

УДК 681.32

**ОЧЕРК ОСНОВНЫХ ПРИНЦИПОВ КВАНТОВОЙ ЛИНГВИСТИКИ****В. А. Широков**

Украинский языково-информационный фонд Национальной академии наук Украины

Статья посвящена определению возможных границ формального подхода к описанию языка. Выявлено, что концептуальное описание квантовомеханических и мыслеречевых систем базируется на общих подходах. Введено понятие семантического состояния языковой единицы, приведено их описание средством теории нечетких множеств. Приведены также формальные аспекты отношения синонимии.

**СУБСТАНЦИЯ ЯЗЫКОВАЯ, ПРИНЦИПЫ КВАНТОВЫЕ, СОСТОЯНИЕ СЕМАНТИЧЕСКОЕ, МНОЖЕСТВО НЕЧЕТКОЕ, ОТНОШЕНИЕ СИНОНИМИИ.**

«Кто говорит, тот кроме имен, взятых с предметов видимых, ничем иным не может слушающим изобразить невидимого».

*Св. Ефрем Сирий*

«Вера же есть осуществление ожидаемого и уверенность в невидимом».

*Св. Ап. Павел, К Евреям, 11, 1.***Введение**

В языкознании время от времени возникают различные течения и движения, направленные на создание формальных схем, призванных, как полагают их адепты, продвинуть внедрение тех или иных математических методов в дело описания языковой субстанции. Такими, в частности, представляются попытки изобразить языковые феномены с помощью случайных процессов, используя для этого аппарат теории вероятности и математической статистики, попытки описать явления языка (в первую очередь, синтаксические) с помощью формальных грамматик, «натянуть» языковую феноменологию на некие логистические исчисления типа логики предикатов и т. п. Одной из последних попыток в этом ряду явилось изобретение нами теории лексикографических систем и лексикографических исчислений [1].

Следует, однако, отметить, что после первых успехов применения названных (а также других) формальных схем у их энтузиастов, как правило, возникает некое разочарование, связанное с тем, что разработанный или освоенный ими формальный аппарат не способен учесть множество существенных деталей даже тех ограниченных областей языковой системы, для которых он с самого начала, как им казалось, был специализирован.

Автор, переживший периоды и энтузиазма, и разочарования в такого рода деятельности, наконец, успокоился, составив определенное отношение к попыткам формализации вещей, имеющих черты (и соответственно, поведение) субстанции, включая и языковую субстанцию. Сразу же оговоримся, что мы не владем формальным и замкну-

тым определением понятия субстанции, полагаясь на интуитивное его понимание — в конце концов, так же поступают и представители других наук (физики, химии, биологии...) по отношению к субстанциям, изучаемым в данных науках.

Несмотря на не очень впечатляющие достижения в области внедрения формальных методов в лингвистику, автор придерживается мнения, что попытки изобретения все новых способов и средств формализации языковой системы оставлять не следует — ввиду хотя бы того, что они помогают яснее очертить возможные границы формального подхода к описанию языка, а в процессе их реализации время от времени возникают полезные идеи, иногда заимствованные из других наук, которые способствуют расширению методологии научного исследования в языкознании.

Примерно в таком логическом ключе следует воспринимать и попытку, предпринятую в данной работе.

Одной из идей, «заимствованных» из других наук, мы считаем попытку применения идеологии квантовой механики к описанию языковой субстанции. Отмеченная экстраполяция находит свое оправдание в неких общих чертах, присущих микромиру, с одной стороны, и миру языка, с другой, которые мы видим в следующем.

**1. Квантовые принципы**

Как известно, квантовая механика была изобретена с целью описания микрочастиц материи в той области взаимодействия, где уже нельзя пренебречь влиянием прибора (или наблюдателя) на наблюдаемый объект. В языковой же области влияние наблюдателя на наблюдаемый объект (а именно: влияние субъекта мыслеречевого процесса на семантические состояния языковых единиц) является, на наш взгляд, основным признаком и основным принципом самого существования языковой системы, поскольку оно представляет квинтэссенцию процесса понимания речевого акта. В этом аспекте квантовомеханические и мыслеречевые системы

действительно конгениальны. Логично предположить, что и концептуальное их описание должно содержать некие общие черты и базироваться на общих подходах. Вообще взаимодействие субъекта и объекта, одним из проявлений которого есть влияние наблюдателя на наблюдаемое, представляется нам настолько фундаментальным фактом, что оно, несомненно, заслуживает того, чтобы стать одним из оснований практически любой рациональной схемы описания в науке.

Вследствие отмеченной особенности квантовомеханических систем возникла необходимость формулировки иной, нежели классическая, парадигмы описания квантовых объектов и процессов. Пришлось отказаться от понятия непосредственно и детерминистски воспринимаемой (или наблюдаемой) траектории частицы, представляющей исчерпывающее описание системы, заменив ее неким «суррогатом» — так называемым *состоянием квантовой системы*. Понятие состояния оказалось, однако, чрезвычайно плодотворным как в эпистемологическом, так и в техническом отношении, поскольку его применение позволило не только создать эффективный математический аппарат, способный достаточно точно описывать явления микромира, но и сформулировать новую философию естествознания, что в конце концов привело и к созданию новой — квантовой научной картины мира.

В плане технического удалось построить формальную теорию квантовомеханических состояний, согласно которой они образуют довольно простой математический объект — линейное бесконечномерное метрическое (гильбертово) пространство — пространство решений так называемого уравнения Шредингера — основного уравнения квантовой механики. Значения наблюдаемых величин при этом представляются через матричные элементы неких (эрмитовых) операторов по данным состояниям, которые служат формально-математическими репрезентантами наблюдаемых. Для самих состояний (в квантовой механике их еще называют волновыми функциями) была со временем предложена прозрачная вероятностная интерпретация. Таким образом, теория состояний квантовомеханических систем предложила — уже на новом, неклассическом уровне — формальную основу для описания соответствующего класса явлений.

В философском плане квантовая теория представила целую гамму новых для науки идей. Во-первых, она естественным образом примирила две фундаментальные физические теории: классическую механику и электродинамику, дав корректное объяснение устойчивости атомных систем, носившей до этого характер парадокса, и определив, кроме того, пределы применимости отмеченных теорий. Точно так же естественно она объяснила дискретно-непрерывный характер материи и корпускуляр-

но-волновой дуализм, дала своеобразное разрешение проблемы тождества и, наконец, предоставила в лице принципа дополнительности Бора новую гносеологическую парадигму.

Одним из фундаментальных следствий квантовой механики стало осознание (и наблюдение на эксперименте) так называемых смешанных состояний, а именно таких, в которых квантовая система обладает одновременно не одним определенным значением наблюдаемой величины, а, например, сразу двумя. Это, в частности, означает, что квантовая система в один и тот же момент может пребывать в двух различных состояниях; допуская некоторую вольность речи, можно утверждать, что с вероятностью  $p$  она находится в состоянии 1 и с вероятностью  $(1 - p)$  в состоянии 2. Такое комбинированное состояние представляется суперпозицией (линейной комбинацией) состояний 1 и 2 — их своеобразной (когерентной) «смесью». Такое положение можно интерпретировать в том смысле, что объект одновременно, в один и тот же момент времени представляет собой и *«нечто»*, и *«нечто иное»*. Принцип суперпозиции состояний сыграл роль основного системообразующего принципа квантовой теории. Один из основоположников квантовой теории П. А. М. Дирак отмечал [2]: «Если состояние образовано путем суперпозиции двух других состояний, то оно будет иметь свойства, которые в некотором, несколько неопределенном смысле слова являются промежуточными между свойствами обоих исходных состояний и которые в большей или меньшей степени приближаются к свойствам одного из них, в зависимости от того, с большим или меньшим «весом» это состояние вошло в суперпозицию».

Следует отметить, что упомянутые феномены в эпоху создания квантовой теории выглядели весьма непривычно, они, казалось, противоречили и обыденному опыту с его наглядными представлениями, и так называемому здравому смыслу, и классической аристотелевой логике. В этой связи приведем еще одно рассуждение П. А. М. Дирака [2]: «...главная задача физической науки заключается не в том, чтобы снабжать нас наглядными картинками, а в том, чтобы формулировать законы, управляющие явлениями, и использовать эти законы для открытия новых явлений. Если наглядная картина существует, то тем лучше; однако существует она или нет — это лишь второстепенный вопрос. В случае атомных явлений нельзя ожидать, что существует наглядная картина в обычном смысле этого слова, в котором под «наглядной» понимается модель, действующая в основном по классическим принципам. Можно, однако, расширить смысл слова «картина» так, чтобы включить в него *любой способ рассмотрения основных законов, при котором их согласованность становится очевидной*».

Заметим, что данное утверждение имеет значение для науки вообще, а не только для физики. И хотя объекты языка могут кому-то представляться весьма наглядными (написанные на бумаге, они могут быть нами осмотрены), но наглядность эта весьма обманчива, поскольку наблюдаемая («наглядная») часть языкового состояния далеко не исчерпывает его сущности. Эффекты такого рода весьма обычны для квантовых систем.

Таким образом, цель данной работы — продемонстрировать ту глубокую аналогию, которая, как нам представляется, существует между квантовой механикой и лингвистикой и простирается так далеко, что позволяет использовать не только понятийный аппарат первой, но и даже ряд конкретных идей квантовомеханических математических конструкций для формального описания языковых феноменов. По ходу дела мы предполагаем сформулировать в терминах состояний интегральный подход к описанию языковой системы, в определенном смысле объединяющий и обобщающий традиционные грамматические и лексикографические подходы.

Сознавая, что аналогии имеют один общий недостаток, а именно — все они неверны, мы тем не менее попытаемся изложить наши соображения по поводу данного соответствия с надеждой, что более проникательные исследователи смогут, вероятно, сделать и более глубокие выводы.

## 2. Языковая субстанция и ее описание

Остановимся на некоторых общеметодологических замечаниях относительно принципов моделирования языковой системы.

Первым вопросом, который возникает при формулировке этих принципов, является вопрос об объектах моделирования, а именно: что представляют собой объекты языка и что, собственно, мы предполагаем моделировать. В качестве исходной посылки мы выбираем положение о том, что собственными объектами языка являются определенные психофизические состояния и процессы, происходящие в мыслеречевом аппарате человека, а устная и письменная ее формы служат элементами инфраструктуры языкового процесса. Описанию и моделированию подлежат как собственно упомянутые психофизические состояния и процессы, так и их инфраструктура. Исходя из данного положения, выясним их роль в процессах моделирования языка.

Очевидно, что мыслеречевой процесс сам по себе является интегрированным, то есть содержащим как языковую, так и ментальную компоненты. В мыслеречевом аппарате он находит выражение в виде динамической системы взаимосвязанных рефлексов, содержание и характер которых исследуется, например, в книге В. М. Бехтерева [3], которая до сих пор не утратила своей актуальности. Согласно теории В. М. Бехтерева, одним из сочетательных

рефлексов, имеющих место в мозге человека, является естественный язык. Учитывая изложенное, отделение языковых процессов от ментальных, принятое многими лингвистами, а также попытки изучения языка «самого по себе» нам представляются неоправданным и методологически некорректным упрощением. Языковую систему следует рассматривать как открытую, что предполагает существенное расширение и феноменологической базы языка, и соответствующего концептуального аппарата. Вспомним в этой связи слова В. Гумбольдта из его письма Ф. Вольфу (1804 г.): «Мне удалось открыть — и эта мысль все более захватывает меня, — что через посредство языка можно окинуть взглядом самые высокие и глубокие сферы и все многообразие мира». При определенном взгляде на вещи язык приобретает черты универсальной субстанции космического масштаба («Вначале было Слово. И Слово было у Бога. И Слово было Бог»).

Устная и письменная формы языка в этом смысле играют роль моделей мыслеречевых процессов и одновременно — коммуникативной для них среды (как отмечалось — инфраструктуры). При такой факторизации можно утверждать, что они представляют языковую периферию.

Сделаем, однако, предостережение относительно возможной недооценки инфраструктурных составляющих языка, которая может возникнуть ввиду того, что «основными», первичными языковыми объектами нами признаются психофизические состояния и процессы мыслеречевого аппарата. Дело в том, что современные данные свидетельствуют, что язык не является врожденным свойством человека. Врожденной является всего лишь *способность* к языку. А процесс «инсталляции» языка у человека с необходимостью предполагает наличие таких инфраструктурных элементов, как так называемые «внешний» язык и «эгоцентрический» язык, которые функционируют уже на ранних стадиях филогенетического развития языковой системы у ребенка и заканчиваются формированием у него «внутреннего» языка, венчающего процесс создания полноценного языкового аппарата [4]. Таким образом, языковая периферия является неотъемлемым элементом языковой системы. Кроме того, она обеспечивает открытость языковой системы (если, разумеется, не принимать во внимание возможность телепатии). Языковая периферия играет также важную роль «классического» объекта, благодаря которому, как отмечал Нильс Бор, только и возможно «квантовое» описание мира.

Следует отметить, что психофизические состояния и процессы (а среди них и мыслеречевые), как правило, не в полной мере достижимы для непосредственного наблюдения, а тем паче — объективной фиксации. Следовательно, устная и письмен-

ная формы языка фактически служат репрезентантами наблюдаемых компонентов состояний языковых объектов и процессов, происходящих в мыслеречевом аппарате. В качестве таковых они и используются в роли основных объектов при концептуальном моделировании языка. Отмеченное обстоятельство еще более подчеркивает значение языковой периферии в изучении системы языка.

Идея наблюдаемости, впервые в истории науки подробно разработанная при обосновании квантовой механики, вообще является весьма продуктивной в методологическом отношении применительно к описанию сложных процессов. Как отмечал А. Эйнштейн в своей дискуссии с В. Гейзенбергом в 1926 г. при обсуждении оснований квантовой механики [5], построить теорию только из одних наблюдаемых величин нельзя (хотя без них никакая научная теория вообще немислима). В конце концов лишь сама теория должна определять, какие из ее величин являются непосредственно наблюдаемыми, а какие — нет. В частности, в квантовой механике состояния, вообще говоря, не являются непосредственно наблюдаемыми величинами. Последним, как отмечалось выше, соответствуют определенному типу операторы, действующие в пространстве состояний, и собственным значениям которых отвечают значения наблюдаемых.

Таким образом, в мире, а значит, и в познании существует «невидимое», то есть то, что не может быть наблюдаемо непосредственно. Это дает нам основания для признания роли веры в когнитивно-интеллектуальных процессах, которая, по выражению апостола Павла, есть уверенность в невидимом.

Следует, однако, признать, что более глубокие эпистемологические причины наблюдаемости не раскрываются и в квантовой теории. По нашему мнению, эти причины кроются в феноменологии сложности проявленного бытия. Любая реальность, будучи объектом наблюдения (иными словами, вступая в отношения «субъект — объект»), проявляет свою сущность в неких ограниченных формах, достижимых для восприятия субъектом. Проявленная сложность отмеченных форм (потенциально она бесконечна — ведь материя не имеет пределов делимости, так сказать, «вглубь» и не допускает контроля всех ее взаимосвязей и опосредствований) оказывается компенсированной, согласованной с «перцептивно-сенсорным» аппаратом воспринимающего субъекта или прибора (что, в принципе, одно и то же). Механизм самокомпенсации сложности представляется нам настолько универсальной «универсалией» отношения «субъект — объект», что его, возможно, стоит отнести к основополагающим принципам системного анализа. Отметим, что сложность здесь мы понимаем как позитивную научную категорию, для которой могут быть сформу-

лированы точные и даже формальные определения типа тех, что имеют место в теории А. Н. Колмогорова [6] и в обширных развитиях этой теории, где понятие сложности оказывается непосредственно связанным с понятием информации.

### 3. Очерк теории семантических состояний

#### 3.1. Понятие семантического состояния языковой единицы

Насколько нам известно, это понятие также впервые было использовано А. Н. Колмогоровым — он глубоко интересовался и лингвистическими вопросами — при попытке формального определения понятия падежа [7]. Применение понятия семантического состояния получило дальнейшее развитие в работах [8] — [14], согласно которым любое слово (вообще говоря — любая единица языка) в контексте или в языковом потоке находится в определенном семантическом состоянии. Для единиц лексического уровня это состояние представляет собой некую сумму признаков грамматической и лексической семантики. Собственно, сам процесс понимания языка с этой позиции выглядит как редукция априорного распределения лексем по тем признакам грамматической и лексической семантики, которые присущи субъективному лексикону реципиента, к одному определенному грамматическому и лексическому значению, характерному именно для того контекста, который находится в поле внимания реципиента и подлежит в данный момент процессу его индивидуальной языковой обработки.

Несмотря на то, что и до сих пор не создан формальный аппарат для определения системы всех возможных семантических состояний любой лексики, естественно считать, что достаточно адекватной моделью данной системы может служить комплекс грамматических и лексических значений — так, как он представлен, скажем, в большом толковом словаре того или иного языка.

Таким образом, при рассмотрении формальных аспектов семантики мы будем исходить из существования соответствия между языковой единицей и ее семантическим состоянием:

$$g : X \rightarrow g(X), \quad (1)$$

где  $X$  — некоторая единица языка;  $g$  — соответствие между  $X$  и  $g(X)$  — формальным объектом, представляющим семантическое состояние единицы  $X$ , который имеет своими детерминантами элементы средств материального выражения семантики. Для любой единицы  $X$  семантические состояния образуют определенное множество — будем считать его конечным, но не ограниченным и обозначим символом  $\{g(X)\}$ . Класс определенных единиц языка  $L$  обозначим символом  $W(L)$  или просто  $W$ , если речь идет лишь об одном конкретном языке; принадлеж-

ность  $X$  к  $W$  обозначим:  $X \in W$ ; множество всех семантических состояний для всех  $X \in W$  обозначим через  $\{g(X); X \in W\} \equiv G$ .

Допустим, что существует оператор  $F$ , действие которого определено на множестве семантических состояний  $G$  и который будем интерпретировать как оператор значения некой семантической категории (ими могут быть, кстати, и категории грамматической семантики — если, например,  $F$  является оператором части речи, то его значениями являются:  $f_1$  — «имя существительное»,  $f_2$  — «глагол»,  $f_3$  — «имя прилагательное» и т. д.). Это означает, что  $F$  представляет собой некий интеллектуальный механизм, который, анализируя семантическое состояние  $g(X)$ , в котором находится единица  $X$ , идентифицирует определенное значение конкретной семантической категории, отвечающей именно данному семантическому состоянию. Понятно, что установление системы операторов типа  $F$  предполагает наличие некой семантической теории, формализованной надлежащим образом.

На математическом языке действие оператора  $F$  может быть выражено следующим образом:

$$Fg_i(X) = f_i g_i(X), \quad (2)$$

где  $f_i$  — некоторое значение семантической категории; функции  $g_i(X)$ ,  $i = 1, 2, \dots$ , представляют семантические состояния единицы  $X$ , маркирующие ее принадлежность значению  $f_i$  категории  $F$ . Значения  $f_i$ ,  $i = 1, 2, \dots$ , будем называть *собственными значениями оператора  $F$* , отвечающими семантическим состояниям  $g_i(X)$ ,  $i = 1, 2, \dots$ . Множество всех семантических состояний, отвечающих собственному значению  $f_i$ , будем называть множеством *частных семантических состояний* и обозначим символом  $G(f_i)$ :

$$G(f_i) := \{g: Fg = f_i g_i\}. \quad (3)$$

По определению множество  $G(f_i)$  состоит только из тех семантических состояний, которые характеризуются определенным значением семантической категории  $F$ , а именно — значением  $f_i$ .

Уравнение (2) в тех случаях, когда множество, которое пробегает индекс  $i$ , состоит более чем из одного элемента, представляет формальное выражение явления семантической многозначности. Для единиц лексического уровня — это полисемия и омонимия (как лексическая, так и грамматическая). В частности, в случае омонимии это отображает ситуацию, когда словоформа  $X$  в одном контексте может находиться в состоянии, скажем, существительного, а в другом — в состоянии, например, глагола. Такая разнovidность омонимии обычно называется *межчастеречной омонимией*.

Возможны случаи, когда в уравнении (2) определенному значению  $f_i$  соответствует не одно семантическое состояние  $g_i(X)$ , а несколько:  $g_j^i(X)$ ,  $j = 1, 2, \dots$ .

Такие состояния будем называть *вырожденными*; о вырожденности сигнализирует наличие верхнего индекса возле символа семантического состояния  $g_j^i(X)$ . Например, в украинском языке словоформа «МАТИ» со значением категории «часть речи» «имя существительное» имеет два грамматических состояния:

$$g_{\text{іменник}}^{\text{жін. рід, одна, наз. відм.}}$$

где слово «МАТИ» имеет лексическое значение «Жінка стосовно дитини, яку вона народила», и:

$$g_{\text{іменник}}^{\text{чол. рід, множина, наз. відм.}}$$

где слово «МАТИ» имеет лексическое значение «Спортивні матраци». Этот пример иллюстрирует явление *внутреннечастеречной омонимии*.

Количество собственных семантических состояний единицы  $X$ , соответствующих определенному значению категории  $F$ , будем называть *кратностью вырождения этого семантического состояния*. Так, в рассмотренном примере состояние  $g(\text{МАТИ})$  имеет кратность вырождения 2:

$$g_{\text{іменник}}^{\text{жін. рід, одна, наз. відм.}} \text{ та } g_{\text{іменник}}^{\text{чол. рід, множина, наз. відм.}}$$

Семантические состояния, в которых оператор  $F$  имеет определенное значение, будем называть *чистыми*. Однако а priori нельзя запрещать существование семантических состояний, для которых оператор  $F$  не приобретает одного определенного значения, но может характеризоваться, например, двумя. Формально такую ситуацию можно изобразить с помощью соотношения:

$$Fg(X) = f_1 \alpha_1(g_1) g_1(X) + f_2 \alpha_2(g_2) g_2(X), \quad (4)$$

где семантическое состояние  $g$  языковой единицы  $X$  при действии на него оператора  $F$  расщепляется на два, а именно:  $g_1(X)$  и  $g_2(X)$ , где  $g_1(X)$  соответствует значению семантической категории  $f_1$ , а  $g_2(X)$  — значению семантической категории  $f_2$ ; лингвистический смысл функций  $\alpha_1(g_1)$  и  $\alpha_2(g_2)$  будет выяснен ниже.

Состояния, для которых оператор  $F$ , действуя на функцию семантического состояния, изображается комбинацией определенного числа чистых семантических состояний, соответствующих различным собственным значениям этого оператора, будем называть *смешанными семантическими состояниями*.

Таким образом, уравнение (4) определяет семантическое состояние, в котором находится единица  $X$  и которое представляет собой своеобразное явление суперпозиции («смеси») чистых семантических состояний  $g_1(X)$  и  $g_2(X)$ , отвечающих собственным значениям  $f_1$  и  $f_2$ , соответственно. Лингвистическая интерпретация уравнения (4) состоит в том, что единица  $X$  имеет семантические признаки одновременно и  $f_1$  и  $f_2$ . Соответствующие показатели —

идентификаторы отмеченной принадлежности содержатся в выражениях для семантических состояний  $g_1(X)$  и  $g_2(X)$ .

Такая ситуация довольно типична в языке. Например, в украинских и русских причастиях соединены свойства глагола и прилагательного. Рассмотрим, например, русские причастные лексемы «ведущий» и «ведомый». Они склоняются по словоизменительной парадигме прилагательного (по шесть падежей в мужском, женском и среднем родах и в множественном числе), имея одновременно в своей структуре глагольный морфологический признак активности/пассивности — он приобретает здесь материальное выражение в виде суффиксов *-ущ* и *-ом*, соответственно. Отмеченный морфологический признак не выглядит сильным с точки зрения принадлежности к глаголу, где он не приобрел статуса словоизменительного; кроме того, он характерен для полной словоизменительной парадигмы, а не только для отдельных ее членов — это дает основание для выделения класса слов с такими свойствами в самостоятельную часть речи — причастие. В украинском же переводе причастие звучит как «дієприкметник», и в этом названии нашли свое отражение оба начала — глагола и прилагательного.

### 3.2. Семантические состояния в формализме нечетких множеств

Подчеркнем принципиальное отличие между описанной ситуацией и обычной омонимией. Явление омонимии также можно описать уравнением типа (4):  $Fg(X) = f_1\alpha_1(g_1)g_1(X) + f_2\alpha_2(g_2)g_2(X) + \dots$ , где разные члены правой части соответствуют различным омонимическим состояниям слова  $X$  — если ему присуща та или иная разновидность омонимии. Но в процессе языковой обработки, когда происходит снятие омонимии, правая часть этого уравнения редуцируется к одному члену, который, собственно, и представляет чистое грамматическое состояние анализируемого слова  $X$  в конкретном контексте. Совсем иная ситуация, когда  $X$  находится в смешанном состоянии — в этом случае никакая языковая обработка не способна уменьшить количество членов уравнения (4), которое и представляет ее конечный результат. Такое положение сигнализирует о существовании единиц, которые в контексте функционируют в смешанных семантических состояниях.

Вывод о существовании единиц, для которых язык допускает функционирование только в смешанных семантических состояниях, выглядит весьма интересным свидетельством невозможности полной формализации языковой системы, точнее, свидетельством пределов этой формализации. Это, в частности, означает, что даже наиболее формальная область теории языка, каковой является грамматика, имеет черты нечеткости, что входит в опре-

деленное противоречие с грамматическим детерминизмом и стимулирует выработку специального языка описания многозначных грамматических ситуаций.

Концептуальную базу такого языка, по нашему мнению, предоставляет теория нечетких множеств Л. Заде [15]. Продемонстрируем, каким образом это достигается.

Обратимся к определению множества  $G$  семантических состояний и множества  $G(f_i) = \{g: Fg = fg_i\}$  частичных семантических состояний. Если бы существовала принципиальная возможность редукции в процессе языковой обработки любого семантического состояния к чистому, то множество  $G$  можно было бы представить в виде объединения множеств  $G(f_i)$ , которые между собой не пересекаются, то есть справедливой была бы формула:

$$G = \cup_i G(f_i); G(f_i) \cap G(f_j) = \emptyset \text{ при } i \neq j. \quad (5)$$

Ситуация совсем иная, если мы учтем описанную возможность существования единиц, одновременно характеризующихся, например, двумя (или более) значениями определенной семантической категории. Семантическое состояние  $g(X)$  такой единицы  $X$  уже не будет принадлежать лишь одному из подмножеств  $G(f_i)$ , но одновременно двум, а возможно, даже и больше.

Формальный механизм описания подобных явлений осуществляется следующим образом. Определим на множестве  $G = \cup_i G(f_i)$  структуру нечеткого в смысле Заде множества. Для этого на каждом из подмножеств  $G(f_i)$  определим функцию принадлежности  $\alpha_i(g)$ , которая для каждого  $g \in G(f_i)$  приобретает определенное числовое значение из интервала  $[0, 1]$ :

$$\alpha_i(g) \in [0, 1]. \quad (6)$$

При этом будем полагать, что если  $\alpha_i(g) = 1$ , то состояние  $g$  является чистым, а если  $\alpha_i(g) < 1$ , соответствующий  $g$  является компонентом смешанного состояния, второй компонент которого  $g'$  (для двухкомпонентных состояний) принадлежит определенному подмножеству  $G(f_j)$ ,  $i \neq j$ , со значением функции принадлежности  $\alpha_j(g')$ , также меньшим 1, но таким, чтобы выполнялось условие:

$$\alpha_i(g) + \alpha_j(g') = 1. \quad (7)$$

Тогда каждое из подмножеств  $G(f_i)$  превращается в нечеткое множество с функцией принадлежности  $\alpha_i(g)$ :

$$G(f_i) \rightarrow \{G(f_i), \alpha_i(g), g \in G(f_i)\}. \quad (8)$$

Элементами определенного таким образом нечеткого множества служат пары  $(g(f_i), \alpha_i(g))$ . Одновременно структура нечеткого множества индуцируется и на всем множестве  $G$  как объединении нечетких множеств (8) при выполнении условия (7).

Функция принадлежности  $\alpha_i(g)$  при этом получает интерпретацию как мера приобретения семантическим состоянием  $g$  свойств значения  $f_i$  категории  $F$ . Максимальное значение, равное 1, функция принадлежности приобретает на чистых состояниях, характеризующих слова с однозначно определенным показателем принадлежности к определенной семантической категории.

Для слов со смешанными семантическими состояниями вида:

$$g(X) = \alpha_1(g_1)g_1(X) + \alpha_2(g_2)g_2(X) \quad (9)$$

величина  $\alpha_1(g_1)$  демонстрирует степень, в которой языковой объект  $X$  проявляет свойства значения  $f_1$ , а  $\alpha_2(g_2)$ , соответственно, — степень, в которой  $X$  проявляет свойства значения  $f_2$ . Условие (7), а именно:  $\alpha_1(g_1) + \alpha_2(g_2) = 1$ , обеспечивает полноту семантических свойств рассматриваемого объекта и замкнутость семантического описания.

### 3.3. Формальные аспекты отношения синонимии и аналогия с определением переводных эквивалентов

В этом пункте рассмотрим некоторые формальные аспекты отношения лексической синонимии, поскольку данное отношение играет весьма важную роль в языке, в частности, в процессах межъязыковой коммуникации.

Будем считать, что лексемы  $x$  и  $y$  находятся в отношении синонимии  $S$ , то есть  $xSy$ , если существуют семантические состояния (состояния лексической семантики)  $c(x)$  и  $c(y)$ , которые являются близкими:  $c(x) \sim c(y)$ . Это означает, что:

$$|c(x) - c(y)| < \varepsilon, \quad (10)$$

где  $\varepsilon$  — некоторая достаточно малая величина.

Понятно, что для оценки величины  $|c(x) - c(y)|$  необходимо иметь такие формальные определения для  $c(x)$  и  $c(y)$ , которые позволяют вычислить разность  $c(x) - c(y)$  и оценить ее абсолютную величину. Это довольно сложная задача, которая обычно решается путем экспертной оценки, выполняемой лингвистом, который, анализируя значения (семантические состояния)  $c(x)$  и  $c(y)$  и полагаясь на свою лингвистическую компетенцию, сам устанавливает их близость.

Не останавливаясь на чисто лингвистических вопросах, отметим некоторые формальные свойства отношения синонимии.

Очевидно, что рефлексивно и симметрично, то есть:  $xSx$  и  $xSy \Rightarrow ySx$ . Однако отношение синонимии не обязательно является транзитивным, то есть из фактов  $xSy$  и  $ySz$ , вообще говоря, не следует  $xSz$ . Это означает, что разности  $|c(x) - c(y)|$  и  $|c(y) - c(z)|$  могут накапливаться так, что  $|c(x) - c(z)| > \varepsilon$ . Мы рассматриваем, однако, лишь такие синонимические ряды (синсеты)  $SX$ , что если  $X = \{x_1, x_2, \dots, x_n\}$  — множество лексем, составляющих некоторый синсет  $SX$ , то  $x_i S x_j$  для всех  $i, j = 1, 2, \dots, n$ . Это дает возмож-

ность обозначить множество значений синсета  $SX$  через  $C(SX)$ . Следовательно, для любых  $c(x) \in C(SX)$  и  $c(y) \in C(SX)$   $|c(x) - c(y)| < \varepsilon$ .

Отметим принципиальную разницу между описанием лексической семантики в толковых и синонимических словарях. А именно, толковый словарь стремится представить, так сказать, абсолютную семантику, то есть для каждой лексемы  $x$  определить и как можно полнее описать множество ее семантических состояний:

$$C(x) = \{c_i(x) \mid i = 1, 2, \dots, k\}, \quad (11)$$

причем таким образом, чтобы для любого контекста  $M(x)$ , содержащего форму  $x$ , нашелся бы элемент  $c_M(x) \in C(x)$ , представляющий значение лексемы  $x$  и в этом контексте. Понятно, что множество  $C(x)$  содержит как чистые, так и смешанные семантические состояния.

В то же время семантика синонимии относительно и в формализме семантических состояний представляется всего лишь разностью  $|c(x) - c(y)|$  семантических состояний  $c(x)$  и  $c(y)$ , удовлетворяющих условию близости в смысле, изложенном выше.

На практике семантика толковых словарей, как правило и к сожалению, не согласована с семантикой синонимических, то есть определение семантических состояний при синонимии, зафиксированное в формуле (10), не всегда согласуется с определениями состояний лексической семантики, представленными в формуле (11). Отмеченный факт создает значительные проблемы при разработке семантических анализаторов и двуязычных переводных словарей и систем лексикографического обеспечения машинного перевода.

Действительно, между определением синонимии и нахождением переводных эквивалентов существует глубокая аналогия. Она базируется на том обстоятельстве, что, если нахождение синонима сводится к поиску близких значений лексем одного языка, то нахождение переводного эквивалента есть тот же поиск близких значений лексем, но разных языков. Введем верхний индекс возле языковой единицы и ее семантического состояния для обозначения языка, в котором эта единица (или состояние) представлены. Тогда синонимия отображается соотношением:  $|c^1(x) - c^1(y)| < \varepsilon$ .

В то же время отношение «переводной эквивалент»:  $x^1 T x^2$  изображается соотношением  $|c^1(x^1) - c^2(x^2)| < \varepsilon$ . Понятно, что для формального определения последней процедуры необходимо привести величины  $c^1(x^1)$  и  $c^2(x^2)$  к какому-то единому представлению, которое может предоставить, например, язык-посредник — обозначим ее верхним индексом 0. Таким образом, сначала имеем отображение  $c^1(x^1) \rightarrow c^0(x)$  и  $c^2(x^2) \rightarrow c^0(y)$ . Тогда получаем эквивалентность:

