

## Issue Statement

John V. Thomson  
SIL International  
john\_thomson@sil.org

Interlinear texts are one of the most common ways for linguists to present their analysis of language data. But, while it is relatively easy to develop a specialized tool for creating interlinear texts, it is more difficult to allow linguists to incorporate the results into more complex documents, for example, when an interlinear display of a sentence is wanted as an example in a paper (whether in a word processor or web page), or when a longer text is to be published with various kinds of front and back matter.

Older interlinear tools use a line-based format for interlinear text files, in which items in related lines are aligned by inserting spaces. While this approach works with typewriter-style fixed width fonts, with proportionally spaced fonts, such data loses the correct appearance when displayed in a tool which does not 'understand' its interlinear nature. Even if it is then hand-adjusted to look right, that adjustment will not survive any change in font or paragraph width.

At least one early interlinear tool (*IT* for Macintosh, 1988) could export data as a 'picture', a captured set of drawing commands which would reproduce the correct appearance within another program. This is somewhat better, but the picture's line wrapping will not adjust to changes in the containing document.

[Bow, Hughes, and Bird, 2003] have proposed an XML representation of interlinear texts and a methodology for presenting this using XSL formatting objects and one flavor of HTML (inline blocks supported by Internet Explorer 5.5 and later). I recently discovered that FireFox has an equivalent (but different) keyword which can be combined with the inline-block element to produce HTML which displays nicely in both browsers. Such displays survive resizing the window (and thus the paragraph width) and will adjust cleanly to changes in the formatting of interlinear elements.

However, getting data into this form from existing tools is a non-trivial problem. (Their approach has three major steps, *after* the data is in their special XML format.) And HTML is not a suitable medium for print publication. Word processors (such as Microsoft Word) which import HTML do not render inline blocks correctly. Not all authors are comfortable with available tools which could produce a publishable document using XSLFO. It is therefore worth investigating whether we can get data into a form which will display well and adjust easily to formatting changes in such popular word processors as Microsoft Word and Open Office Writer.

Preliminary investigations indicate that both tools can arrange a series of blocks containing tables in a paragraph format (text boxes in Word, frames in Writer), and that this arrangement will adjust cleanly to changes in the paragraph width. My preliminary investigation indicates that it may be necessary to develop a macro within the tool to get the individual block sizes to adjust properly when elements within the interlinear are reformatted and edited. Development of such macros, along with the mechanisms for getting data from interlinear tools into the various word processor formats, would be very helpful to linguists seeking to publish this kind of data.