Доклад С.А.Стринюк «Unbalanced Multigenre Corpora» посвящен вопросам обоснования проведения сравнительных исследований на материале несбалансированных разножанровых корпусов.

В качестве основы нашего исследования лежит предположение о возможности сравнения несбалансированных разножанровых корпусов.

Для сбора текстов статей, используемых в качестве образца научной речи и источника интенсивного расширенного чтения для студентов, использованы статьи из ведущих научных журналов из соответственных областей исследования.

Естественно предположить, что тексты Research Project и научные статьи отличаются по объему, но проблема разного размера корпусов снимается нормализацией частотности.

Существует множество исследований, опирающихся на данные BNC в сравнении с небольшими корпусами, скомпилированными учеными в научных целях.

В наших исследовательских корпусах представлены научные работы разных жанров (статьи в корпусе статей и тексты Research Projects, написанные студентами). Мы уверены в жизнеспособности подобного подхода, поскольку научная речь статьи может выполнять роль эталона научной речи как один из наиболее развернутых полнотекстовых жанров.