Document-token and paragraph matrices

