> A paper on Japanese vowel coalescence in Tokyo Japanese was published by Kubozono (2015).

<sup>11</sup> In the history of romanization, the first established Hepburn system was "traditional Hepburn romanization", which underwent slight changes and was renamed to the "modified" or "standard" Hepburn Latin script.

<sup>12</sup> Not counting sequences of the same vowels, there are twenty vowel combinations possible.

<sup>13</sup> First, there is a dispute whether it can be called a sequence of two vowels or whether it is one long vowel. In this research we will distance from this, and treat them phonologically as two vowels. Second, in some Hepburn variations the same vowel sequence is not indicated, e.g. *Tokyo*. That is common for Japanese words that have been adopted into English and is also the convention used in the *de facto* Hepburn used in signs and other English-language information around Japan, mentioned in the paragraph on legal status (Hadamitzky, 2005). Third, a sequence of the same vowel that spread across morpheme boundary is marked by a prime (') in Kunrei Latin script, e.g. sato'oya 'foster parents'.

of other vowels are written with one vowel letter with a macron in case they appear within a morpheme.<sup>14</sup>

- (11) chuui /tʃu.u.i/ chūi 'attention'  
 mizu + umi /mizu.u.mi/ mizuumi 'lake'

Separate attention is usually dedicated to the sequence of two /o/ and /e/ vowels. Besides the regular vowel letter sequence in the *kana* syllabic writing, they also exhibit diachronically evoked sound changes from /o/ + /u/ and /e/ + /i/<sup>15</sup> respectively (see Frellesvig (2010) and Labrune (2012) for details). Hepburn Latin script is consistent with /oo/ when tautomorphic, being 'ō' or 'ou', while for the /ee/ sequence, which is very unstable and with weak dialectal support, 'ei' transliteration is (still?) commonly in use<sup>16</sup>.

- (12) /ko.o.ri/ kōri 'ice'  
 ko + odori /ko.o.do.ri/ koodori 'dance'  
 /ga.k.ko.o/ gakkō 'school'  
 ko + ushi /ko.u.çi/ koushi 'calf'
- (13) /o.ne.e.sa.n/ onēsan 'older sister'  
 nure + en /nu.re.e.n/ nureen 'open veranda'  
 /ke.i.ke.n/ keiken 'experience'  
 tame + ike /ta.me.i.ke/ tameike 'pond'

Finally, Japanese contains glides [j] and [w] that always take a pre-vowel position and may either represent an independent consonant or a means for palatalization of the preceding consonant. In both cases the two glides are written with letters 'y' and 'w' respectively.

- (14) /ki.o.ku/ kioku 'memory'  
 /kyo.ku/ kyoku 'song'  
 /ki.yo.ku/ kiyoku 'cleanly'

## 4.2 Hiatus and Chinese

In Chinese, vowel sequences may appear within a syllable or across syllable boundary as shown in example (15).

- (15) /ai/ ai 'love'  
 /waj/ wai 'outside'  
 /wu.ai/ wu'ai 'loveless'  
 /tʃjaɹ.aʊ/ jiao'ao 'pride'

<sup>14</sup> This distinction was made in the "modified" Hepburn Latin script and is not followed consistently.

<sup>15</sup> Kindaichi (1976) notes that diachronic changes into /oo/ and /ee/ show a time difference, with /o/+/u/ undergoing coalescence before /e/+/i/ did.

<sup>16</sup> In second language teaching material it is usually written as 'ee' though.

As described in Section 3, vowel sequences that cross syllable boundary are called hiatus, whereas those within a syllable are diphthongs.

Toneless Hanyu Pinyin distinguishes 410 syllables with different structure. The target of this research are open syllables which are candidates as V1 in a hiatus structure, and those syllables that begin with a vowel and can represent V2 in the same hiatus structure. Monophthongs and diphthongs are possible in both, V1 and V2 position. Such vowel sequences are marked clearly in Hanyu Pinyin.

When vowels /a/, /ə/, /o/ appear after a syllable break, i.e. at the beginning of syllable, there is an apostrophe before it, unless there are any other visual clues that indicate the break, such as a hyphen or a space.

- |      |                        |           |  |
|------|------------------------|-----------|--|
| (16) | /tʂ <sup>h</sup> aŋ.ʎ/ | chāo'é    | 'to exceed a quota'  |
|      | /pwo.ai/               | Bo'ai     | 'Bo'ai county in Jiaozuo, Henan'                             |
|      | /ou.an.xwei/           | Ōu-Ān-Huì | 'Organisation for Security and Co-operation in Europe, OSCE' |
|      | /wo.tʎ.aŋ.tsi/         | wode erzi | 'my son'   |

There are only four possible diphthongs, and they all begin from the position of either /a/ or /o/. The apostrophe rule applies to them as well (i.e. Bo'ai above).

- |      |    |      |                     |
|------|----|------|---------------------|
| (17) | ai | /ai/ | 'love'              |
|      | ei | /ei/ | 'hey' <sup>17</sup> |
|      | ao | /aŋ/ | 'proud'             |
|      | ou | /ou/ | 'surname Ou'        |

According to the above, vowel letter sequences in Hanyu Pinyin with no apostrophe (or other visual clues) demonstrate monosyllabic sound sequences. They can be either a diphthong or a glide + vowel sequences.

Diphthongs are vowel sequences that start with a vowel quality of higher sonority and end with a vowel of lower sonority. Examples thereof include:

- |      |     |       |             |
|------|-----|-------|-------------|
| (18) | mai | /mai/ | 'to buy'    |
|      | lei | /lei/ | 'tired'     |
|      | dao | /taŋ/ | 'to arrive' |
|      | dou | /tou/ | 'beans'     |

All other vowel sequences within a syllabus are a glide + vowel (GV) sequences. In Standard Chinese, glides are found when high vowels /i/ and /u/, which become realized as [j, ɥ, w], occur before another vowel. Consonant-glide combinations occur before nuclear vowels (Duanmu, 2007, p. 23).

<sup>17</sup> Diphthong /ei/ at the syllable initial position does not appear in any word and is limited to the character 诶, denoting an exclamation *hey*.

(19)	lie	/lje/	‘woodpecker’
	xue	/ɕɥe/	‘to learn’
	huo	/xwo/	‘fire’

→ In CGV syllables, glides are written with letters ‘i’ and ‘u’.

Glides do not contrast with the corresponding high vowels [i, u, y], and the two sets can be treated as variants of each other. (Qiang, 2016, p. 100)

When GV sequences appear at the syllable initial position, Hanyu Pinyin writes the glides with letters ‘y’ and ‘w’. Graphically speaking, Hanyu Pinyin visually indicates the hiatus with the presence of a consonant letter. A reader therefore understands the written sequence of letters as belonging to different syllables, as demonstrated in example (21).

(20)	ye	/je/	‘also’
	yue	/ɥe/	‘moon’
	wo	/wo/	‘I’

→ In GV syllables, glides are written with letters ‘y’, ‘yu’ and ‘w’.

According to the above described details about hiatus in Chinese, the following conclusions can be drawn. All vowel sequences that appear in hiatus are graphically marked in Hanyu Pinyin with punctuation marks, therefore we would expect a Slovene reader to naturally produce hiatus on the right places, without any explicit guidelines. Moreover, in line with the hiatus resolving tendencies in Slovene, we would also expect all other sequences of vowel letters to be pronounced as a diphthong or a glide + vowel.

## 5 Discussion: predicted and actual reading

Several factors motivated this research. They all concern existing Slovene pronunciation of Japanese or Chinese proper names, which can be surprisingly different from what a scholar on the languages would expect, sometimes with two different pronunciations for the same word. In this respect, we have limited our research to vowel sequences that either form a hiatus or look like a hiatus in their written forms. This section will compare phonologies of the source and the borrowing language, and based on the examples discuss the reasons for a discrepancy in pronunciation.

### 5.1 Hiatus resolutions involving second vowel

Consider visiting lexicon *Asai Chu*, a famous oil master from Japan, and Chinese female name *Sai Jinhua*. In this case, Slovene reading better approximates Chinese than Japanese reading.

- (22) a. **Asai** Chu  
 Japanese reading: /sa.i/  
 Slovene reading: /saj̥/  
 b. **Sai** Jinhua  
 Chinese reading: /saj̥/  
 Slovene reading: /saj̥/

### 5.1.1 Japanese

Let us first look at the following sentence that contains visiting words from Japanese.

- (23) Slovenski smučarski skakalci odhajajo na poletne priprave na japonski Hokkaidō, kjer se jima bosta pridružila japonska kolega, domačina Noriaki Kasai in Daiki Itō.

‘Ski jumpers from Slovenia are leaving for their summer camp to Hokkaido, Japan. Their colleagues and a locals there, Noriaki Kasai and Daiki Itō will join them.’

Pronunciation of the northernmost prefecture and the second largest island in Japan, Hokkaidō, which is sometimes slovenized into ‘Hokaido’, has two pronunciation variants, first one with a diphthong (like Slovene words ‘aids’) and second one with a hiatus resolved with an epenthetic /j/ and stress on the third syllable, i.e. vowel /i/ (like Slovene word ‘kokain’). The same description for /a/ + /i/ vowel sequence can be applied to the first name Daiki<sup>18</sup> and to all such sequences in a non-final word position. Cases of a /a/ + /i/ vowel sequence with which a word ends, as the family name Kasai in the above example, fail the heterosyllabic realization due to stress accent position rules in Slovene.

The solution with an epenthetic glide is also undesirable otherwise<sup>19</sup> because it forms a syllable with a glide + vowel /i/ that is, according to the Japanese phonotactic rules, impossible in Japanese. In addition, by placing stress onto the vowel /i/ in Slovene makes the word’s sound shape much different from the original. In the end, there is another simple solution with a diphthong.

- (24) Hokkaidō, Kasai, Daiki  
 Japanese pronunciation: /a.i/, [aj̥]  
 Slovene pronunciation in use:  
 a + i → aj̥ or \*a + i (word non-final) → a.i [a.'ji]

Other examples include: Sendai (city), Daiki (first name), Aiko (first name), Saitō (family name), etc.

<sup>18</sup> Also with common nouns ‘haiku’, ‘aikido’, ‘bonsai’ or slovenized ‘bonsaj’ etc.

<sup>19</sup> Therefore marked with an asterisk in the below description.

Interestingly, this simple rule can be applied to all Japanese vowel sequences with /i/ as the second vowel.

(25) Diphthongization of Japanese V + i sequence:

- a + i → a<sub>i</sub> ex. Aichi (prefecture), Saiki (city), Imai (family name), Kaito (first name)  
e + i → e<sub>i</sub> ex. Shiseidō (company name), Keiyō line (railway line), Ryūhei (first name)  
i + i → i<sub>i</sub> ex. Niigata (prefecture), Miike (family name), Fujii (family name)  
o + i → o<sub>i</sub> ex. Oita (city), Aoi (first name)  
u + i → u<sub>i</sub> ex. Fukui (prefecture), Matsui (family name), Yui (first name)

A note should be made for a rather unusual sequence ij, which could be compared to a Slovene pronunciation of spelling 'ij', such as in 'akvarij' (aquarius) or 'baterijska reja' (battery cage). All other sequence of the same vowels are very often marked unsystematically in the Hepburn Latin script (see section 1.2) but when they are, they may get reduced to a single vowel, or else may undergo epenthesis under the condition that the second of the vowels does not receive stress. Possible outcomes need further investigations.

Diphthongization is a possible solution for a vowel + /u/ sequence in Japanese, besides glide epenthesis. However, due to diachronic sound changes, examples with such a sequence within a phonological word are very rare when we exclude dictionary forms of verbs, which fall out of the scope of visiting lexicon. Most representative is the following sequence.

(26) i + u → i<sub>u</sub> or i + u → i.ju ex. Miu (first name)

Finally, all other vowel sequences in Japanese would keep heterosyllabicity by epenthesis of some kind. Such sequences are not new in Slovene; they appear in loanwords from various languages.

(27) Jōetsu shinkansen (bullet train line) /dʒo.e.tsu/  
Maebashi (city) /ma.e.ba.fi/  
Hayao (first name) /ha.ja.o/

### 5.1.2 Chinese

There is no difference in the predicted and the actual Slovene reading of the family name *Sai*. As described in Section 4.2, Hanyu Pinyin clearly graphically marks hiatus. All other vowel sequences are monosyllabic. Moreover, because of the hiatus resolving tendencies, we do expect a Slovene speaker to pronounce them correctly.

On the other hand, Slovene speakers pronounce the word *laoma* 'mum' counter to expectations, as a hiatus.

- (28) written form:                      laoma    /la<sub>ɔ̃</sub>.ma/  
 expected reading in Slovene:    lao.ma   /la<sub>u</sub>.ma/  
 actual reading in Slovene:        la.o.ma   /la.o.ma/

One might choose either the first or the second vowel for the word stress (aó<sub>ɔ</sub>ta, ká<sub>o</sub>s), however the preliminary testing group of Slovene speakers uniformly decided to stress the letter “o”.

In Chinese, the sonority scale plays a key role in relation to the proper pronunciation. The syllable core is a vowel that is, compared to the surrounding vowels, most sonorous and in writing appears ahead according to the order of the letters a-o-e-i-u-ü. These orthographic guidelines are a reflection of the phonetically grounded sonority scale (Parker 2002, p. 236), which lists the vowels from highest to lowest in the following order:

- (29) Sonority scale (more sonorous → less sonorous)
- |                            |       |
|----------------------------|-------|
| low (open vowels)          | [a]   |
| mid vowels                 | [e o] |
| high vowels (close vowels) | [i u] |
| schwa                      | [ə]   |
| glides (semivowels)        | [j w] |

## 5.2 Hiatus resolutions involving first vowel

Consider different vowel sequences in the following sentence.

- (30) Moja kitajska prijateljica **Mie** se je ravnokar vrnila z Japonske, kjer je dva tedna preživela v prefekturi **Mie**.

‘My Chinese friend **Mie** has just returned from Japan, where she spent two weeks in **Mie** prefecture.’

In the above example hints on how to read the two proper names ‘Mie’ are given (e.g. Chinese friend, Japan), however, without understanding the interrelationships between written and spoken form for each language separately, a Slovene reader will pronounce both as /mi.je/, as in Slovene loanwords ‘Vietnam’, ‘patient’, ‘siesta’ and others. The pronunciation is thus closer to the proper Japanese reading, and not in line with the Chinese reading /mje/.

Languages employ different hiatus-avoiding mechanisms and the basic difference between Japanese and Chinese is that only the latter language uses glide formation.

### 5.2.1 Japanese

Previous example shows that Japanese is not fond of changing the first vowel of a hiatus into a glide to make the vowel sequence tautosyllabic. Glide formation does not

happen neither on phonological nor on phonetic level, and this relates to all potential vowel sequences. Instead, epenthesis is applied. What is interesting though is the accent placement because either of the two vowels may carry accent after the hiatus is resolved. The topic calls for further investigation.

- (31) Uemura (family name) /u.e.mu.ra/  
Mie prefecture /mi.je/  
Shizuoka prefecture /ʃi.zu.o.ka/  
Noriaki (first name) /no.ri.ja.ki/

### 5.2.2 Chinese

In Hanyu Pinyin, sequences of vowel letters within a syllable that are not diphthongs are GV sequences. According to the sonority principle, family name *Mie* from example (30) above is pronounced as /mje/. However, the written form i+e leads a Slovene reader to hiatus.

An interesting “wrong” choice also appears in the proper name *Huawei*. The GV sequence in the syllable *hua* /xwa/ is for a Slovene speaker understood as hiatus.

- (32) written form: huawei /xwa.wei/  
expected reading in Slovene: hua.wei /xwa.wei/  
actual reading in Slovene: hu.a.wei /'xu.a.wei/

Mass media with its frequent advertising can often be the source of a mispronunciation, as in the case of brand name Huawei. The speakers follow the pattern that appears in Slovene word *aktualen*, rather than *donhuan*. GV sequences are not foreign to Slovene language, therefore the discrepancy between predicted and actual reading should be overcome easily with suitable pronunciation guidelines.

## 6 Conclusion

Languages of the world differ in the way and in the extent at which their writing systems reflect actual spoken forms; they generally insert all the information they presume would well reflect relevant information. In this respect, Latin scripts of the languages that otherwise use (at least primarily) other writing systems may be more imprecise. Vowel letter sequences in Japanese and Chinese are yet another such example.

This research focused on vowel sequences in Japanese and Chinese and examined their representation in the Japanese and Chinese official Latin scripts. It placed special attention to the appearance of vowel hiatus in the two languages, on vowel letter sequences that suggest vowel hiatus, and on how each language treats the cases. As Slovene tends to avoid vowel hiatus, the research reviewed hiatus-avoiding mechanisms in Slovene, and discussed the reasons for deviances in the pronunciation



of a visiting lexicon. In the end, the research offered some guidelines on the pronunciation on vowel sequences in words of a visiting lexicon from Japanese and Chinese.

In Japanese, all vowel sequences phonologically function in hiatus, with each monophthong vowel functioning as a syllable core. Hepburn Latin script is consistent with this in case of two different vowels. Entering Slovene as a visiting lexicon, such vowel letter sequences will certainly undergo hiatus-avoiding changes. It is suggested that they follow basic tendencies in the source language that can be summarized as follows.

- a) Diphthongization for  $V + /i/ \rightarrow V_i$ ,  $V + u \rightarrow V_u$
- b) Epenthesis for all the other vowel sequences

Phonetic reduction into one syllable is already observed in Japanese but is limited to the second of the two vowels under the condition that it is either of the two high vowels, e. g. /i/ or /u/. Hiatus resolution that would make a glide + vowel combination is thus not possible because Japanese already has such a sequence. Any other vowel sequence is advised to keep its heterosyllabic property, cf. with an epenthesis of some kind. This kind of solution does not sound strange in Slovene, and could also be applied to the sequences of two equal vowels.

Quite on the contrary, Chinese favors monosyllabic vowel sequences based on the sonority principle. Sonority scale is also incorporated into the Hanyu Pinyin script in the sense that pre-vowel glides /j/ and /w/ are represented by letters 'i' and 'u' respectively in the CGV combinations, which may wrongly lead to the idea of a hiatus. On the other hand, the difference between diphthongs and vowel hiatus in a script is stated clearly. Vowel sequences in Chinese can thus be summarized as follows.

- a) Monosyllabic vowel sequences: diphthong, glide + vowel (/j, w/ + V)
- b) Heterosyllabic vowel sequences: hiatus (graphically marked)

Except for the glide + vowel sequence, which in Hanyu Pinyin falsely gives an impression of a hiatus, there is not much to be careful about hiatus vowel sequences in Chinese. There are some other discrepancies between the written and spoken form of Hanyu Pinyin, which may cause additional disputes in relation to the visiting lexicon in Slovene. For example, in the notation consonant + ui, e.g. *gui*, the vowel sequence of this syllable is phonologically /uei/, not \*/ui/. As we can see, the main vowel /e/ is left out in standard orthography (Tříšková, 2011, p. 99).<sup>20</sup>

---

<sup>20</sup> Note, however, that some Latin scripts are similar in this respect. The same sequence 'ui' is pronounced differently in French words *acquiets* [akjɛʁ], *huit* ['ɥi(t)], *taquin* [takɛ̃], etc. Not being able to produce "proper reading" of foreign words should not be the main argument for slovenization of any Latin script.

Whether a diphthong, a glide before a vowel, or a hiatus, the decision for a metrical organization in Slovene influences word stress placement, and consequently conjugation and declination forms of the visiting lexicon. This calls for a similar investigation on Japanese and Chinese consonants, and the overall examination before final recommendations about the “proper” pronunciation of Japanese and Chinese visiting lexicon are offered.

## References

- Athletes. (n.d.). Retrieved from <https://www.olympic.org/pyeongchang-2018/results/en/general/athletes.htm>
- Breznik, A. (1920). *Slovenski pravopis* [Slovene Orthography]. Ljubljana.
- Carlisle, R. S. (2001). Syllable structure universals and second language acquisition. *International Journal of English Studies*, 1(1), 1-19.
- Casali, R. (1997). Vowel Elision in Hiatus Contexts: Which Vowel Goes? *Language*, 73(3), 493-533.
- Casali, R. (2011). Hiatus resolution. In M. van Oostendorp, C. Ewen, E. Hume & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 1434-1460). Oxford: Wiley Blackwell Publishers.
- Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.
- Detey, S., & Nespoulous, J. L. (2008). Can orthography influence second language syllabic segmentation?: Japanese epenthetic vowels and French consonantal clusters. *Lingua*, 118(1), 66-81.
- Dobrovoljc, H. (2013). Tuja imena vladarjev: prevesti ali ne? *Jezikovna svetovalnica*. Retrieved from <https://svetovalnica.zrc-sazu.si/topic/304/tuja-imena-vladarjev-prevesti-ali-ne>
- Duanmu, S. (2002). Two Theories of Onset Clusters. *Chinese Phonology* [音韻論叢], 11, 97-120. Retrieved from <http://www-personal.umich.edu/~duanmu/CR02.pdf>
- Duanmu, S. (2007). *The phonology of Standard Chinese* (2nd edition). Oxford & New York: Oxford University Press.
- Frellesvig, B. (2010). *A History of the Japanese Language*. Cambridge: Cambridge University Press.
- Hadamitzky, W., & Spahn, M. (2005). *Romanization systems*. Wolfgang Hadamitzky: Japan-related Textbooks, Dictionaries, and Reference Works. Retrieved from [https://www.hadamitzky.de/english/lp\\_romanization\\_sys.htm](https://www.hadamitzky.de/english/lp_romanization_sys.htm)
- Ilc, I. (2009). Slovenjenje v posrednih prevodih japonske književnosti. *Azijske in afriške študije*, 13(2), 39-70.
- International Phonetic Association (1999). *Handbook of the International Phonetic Association: A guide of the use of the International Phonetic Alphabet*. New York, Cambridge: Cambridge University Press.
- Ito, J., & Mester, A. (1999). The Phonological Lexicon. In N. Tsujimura (Ed.), *The Handbook of Japanese Linguistics* (pp. 62-100). Oxford: Blackwell Publishers.

- Jurģec, P. (2004a). Fonologija samoglasniških nizov v slovenščini. *Slavistična revija* LII/2, 119-140.
- Jurģec, P. (2004b). Antihiatki pojavi v knjižni slovenščini [Hiatus Avoiding Processes in Standard Slovene]. *Jezikoslovni zapiski*, 10(1), 125-144.
- Kawahara, S. (2003). On a certain type of hiatus resolution in Japanese. *Phonological Studies* 6, 11-20.
- Kenstowicz, M. (1994). *Phonology in generative grammar*. Cambridge, Mass. & Oxford: Blackwell.
- Kindaichi, K. (1976). *Nihongo no hensen [Historical development of Japanese]*. Tokyo: Kodansha.
- Kubozono, H. (2001). On the markedness of diphthongs. *Kobe Papers in Linguistics* 3, 60-73.
- Kubozono, H. (2015). Diphthongs and vowel coalescence. In H. Kubozono (Ed.), *Handbook of Japanese Phonetics and Phonology* (pp. 215-249). Berlin: De Gruyter Mouton.
- Labrone, L. (2012). *The phonology of Japanese*. Oxford: Oxford University Press.
- Ladefoged, P. (1982). *A course in phonetics*. Los Angeles: University of California.
- Lengar Verovnik, T. (2017). Kako zapisati ime nobelovca Ishigura, Britanca japonskega rodu? *Jezikovna svetovalnica*. Retrieved from <http://svetovalnica.zrc-sazu.si/>
- Lindau, M., Norlin, K., & Svantesson, J. O. (1990). Some cross-linguistic differences in diphthongs. *Journal of the International Phonetic Association*, 20(1), 10-14.
- Mlakar, B. (2009). Pregled sistemov latiničnega zapisa japonskega jezika. *Azijske in afriške študije*, 13(2), 27-38.
- Nobel Prizes 2017. (n.d.). Retrieved from [https://www.nobelprize.org/nobel\\_prizes/lists/year/index.html?year=2017&images=yes](https://www.nobelprize.org/nobel_prizes/lists/year/index.html?year=2017&images=yes)
- Parker, S. (2002). *Quantifying the Sonority Hierarchy*. PhD dissertation. U Mass, Amherst.
- Petrovčič, M. (2009a). Predlogi za zapisovanje in pregibanje besed iz kitajskega jezika. *Azijske in afriške študije*, 13(2), 15-24. Also available at <http://as.ff.uni-lj.si/datoteke/nid/38/zapiskitajscine.pdf>
- Petrovčič, M. (2009b). Pregled načinov latinizacije kitajskega jezika in primerjava pglavitnih sistemov. *Azijske in afriške študije*, 13(2), 83-102. Also available at <http://as.ff.uni-lj.si/datoteke/nid/38/primerjavalatinizacijkitajscina.pdf>
- Petrovčič, M. (2009c). Pregled paragrafov Slovenskega pravopisa s primeri iz kitajskega jezika. *Azijske in afriške študije*, 13(2), 103-113. <http://as.ff.uni-lj.si/datoteke/nid/38/kitajscinaprimeripravopis.pdf>
- Picard, M. (2003). On the emergence and resolution of hiatus. *Folia Linguistica Historica*, 24(1-2), 47-57.
- Pogačnik, A. (2012). Glasovno domačenje lastnih imen iz nelatiničnih pisav. In N. Jakob & H. Dobrovoljc (Eds.), *Pravopisna stikanja: razprave o pravopisnih vprašanjih*, (pp. 73-83). Ljubljana: Založba ZRC.
- Pronounce Names. (n.d.). Retrieved from <https://www.pronouncenames.com/search?name>
- Qiang, L. (2016). An Investigation of Sonority Theory in Mandarin Chinese. In M. J. Ball & N. Müller (Eds.), *Challenging Sonority. Cross-linguistic Evidence* (pp. 97-109). Equinox Publishing, United Kingdom.

- Geodetska uprava Republike Slovenije (2001). *Seznam tujih zemljepisnih imen v slovenskem jeziku*. Ljubljana. Retrieved from [http://www.gu.gov.si/fileadmin/gu.gov.si/pageuploads/publikacije/arhiv\\_publik/seznam\\_tujih\\_zemljepisnih\\_imen.pdf](http://www.gu.gov.si/fileadmin/gu.gov.si/pageuploads/publikacije/arhiv_publik/seznam_tujih_zemljepisnih_imen.pdf)
- Slovenski pravopis. Pravila in slovar*. (2001). Ljubljana: Založba ZRC, ZRC SAZU. Retrieved from [http://bos.zrc-sazu.si/c/sp/sp2001\\_pravila.pdf](http://bos.zrc-sazu.si/c/sp/sp2001_pravila.pdf)
- Šuštaršič, R., Komar, S., & Petek, B. (1999). Slovene. In *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet* (pp. 135-139). Cambridge: Cambridge University Press.
- Takayama, T. (2003). Gendai Nihongo no Onin to sono Kinou [Phonology of modern Japanese and its function]. In Y. Mitahara (Ed.), *Onsei, Onin [Phonetics and Phonology]*. Tokyo: Asakura Shoten.
- 'Ti' or 'chi'? Educators call to unify romanization styles in Japan - The Mainichi. (2017, April 02). Retrieved from <https://mainichi.jp/english/articles/20170402/p2a/00m/0na/003000c>
- Toporišič, J. (2011). *Intervjuji in polemike*. *Linguistica et philologica* 25. Ljubljana: ZRC SAZU.
- Trask, R. L. (1996). *Historical Linguistics*. London: Arnold.
- Třísková, H. (2011). The Structure of the Mandarin Syllable: Why, When and How to Teach it. *Oriental archive*, 79, 99-134.
- Zárate-Sández, G. (2011). Selected Proceedings of the 2009 Second Language Research Forum. In L. Plonsky & M. Schierloh (Eds.), *Cascadilla Proceedings Project*, (pp. 164-181). Somerville: MA.