Testing Tika Format Parsers

This is a test document to ensure that we (naming no names) hasn't broken the Tika format parser again. Strangely it seems that if Tika is broken then the GATE document is created without error but with no content which is a *bit* of a pain.