

Возникает вопрос: а возможно ли вообще выделение семантических отношений для литературного языка, «которые были бы регулярны, стандартны, число которых было бы конечным и не очень большим?» [3, с. 16]. Мы склонны считать, что семантические отношения литературного языка, «вероятно, полностью исчислимы, если заранее условиться о пределе их членения вглубь. В самом деле, как только мы ограничимся отношениями самого общего характера и глобального масштаба, они представятся нам вполне обозримыми» [4, с. 305—306].

Одной из первых проблем, связанных с описанием семантических отношений в лексике, является проблема определения источника их выявления. Естественно предположить, что каждое слово языка связано с каждым другим словом этого языка и способом выявления отношений между ними является попарное сопоставление слов данного языка. Однако этот путь представляется нерациональным. А. П. Клименко пишет: «Описать отношение каждого слова со всеми словами в языке просто невозможно. Если принять словарь, скажем, за состоящий из 100 000 единиц, что не является преувеличением, то надо будет проводить от каждой из этих ста тысяч словарных единиц нити связи ко всем остальным 99 999. Если это теоретически и возможно, то практически нереально» [6, с. 10]. С другой стороны, не все отношения между словами в языке актуальны. Любое отношение может быть установлено при условии наличия общего элемента у его членов по форме или по содержанию. Для семантического отношения этим общим элементом является общность лексических значений его членов, которая может наблюдаться не во всякой произвольно выбранной паре слов.

Другой возможный путь выявления семантических отношений — анализ текстов. Но и здесь вероятность выявления всех типов отношений чрезвычайно мала, так как трудно представить себе текст, в котором встречались бы попарно все лексические единицы языка. «Даже самое частое существительное русского языка по словарю Э. А. Штейнфельдт — слово *год* — имеет частоту 810 из 400 000 словоупотреблений, т. е. вероятность его равна 0,002. Слово же, являющееся тысячным в частотном списке, встретилось 50 раз на 400 000, и его вероятность равна 0,000125. Это означает, что надо прочесть 8000 слов, чтобы встретить это довольно частое слово. Вероятность же его „встречи“ с другим словом во много раз меньше» [6, с. 11]. Таким образом, тексты также не могут служить рациональным источником выявления семантических отношений, они могут быть использованы лишь для подтверждения правильности их выделения из другого источника.

Реальным источником выявления семантических отношений в лексике можно считать, на наш взгляд, словари идеографического типа (собственно идеографические, семантические, ассоциативные, научно-технические тезаурусы), которые в имплицитном или эксплицитном виде содержат эти отношения на массиве слов данного языка или единиц терминологической системы той или иной отрасли знания. Для нашего проблемно-постановочного опыта выявления семантических отношений на русском материале в качестве источника может быть использован «Русский семантический словарь. Опыт автоматического построения тезауруса: от понятия к слову» [7] (в дальнейшем — РСС), специфика которого состоит в том, «что слова в словарной статье группируются вокруг некоторой идеи, или дескриптора, с которым они связаны и смысл которых в какой-то степени раскрывают» [7, с. 3]. Словарные статьи РСС построены таким образом, что в них включаются лексические единицы, которые с определенной долей вероятности связаны с дескриптором тем или иным отношением, что отвечает сущности словаря идеографического типа, который представляет собой «множество смысловыражающих элементов (слов, словосочетаний и т. п.) некоторого языка с заданными смысловыми отношениями» [8, с. 148]. Но в РСС эти отношения присутствуют имплицитно, они не выражены, не