Automatic alignment of hieroglyphic and transliteration (screen shots of the demo)

Mark-Jan Nederhof

2008-07-09

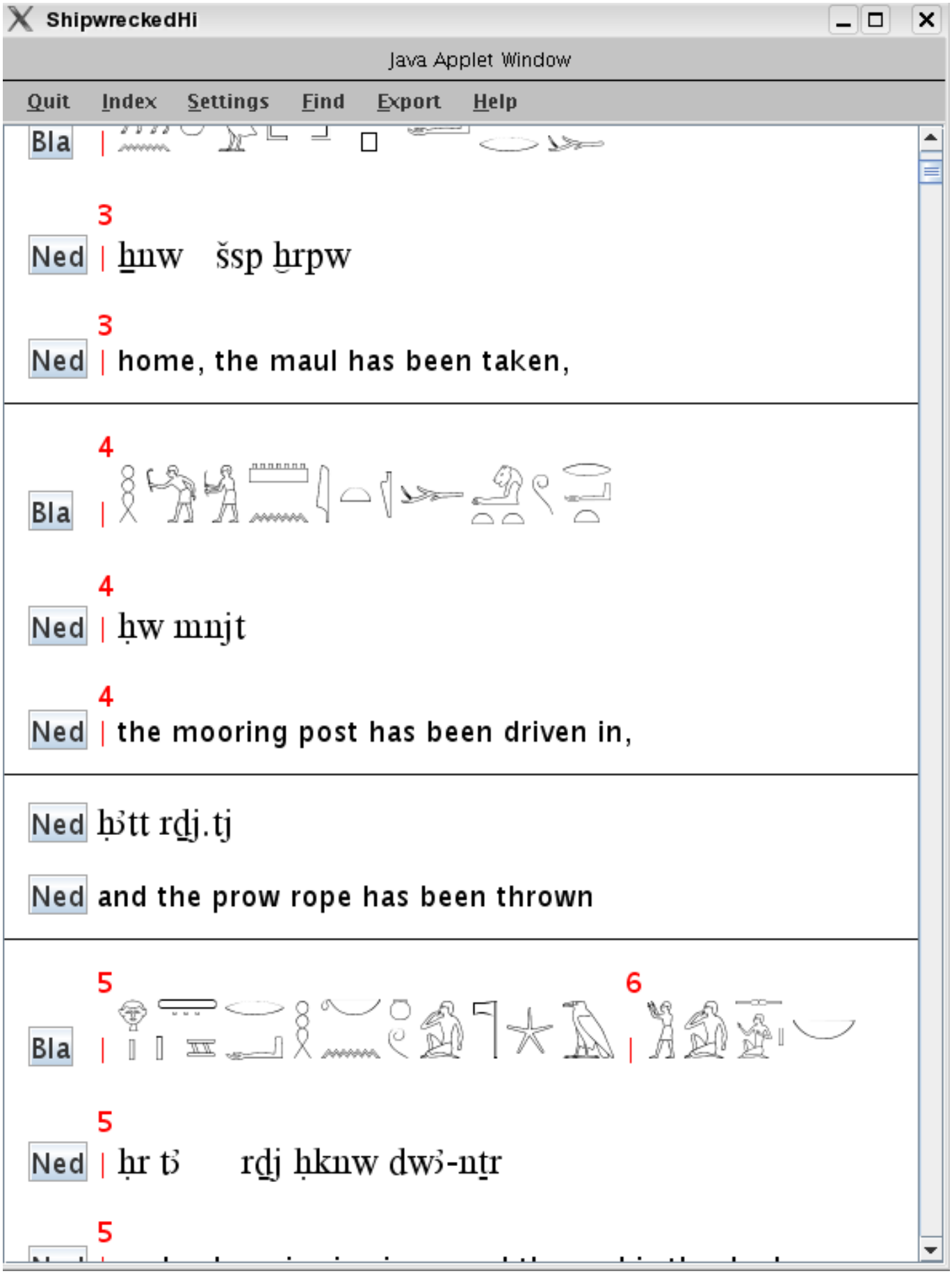

Figure 1: Hieroglyphic and transliteration/translation are taken from two files created by two different people. If alignment can only rely on line numbers of the manuscript (in red), the interlinear representation is often imperfect. We see this for line 4, which has to be broken up because the translation is too wide relative to the width of the window.

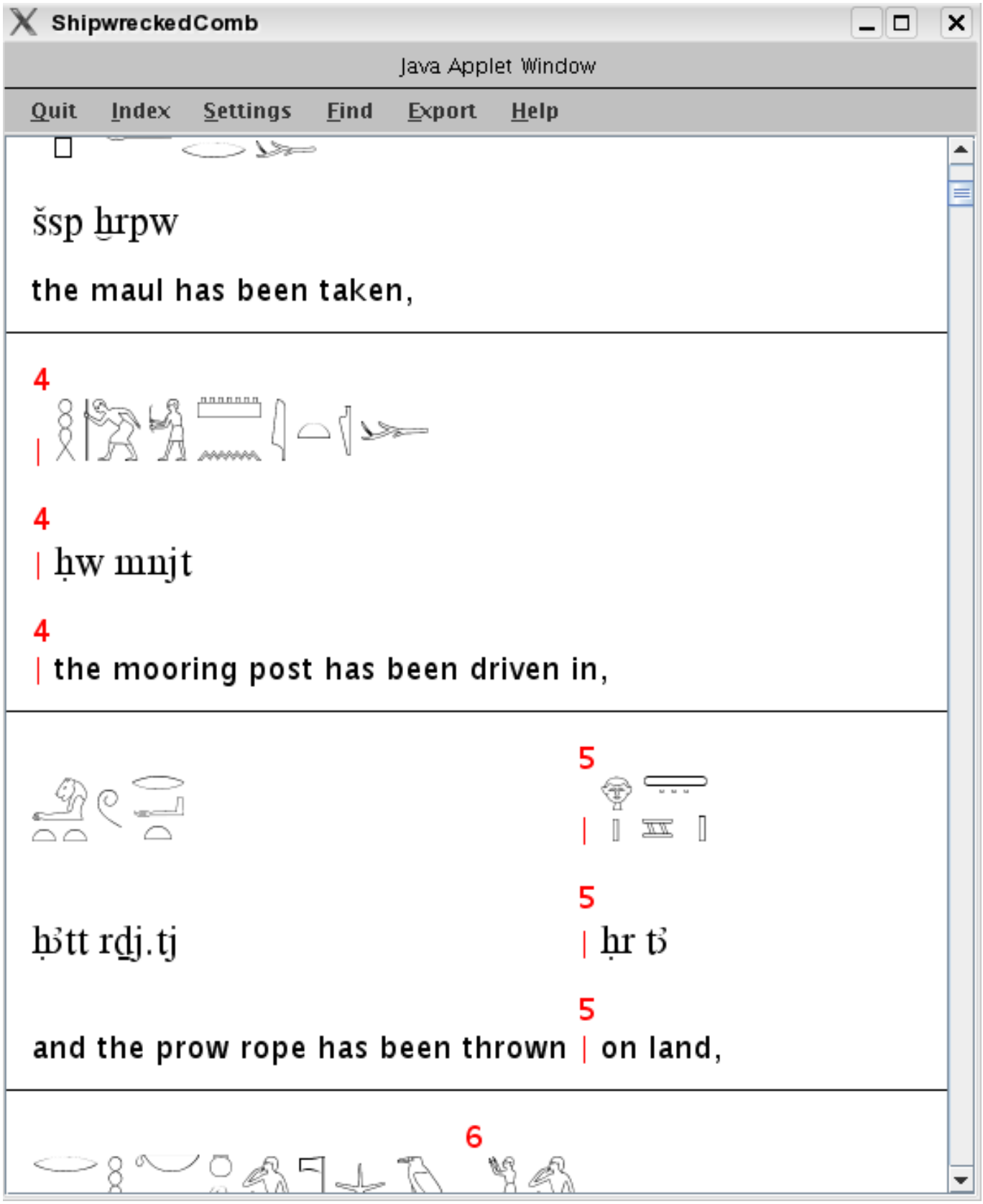

Figure 2: Ideally, we would like the hieroglyphic to be broken up to match the transliteration, as shown. Can we do this automatically, by finding the appropriate sequence of hieroglyphs for each word in the transliteration?

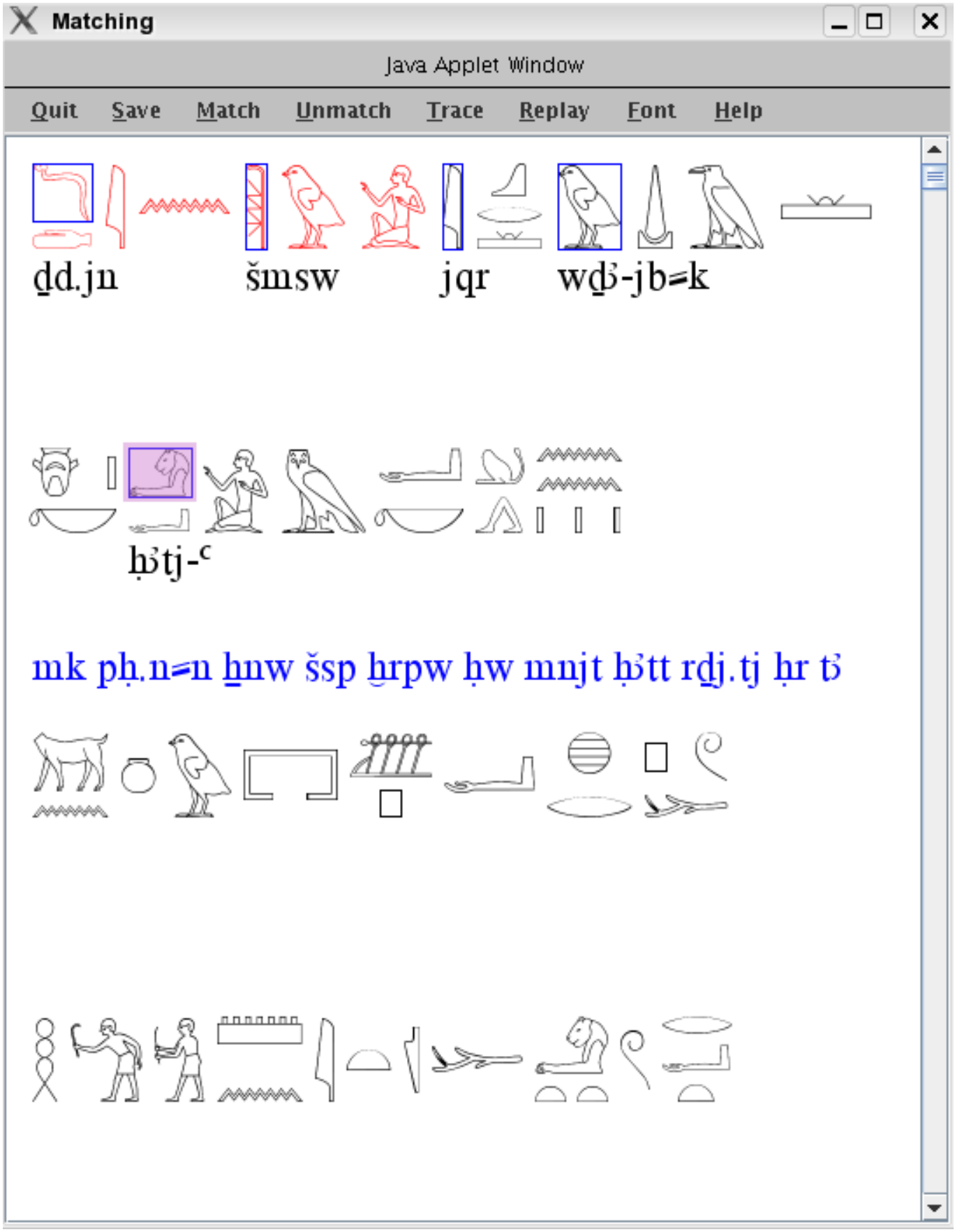

Figure 3: To determine the accuracy of automatic alignment, we first do manual alignment by a mouse click on the first hieroglyph for each word (blue boxes). To help the annotator, the first line of the remainder of the transliteration is shown in blue.

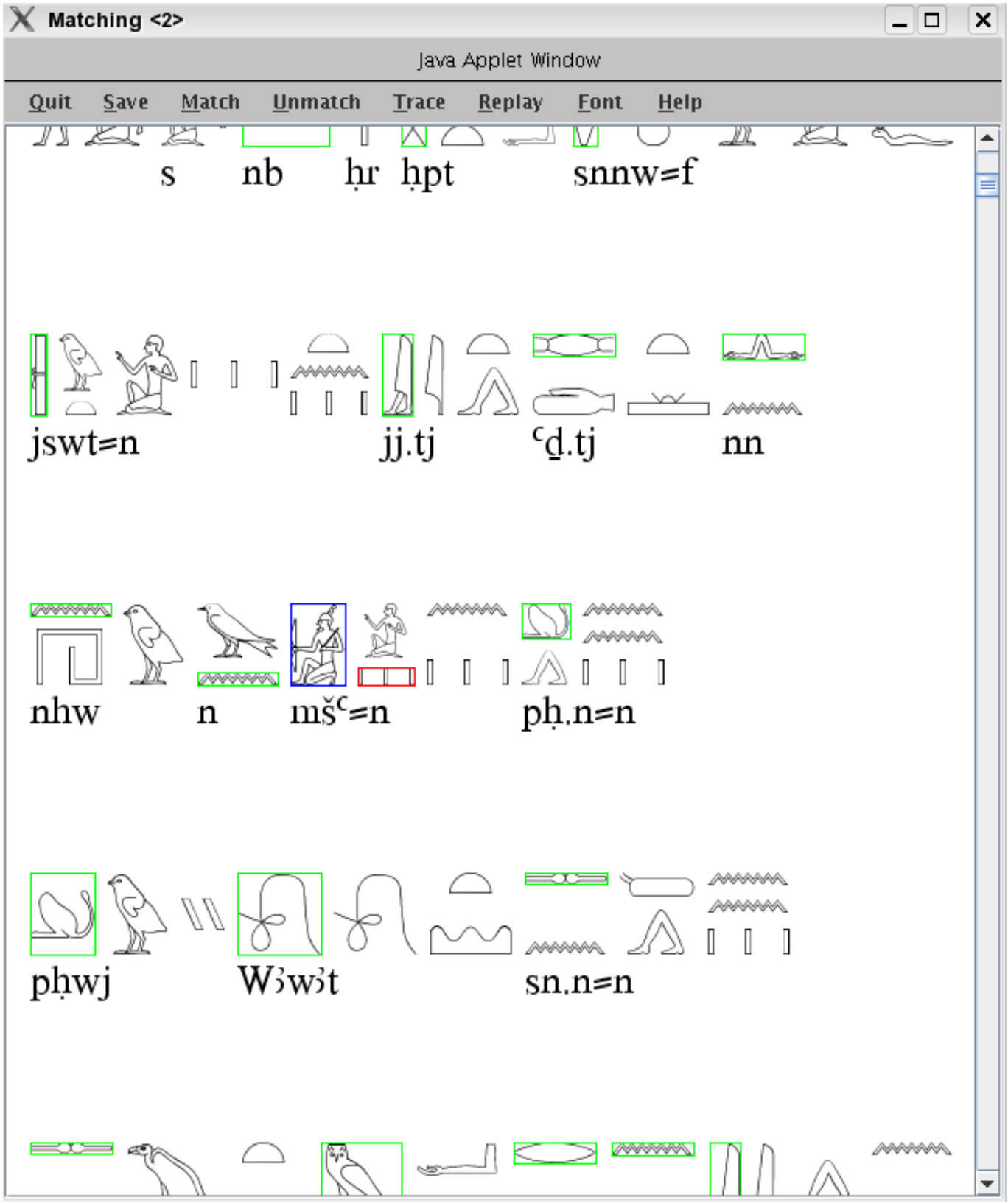

Figure 4: Next, we run the automatic alignment. A green box means that the manual and the automatic annotation concur. Red boxes indicate that automatic annotation has erroneously flagged the beginning of a word. The error shown here is due to a gap in the sign list, which lacks the reading of A12 ('soldier with bow') as ideogram. Because only the reading as determinative is present in the sign list, the incorrect interpretation of A12 and the following sign A1 as determinatives of **n** is easily explained in the light of the crudeness of the used model of hieroglyphic writing.

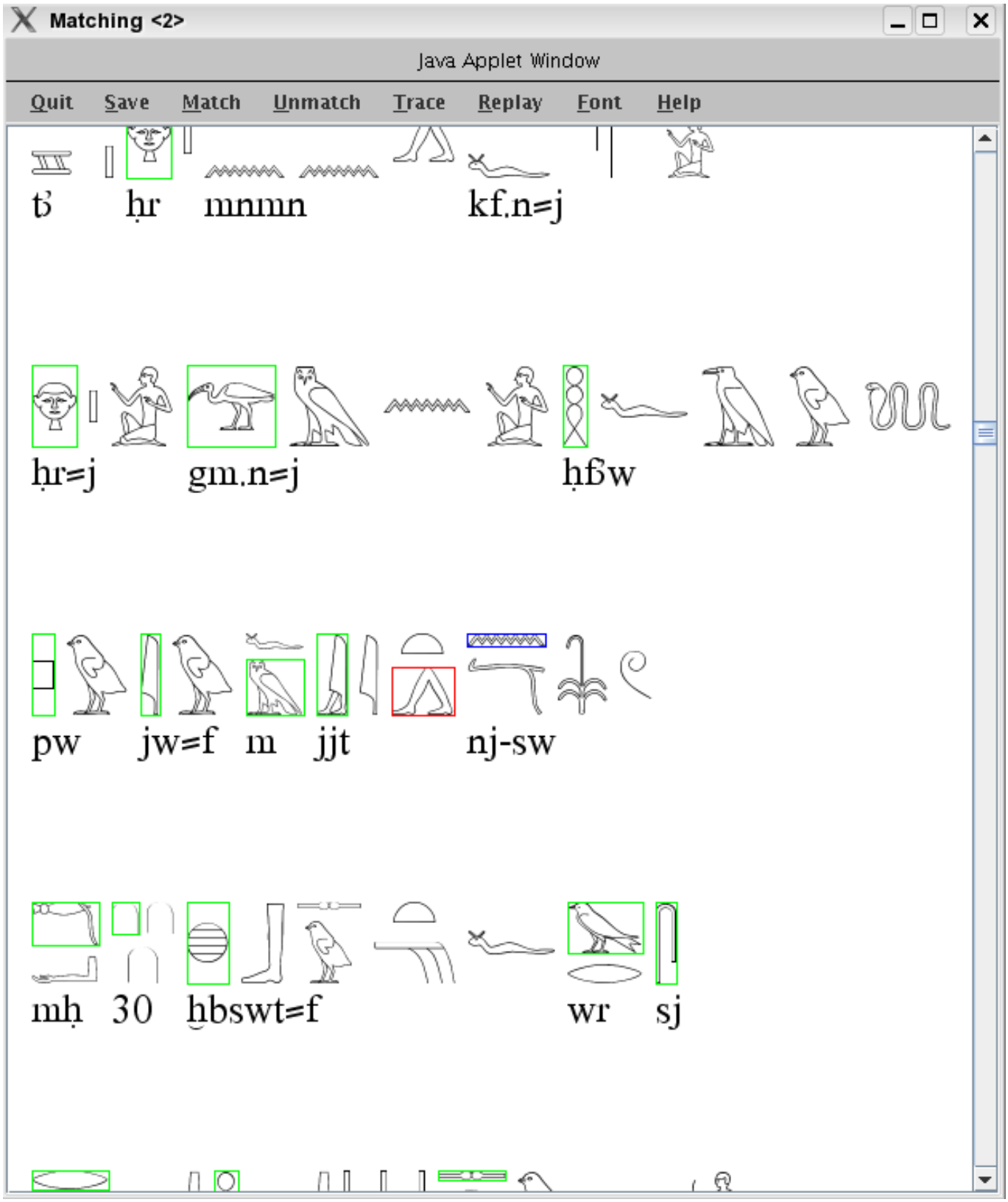

Figure 5: Here again the crudeness of the orthographic model is to blame, being unable to match F20 ('tongue') with reading **ns** against a substring of the transliteration, because the **n** and **s** are separated by other characters.

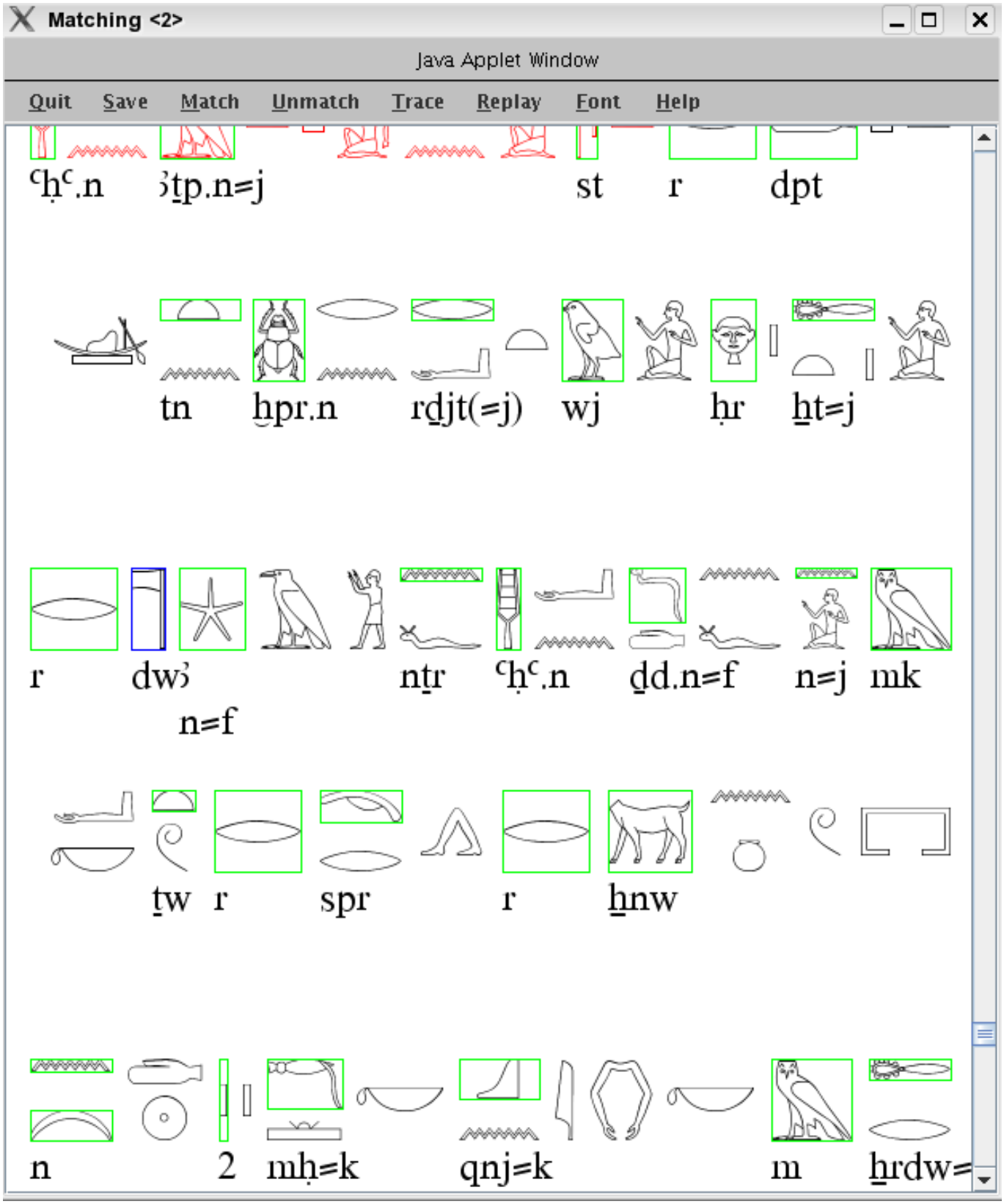

Figure 6: Honorific transposition is difficult to handle in a general way, and the current model ignores the matter altogether.