

Инструкции относно текстовете за корпуса:

1. Текстовете да са word-формат
2. Приемат се само текстове с шрифт **CyrillicaBulgarian10U**, т.е. или текстове набрани с този шрифт, или такива които са конвертирани в него чрез конвертора на г-н Тодор Тодоров
3. На този етап в корпуса няма да фигурират диалектни текстове
4. Приписките ще се явяват като отделен текст
5. Текстовете могат да бъдат придружени от разночетения
6. В корпуса ще се включат както нормализирани, така и ненормализирани текстове
7. Всеки текст следва да е придружен с метаданни. Определянето на метаданните е описано в Протокол от работна среща от 15.06.2011 г.