

СОЗДАНИЕ СЛОВАРЕЙ РАЗЛИЧНЫХ ТИПОВ

Л. Н. Беляева, К. В. Новикова

ПЕРЕВОДНОЙ СЛОВАРЬ В СИНЕРГЕТИЧЕСКОМ АСПЕКТЕ: ПОТЕНЦИАЛ И ПРИМЕНЕНИЕ

Синергетический подход к исследованию текста, вполне принятый в современной лингвистике и дающий интересные результаты, предполагает специальное исследование механизмов организации и самоорганизации текста. Основным условием его применения является понимание текста как сложной системы со своими компонентами, внешней организацией и структурой связей между компонентами, которая отражается на лексическом, синтаксическом и семантическом уровнях. Внешними симптомами работы синергетических механизмов, определяющих существование текста как равновесной системы, являются баланс и зависимости между лексической структурой текста, количественными параметрами отдельных лексических единиц и характеристиками энтропии текста [1].

При этом текст рассматривается как результат решения задачи передачи информации и источник (отправную точку) ее извлечения. Наши знания не только фиксируются и представляются в виде текстов, но и порождаются в языковом смысле как текст [2]. Соответственно, содержание научного текста, то есть та часть его смысла, которая является универсальной и может быть извлечена при минимальном совпадении тезаурусов автора и получателя, определяется в основном информацией об объектах, лингвистически описываемых именами – существительными и именными словосочетаниями. В ситуации перевода возможность извлечения информации из научного текста определяется для читателя корректностью передачи терминов – имен объектов. При этом адекватность понимания текста на лексическом уровне определяется насыщенностью текста именными единицами, степенью компрессии и/или развернутостью номинации объектов, точностью отражения связей между ними.

Если рассматривать это в синергетическом аспекте, то можно предположить, что при «перенасыщенности» текста перевода слишком длинными или чрезмерно свернутыми лексическими комплексами, текст разрушается и его восприятие затрудняется или даже становится невозможным даже для специалиста в соответствующей предметной области. Учет сложности восприятия и понимания подобных лексических образований, их упрощение и уточнение позволяет создавать адекватный текст перевода, то есть от хаоса нагромождения терминологических единиц перейти к новому состоянию, отражающему структуру денотатного пространства текста и терминосистему конкретного языка в том числе [3].

Сегодня можно утверждать как некоторую аксиому, что объективное исследование терминологии конкретной предметной области требует ее предварительного моделирования для того, чтобы установить систему понятий и связей между ними. При этом принято различать терминопле как системное образование плана содержания, представляющее собой организованную совокупность специальных понятий и связей, и терминосистему, т.е. совокупность языковых средств, отражающих исследуемое терминопле. Наиболее традиционным подходом к изучению соотношения терминоплей и отражающих их терминосистем является структурное моделирование на основе тезауруса, создаваемого на основе лексики, определяемой современным состоянием рассматриваемого терминопля и соответствующей ему терминосистемы. При этом терминосистема должна извлекаться из словарного и текстового материала на одном или нескольких языках с учетом постоянного развития различных научных и технических сфер деятельности человека и его отражения в языке. В современном понимании терминопле как концептуальная основа терминосистемы может рассматриваться как онтология [4].

Онтология отличается от тезауруса тем, что содержит только не зависящую от языка информацию и множество семантических отношений, кроме того, она содержит таксономические отношения. Онтология должна задавать понятия для представления значений слова в лексиконе [5] и ее можно рассматривать как пространственное представление семантической сети – сферу, в узлах внешней сетчатой поверхности которой располагаются терминологические единицы разных языков. Эти термины связаны с узлами–понятиями, располагающимися внутри (на следующем уровне) сферы. Узлы–понятия, в свою очередь, могут быть связаны с концептами, отражающими факты внешнего мира и/или нашими представлениями о них и хранящимися на более глубоких уровнях (слоях) сферы.

Синергетический подход к процессу развития терминологии позволяет утверждать, что при активном развитии предметной области в терминосистеме конкретного языка возникают точки дифуркации, которые требуют пересмотра зафиксированной терминосистемы для того, чтобы она реально отражала соответствующую систему понятий. Будем исходить из того, что лексика языка представляет собой систему, что подтверждается наличием структуры связей между ее компонентами и изменением этой структуры при введении или удалении некоторого отдельного компонента или набора компонентов. Тогда лексика языка для специальных целей представляет собой подсистему этой макросистемы, в которой заданы терминологические единицы этого языка (универбы и словосочетания) и отношения между ними, формирующие структуру связей между компонентами.

При создании переводного словаря, ориентированного на узкую предметную область, происходит слияние двух систем, в основе которого лежит единство системы номинируемых понятий, т.е. онтология. Практический смысл создания онтологии заключается в разработке систем для хранения и

поиска фрагментов знания; так, например, онтологическое понимание предмета (сущности) позволяет понять взаимозависимость определенных базовых понятий, а также связанных с ними терминов [6].

При таком подходе переводной словарь также представляет собой сложную систему, которая развивается, отражая изменения в исходных терминосистемах, и достигает точки диффракции, предшествующей его распаду и разбалансировке, то есть приводит к несоответствию задачам переводчика и/или специалиста. Такое состояние словаря характеризуется несколькими параметрами, степень реализации которых должна устанавливаться экспериментально на основе анализа комплекса словарей, оцениваемых специалистами как несоответствующие задачам перевода. Предварительно значения таких параметров можно описать следующим образом:

- отсутствие в словаре определенного, экспериментально устанавливаемого процента слов, встречающихся в текстах предметной области (не менее 20 % терминологических единиц);
- наличие в словаре или в отраслевых глоссариях составных терминов, имеющих более 1 перевода (не менее 10 % состава словосочетаний);
- наличие в словаре или в отраслевых глоссариях терминов-универбов, имеющих более 2 переводов (не менее 10 % состава терминов-универбов).

Предложенные выше параметры и их ориентировочные значения отражают нарушение унификации терминов и терминосистем, что является одним из основных направлений прикладного терминоведения, основной задачей которого и является стандартизация, упорядочение и гармонизация терминологий на различных уровнях описания и фиксации [7]. Результатом такой унификации является создание терминологических переводных словарей, ориентированных на узкие предметные области и языки для специальных целей. Такие словари, как правило, являются бинарными, хотя существуют и многоязычные, для нашей страны основную массу переводных словарей составляют те, в которых исходным языком является английский, словари с русским языком в качестве исходного создаются реже. Теоретически двуязычные переводные словари необратимы, поскольку лексические системы языков, рассматриваемые как множества лексических единиц (ЛЕ), несимметричны. Отношение симметричности работает только на множествах номенов. В случае языков для специальных целей вариативность переводов ЛЕ резко уменьшается и можно говорить о введении локальной симметричности отношения между множествами терминов двух языков. Эта локальная симметричность позволяет разрабатывать методы преобразования словарей. Особое значение при этом приобретают процедуры, позволяющие «перевернуть» словарь, в принятой сегодня терминологии – конвертировать [8].

Современное лексикографическое исследование в области создания переводных словарей (как бумажных, так и электронных) предполагает проведение предварительной терминологической работы для отбора и описания

терминологии на разных языках, осуществление гармонизации этих описаний и согласование терминологических систем разных языков. Однако, в области создания и ведения переводных словарей, представляющих собой естественную основу профессиональной компетентности переводчика, ситуация остается сложной. Дело в том, что в различных областях знаний и переводчиками, и специалистами создается огромное количество глоссариев, которые никак не соотносятся друг с другом, фиксируют разные переводы для одних и тех же терминов, а в качестве терминов используют словосочетания самой разной длины и структуры. Поэтому достижение словарем, глоссарием или системой отраслевых глоссариев точки дифуркации является показателем необходимости коренной перестройки переводного лексикографического ресурса.

Различие терминосистем разных языков, в частности, исходного языка и языка перевода определяет необходимость установления и изучения именно пар терминов вида исходный термин – переводной эквивалент, что позволяет выявить расхождения в терминопле и терминосистемах соответствующей предметной области. Установление таких расхождений и определяет дальнейшее упорядочивание, стандартизацию и унификацию терминологий разных языков, гармонизацию терминосистем этих языков, что и обеспечивает решение проблем перевода терминов и эффективность межъязыковой коммуникации [9].

Международная стандартизация в специальных областях знаний определяет требования к уточнению значений терминов, более того, их точных определений, которые могут использоваться в структуре терминов и соответствующих описаний с учетом номинируемых понятий. Поэтому сегодня в различных областях знаний и переводчиками, и специалистами создается огромное количество глоссариев, которые отражают представления переводчиков и специалистов, но часто не соотносятся друг с другом, а в качестве терминов используют словосочетания самой разной длины и структуры.

Гармонизация терминологии особенно важна для областей повышенного риска, к которым, например, относятся предметные области сейсмобезопасности и сейсмостойкого строительства. В соответствии с ситуацией в этих предметных областях переводные словари должны разрабатываться с учетом терминологии международного стандарта на проектирование конструкций в сейсмоопасной зоне Еврокод8 (Eurocode8: Design of structures for earthquake resistance). Стандарт был утвержден Европейским комитетом по стандартизации (CEN) 23 апреля 2004 года и включает основные положения и требования предыдущих стандартов.

Для решения задачи гармонизации этой терминологии рядом ведущих исследовательских и проектных институтов нашей страны были разработаны и изданы словари, глоссарии и правила.

- Проектирование и строительство. Понятийно-терминологический словарь к Еврокодам EN 1992 ÷ EN 1996, EN 1998, EN 1999. Рекомендации национального объединения строителей. М.: Р НОСТРОЙ, 2014. – 102 с.

- Англо-русский словарь по проектированию строительных конструкций (в порядке возрастания номеров Еврокодов). М.: Межгосударственная научно-техническая комиссия по стандартизации, техническому нормированию и оценке соответствия в строительстве (МНТКС), 2011. – 35 с.

- Англо-русский словарь по проектированию строительных конструкций (в алфавитном порядке). М.: Межгосударственная научно-техническая комиссия по стандартизации, техническому нормированию и оценке соответствия в строительстве (МНТКС), 2011. – 29 с.

- Свод правил. Здания сейсмостойкие и сейсмоизолированные. Правила проектирования. Издание официальное. М.: Министерство регионального развития Российской Федерации, 2014. – 51 с.

- Свод правил. Строительство в сейсмических районах. Правила проектирования. Издание официальное. М.: Министерство регионального развития Российской Федерации, 2014. – 85 с.

- Терминологический словарь для национальных нормативных документов, реализующих Еврокоды. М.: ЗАО «ЦНИИПСК им. Мельникова», 2014. – 199 с.

Рассматриваемые материалы использовались как база для создания исследовательского корпуса текстов, в котором выравнивание осуществляется по ЛЕ, рассматриваемым авторами перечисленных изданий как термины, а также по толкованиям терминов и их переводам. На основе такого выровненного корпуса могут извлекаться и исследоваться кандидаты в термины, которые при последующем анализе могут вводиться в действительно гармонизированный переводной словарь.

Для создания словарей была применена следующая процедура.

- 1) Построение параллельного корпуса на материале объединения глоссариев и терминологических словарей. В основе такого объединения лежат базовые словари, разработанные Межгосударственной научно-технической комиссией по стандартизации, техническому нормированию и оценке соответствия в строительстве (МНТКС), эти словари были дополнены информацией из всех остальных словарей (табл. 1).

- 2) Составление единого словарного массива (индекса) на основе сопоставления английских и русских лексических единиц (табл. 2), выявление кандидатов в английские термины на основе использования программы выделения простых именных групп.

- 3) Получение русско-английского словаря на основе конвертации англо-русского словаря, дополненного кандидатами в термины, дополнение словаря кандидатами в термины – простыми именными группами.

- 4) Работа с экспертами и уточнение списка исходных терминов и их переводов, фиксирование отношений синонимии и нормализация исходных терминов и переводов.

Англо-русский объединенный выровненный словарь (фрагмент)

Термин на английском языке	Номер Еврокода, обозначение	Определение термина на английском языке	Термин на русском языке	Определение термина на русском языке	Источник
adhesion:	EN 1996-1-1	the effect of mortar developing a tensile and shear resistance at the contact surface of masonry units	адгезия: адгезионная прочность сцепления кладки (адгезия)	Сопротивление растяжению или срезу (сдвигу), возникающее на контактной поверхности между раствором и элементами кладки Сопротивление растяжению или срезу (сдвигу), возникающее между раствором и контактной поверхностью кладки из камней или блоков.	Р НОСТРОЙ МНТКС
aerodynamic device:	EN 1993-3-2	A device fitted to the chimney to reduce vortex excitation without increasing the structural damping	аэродинамическое устройство:	Устройство, которое устанавливается на дымовую трубу для уменьшения вихревого возбуждения без увеличения конструкционного демпфирования.	Р НОСТРОЙ
age at loading	EN 1994-1-1, t_0		момент нагружения	Момент нагружения	МНТКС
age of concrete at the time of loading	EN 1992-2, t_0		возраст бетона на момент нагружения	Возраст бетона в момент нагружения	МНТКС
	EN 1992-1-1, t_0		возраст бетона в момент приложения нагрузки	Возраст бетона в момент приложения нагрузки	МНТКС
along-wind vibrations:	EN 1993-3-2		колебания вдоль направления ветрового воздействия:	Продольные колебания дымовой трубы, вызванные эффектом галопирования	Р НОСТРОЙ

Англо-русский словарь (фрагмент)

acceleration of gravity	ускорение свободного падения
	ускорение силы тяжести
accessory masonry unit	дополнительный элемент кладки
accidence	превышение
accidental eccentric	максимально разрешенные случайные эксцентриситеты
accidental eccentricity	случайный эксцентриситет
accidental eccentricity of the mass of one storey from its nominal location	случайный эксцентриситет массы одного этажа от его номинального положения
accidental eccentrics	случайный эксцентриситет
accidental mass eccentricity	случайный эксцентриситет масс
acting shear force	расчетная горизонтальная сила
	действующее усилие сдвига
action	воздействие
	действие
	воздействие, вызывающее изменение напряженно-деформированного состояния конструкции
action effect	эффект расчетного воздействия
	эффект воздействия
	результат расчетного воздействия
action in the seismic design situation	воздействие в расчетной сейсмической ситуации

При анализе словарей установлено большое число случаев, когда в перевод включается экстралингвистическая информация, известная специалисту в конкретной области, но отсутствующая в структуре исходного термина. Так, например, словосочетание *clearance height from roadway surfacing* переводится как *габаритная высота от поверхности дорожного полотна до нижней кромки конструкции моста*. При этом перевод термина *roadway surfacing* дополняется уточнением *до нижней кромки конструкции моста*, что связано с особенностями строительства именно мостовых и/или арочных строений. Для английского словосочетания это является имплицитной информацией, не требующей экспликации. Подобная словарная экспликация терминов, взятых вне контекста, может оказаться некорректной. Так, например, перевод терминологического словосочетания *coefficient of earth pressure at rest*, уточненный как *коэффициент бокового давления грунта в состоянии покоя*, верен только для случая описания объекта, не полностью погруженного в грунт, во всех иных случаях это давление не будет боковым и, следовательно, перевод термина не будет адекватным.

Еще более сложный случай возникает в случае не перевода, а передачи коллокации, описывающей конкретное важное явление или сложный объект, см., например следующее соотношение исходного описания и его перевода на русский язык: *correlation factors to derive the pile resistance from ground investigation results, not being pile load tests* – *неправочные коэффициенты для оценки результатов испытаний свай статической нагрузкой*. При этом в рассматриваемых глоссариях не зафиксированы терминологические слово-

сочетания *pile resistance, ground investigation, pile load, pile load test*, введение которых диктуется особенностями их перевода. Предлагаемый в глоссарии перевод может объясняться тем, что подобные грунтовые испытания свай обычно проводятся именно под статической нагрузкой. Аналогично ЛЕ *diameter of confined core in a circular column* переводится как *диаметр бетонного сечения, замкнутого поперечной арматурой в круглой колонне*.

Количественно такие термины составляют не менее 10 % от объема словаря, что характеризует недостаточную лингвистическую компетентность их составителей. В то же время подобные глоссарии представляют собой важный источник для разработки соответствующего англо-русского переводного словаря, поскольку фиксируют не только современную терминологию, но и экспертное знание.

Исходный вариант русско-английского словаря автоматически сгенерирован из дополненного и отредатированного экспертами англо-русского словаря с использованием основных средств операционных систем, в частности, средств работы с таблицами. Полученный словарь (индекс) представляет собой таблицу русских терминов и их переводов, отсортированных в алфавитном порядке. Процесс редактирования состоял в изменении, удалении или добавлении заглавных единиц, соотнесении и переупорядочении их связей с английскими эквивалентами. При этом (как и в случае работы с англо-русским вариантом) модифицировались конструкции, представляющие собой дефиниции, толкования, описания ситуации, и устанавливались термины, из которых они состоят, определялся весь комплекс переводов конкретных терминов и выявлялась синонимия переводов или многозначность исходного термина (см. табл. 3).

Т а б л и ц а 3

Русско-английский словарь (фрагмент)

Термин на русском языке	Термин на английском языке
поверхность подпорного сооружения	retained earth surface
поворот в пластическом шарнире	plastic hinge rotation
повышенная температура	elevate temperature
повышенная температура	elevated temperature
погружаемость	driveability
погружение	driving
податливые элементы	ductile members
податливый бетонный элемент	ductile concrete member
податливый элемент	ductile member
подкос	strut
подкрепленная пластина	stiffened plate
подкрепляющая стена	stiffening wall
подошва	footing
подошва фундамента	footing
подпорная конструкция или элемент	propped structure or member
подпорная стена	retaining wall
подпорное сооружение	retaining works
поверхность подпорного сооружения	retained earth surface
поворот в пластическом шарнире	plastic hinge rotation
повышенная температура	elevated temperature
погружаемость	driveability
погружение	driving

Полученный комплекс словарей исследовался с точки зрения установленной степени многозначности словарных статей (см. табл. 4). Из таблицы следует, что уровень многозначности выше для англо-русского варианта, который построен на базе исходного объединенного словаря с применением процедуры добавления простых именных групп, выделяемых из многокомпонентных словосочетаний. Русско-английский словарь, полученный в результате конвертации, содержит меньше многозначных многокомпонентных словосочетаний. Результаты показывают, что англо-русский словарь требует дополнительного анализа и обработки.

Т а б л и ц а 4

Количественные характеристики исследуемого комплекса отраслевых словарей

Словарь	Объем словаря	Количество терминов-универбов	Количество многозначных универбов	Количество словосочетаний	Количество многозначных словосочетаний
Англо-русский	2830	382	88 (23,3%)	2448	290 (11,4%)
Русско-английский	3105	304	56 (18,4%)	2801	90 (3,21%)

Полученные таким образом переводные словари должны вручную тщательно редактироваться, но объем этой работы не сопоставим с трудоемкостью создания переводного словаря вручную. При этом следует учитывать, что даже самая изощренная система извлечения терминов не дает окончательного варианта переводного словаря, а предоставляет лишь удобно организованный и оперативно получаемый ресурс для работы терминолога или лексикографа.

Моделирование динамических систем, к которым можно отнести и сам язык, и речемыслительную деятельность в целом, и терминологические системы языков для специальных целей, представляет собой сложную задачу, требующую решения задач категоризации, а также применения качественных и количественных (вероятностных) оценок создаваемых моделей. Кроме того, моделирование требует проверки адекватности разработанных моделей при условии использования при их создании субъективного опыта исследователя.

ЛИТЕРАТУРА

1. *Пиотровский, Р. Г.* Лингвистическая синергетика: исходные положения, первые результаты, перспективы / Р. Г. Пиотровский – СПб : Филологический ф-т СПбГУ, 2006. – 160 с.
2. *Чернявская, В. Е.* Лингвистика текста. Лингвистика дискурса / В. Е. Чернявская. – М. : УРСС-Либроком, 2013. – 224 с.
3. *Беляева, Л. Н.* Синергетика текста и адекватность перевода терминологии / Л. Н. Беляева, О. А. Бородина // Синергетическая лингвистика vs лингвистическая синергетика : матер. Междунар. науч.-практич. конф. – Пермь : Изд-во Пермского гос. технич. ун-та, 2010. – С. 119–126.

4. *Cimiano, P.* Linked Terminologies: Applying Linked Data Principles to Terminological Resources / P. Cimiano [etc.] // Proc. of the fourth biennial conference on electronic lexicography, eLex 2015 : Linking lexical data in the digital age. Sussex, United Kingdom, from 11–13 Aug. 2015. – United Kingdom, 2015. – P. 504–517
5. *Biebow, B.* TERMINAE : a method and a tool to build a domain ontology / B. Biebow, S. Szulman // Benjamins, V.R., Fensel, D. and Gomez-Perez, A. (eds) Proceedings of International Workshop on Ontological Engineering on the Global Information Infrastructure – Dagstuhl Castle, Germany, 1999. – P. 25–30.
6. *Gillam, L.* Terminology and the construction of ontology / L. Gillam, M. Tariq, K. Ahmad // Application-Driven Terminology Engineering. Edited by Fidelia Ibekwe-SanJuan, Anne Condamines and M. Teresa Cabré Castellví. – Terminology, 2005. – P. 55–81.
7. *Лейчик, В. М.* Прикладное терминоведение и его направления / В. М. Лейчик // Прикладное языкознание: учебник; под. ред. А. С. Герда. – СПб. : Изд-во С.-Петерб. ун-та, 1996. – С. 276–286.
8. *Egorova, K.* Editing an automatically-generated index with K Index Editing Tool / K. Egorova // Proc. of the fourth biennial conference on electronic lexicography, eLex 2015 : Linking lexical data in the digital age. Sussex, United Kingdom, from 11–13 Aug. 2015. – United Kingdom, 2015. – P. 268–280.
9. *Беляева, Л. Н.* Лексикографический потенциал современных лингвистических технологий : монография / Беляева Л. Н. [и др.]. – СПб : ООО «Книжный дом», 2014. – 168 с.