Oriya BA, VA, and WA

Michael Everson & Peter Constable

Original Sanskrit initial ν-, pronounced [b] ~ [w], was lost in Oriya as the sound merged with b-: visarga became bisarga, for instance. The sound -ν [w] was preserved after consonants, however, and post-consonantal -bu and -va were both retained, written with different shapes: ग dha [dba] differs from ग dva [dwa].

Because the initial ν- was lost, a consonant letter distinct from ba was not historically preserved. A need to write initial ν-still arose in certain contexts – particularly in Oriya transliteration of Sanskrit words. No conventional letterform was established for this: some authors who write ν- have used the same letterform as ब ba, but often the ba letterform is modified with a dot or circle; compare ग va and ग va. Other glyphs, such as ख and फ, which, in a more modern font style might be ख, have also been observed.

The Unicode Standard preserves the va/ba distinction with two different letters, \U+0B2C ब ORIYA LETTER BA and U+0B35 ग ORIYA LETTER VA, whose usual graphic representations in post-consonantal position are न and न respectively. An exception to this shaping rule occurs with U+0B2E व ORIYA LETTER MA: here ba following ma looks like the glyph for va, although it is pronounced [b] and not [w]. Traditional Oriya text does not give a representation for the syllable mva (though this syllable occurs in Sanskrit).

\begin{verbatim}
\texttt{U+0B15 ka + U+0B4D virama + U+0B35 va → g [kwa]}
\texttt{U+0B17 ga + U+0B4D virama + U+0B35 va → g [gwa]}
\texttt{U+0B1B cha + U+0B4D virama + U+0B35 va → t[ba]}
\texttt{U+0B1C ja + U+0B4D virama + U+0B35 va → d3wa}
\texttt{U+0B1F ta + U+0B4D virama + U+0B35 va → t[wa]}
\texttt{U+0B21 da + U+0B4D virama + U+0B35 va → d[wa]}
\texttt{U+0B23 ta + U+0B4D virama + U+0B35 va → t[wa]}
\texttt{U+0B25 tha + U+0B4D virama + U+0B35 va → t[wa]}
\texttt{U+0B26 da + U+0B4D virama + U+0B35 va → d[wa]}
\texttt{U+0B27 dha + U+0B4D virama + U+0B35 va → d[wa]}
\texttt{U+0B28 na + U+0B4D virama + U+0B35 va → n[ba]}
\texttt{U+0B32 la + U+0B4D virama + U+0B35 va → l[wa]}
\texttt{U+0B36 sa + U+0B4D virama + U+0B35 va → s[wa]}
\texttt{U+0B38 sa + U+0B4D virama + U+0B35 va → s[wa]}
\texttt{U+0B39 ha + U+0B4D virama + U+0B35 va → h[wa]}
\end{verbatim}

But compare:

\begin{verbatim}
\texttt{U+0B26 da + U+0B4D virama + U+0B2C ba → dba}
\texttt{U+0B2E ma + U+0B4D virama + U+0B2C ba → mba}
\texttt{U+0B2E ma + U+0B4D virama + U+0B35 va → mva or mva or mva [mwa]}
\end{verbatim}

The extended Oriya letter U+0B71 व ORIYA LETTER WA is sometimes used in Perso-Arabic or English loan words for [w] at the beginning of words. It appears to have originally been devised as a ligature of ङ o and ग va, but since ligatures of independent vowels and consonants are not normal in Oriya, this letter has been encoded as a single character which does not have a decomposition. It should only be used initially in words or orthographic syllables to represent the foreign consonant; as a native semivowel, virama + va should be used, since that is historically accurate. Glyph variants of wa are ङ, ङ, and ङ – note the dot in the -va of the last glyph.

1 Everson Typography, www.evertype.com
2 Microsoft, www.microsoft.com