

нию частот, то свойства полученного упорядочения практически универсальны для любого языка и текста. Было замечено, что богатство словаря писателя с точки зрения объема его лексикона в основном определяется числом однократно употребленных им слов. Получение количественных характеристик текстовых единиц было несомненно полезным для лингвистики, однако постепенно такого рода подсчеты стали основным содержанием лингвистических работ, где применялись СМ. В этом можно убедиться, просматривая библиографию отечественных работ по лингвистической статистике, составленную Г. В. Ермоленко⁷. Если мы не будем полагаться на заглавия работ, где присутствуют слова «модель», «энтропия», «закон», а обратимся к самим работам, то увидим, как неуклонно растет в общем числе работ доля работ «инвентаризационного» характера. Этот процесс, как можно думать, объясняется, во-первых, кажущейся очевидной ценностью таких числовых данных для решения несформулированных прикладных задач⁸ и, во-вторых, все большей доступностью ЭВМ; последнее обстоятельство иногда действует прямо-таки завораживающе, особенно на начинающих лингвистов. Именно рост публикаций «инвентаризационного» характера, свидетельствующих о превращении задачи измерения количественных показателей в самоцель, и породил распространенное недоверие к осмысленности применения СМ в лингвистике. Измерение было — пусть неявно, но зато прочно — отождествлено с применением СМ, как текстовых.

Действительно, подсчет частоты повторяемости некоторого события — это измерение. Измерять можно разные свойства объектов и пользоваться при этом разными способами производства измерений. Кстати, можно выразить результат измерения числом, но это не обязательно. Если мы установили, что А больше Б, то в основе этой операции также лежит измерение, хотя мы установили между А и Б всего лишь отношение порядка: распространены ситуации, когда мы не знаем и не имеем возможности узнать, на сколько А больше Б, а следовательно, мы не получаем «числа», хотя и производим измерение. Существует специальная наука — теория измерений, предмет которой частично пересекается с математической статистикой, поскольку статистика лежит в основе методов обработки результатов измерений, теории ошибок измерений и т. д. Но неправомерно отождествлять измерение каких-либо характеристик с применением СМ как таковых. В любом научном исследовании измерения, независимо от того, что именно и каким способом измеряется, выступают как средство познания, но не как его цель. Разумеется, задачей какой-либо отдельной работы может быть и измерение как таковое. Частотный словарь составляется с целью дать сводку данных об употребительности слов, т. е. его цель — измерение частоты слов. Общеизвестно, однако, что результаты этих измерений содержат информацию, используемую, например, в работах по изучению закономерностей статистической структуры словаря и текста (теоретическое использование), а эти закономерности в сочетании с данными методики и психологии позволяют использовать частотные словари, в частности, для составления учебных словарей-минимумов, т. е. оптимизировать методы обучения языку (практическое использование). Если же в результате измерений вообще не добывается новое знание, то это просто означает, что исследователь поставил перед собой малоинтересную задачу. Но постановка задачи, как таковая, вообще лежит вне компетенции методов ее решения, будь то статистические или любые другие методы.

⁷ Ермоленко Г. В. Лингвистическая статистика. Краткий очерк и библиографический указатель. Алма-Ата, 1970.

⁸ Необходимо отметить, что ситуация, когда подсчеты, проведенные без ясной цели и вне соответствующих теоретических обоснований, оказались бы непосредственно пригодными для решения прикладных задач, является абсолютным исключением.