

## Zero Width Joiner and Zero Width Non Joiner in Kannada domain names

- Dr. U.B. Pavanaja ([pavanaja@vishvakannada.com](mailto:pavanaja@vishvakannada.com))

Zero Width Joiner (ZWJ) and Zero Width Non Joiner (ZWNJ) have special roles in Kannada.

ZWNJ is used in sequences like Consonant (C) + Halant (U+0CCD) + Consonant kind where the second C should not take the vattu form after (below) the first consonant.

Ex.

1. ಜ (U+0C9C) + ೆ (U+0CCD) + ಕ (U+0C95) = ಜ್ಯ - without using ZWNJ

2. ಜ (U+0C9C) + ೆ (U+0CCD) + ZWNJ (U+200C) + ಕ (U+0C95) = ಜ್ಕ - using ZWNJ

Example of the word not using ZWNJ – ರಾಜ್ಕುಮಾರ್ (Rajkumar)

Example of word using ZWNJ – ರಾಜ್ಕುಮಾರ್ (Rajkumar)

Both words mean the same. But their pronunciations are slightly different. The second form does not originally belong to Kannada, but nowadays due to the influence of English and Hindi, some people are using the form. But this form is needed to write many English words in Kannada like software (ಸಾಫ್ಟ್‌ವೇರ್, using ZWNJ). The word software will become ಸಾಫ್ಟ್‌ವೇರ್ if ZWNJ is not used.

ZWJ has a unique purpose in Kannada where conjuncts are formed with consonants where the first consonant is ರ (Ra) (U+0CB0). When conjunct is formed with Ra + Halant + <consonant>, the default form is second <consonant> followed by arkavattu, which is Kannada equivalent of the reph of Devanagari.

Ex. ರ (U+0CB0) + ೆ (U+0CCD) + ಕ (U+0C95) = ಕ್ಕ

Original Kannada did not have this arkavattu or reph. It was like any other C1+ Halant + C2 where the C2 takes the vattu form, even when C1 = ರ (Ra). To get the vattu form for ರ + Halant + <consonant>, ZWJ is used.

Ex.

ರ (U+0CB0) + ೆ (U+0CCD) + ZWJ (U+200D) + ತ (U+0C95) = ರ್ತೆ –Microsoft's implementation

ರ (U+0CB0) + ZWJ (U+200D) + ೆ (U+0CCD) + ತ (U+0C95) = ರ್ತೆ – as per Unicode's definition. Browsers Firefox and Chrome use this rendering rule.

The original Kannada form (not using arkavattu or reph) is needed when the first letter of the word itself is ರ (Ra, U+0CB0) followed by Halant and a consonant.

Example- ರ (U+0CB0) + Halant (U+0CCD) + ಯ (U+0CAF) = ಯಱ as in ಯಾರಲಿ (rally written in Kannada). This is wrong as per the writing style followed in Kannada. Here the second consonant ಯ (Ya, U+0CAF) must take the vattu form. To get that one must use ZWJ as explained above. Correct form of rally written in Kannada will be ರ್ಯಾಲಿ making use of ZWJ. One such word I can think of will be the company Rallis India. When they go for registering domain name in Kannada, if ZWJ is not allowed in domain name, they will have to go for ಯಾರಲೀಸ್.ಭಾರತ, which is the wrong form in Kannada. They should write it as ರ್ಯಾಲೀಸ್.ಭಾರತ by making use of ZWJ.

Thus both ZWJ and ZWNJ are needed for having proper domain names in Kannada.

### **How to avoid duplicate domain names involving ZWJ and ZWNJ?**

ZWJ and ZWNJ are used mainly to have two display forms of linguistically same word or combination of characters in Kannada. When ZWJ and ZWNJs are allowed in domain names for Kannada, it will create two domain names which have two display forms but linguistically they are same. To make the browsers and DNSs to treat them as equal, we have to ignore ZWJ and ZWNJs for comparing two words. This methodology is followed by the spell-check used in Microsoft Word. Same philosophy can be applied here also.

---\*\*\*---