



В.А. Широков

Украинский языково-информационный фонд НАН Украины, г. Киев, Украина

ЛИНГВИСТИКА И СИСТЕМНЫЙ ПОДХОД. ЧАСТЬ 1

На основе феноменологического анализа свойств языка формулируются принципы системного подхода, основанного на рассмотрении отношений внутри триады «структура – субстанция – субъект». Подчеркивается системообразующий характер отмеченной триады. Излагается сущность лексикографического эффекта в информационных системах. Приведен ряд примеров из лингвистики, иллюстрирующих авторскую системологическую концепцию.

ЛЕКСИКОГРАФИЧЕСКИЙ ЭФФЕКТ, СИСТЕМНЫЙ ПОДХОД, ТЕОРИЯ СЛОЖНОСТИ, СТРУКТУРА, СУБСТАНЦИЯ, СУБЪЕКТ

Введение

Предлагаемый вниманию читателя текст является результатом авторских рефлексий, возникших в процессе исследования свойств словарей, в первую очередь – больших многотомных лексиконов, в которых, как отмечают их создатели, эксплицирована «система» языка. Данное исследование в 90-е годы прошлого столетия привело к определенным обобщениям, результатом которых стало построение теории лексикографических систем¹ – достаточно универсальной информационной конструкции, которая сочетает в себе черты таких формализованных объектов как модели данных и знаний, логико-лингвистические исчисления и тому подобное. Характерно, что в структуре лексикографических систем эксплицированы именно «системообразующие» свойства естественного языка, хотя концептуальный масштаб теории лексикографических систем и возможности ее применения намного превышают то, что сосредоточено в собственном феноменологическом кругу естественного языка.

Отметим, что теория лексикографических систем изначально строилась на определенных системных началах, и среди них прежде всего – на системных свойствах естественного языка. Отмеченная системность языка, который является одним из символов веры современной лингвистики, утверждает такое ее понимание, при котором учитываются не только свойства отдельных элементов языка, но и связей (отношений) между ними, а также и свойства самих этих отношений. Именно указанное обстоятельство, по мнению большинства лингвистов, и сообщает языку свойство системности.

В принципе, далее отмеченной констатации (весьма, впрочем, полезной на практике) языкознание, как правило, не распространяется. Таким образом, свойство системности по существу сводится к понятию структуры, ибо структура объекта, как известно, и определяется набором его элементов и связей (отношений, операций ...) между ними.

¹ В.А. Широков. Інформаційна теорія лексикографічних систем. – К.: Довіра, 1998, 331 с.

Относительно возможных других, дополнительных свойств, которые могут и должны быть положены в основу определения² понятия системы, в свое время также высказывались разные мнения – как собственно в области лингвистики, так и вне ее.

В частности, обсуждалась и возможность учета субстанциальных свойств в определении понятия системы, правда, в смысле скорее отрицательном. Как отмечал Г.П. Мельников в своей известной работе «Системный подход в лингвистике»³: «В работах по системно-структурной методологии и в чисто структуралистских трудах проблема «субстанции», «материальности», «субстрата» или вообще не обсуждается, или упоминается только в связи с тем, чтобы показать, что она не имеет сущностного значения при исследовании объекта как системы. Известный тезис Ф. де Соссюра «язык есть форма, а не субстанция» полностью совпадает с утверждением У. Росса Эшби о том, чем должна заниматься общая теория систем: «Нужно исключить из рассмотрения два фактора, не относящихся к делу. Первым из них является «материальность» – идея о том, что машина должна быть сделана из реальных материалов ... Точно так же не относится к делу любая ссылка на энергию ...». Сейчас, по прошествии значительного времени с момента публикации такого рода взглядов уже стало почти очевидным, что они являются идеализациями, методологическими упрощениями, не вполне адекватно отражающими реальность.

И в самом деле, возьмем, например, понятие информации и ее реальные воплощения в вычислительных системах. Долгое время считалось, что принципиальные свойства вычислительных систем не зависят от «субстанции», из которой составлены преобразователи информации – будь то радиолампы, полупроводники или микроэлектронные чипы. Считалось, что от этой субстанции

² Автор вполне осознает, что точного определения понятия системы быть не может, поэтому здесь слово «определение» употребляется в несколько «пиквикском» смысле.

³ Г.П. Мельников. Системный подход в лингвистике. // Системные исследования. Ежегодник 1972. – М.: Наука, 1973, с. 183–204.

зависят разве что количественные параметры вычислителей — память и быстродействие. Однако ситуация кардинально изменилась с открытием квантовой информации, причем не в смысле банального «перехода количества в качество», а в связи с тем, что квантовая информация обладает свойствами, принципиально отличными от классической (речь идет о квантовой суперпозиции и квантовой запутанности⁴), так что вычислительные системы, создаваемые на основе такой «субстанции», будут иметь и принципиально отличные от классических вычислителей именно системные свойства. Приведенный пример является всего лишь одной из иллюстраций субстанциальной определенности систем.

Г.П. Мельников еще более 40 лет назад предложил расширить определение понятия системы, включив в него субстанциальные свойства⁵. Таким образом, структура и субстанция по Мельникову являются «системообразующими» концептами самого понятия «система», внутренними имманентами его определения.

Автор, солидаризируясь с Г.П. Мельниковым в части реабилитации субстанции, как неотъемлемой концептуальной составляющей определения понятия системы, считает такое расширение все же недостаточным и предлагает дополнить его за счет еще одной концептуальной компоненты.

Имеется в виду следующее.

Начнем с некоторых простых примеров, относящихся как к области языка, так и находящихся вне ее.

Рассмотрим такой относительно простой и хорошо формально определенный объект как система линейных уравнений. В каком смысле и при каких условиях мы можем рассматривать ее именно как «систему» в понимании системного подхода? В стандартных обозначениях система линейных уравнений n -го порядка имеет следующий вид:

$$\sum_{j=1}^n a_{ij}x_j = b_i,$$

где $i, j = 1, 2, \dots, n$; $a_{ij} \in Q$; $x_j \in Q$; $b_j \in Q$. Структура данного объекта сводится к структуре линейных комбинаций и отношений равенства. Уже на этом

первом шаге мы сталкиваемся с невозможностью полного формального определения системы без информации о «субстанции» величин, составляющих эти линейные комбинации. Действительно, в приведенной формуле фигурируют операции сложения и умножения, которые по-разному определяются для различных алгебраических объектов (можно сказать, разных алгебраических субстанций). То же касается и равенства — операции (или отношения) сравнения. Зафиксировав, например, элементы из этой формулы, как принадлежащие к полю рациональных чисел, мы получаем систему, свойства которой отличаются от той, как если бы в данной формуле фигурировали объекты иной алгебраической природы. Следовательно, даже в этом простейшем случае свойство системности предполагает определение субстанциальных свойств элементов структуры системы. Таким образом, «субстанциальная» инициатива Г.П. Мельникова обретает здесь вполне естественную иллюстрацию.

Но этого, как мы полагаем, еще недостаточно. Ведь объект, определенный этой формулой, приобретает свойства системы тогда и только тогда, когда все его образующие элементы и отношения получают некую достаточно определенную интерпретацию. А для этого необходимо иметь внятный механизм или инструмент, способный «распознавать» эти элементы и отношения, идентифицировать их как элементы определенных совокупностей, выполнить с ними какие-то разрешенные для этих совокупностей действия согласно определенным правилам (алгоритмам) и, наконец, получить совокупность решений «системы» или убеждение, что таковых не существует. Именно эта идентификационно-интерпретационно-алгоритмическая компонента, действующая как своего рода *субъект* — «аналитическое» и «активное» начало и образует последний член триады «структура — субстанция — субъект», которая, как мы считаем, является системообразующей для определения понятия системы.

Таким образом, определение понятия системы можно изобразить в виде символического равенства: $C = C + C + C$, где «C» левой части обозначает понятие «система», а правая часть демонстрирует наличие и взаимодействие основных образующих компонент этого понятия, а именно «структуру», «субстанцию» и «субъект»⁶.

Сделаем некоторые уточнения и разъяснения касательно этого последнего члена системной триады. Во-первых, выясним, является ли свойство субъектности внутренним свойством системы или оно есть результат действия какой-то причины, внешней по отношению к самой системе.

⁶ Автор также осознает те трудности, которые встретятся на пути определения понятий структуры, субстанции и субъекта. Поэтому данным вопросом, которому, впрочем, посвящена колоссальная библиография, в настоящей работе мы заниматься не будем, апеллируя к интуиции читателя.

⁴ Прескилл Дж. Квантовая информация и квантовые вычисления. Том 1. — М.—Ижевск: НИЦ «Регулярная и хаотическая динамика»; Институт компьютерных исследований, 2008. — 464 с. К. А. Валиев. Квантовая информатика. Компьютеры, связь и криптография. Вестник РАН, том 70, № 8, с. 688–695 (2000). С.И. Чивилихин. Квантовая информатика. Учебное пособие. Санкт-Петербург, 2007. С. Eric S. Piotrowski, Toward a Quantum Linguistics: Possibilities for Change in the Delta Zone, 2000. Stapp, Henry P. “Quantum Nonlocality and the Description of Nature,” in James T. Cushing and Ernan McMullin, eds., Philosophical Consequences of Quantum Theory (Notre Dame: University of Notre Dame Press, 1989). Strehle, Susan. Fiction in the Quantum Universe. Chapel Hill: University of North Carolina Press, 1992.

⁵ Отметим также его работу «Системная лингвистика и ее отношение к структурной» // Проблемы языкознания: Докл. и сообщ. советских ученых на X Международном конгрессе лингвистов. — М.: Наука, 1967, с. 98–102.

1. Информационное введение в понятие лексикографического эффекта в информационных системах

1.1. Начальные представления о лексикографическом эффекте в информационных системах

Некоторые свойства субъектности, как третьего члена триады «структура — субстанция — субъект», можно прояснить, анализируя механизмы лексикографического эффекта в информационных системах, сформулированного нами в 90-х годах прошлого века⁷. Для этого нам понадобится понятие лексикографического описания систем, чему и будет посвящен данный раздел, изложение которого в основном следует работам⁸. При этом под лексикографическим описанием мы понимаем гораздо более общий тип описания систем, чем принятый в классической лингвистической лексикографии и трактующий вопросы составления словарей. Коротко говоря, лексикографическое описание изучаемого объекта состоит в:

а) выделении в нём дискретных спектров неких единиц, относящихся к его онтологической природе, а именно, таких дискретных, относительно устойчивых сущностей, которые, комбинируясь, эксплицируют феноменальные проявления исследуемого объекта;

б) построении интерпретаций упомянутых спектров как со стороны форм их проявления, так и со стороны внутреннего содержания, связанного с данными формами и в них проявляющегося.

Базовые принципы лексикографического описания в нашем понимании основываются на информационной теории систем⁹, благодаря чему данное описание, построенное на информационной интерпретации, приобретает намного более общие черты, чем относящееся непосредственно к языковой системе, и, фактически, касается любых объектов, где происходят информационные процессы с системными признаками, лишь в некоторой степени подобными соответствующим признакам языковой субстанции. По своей типологии принятый нами подход мы характеризуем как феноменологический, поскольку он базируется на общих информационных проявлениях систем и не является специфическим следствием их конкретного строения. Одновременно подчеркнем наше сознательное стремление при рассмотрении информационных аспектов реальности быть максимально приближенным к тому кругу феноменов, которые близки к естественноречевым либо могут квалифицироваться как таковые.

⁷ Первое регулярное изложение теории лексикографического эффекта осуществлено в нашей книге «Інформаційна теорія лексикографічних систем» (К.: Довіра, 1998).

⁸ В.А. Широков. Феноменологія лексикографічних систем. — К.: Наукова думка, 2004. В.А. Широков. Елементи лексикографії. — К.: Довіра, 2005. В.А. Широков. Комп'ютерна лексикографія. — К.: Наукова думка, 2011.

⁹ Шилейко А.В., Кочнев В.Ф., Химушин Ф.Ф. Введение в информационную теорию систем. — М.: Радио и связь, 1985. — 278 с.

Общей чертой всех процессов информационного обмена является преобразование информации из одной формы в другую, причем современные естественнонаучные теории вполне определенно подтверждают принципиально дискретный („квантовый”) характер процессов взаимодействия и обмена и, следовательно, принципиальную дискретизацию процессов восприятия и интерпретации (описания) действительности. Отмеченная дискретизация имеет по крайней мере одну общую для всех известных процессов черту, как нам представляется, фундаментального характера. А именно: наблюдая и обобщая поведение различных систем, заключаем, что в *процессе эволюции (динамики, саморазвития...) системы любой природы в ее структуре, при ее взаимодействии (наблюдении) с неким субъектом проявляется подсистема относительно устойчивых дискретных сущностей* („подсистема порядка”), которые играют роль ее *элементарных информационных единиц*, так что все другие феномены системы представляют собой определенным образом организованные комбинации этих элементарных информационных единиц.

Отмеченная подсистема обладает свойствами, некоторым образом родственными свойствам лексической системы естественного языка: она „генерирует” в своей структуре нечто наподобие тезауруса и грамматики с присущими этим конструкциям свойствам знаковости, значения, содержания, полиморфности и т. д.; она является носителем как „плана выражения”, так и „плана содержания”. Данное обстоятельство поясняет употребление нами термина «лексикографический эффект».

Совокупностям элементарных информационных единиц присуще свойство „субстанциальности”, как и другим агрегациям, обусловленным объективными процессами, вследствие чего отмеченные совокупности, как правило, обладают относительной устойчивостью своих характеристик, обеспечивающей их локализацию в соответствующих областях системных параметров. Точно также мы заключаем, что любой лексикографический эффект развивается в среде отношений «субъект — объект», причем проявления субъектности варьируют в чрезвычайно широких границах, начиная с универсального свойства отражения, присущего всем вещам, и заканчивая психическими и когнитивными реакциями и отправлениями интеллектуальных сущностей.

Описанная совокупность явлений составляет содержание лексикографического эффекта. Можно утверждать, что при изучении любых предметных областей специалисты фактически исследуют лексикографические эффекты, происходящие в данных областях. Таким образом, лексикографический эффект обладает не только феноменологической составляющей, но и методологической, поскольку он имеет определенный „потенциал

операциональности”, стимулируя в процессе моделирования тех или иных систем устанавливать и определять соответствующие комплексы элементарных информационных единиц, учитывая, конкретизируя и репрезентируя их свойства. В этой ипостаси концепция лексикографического эффекта выступает как метод абстрагирования данных.

Поскольку лексикографический эффект выражается в представлении некоего, в сущности, континуального универсума через посредство дискретных множеств, появляется привлекательная возможность его формализации, используя известную в теории моделей теорему Лёвенгейма-Сколема¹⁰. Последняя состоит в том, что при определенных ограничениях можно установить некий изоморфизм между несчетным и счетным множествами; в некотором смысле, потенциальная бесконечность может иметь конечную интерпретацию, по сути, строится финитная модель бесконечности. Теорема Лёвенгейма-Сколема утверждает, что любая непротиворечивая теория первого порядка, имеющая несчетную модель, имеет также и счетную модель. Это означает, что если некое структурированное множество задано счетным множеством правил, то существует счетное множество (то есть собственное подмножество множества натуральных чисел), на котором можно построить точную модель этого структурированного множества, где будут выполняться все исходные аксиомы. Следовательно, имеет место такое представление бесконечного описания объекта через конечное, которое содержит в себе всю информацию о бесконечном объекте. Таким образом, можно утверждать, что теорема Лёвенгейма-Сколема на самом деле выступает в роли формального коррелята лексикографического эффекта¹¹.

В дальнейшем комплекс элементарных информационных единиц системы D , индуцированный вследствие развития лексикографического эффекта Q , обозначим через $I^Q(D)$ или просто $I(D)$, если ссылка на тип лексикографического несущественна. Система элементарных информационных единиц обладает определенной структурой. В частности, в ней можно выделить ядро — некую собственную подсистему $I_0^Q(D) \subset I^Q(D)$ и определить порождающую процедуру π :

$$\pi: I_0^Q(D) \rightarrow I^Q(D). \quad (1)$$

Тройку $(I^Q(D), I_0^Q(D), \pi)$ будем также отождествлять с системой элементарных информационных единиц и использовать наряду с $I^Q(D)$, $I(D)$, $I_0^Q(D)$, $I_0(D)$ как эквиваленты, полагая, что порождающая процедура определена, известна и понятна из контекста.

Подытоживая содержание данной феноменологии, констатируем, что процессы, аналогичные описанным, происходят во всех достаточно сложно устроенных естественных и социотехнических системах и вообще в системах любой природы, в которых действуют источники, преобразователи и потребители информации и, значит, происходят аналоги перцептивно-сенсорных актов и интеллектуальных процессов, реализующихся в среде отношений системной триады «структура — субстанция — субъект». Сказанное сообщает дополнительное основание для справедливости вывода о существовании и универсальности лексикографического эффекта в информационных системах с характеристиками *знаковости, значения, содержания и полиморфности*, роднящего его с естественным языком, которому также присущи указанные свойства¹². Отмеченные факторы усиливают тенденцию построения как можно более формально определенных структур, являющихся представителями лексикографических эффектов, и их применения к описанию феноменов любой природы, в частности, естественных языковых.

Многочисленные подтверждения изложенного мы наблюдаем в целом ряде природных явлений. Несмотря на то, что по современным представлениям всё мироздание помещено в четырёхмерный пространственно-временной континуум, наблюдаемые величины зависят от небольшого числа так называемых «мировых констант» (скорости света, заряда и массы электрона, постоянной Планка и т.п.), которые вместе со множеством целых чисел составляют своеобразный «алфавит» физики, в котором выражаются все содержательные утверждения о поведении физических систем (значения наблюдаемых величин). Аналогичная ситуация и с научным описанием иных системы: все химические вещества являются некими «словами» в алфавите химических элементов, а их взаимопревращения — «предложениями» этого языка; белки в основном состоят из 20 аминокислот, молекулы ДНК строятся на основе четырех типов нуклеотидов и т.д. Сами химические элементы являются элементарными информационными единицами относительно лексикографического эффекта, определяемого принципом Паули, который допускает нахождение в одном квантовом состоянии не более одного электрона, регламентируя таким образом порядок заполнения электронных орбит атомов и, следовательно, определяя возможный набор химических элементов. Примеры такого рода можно приводить десятками. Показательно, что данное поведение характерно не только для, так сказать, реальных объектов мира — природных и технических, но и для образований концептуального уровня, оперирующих идеальными объектами, абстракциями и ментальными конструкциями.

¹² Соломоник А. Семиотика и лингвистика. — М.: Молодая гвардия, 1995. — 352 с.: ил.

¹⁰ Бруно Пуаза. Курс теории моделей., Алматы, 2001. — 460 с. (электронная книга в PDF).

¹¹ Г.Г. Головкин, В.А. Широков. Теорема Лёвенгейма-Сколема как формальный коррелят лексикографического эффекта в информационных системах. MegaLing'2006. Горизонти прикладної лінгвістики та лінгвістичних технологій// Доповіді Другої Міжнародної наукової конференції, 20-27 вересня, 2006, Україна, Крим, Партеніт. ISBN: 966-8180-44-5. С. 201-202.

1.2. Информационная сущность отношения «форма — содержание»

Для проведения анализа деталей разворачивания отношения „форма — содержание” (ОФС) рассмотрим диаграмму, символически изображающую процесс восприятия некоего объекта неким субъектом:

$$S: D \longrightarrow V(D). \quad (2)$$

Здесь буквой D обозначено “нечто” из реального (или воображаемого) мира, выступающее в качестве объекта процесса восприятия (наблюдения, изучения, внимания, переживания ...) со стороны некоего S , которое мы считаем субъектом данного процесса; через $V(D)$ обозначим результат этого процесса. Отметим, что в роли S может выступать человек или сконструированный человеком прибор, или человекомашина, или что-либо другое, наделенное свойствами восприятия и ощущения (“отображения”); S может быть и “коллективным субъектом” — группой людей, социальным сообществом, этносом, нацией, народом, совокупностью народов или даже человечеством в целом.

Здесь мы не детализируем принципы и способы устройства субъекта S , за исключением одного: S имеет два определяющих свойства: а) восприятие и ощущение внешних сигналов (“перцептивно-сенсорный” аппарат) и б) способность к их внутренней обработке (осознанию и интерпретации). Таким образом, в результате физической, психической, интеллектуальной и другой ограниченности субъекта S вся совокупность свойств объекта D для его восприятия делится на две, не очень четкие, неоднозначные, меняющиеся и нечетко отделяемые части. К первой из них относим те свойства D , которые более непосредственно воспринимаются “перцептивно-сенсорным” аппаратом S — обозначим эту часть через $F(D)$ и трактуем ее как совокупность свойств D , соотношенных к его форме с точки зрения субъекта S , воспринимающего D . Ко второй части относим свойства D , непосредственно не воспринимаемые перцептивно-сенсорным аппаратом S , но отражающиеся в нем опосредованно через средства внутренней обработки. Обозначим эту часть через $C(D)$ и будем рассматривать ее как совокупность содержательных свойств D — опять-таки с точки зрения восприятия субъекта S . Таким образом, формула (2) приобретает следующий вид:

$$D \xrightarrow{S_F} F(D) \xrightarrow{H} C(D) \quad (3)$$

где символом S_F обозначено действие “перцептивно-сенсорного” аппарата субъекта S , результатом которой является совокупность формальных (с точки зрения S) свойств D ; символом H обозначен механизм, осуществляющий интерпретацию (связь между формой и содержанием) и обеспечивающий целостность восприятия объекта D субъектом S (если ему действительно удастся обеспечить указанную целостность). В то же время, предположив существование механизма, который позволяет

переход от D к $C(D)$ — обозначим этот механизм через S_C , — получаем следующую трансформацию диаграммы (3):

$$\begin{array}{ccc} & S_F & \\ D & \nearrow & F(D) \\ S_C & \searrow & \\ & C(D) & \end{array} \quad (4)$$

где, как видим, произошла “декомпозиция” субъекта S на его “составляющие” S_F и S_C , отвечающие за реконструкцию формальных и содержательных свойств D соответственно. Итак, в приведенной декомпозиции субъект выступает имеющим дуалистическую природу: он обладает механизмом для непосредственного («перцептивно-сенсорного» — учитывая форму вещей) восприятия объективной действительности D , а также наделен аппаратом для «интеллектуальной» обработки полученной информации путем толкования ее «содержания». Кроме того, он снабжен и определенным «интерфейсом» между этими двумя способами обработки информации, который в схеме представлен элементом H .

Мы не склонны абсолютизировать описанную схему. Между $F(D)$ и $C(D)$ нет четкой границы, как ее вообще нет между формой и содержанием. Также почти никак не детализировались свойства S , хотя из общих соображений и была осуществлена его декомпозиция на S_F и S_C . Следовательно, этот подход действительно может квалифицироваться как феноменологический, поскольку он не опирается на предположение о возможной “конструкции” S и механизмов его функционирования. Из этих соображений можно утверждать, что изложенная схема является достаточно общей — в нее не заложено никаких специфических “анзатцев”. Единственным, характерным для языка, мы считаем предположение, что $F(D)$ должен иметь линейный характер, то есть изображаться линейными последовательностями дискретных объектов, источником которых выступает некое конечное множество. Учитывая сказанное, даже сама возможность существования такого явления как язык выступает следствием фундаментального свойства S “быть субъектом”, то есть таким, для которого любой объект, представляется имеющим свою внешнюю сторону (форму) и внутреннюю (содержание). Отношение между этими различными аспектами восприятия, символически изображенные величинами S_F , S_C , H , отличаются большим разнообразием, источником которого является фундаментально присущие (то есть такие, от которых в принципе нельзя избавиться) свойства воспринимающего субъекта S : изменчивость, нерегулярность, разнообразие, ограниченность, нечеткость и тому подобное.

Отметим еще одну черту процесса разворачивания ОФС. Современная культурология склонна трактовать ее как символ постмодернизма, а именно — влияние субъекта на объект, то есть возмож-

ность изменения состояния объекта D в процессе его восприятия (наблюдения, исследования ...) субъектом S . Дело в том, что для того, чтобы произошёл процесс, символически изображённый диаграммами (2) – (4), во многих случаях необходимо “активизировать” объект D , чтобы он “проявил” свои свойства, которыми “интересуется” S . В классической научной парадигме считалось, что такое возбуждение объекта можно сделать сколько угодно малым и пренебречь им, считая что оно существенно не меняет состояния объекта. Однако развитие науки обнаружило, что на самом деле это, вообще говоря, не так. Исторически первой научной дисциплиной, учитывающей влияние субъекта (прибора) на объект исследования, оказалась квантовая механика.

То, что было изложено выше, оказывается созвучным с совсем другими явлениями и их формальными моделями, которые были построены независимо, по другому поводу и для других целей. Речь идет об определении информации А.Н. Колмогоровым¹³.

Введение колмогоровской информационной меры, как известно, имеет целью уточнение понятия информации, во-первых, без привлечения вероятностного подхода, характерного для информационной меры К. Шеннона, а во-вторых, предоставление возможности применения этой меры к индивидуальным объектам.

Основная идея колмогоровского подхода заключается в том, что информация об объекте считается полученной, когда возможно воспроизвести (реконструировать) этот объект (модель объекта) по определенному его конечному описанию (набору признаков). Построение меры Колмогорова базируется на таких фундаментальных понятиях, как алгоритм, машина Тьюринга, рекурсивная функция, и восходит к идеям теории сложности вычислений (сложности алгоритмов), которая, собственно, и является источником интерпретации информации как степени сложности, неоднородности, нерегулярности и структурированности объектов, а также уверенности в универсальности такой характеристики как сложность, поскольку любой объект, независимо от его природы, характеризуется определенной сложностью, неоднородностью и имеет определенную структуру, хотя бы и тривиальную.

Соответствующая, не перегруженная деталями математическая конструкция, которую мы здесь приводим, следуя указанным трудам А. Н. Колмогорова, формулируется следующим образом.

Рассмотрим определенное счетное множество $X = \{x\}$. Будем считать, что существует взаимно-однозначное соответствие между X и множеством D двоичных слов, начинающихся с единицы – то есть, пусть задано биективное отображение:

$$n: X \rightarrow D, \quad (5)$$

такое, что каждому $x \in X$ однозначно отвечает некоторый $d = n(x)$, $d \in D$, и наоборот. Будем считать, что:

1. $n(x)$ – общерекурсивная функция на D . Обозначим через $l(d)$ длину двоичного слова $d \in D$, то есть число нулей и единиц, в нем содержащихся. Тогда $l(n(x)) = l(x) + C$, где C – некая константа.

2. существует однозначное отображение $\chi: X^2 = X \times X \rightarrow X$ такое, что для $\forall x \in X, y \in X \exists z \in X$, что $z = \chi(x, y) \equiv (x, y)$ и $n(z) = n(x, y)$ – общерекурсивная функция от $n(x)$ и $n(y)$, причем: $l(x, y) \leq C_x + l(y)$, где константа C_x зависит лишь от x .

Полагаем изоморфизм (5) установленным, так что множество X также рассматриваем как множество двоичных слов. Допустим, что существует частично рекурсивная функция $\phi(p, x)$, ставящая в соответствие двоичному слову x двоичное слово y , причем $p \in D$, интерпретируется как алгоритм (или программа), „перерабатывающая” x в y :

$$p: x \rightarrow y, \quad (6)$$

а ϕ представляет при этом метод (язык) программирования. Без утраты общности полагаем, что p для данного x задается неким двоичным словом.

Обозначим:

$$K_\phi(y|x) = \begin{cases} \min_p l(p), & \text{если } \phi(p, x) = y, \\ \infty, & \text{если не существует конечного } p \\ & \text{такого, что } \phi(p, x) = y. \end{cases} \quad (7)$$

Таким образом, $K_\phi(y|x)$ является длиной минимальной программы p , перерабатывающей x в y при заданном методе программирования ϕ . Эта величина называется сложностью y относительно x при данном ϕ . На интуитивном уровне понятно, что «сложность y относительно x при данном ϕ » связана с определенной неоднородностью объектов x и y , то есть с тем, насколько нерегулярно размещены нули и единицы в них, насколько это распределение является неоднородным, таким, которое не может быть существенно упрощено и, следовательно, не допускает более короткого описания. Разумеется, зависимость величины сложности от ϕ – недостаток описанного метода, но существует теорема¹⁴, утверждающая существование „наилучшего” метода программирования A такого, что для любой частично рекурсивной функции ϕ справедливо неравенство:

$$K_A(y|x) \leq K_\phi(y|x) + C_\phi, \quad (8)$$

где постоянная C_ϕ зависит только от ϕ и не зависит от x и y .

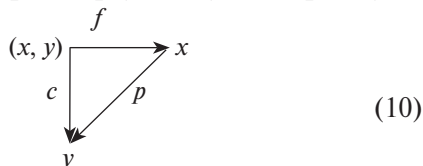
Величину $K_A(y) \equiv K_A(y|1)$, „отнормированную” относительно единичного элемента $x = 1$, естественно считать сложностью элемента y . При этом количество информации в объекте x относительно объекта y определяется как разность:

$$I_A(x|y) = K_A(y) - K_A(y|x). \quad (9)$$

¹³ Колмогоров А.Н. Три подхода к определению понятия „количество информации”. В кн. Теория информации и теория алгоритмов. – М., Наука, 1987. – С. 213–223.

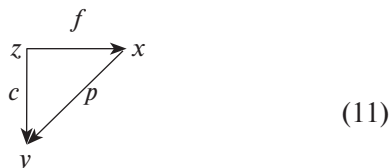
¹⁴ А.Н. Колмогоров. Три подхода к определению понятия количества информации. В кн. Теория информации и теория алгоритмов. – М.: Наука, 1987. – С. 220.

Последняя формула и определяет меру информации — так называемую алгоритмическую меру информации Колмогорова. Формула (9) и весь описанный подход определения информации через алгоритмическую сложность могут быть проинтерпретированы несколько иным способом. С этой целью построим треугольную диаграмму:



где x и $y \in X$; элемент $(x, y) \in X^2$, причем благодаря существованию отображения $\chi: X^2 \rightarrow X$, $\chi(x, y) = z$, $z \in X$. В диаграмме (10) отображения f и c реализуют проекции элемента (x, y) на первый и второй сомножители, соответственно, причем справедливыми остаются формулы (6) — (9) и интерпретация сложности и информационной меры. Диаграмму (10) полагаем коммутативной: композиция отображений f и p равняется c : $p \circ f = c$.

Предположим теперь, что объект z , который до этого представлял образ декартова произведения (x, y) при отображении χ , на самом деле является самостоятельным, в известной степени независимым от x и y объектом внешнего мира. Это предположение позволяет построить такую коммутативную диаграмму:



где $z \in Z$, Z — множество-источник объектов z , $x \in X$, $y \in Y$. Тогда отображения f и c определяют различные интерпретации объекта z , и, кроме того, отображение p интерпретирует x через y ; $p \circ f = c$. Естественно предположить, что объект x отражает “формальные” свойства объекта z , а y — его “содержательные” свойства, причем связь между “формой” и “содержанием” берет на себя p . Требование минимальности p здесь вполне понятно, ведь “толкование” формы (а его результат — это и есть содержание) не должно содержать “лишних”, случайных относительно x (а также и z) элементов. Построенная таким образом конструкция, основанная на тройке объектов (z, x, y) и тройке отображений (f, c, p) , формирующих диаграмму (11), допускает естественную интерпретацию как комплекс, который является носителем отношения “форма-содержание”.

Итак видим, что ОФС таким своеобразным способом оказывается “зашифрованным” в алгоритмической теории сложности и информации, а значит, это отношение (равно как и отношение «субъект — объект») тесно связано с информационными процессами; даже больше — оно является их системообразующим отношением. Между тем, оно также является одним из основных системообразующих отношений естественного языка, что настойчиво утверждает лингвистическая теория.

Очевидное сходство, «сродство» диаграммы (4) с определением информации по Колмогорову, а также дальнейшие рассуждения, которые привели к построению диаграммы (11), наводят на мысль о том, что в основе и того и другого лежат одинаковые закономерности. Сама форма представления информационной меры свидетельствует об определенном процессе, результатом которого стала генерация “алфавита” — знаковой системы представления объекта. Развертывание отображений (f, c, p) в диаграммах (10) — (11), которым сопоставляются элементы ОФС из диаграммы (4), побуждает к сопоставлению последних с составляющими информационных процессов, которые в теории Колмогорова сводятся к математическим отношениям (алгоритмам, рекурсивным функциям, ...), определенным на дискретных множествах.

Таким образом, лексикографический эффект можно рассматривать в качестве феноменологического основания теории сложности и соответствующей теории информации, а также наоборот: теорию сложности, колмогоровскую информацию и теорему Лёвенгейма-Сколема естественно рассматривать в качестве формальных коррелятов лексикографического эффекта в информационных системах.

2. Субъект в структуре лексикографического эффекта

Из формул (2) — (4) и конструкции информационной меры Колмогорова следует, что субъекту присуща дуалистическая природа: он обладает механизмом для непосредственного («перцептивно-сенсорного», отображающего форму вещей) восприятия объективной действительности D , а также наделен аппаратом для «интеллектуальной» обработки полученной информации путем толкования ее «содержания». Кроме того, он наделен и неким «интерфейсом» между этими двумя способами обработки информации, который в схеме представлен элементом H .

Как было показано, содержание формул (5) — (9) и диаграмм (10) — (11) невозможно понять без введения субъективного начала, которое абстрагируется из фундаментальных, имманентных свойств материи. В формулах (2) — (4) указанная констатация изначально вводится как нечто, само собой разумеющееся. При этом необходимо подчеркнуть, что несмотря на то, что информационные характеристики (свойства, связанные со сложностью объекта) проявляются (актуализируются, эксплицируются ...) в процессе взаимодействия «субъект — объект», это может произойти только тогда, когда эти свойства в «свернутом» виде скрыты в структурах как субъекта, так и объекта.

Таким образом, «субъектность» выступает не только внешней по отношению к объекту чертой, но и, по сути, внутренней способностью объекта к «отражению», его имманентным свойством. Изложенное понимание системности имеет тес-

ную связь с квантовыми принципами описания реальности, согласно которым фундаментальной характеристикой объекта является его состояние, которое в теории, то есть в процессе научного описания приобретает черты основного концептуального объекта. В этой связи приведем некоторые общенаучные соображения относительно концепции состояний системы. Это понятие, которое используется во многих естественных, социогуманитарных и технических дисциплинах, по нашему мнению, наиболее глубоко теоретически и практически разработано именно в квантовой механике, где оно является основополагающим.

Согласно канонической доктрине квантовой механики, каждая система в определенный момент времени находится в определенном состоянии. Состояние системы формализуется как решение уравнения Шредингера для данной системы. Поскольку уравнение Шредингера является определенного типа дифференциальным уравнением в частных производных, множество его решений, которые отождествляются с состояниями рассматриваемой системы, образует бесконечномерное гильбертово пространство. Следовательно, число состояний квантовомеханической системы теоретически бесконечно.

Состояние системы представляет максимально полное ее описание в теории и определяет вероятностную интерпретацию, но само оно, вообще говоря, не является непосредственно наблюдаемой величиной. Наблюдаемые величины представляются в квантовой механике эрмитовыми операторами, которые действуют в гильбертовом пространстве состояний, а возможные значения наблюдаемых величин вычисляются как матричные элементы этих операторов в пространстве состояний. Однако в некоторых других теориях состояния системы являются наблюдаемыми величинами. Скажем, в классической механике состояние материальной точки задается парой координата-импульс в определенный момент времени: $(x(t), p(t))$, которые являются наблюдаемыми — как отдельно, так и вместе. В квантовой же механике существует фундаментальное ограничение на одновременное измерение координат и импульсов, которое определяется принципом неопределенности Гейзенберга.

Следовательно, понятие и статус наблюдаемой величины неинвариантны и по-разному определяются в разных естественнонаучных (и других) теориях. Это придает определенную пикантность использованию понятия состояния в теории исчислений, которая в ее нынешнем виде явление наблюдаемости, по нашему мнению, вообще игнорирует.

Можно было бы потребовать для теории оперировать только наблюдаемыми величинами, но данный вопрос не прост. Он широко дискутировался во времена возникновения квантовой теории и не потерял актуальности в наше время. До-

стижения этой области теоретического знания на самом деле содержат настолько общие методологические уроки и установки, что они с пользой могут и должны быть усвоены любой наукой, которая имеет амбиции касательно теоретического осознания природы изучаемых ею вещей.

Первым и самым важным из них является, пожалуй, то, что для характеристики состояний объектов действительно используются как наблюдаемые, так и непосредственно не наблюдаемые величины. Причем, по убеждению большинства ученых, строить теорию только из одних наблюдаемых величин невозможно. Но, в то же время понятно, что без наблюдаемых величин никакая научная теория и вообще наука немислимы в принципе. Наблюдаемые и непосредственно не наблюдаемые величины должны иметь разный логический и онтологический статус, однако, насколько нам известно, общая теория этого вопроса до сих пор в деталях не разработана.

В свете изложенного напрашивается такая интерпретация взаимоотношений между наблюдаемыми и непосредственно не наблюдаемыми величинами теории: *они представляют, соответственно, “формальную” и “содержательную” стороны исследуемого объекта и, следовательно, могут быть формализованы как реестровая и интерпретационная части некоторой гипотетической лексикографической системы.*

В применении к объектам языка такая интерпретация может быть детализирована в том смысле, что состояние любой языковой единицы допускает разложение на формальную часть (достижимую для непосредственного восприятия субъектом — будь то звук или графическое изображение), а содержательная представляется совокупностью “всех контекстов”, в которых может функционировать данная языковая единица — это обстоятельство, собственно, и делает непосредственно не наблюдаемой указанную часть состояния

В научной дискуссии о логических и психологических основах феномена наблюдаемости стоит упомянуть о такой философской установке как принцип Маха¹⁵, согласно которому чувственные впечатления упорядочиваются в мышлении человека способом, который предусматривает максимально экономную компоновку этих впечатлений в устойчивые комплексы. Характерно, что А. Эйнштейн¹⁶, считая этот принцип слишком банальным для того, чтобы он был способен сыграть роль универсального гносеологического закона, отмечал особую роль языка в онтолого-логико-психологическом развертывании процесса познания. Языковые конструкции он считает не только способом фиксации чувственных комплексов, но и отражением того, что существует (или даже только может существовать) вне этих комплексов и

¹⁵ Принцип Маха — принцип «экономии мышления».

¹⁶ Гейзенберг В. Физика и философия. Часть и целое. Пер. с нем. — М.: Наука, 1989. — С.191–196.

без связи с ними. По нашему мнению, замечания Эйнштейна (а он был очень чувствителен к вопросам философии познания) о роли языка является неслучайными — они подчеркивают наш тезис об универсальности культурно-информационных процессов на всех уровнях реальности. Вспомним в этой связи сформулированный выше вывод, что даже сама возможность существования такого феномена как язык является следствием свойства „быть субъектом”.

Следующие замечания касаются обсуждения критерия простоты научной теории — его совсем необязательно связывать с принципом Маха. Простота научной теории приобретает для большинства исследователей эстетическую окраску — простота и красота математической схемы, подсказанной природой, имеет для них большую убеждающую силу.

Заметим, что во времена построения квантовой теории понятие простоты (и антонимического, и следовательно — родственного с ним понятия сложности) были общеязыковыми; тогда еще не была сформулирована теория сложности — она, как известно, появилась только в пятидесятых годах прошлого века. Не была также выяснена и связь такой характеристики как сложность объектов и их описаний (а следовательно — и их простота!) с информацией, и не были известны количественные меры для оценки этих величин и их связей. То, о чем говорилось в отношении разработанного А.Н. Колмогоровым и другими учеными понятия сложности, его связи с информационными аспектами описания действительности и собственно с понятием информации и ее количественной меры, имеет глубокую связь с критерием простоты и красоты научной теории. Минимальность описания изучаемого объекта, которая по А.Н. Колмогорову является объективной мерой количества информации об этом объекте, побуждает ученых (хотя бы на уровне подсознательной установки) к нахождению описаний именно такого типа, хотя и не указывает путей и не дает рецептов, поскольку, вообще говоря, относится к классу алгоритмически неразрешимых проблем. Однако отсутствие путей и рецептов не отрицает объективности существования минимального описания — оно является лишь свидетельством того, что не существует формулы или алгоритма для получения новых научных истин. И когда такое описание найдено, то оно, очевидно, должно выглядеть как самое простое — таким оно, по сути, есть. Следовательно, критерий простоты (или красоты) научной теории, по нашему убеждению, является не столько следствием принципа экономии мышления (который Эйнштейн квалифицирует как “подозрительно коммерческий” и который, собственно, имеет лишь очень опосредованное отношение к сути дела, поскольку здесь, скорее, говорится о фундаментальном информационном свойстве объективно существующих вещей, чем о черте мышления как

субъективного процесса), сколько вытекает из общей природы информации и корреспондируется с формальным определением меры ее количества по А.Н. Колмогорову.

Действительно, когда получено описание изучаемого объекта (процесса, системы и т.п.), которое наиболее адекватно соответствует его сущности, то это описание обязано быть минимальным, поскольку оно представляет только существенную информацию об исследуемом объекте и не содержит описания случайных, несущественных деталей, которые “загромождают” существенное “лишними” элементами. Ученый, так сказать, инстинктивно стремится получить именно такое описание исследуемого объекта, которое согласуется с определением информационной меры А.Н. Колмогорова, базирующейся на минимальности описания — этим, по нашему мнению, объясняется и та психологическая уверенность, которую исследователь чувствует, когда ему удается получить простую (красивую!) формулу, уравнение, вывод и т.п.

Формализм теории сложности является одновременно и прозрачным, и глубоким, его следует воспринимать онтологически, как объективное свойство вещей. Одним из нетривиальных проявлений отмеченной черты является то, что сложность составного образования, вообще говоря, не равна сумме сложностей формирующих его сущностей. Выражаясь точнее, сложность не является аддитивной функцией системы. Иными словами, если имеется некая система, состоящая из других, “меньших” подсистем, являющихся ее конституентами, т.е. если:

$$D = \bigcup_i D_i; D_i \cap D_j = \emptyset, i \neq j,$$

где символом D обозначена рассматриваемая система, а D_i — ее составляющие, то:

$$K(D) \neq \sum_i K(D_i), \quad (12)$$

где $K(D)$ — количественная мера сложности системы D , а $K(D_i)$, соответственно, — количественные меры сложности ее конституентов D_i (обычно $K(D) < \sum_i K(D_i)$). Данные представления, разумеется, распространяются и на отдельные $K(D_i)$, а также на их составляющие.

В процессе образования, функционирования и взаимодействия составных систем происходит такое явление, которое мы квалифицируем как “самокомпенсацию сложности”. Содержание этого феномена сводится к следующему. Характер взаимодействия конституентов, образующих некое единство (целостность), идентифицируемое как составной объект, является таким, что они проявляют в “связанном” состоянии лишь определенную часть их полной, “имманентной” сложности. Необходимость такого поведения можно трактовать как свойство, обеспечивающее принципиальную возможность познания “проявленного”

бытия, а может даже и его существования. В противном случае сложность любого объекта была бы актуально бесконечной (потенциально она такой и есть), а так — сложности отдельных компонентов как бы “самокомпенсируются” в процессе формирования целого. Так что можно утверждать, что потенциально сложность любой вещи бесконечна, поскольку сегодня мы не видим границ делимости материи и каждый нижестоящий структурный уровень имеет свою ненулевую сложность. Но “одномоментно” все разновидности сложности компонентов не “проявляются” в целом, они выявляются лишь “поуровнево”. Поэтому, сложность в каждом случае подлежит “перенормировке”, если идти от аналогии с квантовой электродинамикой, где для устранения расходимостей также приходится применять процедуру “вычитания бесконечностей”. Наглядный пример самокомпенсации сложности нам предоставляет язык. Так, например, мерой сложности конкретного слова можно считать длину соответствующей словарной статьи толкового словаря, где учтены эффекты грамматической и лексической семантики, в том числе, множественность грамматических значений, лексическая полисемия, фразеологическая структура лексемы и т. д. Между тем, слово в предложении (конкретном контексте) функционирует только в определенном значении — одном или своеобразной “смеси” из нескольких возможных значений для полисемических лексем и, следовательно, мера его сложности в конкретном контексте определяется лишь частью словарной статьи, причем в отдельных случаях она может составлять лишь десятки, а то и сотые доли от полной сложности лексемы. Таким образом, сложность целого предложения может оказаться меньшей, чем полная сложность отдельного слова, которое является его составной частью.

Конструкция бытия оказывается парадоксальной! Феноменологический подход подсказывает: сложные вещи в действительности состоят из еще сложнее. В этом смысле “большее” является меньшим, чем “меньшее”. Нетривиальным, на наш взгляд, подтверждением этого тезиса является известный эффект, который имеет как онтологическое, так и гносеологическое и даже психологическое измерение — он касается сложности научных теорий: теория атомов, например, не выглядит проще, чем теория молекул, теория ядер не кажется проще теории атомов, теория элементарных частиц не проще теории ядра и так далее. В лингвистике, например, теория слова (“лексикология”) также не выглядит проще теории предложения (“синтаксис”). В свете сказанного принцип редукционизма, по которому сложные вещи должны состоять из более простых, выглядит не только не очевидным, но даже сомнительным, что побуждает к ревизии основ стандартного системного анализа, предпринятой нами в данной статье. На этом уров-

не мы предлагаем учитывать эффекты, описываемые теорией сложности, так что последняя приобретает черты и статус естественно- и общенаучной, а не только чисто математической доктрины.

Выводы

Автор развивает понимание системного подхода, основываясь на рассмотрении отношений триады «структура — субстанция — субъект», которая полагается системообразующей. Именно выделение структурных, субстанциальных и субъектных свойств, а также отношений между ними сообщает анализируемой реальности свойство «быть системой». Исходя из данных предположений и общих теоретико-информационных представлений вводится понятие лексикографического эффекта в информационных системах. Проведен анализ системной триады «структура — субстанция — субъект», а также информационной сущности отношения «форма — содержание» в системной среде лексикографического эффекта. Выявлено, что «субъектность» выступает не только внешней по отношению к объекту чертой, но и его внутренней способностью к «отражению». Рассмотрены приложения понятий колмогоровской сложности к анализу свойств научной теории; отмечены условия, при которых теория сложности приобретает черты и статус естественно- и общенаучной доктрины.

Некоторые лингвистические эффекты, интерпретируемые в духе сформулированной системной доктрины, будут изложены во второй части работы.

Поступила в редколлегию 14.01.2015

УДК 658.012.011.56

Лінгвістика та системний підхід. Частина 1 / В. А. Широков // Біоніка інтелекту: наук.-техн. журнал. — 2015. — № 1 (84). — С. 3–12.

Проведено аналіз системного підходу до дослідження природної мови. Визначено поняття лексикографічного ефекту в інформаційних системах, а також проаналізовано роль суб'єкта в його структурі. Авторський системологічний підхід в комплексі з ідеями теорії складності застосовано до аналізу концептуальних підвалин наукової теорії.

Бібліогр.: 16 найм.

UDC 658.012.011.56

Linguistics and systematic approach. Part 1/ V.A. Shirokov // Bionics of intelligence: Sci. mag. — 2015. — № 1 (84). — P. 3–12.

Analysis of systematic approach to researching of language has been done. The concept of lexicographical effect in information systems has been defined and also the role of subject in structure of lexicographical effect has been analyzed. Author's systemological approach in conjunction with the ideas of complexity theory has been applied to the analysis of the basis of scientific theory.

Ref.: 16 items.