

81.2 Ан

К 181

МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО ОБРАЗОВАНИЯ БССР
МИНСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ ИНСТИТУТ
ИНОСТРАННЫХ ЯЗЫКОВ

На правах рукописи

КАМЕНЕВА Зинаида Михайловна

О МЕЖЪЯЗЫКОВОЙ ИДИОМАТИЧНОСТИ /НА МАТЕРИАЛЕ ФОРМАЛЬ-
НОГО АНАЛИЗА ПЕРЕВОДОВ АНГЛИЙСКИХ СОСТАВНЫХ ТЕРМИНОВ/

10.02.04 - германские языки

/Диссертация написана на русском языке/

А в т о р е ф е р а т
диссертации на соискание учёной степени
кандидата филологических наук

Минск 1973

МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО ОБРАЗОВАНИЯ БССР
МИНСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ ИНСТИТУТ
ИНОСТРАННЫХ ЯЗЫКОВ

На правах рукописи

КАМЕНЕВА Зинаида Михайловна

О МЕЖЪЯЗЫКОВОЙ ИДИОМАТИЧНОСТИ /НА МАТЕРИАЛЕ ФОРМАЛЬНОГО АНАЛИЗА ПЕРЕВОДОВ АНГЛИЙСКИХ СОСТАВНЫХ ТЕРМИНОВ/

10.02.04 - германские языки

/Диссертация написана на русском языке/

А в т о р е ф е р а т
диссертации на соискание учёной степени
кандидата филологических наук

Минск 1973

7612* 1821

БІБЛІЯТЭКА
М Д Л У

Работа выполнена на кафедре общего языкознания Минского государственного педагогического института иностранных языков.

Научный руководитель - доктор филологических наук,
профессор Р.Г. ПИОТРОВСКИЙ.

Научный консультант - кандидат филологических наук
А.Д. БОРИСЕВИЧ.

Официальные оппоненты :

1. Доктор филологических наук, профессор И.Б. ХЛЕБНИКОВА.
2. Кандидат филологических наук, доцент А.Н. ШАРАНДА.

Ведущее научное учреждение - кафедра иностранных языков Ленинградской орденов Ленина и Ушакова военно-морской академии.

Автореферат разослан "20" *ноября*..... 1973 г.

Защита кандидатской диссертации состоится "3" *авг.*
авг. 1974 г. на заседании Совета по присуждению учёных степеней Минского государственного педагогического института иностранных языков /Минск, ул. Захарова, 21/.

С диссертацией можно ознакомиться в библиотеке института.

УЧЁНЫЙ СЕКРЕТАРЬ СОВЕТА

Естественный язык полнее и лучше всего может быть описан при его конфронтации с другой семиотической системой. Примером такой конфронтации может служить перевод с одного языка на другой, причем выходной язык выступает в роли метаязыка.

При осуществлении перевода наиболее сложной задачей является распознавание актуального /текстового/ значения входного лингвистического знака, в результате чего достигается устранение его асимметрии. Если распознавание осуществляется с помощью средств выходного языка, то здесь следует считаться и с неконгруэнтностью¹ /неизоморфностью, взаимной идиоматичностью/ знаков входного и выходного языков, являющейся результатом несоответствия смысловых объемов и значимостей² единиц конфронтруемых языков.

Реферлируемая диссертация посвящена исследованию межъязыковой /англо-русской/ идиоматичности — одной из сложных и актуальных проблем современной науки о языке.

У НАСТОЯЩЕЙ РАБОТЫ ДВЕ ЗАДАЧИ : во-первых, исследование лексико-грамматической межъязыковой идиоматичности на материале терминологической фразеологии подязыка судовых механизмов ; во-вторых, построение системы иконического и алгоритмического устранения этой неконгруэнтности. Обе эти задачи решаются в русле теории детерминистского распознавания смыслового образа³.

¹ Н. Д. Андреев. Статистико-комбинаторные методы в теоретическом и прикладном языковедении. Л., 1967, стр. 151.

² Ф. де Соссюр. Курс общей лингвистики. Перевод с французского. М., 1933; стр. 109 и сл.

³ I. Palibina, R. Piotrowski. Automatic Pattern Recognition Applied to Semantic Problems. 1973 Congress of the Computational Linguistics. Pisa, 1973 (Preprint).

Исследование проводилось по следующим этапам :

1. Подготовительный отбор двужанчной идиоматики для словаря, кодирование и перфорация отобранных оборотов ;
2. Введение списка оборотов в ЭВМ для осуществления иконического перевода ;
3. Классификация именных оборотов терминологического значения /составных терминов/ с точки зрения их глубинных структур, в качестве которых рассматриваются структуры русских эквивалентов ;
4. Статистическое упорядочение составных терминов и их глубинных структур с целью выявления наиболее типичных ситуаций;
5. Алгоритмизация наиболее типичных схем с целью дальнейшей реализации этих схем на ЭВМ.

Во введении обосновывается выбор единицы исследования, "единицы перевода". В качестве таковой взято словосочетание, в частности, терминологическое сочетание, являющееся основным носителем информации научно-технического текста¹.

Идиоматичность в теоретической и прикладной лингвистике рассматривается в первой главе диссертации. Если проблема внутриязыковой идиоматичности глубоко и интенсивно исследовалась уже давно /Н.Н.Амосова, О.С. Ахманова, А.М. Бабкин, В.В.Виноградов, А.В. Куния, А.П.Мордвилко, А.И. Молотков, В.Н. Телия, Н.М. Шанский и др./, то проблема межъязыковой идиоматичности стала изу-

¹ Ср. Г.Г.Белоголов. Об использовании терминологических сочетаний в автоматизированных ИПС. В кн.: "Цифровая вычислительная техника и программирование". М., 1967, стр. 172 ; О.Кадэ. Zufall und Gesetzmäßigkeit in der Übersetzung. Leipzig, 1960, S.90 ; S.Perschke. Machine Translation. The Second Phase of Development, "Endeavour", vol. XXVII, N 101, 1968, p.98.

чаться сравнительно недавно /И.Бар-Хиллел, М.м. Копыленко, И.А.Мельчук, М.А. Деклер, Р. Якобсон, У. Вайнрайх и др./.

В условиях автоматической переработки текста при двуязычной ситуации¹ идиоматичность рассматривается с точки зрения возможности или невозможности перевода данного словосочетания с помощью автоматического словаря /АС/ словоформ и работающих вместе с этим словарем лексико-грамматических алгоритмов. Сопоставительный анализ входных терминологических словосочетаний и их русских соответствий позволил сделать вывод о том, что проблема межъязыковой идиоматичности не может быть полностью разрешена с помощью АС словоформ и лексико-грамматических алгоритмов, т.к. АС словоформ дает возможность осуществить только построчный лексический перевод. Такой перевод недостаточно эффективен, поскольку в нем потеряна значительная часть информации.

Первый шаг для улучшения качества лексического перевода состоит в том, чтобы перейти от построчного к "пооборотному" переводу. В связи с этим используется следующее определение машинного оборота: машинным оборотом является словосочетание, которое невозможно перевести путем замены его компонентов по соответствующим двуязычным АС или с помощью работающих лексико-грамматических алгоритмов.

Именные терминологические словосочетания можно разделить

¹ Двуязычная ситуация представляет собой суперструктуру, возникающую при сопоставлении структур и элементов двух языков в процессе перевода, см.: Р.Г.Пиотровский, в.А.Чихачовский. О двуязычной ситуации. В сб.: "Статистика речи и автоматическая переработка текста". Л., 1974 г.

с точки зрения всего сказанного на две группы : свободные терминологические словосочетания и терминологические словосочетания фразеологического характера / терминологические обороты/. Неконгруэнтность первых устраняется при помощи алгоритма пословного перевода. Неконгруэнтность терминологических словосочетаний второго типа устраняется с помощью предварительно введенного в ЭВМ двуязычного автоматического словаря оборотов /АСО/ исследуемого подъязыка.

Методика составления АСО описывается во второй главе.

Методика составления АСО и его работа в рамках бинарного вероятностного перевода часто обсуждалась в научной литературе ¹, поэтому мы ограничимся лишь указанием на общие принципы, которым мы следовали при построении АСО для перевода английских текстов по судебным механизмам.

Исходным материалом при построении АСО послужили 20502 трехсловных именных сегмента, извлеченных с помощью ЭВМ из английских текстов указанной тематики², а также 17195 именных сегментов нефиксированной длины, автоматически полученных из тех же текстов с помощью специальной программы, разработанной в группе "Статистика речи"³.

¹ См. В.А.Вертель, Е.В.Вертель, В.С.Крисевич, Р.Г.Пиотровский, Л.И.Трибис. Автоматические словари в системе вероятностного МП. Сб.: "Инженерная лингвистика". Л., 1971; А.Д.Борисевич. Англо-русский автоматический словарь оборотов. /К проблеме идиоматичности при обращении текстов в системе "человек-машина-человек"/ АКД. Минск, 1972 г.

² К.Ф.Лукьяненко. Лексико-статистическое описание английского научно-технического текста с помощью электронно-вычислительной машины /подъязык судебных механизмов/. Приложение канд. дисс., Минск, 1969 г.

³ М.В.Данейко, В.М.Петровская. Алгоритм автоматической сегментации английского предложения. В сб.: "Статистика текста", т. II., АПТ, Минск, 1970г., стр.215-290.

Эти сегменты с помощью приемов напластования и отсеечения лиш-
них словоформ преобразовывались в машинные обороты.

В АСО отбирались в основном машинные обороты терминологи-
ческого значения. Русские эквиваленты этих оборотов образуют
выходной словник АСО, который состоит из словоформ и словосо-
четаний, добытых путем статистико-дистрибутивного исследования
аналогичных по тематике научно-технических текстов из русских
периодических изданий /50.000 словоупотреблений/, а также взя-
тых из двуязычных словарей.

При построении словаря оборотов следует учитывать то, что
он стыкуется со словарем словоформ, а это значит, что следует
принимать во внимание те ограничения, которые накладывает на
словарь оборотов алгоритм работы АС. Ввиду того, что алгоритм
работы АС построен с таким расчетом, чтобы соответствовать ос-
новному требованию работы с ЭВМ- экономии машинной памяти, то
ограничение, которое положено в основу составления АСО, каса-
ется, в первую очередь, количественного представления входных
и выходных оборотов в машине. Каждый входной оборот строится
не более чем из шести словоформ, а выходной эквивалент-не более
чем из шестнадцати русских словоформ. Всего АСО включает 6600
машинных оборотов.

Нерегулярная переводимость отобранных оборотов обусловле-
на морфологической, синтаксической и семантической неоднознач-
ностью компонентов оборота, что и приводит соответственно к по-
явлению морфологических, синтаксических и семантических несоот-
ветствий и их комбинаций. Наиболее сложными типами несоответ-
ствий являются семантико-синтаксические и семантико-синтаксико-

морфологические несоответствия. Идиоматичность оборотов этого типа разрешается иконическим способом – включением в АСО.

Перевод сегментов текста с помощью АСО является примером решения трудных вопросов МП с позиции грубой силы. Идя только по этому пути нельзя добиться качественного перевода уже хотя бы потому, что число возможных входных сегментов во много раз превышает потенции запоминающих устройств ЭВМ. В том случае, когда мы получаем возможность использовать более мощные ЭВМ, а также на более продвинутой стадии автоматизации лингвистических исследований появляется возможность разрабатывать программы и подпрограммы алгоритмического устранения неконгруэнтности.

Методика алгоритмического разрешения межъязыковой идиоматичности рассматривается в третьей главе.

Условием составления работающего алгоритма является всестороннее исследование структур именных терминологических словосочетаний. С этой целью проводится типизация входных сегментов, как с точки зрения их собственной поверхностной структуры, так и с точки зрения лежащих в их основе глубинных структур. Обсуждаемая в лингвистической литературе последних десяти лет проблема глубинных структур не имеет ясного и недвусмысленного определения¹. Поэтому мы пользуемся рабочим подходом, подсказанным двуязычным направлением нашего исследования,

¹ Г.Б.Колшанский. О понятии глубинной и поверхностной структуры в языке. Тезисы научной конференции "Глубинные и поверхностные структуры в языке". М., 1972, стр.25-27; там же Ю.С. Мартынялов. Глубинные структуры в описании языковой семантики, стр.28-35; там же Н.Д.Арутюнова. Глубинная структура и недвусмысленность предложения, стр.3-7.

согласно которому глубинными структурами английских именных групп мы считаем поверхностные отрутуры их русских переводных эквивалентов.

Проводится классификация именных терминологических словосочетаний с точки зрения их поверхностной структуры. В качестве критерия берется, с одной стороны, количество словоформ, образующих словосочетание, а с другой – их принадлежность к тому или иному грамматическому классу¹. Морфологический анализ словосочетаний является предварительным этапом для установления структуры словосочетания, которое выступает в функции определенной синтаксической единицы². В исследуемом материале выделяется пять типов словосочетаний с точки зрения поверхностной структуры, см. табл.1.

Таблица 1

| Типы словосочетаний | Английские словосочетания | | Русские эквиваленты | |
|---------------------|---------------------------|-------|---------------------|-------|
| | Количество | в % | Количество | в % |
| однословные | — | — | 419 | 6 |
| двухсловные | 3728 | 57 | 2468 | 37 |
| трехсловные | 2200 | 33,3 | 1916 | 30 |
| четырёхсловные | 526 | 8 | 1034 | 16 |
| пятисловные | 92 | 1,4 | 418 | 6 |
| шестисловные | 24 | 0,3 | 345 | 5 |
| Итого: | 6600 | 100 % | 6600 | 100 % |

¹ Мы ориентируемся на ту классификацию, которая используется в группе "Статистика речи", см. А.Д.Борисевич, В.В.Гончаренко, Э.М.Добрускина, Б.К.Кочеткова, В.С.Крисевич, О.А.Нехай, Р.Г.Пиотровский, Т.А.Штирбу, С.В.Ястребова. Кодирование грамматической информации в машинном словаре. "Статистика текста", об. статей, том II, Минск, 1970, стр. 331-614.

² Д.Варга. Построение новой анализирующей системы предложения. ИТИ, 1968, № 4, сер. 2, стр. 17-23.

По частоте употребления доминантной является модель двух-словных именных терминологических сочетаний. Последующие типы именных словосочетаний располагаются в порядке убывающей частотности /см. табл. 1/. Установлено, что с увеличением количества определяющих компонентов в препозиции удельный вес словосочетаний уменьшается.

В каждом типе именных терминологических словосочетаний отмечен ведущий структурный вариант, который имеет наибольшую частотность среди вариантов данного типа словосочетаний. Таким оказался структурный вариант, составные компоненты которого выражены существительным.

Количество структурных вариантов именных терминологических словосочетаний возрастает с увеличением числа их зависимых компонентов. Выявлено 37 структурных вариантов, которые представляют собой семантически связанные, грамматически оформленные именные группы. Эти варианты организованы в частотный список /ЧС/ входных структур, верхняя часть которого представлена в таблице 2.

Таблица 2

| | Структурный вариант | f |
|----|---------------------|------|
| 1 | 2 | 3 |
| 1. | $N_1 + N_2$ | 3105 |
| 2. | $N_1 + N_2 + N_3$ | 1552 |
| 3. | $A_1 + N_2 + N_3$ | 506 |
| 4. | $A_1 + N_2$ | 290 |

¹ Условные обозначения даны на 13 стр.

| 1 | 2 | 3 |
|-----|-------------------------------------|-----|
| 5. | $N_1 + N_2 + N_3 + N_4$ | 228 |
| 6. | $P_1 + N_2$ | 185 |
| 7. | $P_2 + N_2$ | 148 |
| 8. | $P_1 + N_2 + N_3$ | 104 |
| 9. | $P_2 + N_2 + N_3$ | 100 |
| 10. | $A_1 + N_2 + N_3 + N_4$ | 87 |
| 11. | $N_1 + N_2 + P_1 + N_4$ | 39 |
| 12. | Num+ $N_2 + N_3$ | 36 |
| 13. | $P_1 + N_2 + N_3 + N_4$ | 33 |
| 14. | $A_1 + N_2 + A_3 + N_4$ | 30 |
| 15. | $A_1 + N_2 + N_3 + N_4 + N_5$ | 26 |
| 16. | $N_1 + N_2 + N_3 + N_4 + N_5$ | 22 |
| 17. | Num+ $N_2 + N_3 + N_4$ | 20 |
| 18. | $N_1 + A_2 + N_3 + N_4$ | 20 |
| 19. | $A_1 + N_2 + P_2 + N_4$ | 17 |
| 20. | $P_2 + N_2 + N_3 + N_4$ | 16 |
| 21. | $A_1 + A_2 + N_3 + N_4$ | 16 |
| 22. | $N_1 + P_2 + N_3 + N_4$ | 14 |
| 23. | Num+ $N_1 + N_2 + N_3 + N_4$ | 10 |
| 24. | $N_1 + N_2 + P_2 + N_4$ | 10 |
| 25. | $A_1 + N_2 + A_3 + N_4 + N_5$ | 10 |
| 26. | $N_1 + N_2 + A_3 + A_4 + N_5 + N_6$ | 8 |
| 27. | $P_2 + N_2 + N_3 + N_4 + N_5 + N_6$ | 6 |
| 28. | $N_1 + A_2 + N_3$ | 6 |
| 29. | $A_1 + N_2 + P_1 + N_4 + N_5$ | 4 |
| 30. | $A_1 + A_2 + N_3 + N_4 + N_5$ | 4 |

| 1 | 2 | 3 |
|-----|-------------------------------------|---|
| 31. | $A_1 + N_2 + N_3 + N_4 + A_5 + N_6$ | 4 |
| 32. | $N_1 + P_2 + N_3 + N_4 + N_5$ | 4 |
| 33. | $Num + N_2 + P_1 + N_4 + N_5$ | 4 |
| 34. | $A_1 + P_2 + N_3 + N_4 + N_5$ | 2 |
| 35. | $D_1 + P_2 + N_3 + N_4$ | 2 |
| 36. | $N_1 + D_2 + P_2 + N_4$ | 2 |
| 37. | $N_1 + N_2 + A_3 + N_4$ | 2 |

Проведена типизация выходных структур для доминантных входных структурных вариантов оловосочетаний. Результатом типизации являются установленные по структурно-семантическому принципу формулы англо-русских соответствий.

Все формулы соответствий расклассифицированы следующим образом : например 3.4.5 - цифра 3 указывает тип словосочетания на входе / количество компонентов/, цифра 4 - тип эквивалентного оловосочетания на выходе / количество компонентов/, 5 - порядковый номер подтипа соответствия на выходе.

Понятие составляющих словосочетания дополняется понятием синтаксического подчинения для формализации способов передачи информации "о степени близости подчиненных подчиняющимся". Такая информация вполне удобно передается при помощи "стрелок"¹.

¹ И.А. Мельчук. Автоматический синтаксический анализ, т. I, Новосибирск, 1964, стр. 18 ; Л.Н. Иорданская. Автоматический синтаксический анализ, т. II, Новосибирск, 1967, стр. 18-40 ; С.Н. Фиталов. О моделировании синтаксиса в структурной лингвистике. В кн.: "Проблемы структурной лингвистики". М., 1962, стр. 100-114 ; ср. также использование стрелочной записи в работах : Э.Ф. Скороходько. Вопросы перевода английской технической литературы. Киев, 1963 ; К.Н. Качалова. Грамматика английского языка. М., 1964, стр. 17.

Формулы соответствий рассматриваются последовательно по типам словосочетаний, согласно таблице I. Формулы соответствия определяются семантикой компонентов словосочетания.

Доминантным структурным вариантом двухсловных именных сочетаний является схема $N_1 + N_2$ на входе, которая на выходе выражена различными эквивалентами.

2.1.0.

Схема I.

| В | Х | О | Д | В | Ы | Х | О | Д |
|---|---|---|---|---|---|-------|-------|---|
| | | | | | | N_1 | | |
| | | | | | | | N_m | |
| | | | | | | N_n | | |
| | | | | | | | | |

2.1.1. $N_1 + N_2 \rightarrow N_1$

Ядро входного словосочетания представляет собой существительное обобщенного значения типа **room, set, system, unit**, которое в данном словосочетании не переводится на русский язык. Таким образом, в выходном сегменте происходит перемещение семантического центра с ядра на определяющий компонент, принадлежащий к лексико-семантической группе "устройство", который и актуализирует значение эквивалента.

Ср. \downarrow blower unit 'вентилятор'.

I а) для английского языка :

A - прилагательное ; N - существительное ;
 A_1, A_2, A_3, A_4 ; A_5 - определяющие компоненты оборотов ;
 N_6 - ядро оборота ; P_1 - причастие I /герундий ;
 P_2 - причастие II.

б) для русского языка :

A_x - эквивалент прилагательного,
 A_x - прилагательное, отсутствующее в английском языке,
 A_1, A_2, A_3, A_4, A_5 - определяющие компоненты оборота, выраженные прилагательным,
 N_6 - эквивалент существительного,
 N_1, N_2, N_3, N_4, N_5 эквиваленты определяющих компонентов оборота,
 N_6 - эквивалент ядра оборота,
 N_x - существительное, отсутствующее в английском обороте,
 x - предлог, C - сложное слово.

$$2.1.2. \quad N_1 + N_2 \rightarrow N'_m$$

Компоненты входного сочетания имеют семантические характеристики "вещество-устройство", в то время как его эквивалент, выводимый из суммы значений входных компонентов, характеризуется семантикой "устройство".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{heat exchanger} \end{array}$ 'теплообменник'.

$$2.1.3. \quad N_1 + N_2 \rightarrow N'_n$$

Компоненты входного словосочетания характеризуются семантикой "вещество-устройство", на выходе же эквивалент характеризуется семантической характеристикой "устройство".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{air heater} \end{array}$ 'калорифер'

2.2.0.

Схема 2

| В х о д | В ы х о д |
|-------------|---|
| $N_1 + N_2$ | $1) \begin{array}{l} N'_2 + N'_1 \quad 1) \\ N'_2 + N'_1 \quad 2) \end{array}$ $2) A'_1 + N'_2$ |

$$2.2.1. \quad N_1 + N_2 \rightarrow N'_2 + N'_1 \quad 1)$$

а) В одном случае, определяющий компонент входного словосочетания характеризуется семантикой "деталь", а ядро - семантикой "часть от целого", выходные эквиваленты же характеризуются отношением "часть от целого" - "деталь".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{blade root} \end{array}$ 'корень лопатки'.

1 - эквивалент существительного в родительном падеже.

2 - эквивалент существительного в творительном падеже.

б) в другом случае, определяющий компонент характеризует-ся семантикой "назначение", а ядро - "устройство", а выходные эквиваленты характеризуются семантикой "устройство-назначение".

Ср. $\overbrace{\text{feed heater}}^{\text{'коллектор питания'}}$

$$2.2.2. N_1 + N_2 \rightarrow N_2' + N_1' \quad **$$

Компоненты входного словосочетания характеризуются семантикой "устройство-процесс", а выходной эквивалент - "процесс-устройство".

Ср. $\overbrace{\text{motor control}}^{\text{'управление электродвигателем'}}$

$$2.2.3. N_1 + N_2 \rightarrow A_1' + N_2'$$

Определяющий компонент характеризуется семантикой "назначение", а ядро - "устройство", выходные эквиваленты передаются соотношением "назначение-устройство".

Ср. $\overbrace{\text{sea chest}}^{\text{'кингстонная коробка'}}$

$$2.3.1. N_1 + N_2 \rightarrow N_2' + N_1' + N_x'$$

В тех случаях, когда определяющий компонент выражен сложным существительным с семантикой "устройство-процесс", "деталь-часть", а ядро выражает "процесс", выходные эквиваленты характеризуются следующим соотношением семантических характеристик "процесс-процесс-устройство".

Ср. $\overbrace{\text{boiler-blowdown procedure}}^{\text{'процедура продувания котлов'}}$

$$2.3.2. N_1 + N_2 \rightarrow N_2' + A_x' + N_1'$$

Определяющий компонент выражен сложным существительным с семантикой "вещество-деталь", а ядро - "локальность", в то вре-

мя как выходные эквиваленты характеризуются семантикой "локальность-вещество-деталь".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{smoke-box} \quad \text{end} \end{array}$ 'дно дымовой коробки'.

2.3.3. $N_1 + N_2 \rightarrow A_2' + N_x' + N_1'$

Когда определяющий компонент выражает "процесс", а ядро "деталь", выходные эквиваленты характеризуются семантическими характеристиками "назначение-деталь-процесс".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{engine camshaft} \end{array}$ 'кулачковый вал двигателя'.

2.3.4. $N_1 + N_2 \rightarrow N_2' + x + N_1'$

а) "деталь-процесс" "процесс-деталь".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{bearing pressure} \end{array}$ 'давление в подшипнике',

б) "процесс-результат действия" "результат действия - процесс".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{compression ignition} \end{array}$ 'возпламенение от сжатия'.

в) "деталь-устройство" "устройство-деталь".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{gear pump} \end{array}$ 'насос с редуктором'.

г) "локальность-физическая характеристика" "физическая характеристика-локальность".

Ср. $\begin{array}{c} \swarrow \quad \searrow \\ \text{inlet temperature} \end{array}$ 'температура на входе'.

Подобные несоответствия обусловлены различной сочетаемостью единиц во входном и выходном языках.

Трехсловные именные сочетания. Комбинаторные возможности трехсловных сочетаний как на уровне структуры, так и на уровне семантики гораздо выше, чем у двухсловных.

Наиболее типичными для трехсловных советаний являются следующие схемы несоответствий :

$$3.1.1. \quad N_1 + N_2 + N_3 \rightarrow N'$$

Ср. expansion gas turbine

'турбодетандер'

$$3.2.1. \quad N_1 + N_2 + N_3 \rightarrow N'_1 + N'_3$$

Ср. ventilation supply fan

'нагнетательный вентилятор'

$$3.2.2. \quad N_1 + N_2 + N_3 \rightarrow N'_3 + C'$$

Ср. heat transfer rate

при $C' = N_1 + N_2$
'скорость теплопередачи'

$$3.2.3. \quad N_1 + N_2 + N_3 \rightarrow C' + N'_1$$

Ср. turbine reduction. gear

при $C' = N_2 + N_3$
'редуктор турбины'

$$3.3.1. \quad N_1 + N_2 + N_3 \rightarrow A'_1 + A'_2 + N'_3$$

Ср. propulsion gear train

'главная зубчатая передача'

$$3.3.2. \quad N_1 + N_2 + N_3 \rightarrow A'_2 + N'_3 + N'_1$$

Ср. boiler water test

'гидравлическое испытание котла'

$$3.3.3. \quad N_1 + N_2 + N_3 \rightarrow A'_1 + N'_3 + N'_2$$

Ср. design condenser surface

'расчетная поверхность конденсатора'

Чаще всего такое несоответствие возникает в сочетаниях со входной структурой $A_1 + N_2 + N_3$,

Ср. high propeller efficiency 'высокий КПД винта'

3.3.4. $N_1 + N_2 + N_3 \rightarrow N'_3 + A'_1 + N'_2$

Ср. bilge pump capacity
'производительность трюмного насоса'

3.3.5. $N_1 + N_2 + N_3 \rightarrow N'_3 + N'_2 + N'_1$

Ср. engine valve spring 'пружина клапана двигателя'

3.3.6. $N_1 + N_2 + N_3 \rightarrow N'_3 + x + O'$ при $C' = N_1 + N_2$

Ср. back pressure turbine 'турбина с противодавлением'

3.4.1. $N_1 + N_2 + N_3 \rightarrow A'_2 + N'_3 + A'_x + N'_1$

Ср. sea suction valve 'приемный клапан заборной воды'

3.4.2. $N_1 + N_2 + N_3 \rightarrow N'_3 + x + A'_2 + N'_1$

Ср. carbon steel nut 'гайка из углеродистой стали'

3.4.3. $N_1 + N_2 + N_3 \rightarrow A'_2 + N'_3 + x + N'_1$

Ср. explosion relief valve
'предохранительный клапан с сигнализацией'

$$3.4.4. \quad N_1 + N_2 + N_3 \rightarrow N_3' + N_1' + x + N_2'$$

Ср. water inlet temperature

'температура воды на входе'

$$3.4.5. \quad N_1 + N_2 + N_3 \rightarrow N_3' + x + N_2' + N_1'$$

Ср. reactor input condition

'параметры на входе реактора'

Такое несоответствие вызвано семантикой ядра оборота, которое обозначает "устройство" или "абстрактное существительное".

$$3.5.1. \quad N_1 + N_2 + N_3 \left\{ \begin{array}{l} N_3' + x + N_2' + A_x' + N_1' \\ N_3' + x + N_2' + x + N_1' \end{array} \right.$$

Ср. bucket friction loss

'потери при трении активных лопаток'

compressor discharge pressure

'давление при выходе из компрессора'

В результате анализа четырёхсловных именных терминологических сочетаний выявлены 17 структурных варианта. Анализу были подвергнуты словосочетания со структурой $N_1 + N_2 + N_3 + N_4$, которые выявляют следующие наиболее типичные несоответствия в выходном лямке :

$$4.2.1. \quad N_1 + N_2 + N_3 + N_4 \rightarrow C' + N_1' \text{ при } C' = N_2 + N_3 + N_4$$

Ср. turbine reduction gear unit

'редуктор турбины'

$$4.3.1. \quad N_1 + N_2 + N_3 + N_4 \rightarrow A_3' + N_4' + A_1' + N_2'$$

Ср. fuel oil settling tank

'отстойная цистерна жидкого топлива'

$$4.4.1. \quad N_1 + N_2 + N_3 + N_4 \rightarrow N'_4 + A'_1 + A'_2 + N'_3$$

Ср. two-stage gas turbine engine

'двигатель двухступенчатой газовой турбины'

В тех случаях, когда в качестве первого компонента входного терминологического словосочетания употребляется прилагательное или P_2 , поверхностные структуры принимают следующую форму :

$$A_1 + N_2 + N_3 + N_4 \rightarrow N'_4 + A'_1 + A'_2 + N'_3$$

Ср. naval gas turbine project

'проект судовых газовых турбин'

$$4.4.2. \quad N_1 + N_2 + N_3 + N_4 \rightarrow N'_4 + N'_3 + A'_1 + N'_2$$

Ср. fuel oil service pump

'насос подачи жидкого топлива'

Также переводятся и входные словосочетания, которые имеют в своем составе прилагательное в качестве первого компонента, ср. high pressure pump housing 'корпус насоса высокого давления'.

$$4.5.1. \quad N_1 + N_2 + N_3 + N_4 \rightarrow N'_4 + N'_3 + N'_2 + x + N'_1$$

Ср. laboratory boiler test condition

'параметры испытания котла в лаборатории'

Пятисловные именные терминологические сочетания. Чем

больше количество компонентов в обороте, тем меньше его частотность¹. Выявлено 92 пятисловных терминологических сочетаний, которые распределены по 8 структурным типам.

Уже при анализе структурных типов четырехсловных сочетаний можно отметить тот факт, что сочетания не делятся по способу непосредственно составляющих /НС/, а согласно грамматике зависимостей путем сцепления или включения².

$$5.4.1. N_{10} + N_2 + N_3 + N_4 + N_5 \rightarrow N'_5 + C' + A'_3 + N'_4$$

two stage gas turbine engine

'Двигатель двухступенчатой газовой турбины'

$$5.4.2. N_1 + N_2 + N_3 + N_4 + N_5 \rightarrow C' + A'_2 + N'_3 + N'_1$$

Sp. turbine lube oil pressure gage

'манометр системы смазки турбины'

¹ К.Е. Харпер, Д.Г. Хейс. Использование машин при построении грамматики и программа для структурного анализа. В сб.: "Автоматизация в лингвистике". М.-Л., 1966, стр. 132-143.

² U. Weinreich. On the Semantic Structure of Language. "Universals of Language", Cambridge, Mass., 1966, p.130.

$$5.4.3. N_1 + N_2 + N_3 + N_4 + N_5 \rightarrow N_5' + A_1' + C' + N_4'$$

$$\text{при } C' = N_2 + N_3$$

Ср. ship's fuel oil storage capacity

'емкость корабельного топливного бака'.

$$5.4.4. A_1 + N_2 + N_3 + N_4 + N_5 \rightarrow C' + N_3' + A_1' + N_2'$$

$$\text{при } C' = N_4 + N_5$$

Ср. high pressure oil pipe line

'трубопровод масла высокого давления'.

$$5.5.1. A_1 + N_2 + A_3 + N_4 + N_5 \rightarrow N_5' + A_3' + N_4' + A_1' + N_2'$$

Ср. large diameter main bearing journal

'шайка коленчатого вала большого диаметра'.

$$5.5.2. A_1 + N_2 + N_3 + N_4 + N_5 \rightarrow A_4' + N_5' + A_1' + A_2' + N_3'$$

Ср. naval gas turbine power plant

'энергетическая установка военно-морских газовых турбин'.

$$5.5.3. A_1 + N_2 + N_3 + N_4 + N_5 \rightarrow A_3' + A_4' + N_5' + A_1' + N_2'$$

Ср. variable output speed gas turbine

'скоростная газовая турбина переменной мощности'.

$$5.5.4. A_1 + N_2 + N_3 + N_4 + N_5 \rightarrow A_4' + N_5' + N_3' + A_1' + N_2'$$

Ср. low pressure reheater drain pump

'водостливный насос подогревателя низкого давления'

5.6.I. $A_1 + A_2 + N_3 + N_4 + N_5 \rightarrow A'_1 + A'_2 + N_5 + x + A'_3 + N'_4$

Ср. powerful marine oil engine unit

'мощная судовая установка на жидком топливе'

Шестисловные именные терминологические сочетания. Выявлено 24 шестисловных именных терминологических сочетаний.

6.5.I. $P_2 + N_2 + N_3 + N_4 + N_5 + N_6 \rightarrow N'_6 + N'_5 + C' + N'_3 + N'_4$

Ср. forced draft blower turbine control mechanism

'механизм управления нагнетательным вентилятором турбины'

6.6.I. $N_1 + N_2 + N_3 + N_4 + N_5 + N_6 \rightarrow N'_6 + A'_4 + N'_5 + N'_3 + A'_1 + N'_2$

Ср. screen wall header bottom blow valve

'клапан нижнего продувания коллектора водяной защиты'

6.7.I. $A_1 + N_2 + N_3 + N_4 + A_5 + N_6 \rightarrow A'_5 + N'_6 + A'_3 + N'_4 + A'_1 + N'_2$

Ср. constant volume air cycle specific output

'удельная мощность воздушного цикла при постоянном объеме'

Итак, проведена систематизация наиболее типичных поверхностных и глубинных структур по формальному, морфологическому и семантическому признакам, на их основе составлен алгоритм обработки трехкомпонентных структур словосочетаний, имеющих пять типов глубинных структур на выходе.

Так как составление действующих алгоритмов — задача весьма трудоёмкая, мы вынуждены были ограничиться лишь алгоритмической обработкой наиболее частых входных именных терминологических сочетаний со схемой $N_1 + N_2 + N_3$. В результате работы составленного алгоритма часть именных терминологических словосочетаний, которые при иконическом подходе считались машинными оборотами и обрабатывались машиной как готовые лингвистические единицы, рассматриваются как свободные словосочетания. Алгоритмический подход позволяет перерабатывать эти словосочетания, не задавая их в памяти машины; при этом объем АСО заметно сокращается /примерно на 18 %/. Остальные структурные типы обрабатываются пока иконическим путем и включаются в виде готовых лингвистических единиц в АСО.

Реализация составленного АСО освещается в четвертой главе. Корректность лингвистических разработок, построенных автором реферируемой диссертации, была проверена в ходе МП английских текстов по судовым механизмам. При переводе использовались построенные им алгоритмы иконического /гл. II/ и алгоритмического подхода /гл. III/.

При иконическом подходе текстовые сегменты, состоящие из нескольких контактирующих словоупотреблений, отождествляются с аналогичным сегментом, значение которого заранее было взе-

дено в ЭВМ в виде словаря оборотов. Принцип работы алгоритма отождествления единиц текста с элементами словаря /словоформ и оборотов/ и выдачи выходных результатов довольно прост, зато единицы, введенные в память ЭВМ, занимают значительную часть её.

При алгоритмическом же подходе заложенная в ЭВМ информация невелика. Алгоритм МП именных терминологических свободных сочетаний оставлен на основе разработанной в третьей главе лингвистической типологии трехловных составных терминов со входной структурой $N_1 + N_2 + N_3$, которым соответствует в выходном языке пять конструктивных моделей перевода: $A'_1 + A'_2 + N'_3$; $A'_1 + N'_3 + N'_2$; $A'_2 + N'_3 + N'_1$; $N'_3 + A'_1 + N'_2$; $N'_3 + N'_2 + N'_1$ /пункты 3.3.1. - 3.3.5. гл. III/.

Полученный на ЭВМ экспериментальный перевод¹ правке не подвергался и приводится полностью в том виде, в каком он был выдан на печать /см. стр. 26 в/.

Ошибки, полученные при переводе, не превышают 3% всего экспериментального материала. Эти ошибки могут быть отнесены к следующим типам:

- а) технические ошибки, объясняющиеся неправильной перфорацией или кодированием того или иного компонента оборота;
- б) лингвистические ошибки, являющиеся результатом неверного подбора выходных эквивалентов;

¹ Программа пословного перевода и машинного алгоритма перевода трехловных терминологических свободных сочетаний были составлены старшим инженером - программистом ВЦ АН БССР канд. техн. наук Б.С. Кризевичем и старшим научным сотрудником НИЭЛ Мин. Аэропром. СССР Л.А. Славинной, которым автор приносит свою искреннюю благодарность.

в) ошибки, связанные с несовершенством программы.

К последнему виду ошибок следует отнести неправильный перевод оборота *by reducing valve* через эквивалент 'сокращая', 'клапан', 'лампа' вместо соответствующего 'детандером'. Эта ошибка возникла в связи с тем, что в программе недостаточно продумана обработка таких лингвистических ситуаций, когда в составе одного оборота содержится другое словосочетание /в нашем случае *by reducing*, являющееся самостоятельным оборотом/.

Для оценки восприятия живым приемником информации, полученных ЭВМ переводов с помощью АС словоформ и АСО, был проведен психологический эксперимент. Двум информантам /А и Б/ было предложено отредактировать: первому /А/ - перевод, полученный при помощи АСОЛ^I и АС словоформ отраслевого характера; второму /Б/ - перевод, полученный при помощи АСОЛ; АС словоформ и АСО отраслевого характера. Третий информант, обладавший той же квалификацией, что и первые два, должен был делать перевод этого же текста вручную, пользуясь обычным двуязычным и специальным словарями. Эксперимент показал, что информантам потребовалось равное количество времени для осуществления поставленной задачи. Редакция перевода, полученного ЭВМ с помощью АСОЛ, АС словоформ и АСО является наиболее близкой к переводу, выполненному квалифицированным переводчиком вручную. Таким образом МП научно-технических текстов /ор. данные табл. 3/ дает заметный выигрыш во времени.

^I С.В. Ястребова. Словарь общеупотребительной лексики английского языка. Сб.: "Статистика текста", т. 2, Минск, 1970 г., стр. 615-640.

A WOODWARD
VARIABLE SPEED
GOVERNOR

THE
SPEED
SETTING
LEVER
BEING
CONTROLLED
BY

A PNEUMATIC
ACTUATOR

THE
ENGINE
CONTROLS
FOR
LOCAL
OPERATION
CONSIST
OF

A LEVER-OPERATED
REVERSING
VALVE

A STARTING
LEVER

A FUEL
PUMP
BACK
CONTROL
LEVER

РЕГУЛЯТОР ИЗМЕНЕНИЯ СКОРОСТИ

РЫЧАГ СКОРОСТИ

КОНТРОЛИРУЕТСЯ

ПНЕВМАТИЧЕСКИМ РЕГУЛЯТОР

ОРГАНЫ УПРАВЛЕНИЯ ДВИГАТЕЛЕМ

ДВА-ИЗ-НА
МЕСТНЫХ
ДЕЙСТВИЕ
СОСТОИТ
ИЗ-

РЕВЕРСИВНЫЙ КЛАПАН С РЫЧАЖНЫМ ПРИВОДОМ

ПУСКОВОЙ РЫЧАГ

РЫЧАГ УПРАВЛЕНИЯ ТОПЛИВНЫМ НАСОСОМ

| Информанты | Время, необходимое для выполнения задания | Оценка качества перевода. | Длина переводимого текста | Количество ошибок |
|------------|---|---------------------------|---------------------------|-------------------|
| А | 60 мин. | 3 | 1000 с/у | 11 |
| Б | 20 мин. | 4 | 1000 с/у | 4 |
| В | 150 мин. | 5 | 1000 с/у | 2 |

Приложения /1,2,3/ содержат экспериментальные материалы в виде АСО по судовым механизмам в объеме 6600 английских составных терминов /1/ с их соответствиями в русском языке /2/ и научно-технического текста в объеме 2000 словоупотреблений с переводом именных терминологических оборотов /конечный подход/ на ЭВМ "Минск-22" /3/.

+ +
- +

Основные положения диссертации освещены в следующих работах :

1. Сопоставительное изучение семантической структуры слова при составлении двуязычного словаря для машинного перевода. Тезисы докладов Всесоюзной конференции "Актуальные проблемы лексикологии". Минск, 1970 / в соавторстве с А.Д.Борисевич и др./.
2. Проблема лексической сочетаемости при составлении словарей оборотов для машинного перевода. Тезисы докладов Всесоюзной конференции "Актуальные проблемы лексикологии". Минск, 1970, /в соавторстве с А.Д.Борисевич, В.С. Кризевичем/.
3. Статистическое описание лексики подязыка судовых меха-

тивов для построения автоматического словаря. Тезисы докладов и сообщений к республиканской научной конференции "Статистическое изучение стилей языка и стилей речи", вып. I, Горький, 1970 /в соавторстве с Т.А. Штирбу/.

4. Предсказуемый анализ при вероятностном МП. Материалы Всесоюзной конференции "Автоматическая переработка текста методами прикладной лингвистики". Кишинев, 1971, /в соавторстве с Э.Ф. Беляевой и др./.

5. Система лексического МП английских научно-технических текстов. Материалы Всесоюзной конференции "Автоматическая переработка текста методами прикладной лингвистики". Кишинев, 1971 /в соавторстве с А.Д. Борисевич, В.В. Гончаренко и др./.

6. К вопросу об идиоматичности при бинарном вероятностном машинном переводе /МП/. "Вопросы романо-германской филологии и общего языкознания". Минск, 1972 /в соавторстве с К.Ф. Лукьянковым и В.М. Петровской/.

7. Формулы межязыковых /англо-русских/ соответствий именных оборотов подъязыка судовых механизмов. "Вопросы романо-германского и общего языкознания". Минск, 1973 .

Подписано к печати 15.II.1973 г. Бумага 60x84^I/16. Печ. л. I, 7.
Уч. изд. I, 19. Заказ 64. Тираж 200 экз. Бесплатно. Отпечатано на
роталпринте Минского государственного педагогического института
иностраных языков. Минск, Захарова, 21.