

## Question for Malay Jawi letter in Unicode - three quarter hamza

Hello everyone, I'm new on this list.

I've got question for Jawi (Malay in Arabic script) in Unicode.

### **The codepoint for Jawi Letter Hamza Three Quarter High?**

I have been seeing "Jawi Letter Hamzah Three Quarter High" (sic) mentioned on many local documents, that it was not in Unicode but said to be proposed back then to be included after Unicode 5.0, but I could not find it on Unicode even on the current version 13.0. Is "Jawi Letter Hamzah Three Quarter High" even formally proposed and encoded yet? If so, what is the actual codepoint for it in Unicode? or did no one brought it to Unicode's attention in the first place?

*Note: The spelling "Hamzah" in all those documents are influenced from Malay (and also pronounced as such in Malay, with an H sound at the end), seemed like none of them realized that the final H does not exist in the English spelling "Hamza", and they just carried over the Malay spelling "Hamzah" in all of their English documents. For consistency, I'm using the proper English spelling "Hamza" instead of the actual spelling being used over here "Hamzah".*

While most of the documents are local, some of them do exist online, such as this document (linked after quote) in 2009 from Malaysia Network Information Center (MYNIC) to Internet Assigned Numbers Authority (IANA) for inclusion in their repository of Internationalized Domain Names (IDN) tables for .my Malay (macrolanguage) (Malaysia) entry, in which the document ends with (sic)

This character is not in the Unicode Table 5.0. The linguist came up with the decision to propose the inclusion of Jawi Letter Hamzah Three Quarter into the Unicode table.

Link: [https://www.iana.org/domains/idn-tables/tables/my\\_ms-my\\_1.0.pdf](https://www.iana.org/domains/idn-tables/tables/my_ms-my_1.0.pdf) (via <https://www.iana.org/domains/idn-tables>)

The Jawi letter Hamza Three Quarter High is part of everyday use words, for example in the Jawi spelling of the word "air" which mean "water" or "drink" (noun) which is **اير**. Another example would be in the Jawi spelling of the word "perduaan" which mean "binary" (term in computing, science and mathematics) which is **قردوان**.

These are the most common usage of Hamza Three Quarter High in Malay Jawi:

[A] consecutive vowels for au, ai, and ui, mostly in native Malay words. Example:

- "laut" [la·ut] (meaning: sea) is spelt **لاوت**
- "baik" [ba·i?] (meaning: good / nice) is spelt **بايق**
- "buih" [bu·ih] (meaning: bubble) is spelt **بويه**

[B] diphtongs for au and ai, mostly in Malay words loaned from English. Example:

- "audio" [au·dio] (meaning: audio) is spelt **اوديو**
- "aising" [ai·sin] (meaning: icing) is spelt **ايسينغ**

[C] suffix -an after vowel a, and suffix -i after vowels a and u, as part of Malay grammar rule which attaches affixes to modify word form. Example:

- "kenyataan" [kə·ɲa·ta·an] (meaning: statement) is spelt **کپاتان**, it got affixed from root word "nyata" [ɲa·ta] (meaning: to state)
- "cintai" [tʃin·ta·i] (meaning: love (conjugated verb)) is spelt **چينتاي**, it got affixed from root word "cinta" [tʃin·ta] (meaning: to love)
- "melalui" [mə·la·lu·i] (meaning: via / through) is spelt **ملالوي**, it got affixed from root word "lalu" [la·lu] (meaning: to pass through)

[D] spelling of Malaysian Chinese names in Malay. Example:

- "Ng" [əŋ] (multiple family names including 吳) is spelt **غ**
- "Ong" [oŋ] (multiple family names including 黃) is spelt **وغ**

Sample sentences:

Sample sentences with both **regular hamza** and **three-quarter high hamza** in used

سيلا باچ دعاء سبلوم مينوم اير.

Sila baca doa sebelum minum air.  
(Please recite the prayer before you drink the water.)

لبيه بايق جك وقتو دانتارا صلاة مغرب دان  
صلاة عشاء دايسي دغن باچان القران.

Lebih baik jika waktu di antara solat Maghrib dan solat Isyak diisi dengan bacaan al-Quran.  
(It is better to fill the gap between Maghrib prayer and Isha prayer by reciting al-Quran.)

ددنيا فنوه كظاليمن اينم ماسيه اد اورغ يثغ برقریکماءنسان.

Di dunia penuh kezaliman ini, masih ada orang yang berperikemanusiaan.  
(In this world full of cruelty, there still exist humane people.)

I'm seeing people using either superscripted version of Arabic Letter Hamza (U+0621) or abuses Arabic Letter High Hamza (U+0674) but those doesn't help in plain text where none of the abused formatting would work in the first place. Some people even gave up and just create image with the correctly positioned letters and uses that image in place of text. Some font, notably Amiri, actually displayed Arabic Letter High Hamza on three-quarter high, probably (but not necessarily the case) to accommodate Jawi users who abuses that letter for their Hamza Three Quarter High, but that would lock users to use only that font as other fonts still display the original height.

Link to Amiri font: <https://www.amirifont.org/>

The letter is the same size as regular Hamza, not any smaller like the size of High Hamza or the Hamza above/below Alef. It is positioned at three-quarter height of the writing line, unlike Arabic Letter High Hamza that displays on the highest position on the line nor regular Hamza that displays on the baseline.

The letter is also separate from regular Hamza, which is mostly used for Arabic or Sanskrit loanwords in Malay. Regular Hamza and Three Quarter High Hamza co-exist in Malay Jawi, could exist in the same sentence or even the same word, but should not be shown at the same height in all cases including plain text, otherwise it could cause confusion when reading since it could signal different sound, and it is grammatically wrong as well.

To recap the questions from first paragraph in case you are still unclear on the actual questions: Is "Jawi Letter Hamzah Three Quarter High" even formally proposed and encoded yet? If so, what is the actual codepoint for it in Unicode? or did no one brought it to Unicode's attention in the first place?

I'm looking forward for more information regarding this.

Best regards,

[Yaya]

Yaya MNH48