

В соответствии с двояким характером исследования в целом настоящий том включает две части. В части 4 дистрибутивно-статистический анализ (ДСА) проводится на материале 15 миллионов слов, непосредственно примыкающих друг к другу в текстах русской прозы 1850-1870-х гг. Тот же формальный метод прилагается к английскому корпусу (4 миллиона слов романов Диккенса). И в том, и в другом случае ДСА приводит к открытию дистрибутивных классов и некоторых грамматических конструкций. Часть 5 представляет собой частотный словарь русского корпуса, показывающий распределение слов по основным жанрам, по микрожанрам и по текстам 25 писателей.

В существенно расширенном виде том 2 представлен на компакт-диске.