

NGEfreak
12th June 2007, 17:25

ISO with modified particles rules?

massin
21st June 2007, 06:13

I prefer transcriptions that represent pronunciation rather than cryptic phonetics. So using o, e, wa, etc is the way to go. Plus, we don't live in the past, and modern Japanese follows these rules in present day. Maybe it would be more reasonable to have different naming for different time periods but my guess is that "modern" refers to the last 100-200 years rather than something recent which could affect console naming. And finally, from my experiences, I've found it to be easier on both Japanese and non-Japanese readers to read romanized names that have a bias for speech.

NGEfreak
25th June 2007, 06:14

I prefer wo/ha/he more than o/wa/e because

- wo/ha/he is based on writing, o/wa/e is based on spelling
- the furigana uses wo/ha/he
- Japanese use wo/ha/he to input characters

If you have a bias towards speech then why don't you write for example "desu" just as "des"?

massin
30th June 2007, 17:21

Two reasons:

1. Writing with a bias toward speech doesn't mean you necessarily write everything as it sounds. It would be nice if one could simply take a Japanese character and say it has one sound and that's that. But I'm sorry, Japanese does have exceptions just like every other language. Why are we being ignorant to this point? Keep in mind that this is a transcription. With that said, I will say this again: In my encounters with Japanese people, they say that its easier to read wa, e, o. However, we and wo are generally understood while using ha is more often than not confusing compared to using wa. Furthermore, those changes are accepted changes to the language. If Japanese was written originally using romanized characters we wouldn't be having this debate. But the art of transcription is that we are continuously refining our methods and the ISO + Exceptions is what I consider to be the most refined method of transcription based on pronunciation of words and based ease of reading for the Japanese people I know. Although I disagree with Lukeage that we should be making Dats biased to English speaking people, this specific choice is actually more beneficial for English speaking people because they will be able to read the titles more accurately simply because they would be written as they sound. Using something like MS IME as a reason to use ha/we/wo is kinda silly. The only reason they use those input characters is because there is no other way to distinguish between "o" as a particle and "o" as a vowel (or whatever its called for Japanese).

2. Now my second point is simple. "desu" is not pronounced "des". It just sounds that way and that's the simplest way to teach a non-Japanese speaker. I've had this discussion with a number of Japanese people. The "u" is "pronounced" but its necessarily vocalized.

And the truth of the matter is while we squabble about this silly issue of three or some odd characters, there are much more important issues to do with romanization that we haven't discussed. For example:

1. How do we deal with names that have an "official" romanization that does not follow our standard?
2. How do we chose when a "word" is a "word". I could write the same thing a few different ways but it would still mean the same thing. One particular situation comes to mind where I was trying to fix "Dai 1 Kan" and I could write it as "daiikkan" or "dai ikkan" or "dai-ikkan". Technically all of them are right but that leaves a lot of ambiguity in our naming.
3. How do we deal with situations like the example above where you see something like Dai 1 Kan or 10-tsu. I've made the choice to convert them all the written Japanese but no one else really considers it one way or the other.
4. What do we do about these long Japanese titles? Some of these title are almost like small paragraphs with things like "This is a hard game. Do your best" (I just made that up but there are many games like that). Do we try to extract a shortened, to the point, title or do we leave in these long paragraph titles? In my opinion its silly to have such long titles but then on the other hand I've been told that those are really part of the title. Its simply a cultural difference.
5. How do we handle games that have both an English title and a Japanese title? Include both or what?
6. Do we try to translate Katakanaized foreign words (from a Japanese point of view) to their original language. At this point I do that, and for the most part the GBA dat has that. But what's the rule if there is one?

Lukeage

30th June 2007, 18:45

Furigana is the use of Hiragana to represent Kanji.

The three main ones given on wiki (of which 2 we are considering) all use o/wa/e

Addressing some of massin's concerns:

- 1 - Personally I feel we just convert the title into our Romaji system to keep with conventions. While it might be annoying searching for a title which has slightly different romaji, given our naming system will be standardised, 'converting' the title is practically a non issue. I think it would be worse searching for a title which doesn't conform to our conventions since you don't know what it conforms to anyway.
- 2 - Good question. In the case given, I'd suggest as a single word given dai is a prefix and kan is a suffix, that is, dai 'converts' the 1 to 'first' then kan, being a counter, changes it to what the object is (in this case, a counter for volumes). Perhaps this is how we should handle these cases? if they aren't suffix or prefix words then they should be separate.
- 3 - I'm undecided at this point (despite previous discussions.) Is a non-issue if we go with the ha/he/wo particle romanisation, as this will need to follow how it is written too.
- 4 - Big issue, especially now that we have a game which the title contains more characters than any common file system can handle. While I'd probably agree with writing the entire title, at a minimum, we need to come up with something to handle the 256+ character titles.
- 5 - Wasn't there a decision on this in regards to megadrive titles?
- 6 - Another non-issue if we go with the ha/he/wo particles. In the past when they have been changed, has it only been done for the one's originating from English, or for all languages? I personally prefer the change, however, when spoken, sometime the Japanese are completely dumbfounded if the word isn't 'Japanised' when pronouncing it.

massin
30th June 2007, 18:55

Can you explain what you mean by 4 and 6 being a non-issue with the use of ha/he/wo? I'm not quite seeing what you mean?

When I mean for 5 may or may not be different from the mega drive discussion. What I'm referring to are games that on the box or screen have both titles. Some NDS games are like this. What I'm not referring to are games that have the same rom released under different names. Is this considered the same issue as before?

Lukeage
1st July 2007, 00:52

If we are using ha/he/wo because that is how they are written, rather than how they are pronounced, then I don't see why this rule won't be applied elsewhere too. Hence, if 5tsu is written on the box then 5tsu needs to be in the dat, not itsutsu. As for 6, the original word is possibly neither how it is spelt or pronounced in Japanese (obviously words like Banana translate nicely).

We also need to discuss long vowels as we are restricted to 7-bit ASCII values which do not contain any of the accents above the vowels.

massin
4th July 2007, 06:33

So far I follow this:

i- = ii
o- = ou
u- = uu
e- = ei (i think this is how its supposed to go, i forget)
a- = aa (same as above)

Not to downplay the issue, but I don't see how this is any different than how we have been doing it before? Seems like a simple solution. Is there any conflict in particular?

Lukeage
4th July 2007, 07:30

Wasn't as much an issue, rather than actually formalising the rule as the romaji standards both use accents.

NGEfreak
5th July 2007, 08:18

First a general overview of all topics regarding Japanese romanization, if you have more please add them.

1. Romanization system:

Please read this first:
http://www.hadamitzky.de/english/lp_romanization_sys.htm

Comparison table:

Hiragana	Hepburn	Kunrei (ISO)	Kunrei exceptions
あ	a	a	
か	ka	ka	
さ	sa	sa	
た	ta	ta	
な	na	na	
は	ha	ha	
ま	ma	ma	
や	ya	ya	
ら	ra	ra	
わ	wa	wa	
が	ga	ga	
ざ	za	za	
だ	da	da	
ば	ba	ba	
ぱ	pa	pa	
い	i	i	
き	ki	ki	
し	shi	si	shi
ち	chi	ti	chi
に	ni	ni	
ひ	hi	hi	
み	mi	mi	
り	ri	ri	
ゐ	wi	i	
ん	n	n	
ぎ	gi	gi	
じ	ji	zi	ji
ぢ	ji	zi	di
び	bi	bi	
ぴ	pi	pi	
う	u	u	
く	ku	ku	
す	su	su	
つ	tsu	tu	tsu
ぬ	nu	nu	
ふ	fu	hu	fu
む	mu	mu	
ゆる	yu	yu	
る	ru	ru	
ぐ	gu	gu	
ず	zu	zu	
づ	zu	zu	du
ぶ	bu	bu	
ぷ	pu	pu	
え	e	e	
け	ke	ke	
せ	se	se	
て	te	te	
ね	ne	ne	
へ	he	he	
め	me	me	
れ	re	re	
ゑ	we	e	
げ	ge	ge	
ぜ	ze	ze	
で	de	de	
べ	be	be	
ぺ	pe	pe	
お	o	o	
こ	ko	ko	
そ	so	so	
と	to	to	
の	no	no	
ほ	ho	ho	

も	mo	mo	
よ	yo	yo	
ろ	ro	ro	
を	wo	o	wo
ご	go	go	
ぞ	zo	zo	
ど	do	do	
ぼ	bo	bo	
ぽ	po	po	
き	kya	kya	
しゃ	sha	sya	sha
ちゃ	cha	tya	cha
に	nya	nya	
ひ	hya	hya	
み	mya	mya	
り	rya	rya	
ぎ	gya	gya	
じゃ	ja	zya	ja
ぢゃ	ja	zya	dya
び	bya	bya	
ぴ	pya	pya	
き	kyu	kyu	
しゅ	shu	syu	shu
ちゅ	chu	tyu	chu
にゅ	nyu	nyu	
ひゅ	hyu	hyu	
みゅ	myu	myu	
りゅ	ryu	ryu	
ぎゅ	gyu	gyu	
じゅ	ju	zyu	ju
ぢゅ	ju	zyu	dju
びゅ	byu	byu	
ぴゅ	pyu	pyu	
き	kyo	kyo	
しょ	sho	syo	sho
ちよ	cho	tyo	cho
にょ	nyo	nyo	
ひょ	hyo	hyo	
みょ	myo	myo	
りょ	ryo	ryo	
ぎょ	gyo	gyo	
じょ	jo	zyo	jo
ぢょ	jo	zyo	djo
びょ	byo	byo	
ぴょ	pyo	pyo	
く			kwa
ぐ			gwa

Example:

Kanji: 筋肉番付 GB 挑戦者はキミだ!
 Kana: きんにくばんづけ ちょうせんしゃはきみだ

Hepburn: Kinniku Banzuke GB - Chousensha ha Kimida!
 Kunrei: Kinniku Banzuke GB - Tyouzensya ha Kimida!
 Mod. Kunrei: Kinniku Banzuke GB - Chousensha ha Kimida!

2. Particles は, へ and を

Particle は as wa

- better represents modern pronunciation

- common practice

Particle は as ha

- better represents the writing

Particle へ as e

- better represents modern pronunciation
- more common in anime/videogame communities (球へ... (Terra e...), etc.)

Particle へ as he

- better represents the writing

Particle を as o

- better represents modern pronunciation

Particle を as wo

- better represents the writing
- more common in anime/videogame communities (満月をさがして (Full Moon wo Sagashite), エースをねらえ! (Ace wo Nerae!), etc.)

3. っ when geminate consonant

Usually geminate consonants are marked by doubling the consonant following the sokuon. However this simple rule gets complicated by one particular case. For っち/っちや/っちゆ/っちよ the cluster tch is maybe a better transcription than cch. Though, making a general rule is very difficult.

Examples:

たまごっち is better known as Tamagotchi than Tamagocchi.

On the other hand:

えっち is better known as Ecchi than Etchi

4. っ when exclamation

っ can be used as exclamation mark:

Popular example:

ああっ女神さまっ

Aa! Megami-sama!

Alternatives:

- discard っ at the end of a word
- transcribe っ at the end of a word as 'h'

5. Long vowels

Tōkyō: indicated with macrons. This follows the rules of the traditional and revised Hepburn systems, and is considered to be standard.

Tokyo: not indicated at all. This is common for Japanese words that have been adopted into English. This 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.

T?ky?: indicated with circumflexes. Circumflexes are how long vowels are indicated by the alternative Nihon-shiki and Kunrei-shiki romanizations. Circumflexes are often used when a word processor does not allow macrons. With the spread of Unicode, this is becoming rare.

Tohkyoh: indicated with an "h". This is sometimes known as "passport Hepburn", as the Japanese Foreign Ministry has authorized (but not required) this usage in passports. [5]

Toukyou: written using kana spelling: ō as ou or oo (depending on the kana) and ū as uu. This is sometimes called wāpuro style, as this is how text is entered into a Japanese word processor (wādo purosessā) using a keyboard with Roman characters. This method most accurately represents the way that vowels are written in kana, differentiating between おう (as in とうきょう (東京), written Toukyou in this system) and おお (as in とおい (遠い), written tooi in this system).

Tookyoo: written by doubling the long vowels. This follows the rules of the modified Hepburn system, but is also common when writing words of foreign origin without reference to any particular system, i.e. paatii for パーティー ("party") instead of pātī. This is also used in the JSL form of romanization.

6. Capitalization

Commonly the first letter of all words are capitalized, particles and subfixes are not.

There exists also other common styles like for example writing all katakana in ALL CAPS and other words lowercased, etc.

7. Honorifics and subfixes

Honorifics (san, sama, chan, etc) that are no standalone words are usually added with '-'. Otherwise just seperated by space.

Chibi Maruko-chan
Mononoke Hime

Usually Japanese suffixes are seperated by '-' if they are included with a foreign word.

Lodoss-tou Senki

たち (tachi) is usually split from the main word.

Bokujou Monogatari - Mineral Town no Nakama-tachi

8. Loan words and wasei eigo terms

Many words from other languages are used in Japanese. If possible they are usually written as in their original language.

There are also several abbreviated and contracted words:

Example:

ファミコン famikon -> Famicom -> Family Computer

プロレス puroresu -> Pro Wres -> Pro Wrestling

All three forms could be used for romanization.

More examples: http://en.wikipedia.org/wiki/List_of_Gairaigo_and_Wasei-eigo_terms

9. Numerals and latin letters

Numbers in words are often abbreviated with numerals:

Example:

第3次スーパーロボット大戦 (Dai-3-ji Super Robot Taisen)

第3次 could be written

- Dai 3 Ji

- Dai-3-ji

- Daisanji

- Dai-Sanji

etc

Sometimes they may have special reading:

Example:

ボボボーボ・ボーボボ 奥義 87.5 爆烈鼻毛真拳 (Boboboubo Boubobo - Ougi 87.5 Bakuretsu Hanage Shinken)

87.5 is spelled はなげ (hanage)

Also, letters could have a special reading:

Examples:

新機動戦記ガンダム W (Shin Kidou Senki Gundam W)

The W is spelled ウイング (wingu)

vs (versus) is often spelled たい (tai).

10. Official romaji and reimported character names

Sometimes there exists also an official romaji title. All 3 major romanization systems with all variations are often used for them.

Sometimes fucked up Americanized names get reimported:

フリーザ (Fuuriza) Dragon Ball character

Wrong: Frieza
Wrong: Freeza
Correct: Freezer

孫悟空 (Son Gokuu) Dragon Ball character:
Wrong: Son Goku
Wrong: Son Gokuh
Wrong: Son Gokou
Correct: Son Gokuu

11. Special readings

Some titles may have special readings (furigana) for some kanjis:

法騎士レイアース -> Magic Knight Rayearth, not Mahou Kishi Rayearth

This is sometimes only indicated on the box or ingame title, not both.

Sometimes it's not indicated at all:

ドラゴンボールZ 超サイヤ伝説 -> Dragon Ball Z - Super Saiya Densetsu

In this case both, the ingame title and box, do not have any furigana clues.
Though, 超サイヤ is always spelled as Super Saiya and not Chou Saiya.

12. English extra titles

Sometimes, an English title is included on the box and/or ingame titlescreen. This may be a translation or completely different:

スーパー人生ゲーム (Super Jinsei Game) also has the translation "The Game of Life" included.

トップをねらえ! (Top wo Nerae!) also has "Gunbuster" included.

NGE freak
5th July 2007, 08:18

1)

I think that Hepburn is the best choice. It's the most common used system in Japan and internationally and thus the de-facto standard. I dislike Kunrei and Nippon system because they hardly reflect pronunciation.

Kunrei with modifications may be the second best "system". However, for a convention it is the worst choice, because it's a fucked up mix of all three major systems, Kunrei, Hepburn and Nippon.

A comparison of the three Romanization systems shows that Hepburn romanization is the only one that is suitable for all three purposes of a transcription system: as an aid for pronunciation, data retrieval, and keyboard input.

2)

I may accept o/wa/e, but only if we use Hepburn. Otherwise the reasoning that it better reflects pronunciation would be void.

3)

I don't know. This really sucks.

I guess it's best to allow both, but with general bias towards cch.

4)

I prefer to use an exclamation mark.

5)

Letters with macrons and circumflexes can't be used because High ASCII seems to be problematic on certain localized systems.

I think the best choice is to use kana spelling. However for certain names and places the long vowel should be just ignored (ex. Toukyou -> Tokyo, Nintendou -> Nintendo).

6)

I prefer the common standard.

7)

Yes, prefixes should be added with a "-" and lower cased.

8)

If possible they should be written as in their original language.

For abbreviated words I think all three forms should be allowed.

Famicom for ファミコン is alright, but Pro Wres or Pro Wrest for プロレス is just odd. Pro Wrestling should be used. Sometimes even kana spelling may be better.

9)

I prefer to not write out the numerals. We don't do that in titles from other languages too (i.e. we don't write Final Fantasy Six instead of Final Fantasy VI).

All elements should be separated by "-" and suffixes lowercased.

Example:

Wrong Dai 3 Ji

Wrong Dai-3-Ji

Ok Dai-3-ji

I think that Dai-3-ji is much more recognizable by everyone than Daisanji.

Also, special reading should be not applied for numerals and latin letters.

10)

Ignore official Romanization and fucked up names.

11)

Yes, the furigana should have the highest priority, no matter from which source (box, title, etc), and should be applied (except for numerals and latin letters).

12)

Ignore them, except if they represent the reading.

Yakushi~Kabuto
5th July 2007, 08:49

Just want to point one thing you seemed to have overlooked.
We are allowed to use ISO- permitted exceptions but they are not mandatory.
There is no rule saying that once you use one exception, you have to use only exceptions.

Therefore your example:

Kanji: 筋肉番付 GB 挑戦者はキミだ!
Kana: きんにくばんづけ ちょうせんしゃはきみだ

Hepburn: Kinniku Banzuke GB - Chousensha ha Kimida!
Kunrei: Kinniku Banzuke GB - Tyousensya ha Kimida!
Mod. Kunrei: Kinniku Banduke GB - Chousensha ha Kimida!

should be

Kanji: 筋肉番付 GB 挑戦者はキミだ!
Kana: きんにくばんづけ ちょうせんしゃはきみだ

Hepburn: Kinniku Banzuke GB - Chousensha ha Kimida!
Kunrei: Kinniku Banzuke GB - Tyousensya ha Kimida!
Mod. Kunrei: Kinniku Banduke GB - Chousensha ha Kimida!
or Kinniku Banzuke GB - Chousensha ha Kimida! (note that it matches Hepburn)
or Kinniku Banzuke GB - Tyousensha ha Kimida!
...
...

...there are up to 2^3 ways to write it using permitted exceptions.

While it may appear chaotic, I think it is a system that gives datters a lot of initiative. Also, I like how it distinguishes characters better (じゆ & ぢゆ, じよ & ぢよ...)

NGEfreak
19th September 2007, 13:50

That's exactly what we don't want. The goal is to have a consistent style and not simply allow everything.

Lukeage
5th July 2007, 08:53

1) I agree with NGE. Hepburn is the best choice. Kunrei with modifications may work, but to me it is almost turning it into Hepburn anyway, so why not just go with Hepburn?

2) o/wa/e

3) Other than Tamagotchi, can you give any good examples? I feel that Tamagotchi could fall under official romajianisation rule (see 10). I'd probably prefer the cchi in the majority of cases.

4) Exclamation

5) If we aren't using High ASCII, then the o->ou etc.. rules. Tokyo and Nintendo can fall under rule 10.

6) Common. I guess it feels more like English then.

- 7) - with lower case is how I've always done it for my friends names.
- 8) I prefer original language.
- 9) Still undecided.
- 10) If there is an official or reimported name, that to me defines the word in English. Hence it now falls under rule 8. I'm not holding hard and fast to this decision however.
- 11) If it is obviously using the Kanji as a short hand for an English word(s), then rule 8 applies
- 12) Ignore it, however use it as an Official romajinisation or reference if needed.

massin

6th July 2007, 04:47

- 1) The reason I put ISO+Exceptions higher than Hepburn is mainly because its an international standard and believe it or not but it is the Japanese Government standard, regardless of if its used in full practice or not. I don't like strict ISO translation because it is completely wrong for speech. つ is never ever pronounced "tu." The second reason is because I don't agree with つち being spelt as "tchi" since I feel "cchi" is more accurate. If it came down to it, either ISO+Exceptions or Hepburn w/ wa/o/e.
- 2) As above, wa/o/e. However, in my experience, wo and o are used interchangeably in speech, even by the same person. I'm more or less indifferent on wo and o but for the sake of choosing one or the other I choose "o." It might be worth throwing around the idea of using "wo" when the Japanese related to something from an older time when "wo" was considered normal and use "o" for modern situations. The only issue is that none of us would be able to tell that easily. We would need Japanese people to help us with that.
- 3) The only example I've seen is Tamagotchi. So as Lukeage, I prefer cchi.
- 4) I'd like to point out that the use of the "つ" in those situations is not considered proper or "real" Japanese. If it were my choice, I'd ignore it since it may or may not stay in Japanese culture. However, I'd be just fine putting in the exclamation.
- 5) use the rules I posted in the other thread. I think they follow the kana method. o->ou, e->ei, i->ii, u->uu, a->aa. These more closely represent speech over o->oo or e->ee.
- 6) First common method.
- 7) Not much to say. Looks fine to me.
- 8) My rule is that if the katakana directly represents the foreign word then use the foreign word. If its a shortened or modified foreign word (like "Japalish" words) then transcribe normally. Accepted exceptions like Famicom are acceptable.
- 9) If the number has a modifier (like in your first example) then I prefer to write it out in full. Otherwise as in the second example, leave it as a number. Main reason being that although it might make sense to an English person to leave the number, but in Japanese it is not read that way. Now you might say well 1st is not read as 1-"st" but 1st is an English rule for an English abbreviation. But writing Dai-2-ji we are applying an English rule to a Japanese abbreviation. To avoid that issue, I say write them out in full when they have modifiers.

10) Whichever works best. As I've said before, I'd prefer this to be as readable by Japanese people as possible. So whichever one is less confusing for a Japanese person is the one I'm for. Probably ignoring reimported names is where I'm headed.

11) As Lukeage said.

12) I'd say keep them for the sake of preservation but I don't really care either way.

NGEfreak
19th September 2007, 13:50

@ Lukeage:

3) Maybe ガッチャマン Gatchaman / Gacchaman

10) Most reimported names are simply wrong, like Frieza / Freezer.

@ Massin:

1) Hepburn is the international standard and most common used Romanization.

5) Written using kana spelling is the most common style. ô is not always ou; ex: 大神 Ookami

9) It's much easier for everyone to read Dai-2-ji than Dai-Sanji. How many people are able to find the 2nd Super Robot Taisen game if we name it with Dai-Sanji?

Lukeage
19th September 2007, 15:38

@ NGEfreak:

3) While googling about, it appears that ~tch~ is a Hepburn standard. If we go with Hepburn then we probably should follow this. The link I've posted below states this as well as some wiki pages.

10) As I said, I wasn't hard and fast on my decision. I'm more than happy for use to follow a standard rather than the reimport.

1) To back that statement up further: <http://www.indopedia.org/Romaji.html>
Hepburn is the most common romanization system in use today, especially in the English-speaking world. Japanese school children now learn Hepburn when they first begin to learn the English alphabet in junior high school.

5) Agreed, it should match the kana.

9) After giving this more thought, I have to say that I agree with using the numbers if that is what the boxart/title has. After all, that is what we use to document every other release.

gigadeath
21st September 2007, 08:48

Hepburn has been the standard code since ages, AND it's more accurate.

Yakushi~Kabuto
21st September 2007, 15:55

Poll results:

4 votes: Hepburn ISO (strict Kunrei)

2 votes: ISO w/ Exceptions (exceptions are "permitted" and therefore optional, I don't think they were meant to be forced)

0 votes: ISO (strict Kunrei)