Ч
П-71
70
б-ка ИРЯ

ИНСТИТУТ РУССКОГО ЯЗЫКА АН СССР

Проблемная группа по экспериментальной
и прикладной лингвистике

ПРЕДВАРИТЕЛЬНЫЕ ПУБЛИКАЦИИ

Выпуск 70

I. А. П. ВАСИЛЕВИЧ

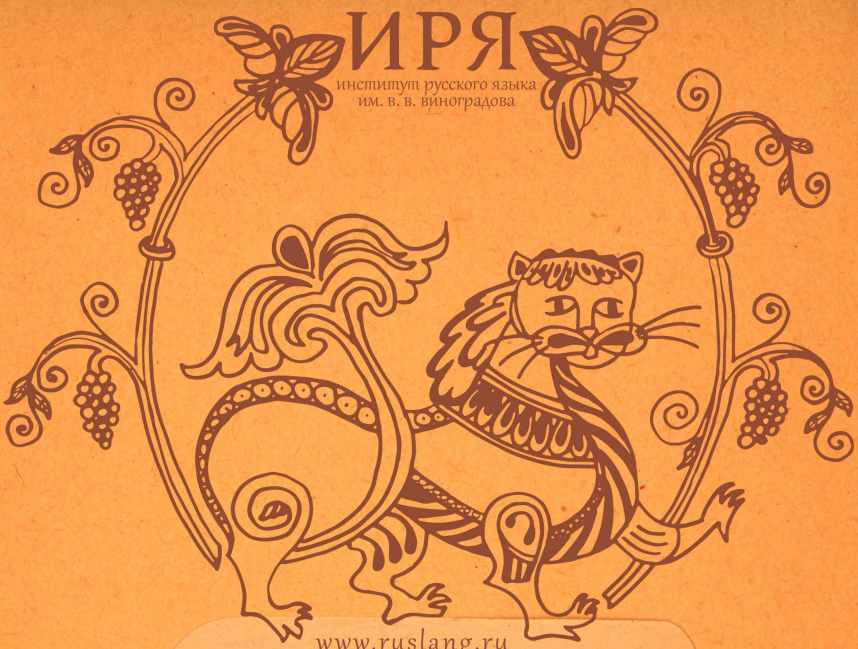
ПЕРИФРАЗИРОВАНИЕ: ОПЫТ СОПОСТАВЛЕНИЯ ДАННЫХ МОДЕЛИ С ЭКСПЕРИМЕНТАЛЬНЫМИ ДАННЫМИ

II. Е. Н. ГЕРГАНОВ, А. П. ВАСИЛЕВИЧ,
Р. М. ФРУМКИНА

К ПРОБЛЕМЕ ПОЛУЧЕНИЯ ОЦЕНОК АССОЦИАТИВНОЙ СИЛЫ БУКВОСОЧЕТАНИЙ

Москва — 1975

ИРЯ
институт русского языка
им. в. в. виноградова



www.ruslang.ru

ИНСТИТУТ РУССКОГО ЯЗЫКА АН СССР
Проблемная группа по экспериментальной и
прикладной лингвистике

ПРЕДВАРИТЕЛЬНЫЕ ПУБЛИКАЦИИ

Выпуск 70

I. А.П.Басилевич. Перифразирование: опыт сопоставления данных модели с экспериментальными данными.

II. Е.П.Герганов, А.П.Басилевич, Р.М.Фрумкина. К проблеме получения оценок ассоциативной силы буквосочетаний.

Москва - 1975.

К семидесятипятилетию
Александра Александровича
Реформатского
в знак признательности за
неизменную благожелатель-
ность к автору и его ра-
боте

А.П.Василевич

Перифразирование: опыт сопоставления данных модели с экспериментальными данными

4
П-71 70
БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР
70327

Постановка задачи

В лингвистической литературе до сих пор продолжаются дискуссии о том, что является предметом лингвистики как науки. Различие точек зрения по этому вопросу обуславливается как различным подходом к самому понятию "язык", так и различным пониманием того, какие "смежные" области исследования должны включаться лингвистом в круг рассматриваемых им явлений.

Разумеется, к настоящему времени имеется уже значительное число идей, относительно которых в науке достигнуто единодушие. Так, вряд ли кто-либо из лингвистов будет всерьез оспаривать точку зрения, согласно которой постулируется существование двух фундаментальных понятий: смысл и текст. Неудивительно, что все большее число сторонников завоевывает себе и вытекающий отсюда взгляд на естественный язык как на механизм, осуществляющий отображение множества смыслов на множество текстов и наоборот. Это отображение, по-видимому, реализуется совокупностью правил, закодированных в мозгу носителей языка. Знание этих правил (читай: владение языком) и позволяет индивиду строить текст, соответствующий данному смыслу (синтез, или порождение текста), или извлекать из данного текста некоторый смысл (анализ текста).

Соответственно перед лингвистами, описывающими язык (или строящими модель языка) стоят две важных задачи: описание (модель) синтеза текста и описание анализа текста.

Рассмотрим первую из этих задач. Среди лингвистов, разделяющих модельный подход к описанию языка, нет единодушия относительно того, какого рода "тексты" должна порождать модель типа "смысл-текст". Весьма распространена точка зрения, согласно которой задача синтеза формулируется как задача моделирования языковой способности (language competence) носителя языка, причем сама языковая способность понимается как способность порождать все правильные, "грамматичные" (grammatical) предложения данного языка и только такие предложения. Авторы подобных моделей вынуждены отвлекаться от "живых" текстов устной или письменной

речи, поскольку в них сплошь и рядом встречаются "неправильные", или как одно время говорили "неотмеченные" предложения.

Итак, многие создатели моделей типа "смысл-текст" ориентируются на так называемого "идеального" носителя языка, интересуясь не столько тем, какие тексты порождает носитель языка in vivo , сколько тем, какие тексты он должен порождать в идеале.

По нашему глубокому убеждению исследование функционирования языка также входит в компетенцию лингвистики. Мы полагаем, что существенный интерес представляет, в частности, изучение соответствия между задаваемым моделью перечнем "правильных" текстов и реально порождаемыми текстами.

В настоящей работе будут рассмотрены два аспекта названного соответствия:

I) в какой мере спонтанно порождаемые тексты являются нормативными (если считать эталоном нормативности предсказания какой-либо определенной модели смысл-текст)?

2) в равной ли мере реализуются в "живых" текстах все предсказания модели?

Наше исследование начинается с того этапа порождения текста, когда смысл уже задан, т.е. выражаясь разговорным языком, когда человек "знает, что он хочет сказать". Совершенно очевидно, что человек, свободно владеющий данным языком, в большинстве случаев может выразить заданный смысл несколькими способами, хотя реально каждый раз пользуется каким-то одним, порождая одну конкретную фразу, передающую заданный смысл. Естественно предположить, что в большинстве случаев выбор того или иного способа выражения будет неслучайным. Чем обуславливается этот выбор? Вероятно, существуют некоторые факторы, по которым различаются возможные способы выражения. Мы предполагаем, что в сознании человека имеется определенный механизм, в функцию которого входит оценка этих факторов в каждой конкретной ситуации и выбор способа выражения заданного смысла на основе произведенной оценки. Проблема выделения факторов этого рода также рассматривается в настоящей работе.

Исследование проводилось на уровне предложения.

Для решения поставленных задач нам необходимо было иметь в своем распоряжении два массива данных:

А. Перечень средств, существующих в языке для выражения некоторого смысла (делается на основе предсказаний какой-либо достаточно разработанной модели смысл-текст);

Б. Множество реальных текстов, выражающих тот же смысл.

Получение данных обоих типов представляет собой самостоятельную задачу, поэтому мы подробно обсудим каждую из них в отдельности (см.раздел Метод). Результаты сопоставления данных и собственно обсуждение поставленных задач см. в разделе Эксперимент. При этом следует иметь в виду, что начатое нами исследование еще не завершено в полном объеме, поэтому в настоящей работе мы делаем упор скорее на методику, чем на сами результаты.

Часть I. Метод

1. Получение нормативного перечня синонимичных средств выражения некоторого смысла.

Что мы имеем в виду, говоря о "разных способах выражения некоторого смысла"?

Рассмотрим предложения:

"Петров стреляет по лодке" (I)

"Петров ведет стрельбу по лодке" (2)

"Лодка обстреливается Петровым" (3)

"Петров делает выстрелы по лодке" (4)

Большинство грамотных носителей русского языка вполне согласится в том, что все эти фразы "означают примерно одно и то же", "передают примерно один и тот же смысл". Говоря на языке лингвистики, эти фразы "синонимичны". Мы примем, что "синонимичными" будут такие предложения, которые выражают один и тот же смысл. Поскольку мы намерены изучать разные способы выражения одного и того же смысла, следует указать, какие именно синонимичные предложения мы будем считать "разными: к "разным" обычно относят те предложения, которые отличаются друг от друга не менее, чем одним словом или порядком слов. При таком определении круг синонимичных предложений в некоторых случаях может быть весьма широк.

Обозначив символом S инвариант приведенной выше группы синонимичных предложений I-4, т.е. то (семантически) общее, что в них есть, мы можем сказать, что S есть некоторый "смысл", а фразы (I-4) - различные способы выражения смысла[x].

Теперь задача получения перечня языковых средств, выражающих один и тот же смысл, может быть сведена к получению полного списка членов какой-либо группы синонимичных предложений. В свою оче-

x) Ср. определение понятие "смысла высказывания" у И.А.Мельчука [2].

редь, эту задачу можно упростить следующим образом: найти для заданного предложения все предложения, синонимичные ему.

Сформулированную таким образом задачу наиболее удовлетворительным образом можно решить с помощью системы синонимичных преобразований, созданной И.А.Мельчуком и А.К.Жолковским в рамках разрабатываемой ими модели "Смысл - Текст". Данный выбор обусловлен несколькими соображениями. Во-первых, эта модель находится на такой стадии разработки, что она позволяет получать полные перечни синонимичных предложений для достаточно большого числа смыслов. Во-вторых, предложенный авторами аппарат лингвистического анализа структуры предложения представляется нам перспективным в плане возможного объяснения тех или иных предпочтений в употреблении информантами различных предложений из одной и той же синонимической группы.

Мы сочли необходимым кратко изложить некоторые положения модели "Смысл - Текст" (далее - Модель), главным образом, в той ее части, которая относится к системе синонимичных преобразований[x]. Этот шаг обусловлен не столько тем, что мы намереваемся изложить нечто лучше чем это сделано в оригинале, сколько тем, что мы сделали определенные модификации.

Модель "Смысл-Текст". Некоторые замечания.

Как мы уже говорили, модель представлялась нам весьма полезной, прежде всего в той ее части, которая называется "правила синонимических преобразований" (или "правила перифразирования"). Указанные правила позволяют путем несложных (и чаще всего формальных) операций получить для любого заданного предложения множество фраз, которые, как утверждают авторы модели, в идеале обладает следующими важными свойствами: а) каждая фраза множества выражает в точности тот же смысл, что и заданное предложение (т.е. синонимична этому предложению); б) указанное множество полно: вне его не существует фраз, синонимичных данному предложению; в) все фразы множества будут (грамматически) правильными предложениями.

---

x) Основные идеи Модели были впервые изложены в совместных публикациях И.А.Мельчука и А.К.Жолковского (из которых нам в первую очередь полезна была работа [1]), однако в основу описания будет положена и монография И.А.Мельчука [2], где Модель изложена наиболее полно.

Прежде, чем дать конкретное описание правил перифразирования, мы сделаем несколько замечаний общего характера.

I. Правила делятся на лексические и синтаксические (к настоящему времени их число составляет соответственно около 50 и около 20). При этом между обоими типами правил имеется весьма "жесткая" связь: применение лексического правила А сопряжено с применением вполне определенного синтаксического правила Б (например, при применении правила конверсии естественно следует изменить порядок следования субъекта и объекта и т.д.). Для удобства мы в дальнейшем опускаем из рассмотрения синтаксические правила, полагая, что их выбор однозначно определяется выбором лексического правила.

2. Существующие пока правила рассчитаны на простые предложения. Так, фраза Он обвинил ее в происшедшем может быть с помощью правил преобразована во фразу Он считал ее виновной в происшедшем, но не может быть преобразована во фразу Он считал, что она виновна в происшедшем.

3. Правила обладают свойствами симметричности.

4. Если некоторая фраза C_1 получена по правилу А из некоторой другой фразы C_2 и к ней применимо правило Б (т.е. по правилу Б можно получить C_3, эквивалентную C_2), то фраза C_3 будет перифразой C_1, хотя в списке правил может и не быть правила, непосредственно преобразующего C_1 и C_3. Это свойство правил (назовем его цепочечностью) для нас весьма важно, поскольку позволяет сопоставлять перифразы, например, по "длине" цепочки.

Как мы говорили, число построенных к настоящему времени правил перифразирования достигло 50 и продолжает расти. Однако многие из них приложимы лишь к отдельным словам русского языка. На нынешнем этапе исследования мы решили исключить такие правила из рассмотрения. Так, мы не рассматриваем правил с функциями S_1, S_2, A_1 и A_2, поскольку, во-первых, сами эти параметры реализуются довольно редко, а, во-вторых, далеко не все слова, имеющие эти параметры, допускают применение соответствующих правил.

В результате в число рассматриваемых нами правил попали только правила, оперирующие функциями Oper, Func (кроме $Func_0$) Labor, Conv, Syn и S_0. Смысл этих функций исчерпывающе описан в работе [2]. Заметим только, что авторы Модели считают наиболее регулярным способом образования конверсива (Conv) пассив (возможный,

разумеется, лишь при переходных глаголах - ср. Conv (победить) = быть побежденным). Однако, нам представляется удобным выделить конверсивы, образующиеся простым изменением залога, в отдельную группу; такие конверсивы мы будем далее называть Пассивами.

Приведем теперь примеры нескольких правил синтаксических эквивалентных преобразований по Модели.

I $i_0 \Longleftrightarrow Syn\ (i_0)$

2 $i_0 \Longleftrightarrow Conv\ (i_0)$

3 $i_0 \Longleftrightarrow S_0(i_0) + Oper_j\ (j=1,2)$

4 $i_0 \Longleftrightarrow S_0(i_0) + Func_j\ (j=1,2)$

5 $i_0 \Longleftrightarrow S_0(i_0) + Labor$

Напомним, что смысл правил состоит в том, фразы, соответствующие левой и правой части каждого правила, объявляются синонимичными по смыслу. Другими словами, если у некоторого слова i_0 во фразе C_1 имеется синоним, то, заменив во фразе C_1 слово i_0 на его синоним, мы получим фразу C_2, синонимичную C_1 по правилу I; аналогично, если у i_0 есть S_0 (например, S_0 (победить)= победа), а у этого S_0, в свою очередь, определена функция $Oper_1$ ($Oper_1$ (победа) = одержать), то по правилу 3 мы получим вместо победить - одержать победу, что превращает фразу А победил Б в синонимичную ей фразу А одержал победу над Б).

Для целей нашего исследования список правил был модифицирован следующим образом.

I. Мы ввели правило $i_0 \Longleftrightarrow$ Пассив (i_0), отделив, как мы говорили выше, данный способ конверсива от прочих. Это было сделано потому, что, как нам представлялось, при рассмотрении вопроса о предпочтениях правил информантами группа правил с конверсивами была бы неоднородна; кроме всего прочего, к структурам с Oper'ами Func'ами и Labor'ами может быть применен только Пассив.

2. Мы добавили правила с $Oper_3$ и $Func_3$, поскольку трехактантные предикаты более распространены и именно из их числа естественнее всего было брать стимулы.

3. Мы расчленили Labor (правило 5) на несколько Labor'ов ($Labor_{12}$, $Labor_{13}$ и т.п.), сформулировав и соответствующее число правил. Этим мы опять-таки стремились создать условия, в которых те или иные возможные предпочтения в использовании правил, могли бы проявиться в максимальной степени.

4. Наконец, в работе [1] говорится о возможности возникновения "цепочек" правил; нам же было удобнее провести эту возможность последовательно до конца, эксплицитно выписав все возможные цепочки.

В результате система правил перифразирования была представлена нами в следующем виде (рис.I):

2. Получение данных о выборе средств выражения некоторого смысла in vivo.

Имеет ли смысл попытка собрать необходимый массив данных путем обследования текстов? Легко видеть, что такой путь проведения исследования крайне нерационален (если вообще возможен), т.к. для получения данных об употребительности того или иного способа выражения некоторого смысла S необходимо, чтобы смысл S многократно встречался в обследуемом корпусе текстов.

Последнее, однако, вполне достижимо, если обследуются тексты, порожденные информантами в некоторых специальных, заданных исследователем, условиях. Другими словами, мы приходим к необходимости моделировать интересующую нас ситуацию в эксперименте.

В самом деле, именно в эксперименте можно создать такие условия, при которых у информанта возникнет необходимость выразить как раз тот смысл, какой нужен исследователю; тогда отобрав группу информантов и поставив их в одни и те же условия, исследователь получит в свое распоряжение необходимое множество выражений одного и того же смысла.

Каким же образом создать такие условия? Нам представляется, что существует три основных способа побудить информанта выразить требуемый смысл, пользуясь средствами родного языка:

1) задать смысл на некотором языке, отличном от родного;

2) задать смысл неязыковыми средствами;

3) задать смысл, сообщив информанту какую-либо реальную фразу на исследуемом языке и прося его "выразить то же самое другими словами".

За недостатком места мы не будем подробно обсуждать названные способы. Отметим только, что решающим недостатком первых двух является тот факт, что они не позволяют добиться однозначного задания смысла. Задание же смысла путем предъявления информантам обычного предложения на родном языке не имеет этого недостатка: мы всегда можем выбрать для предъявления такую фразу, которая

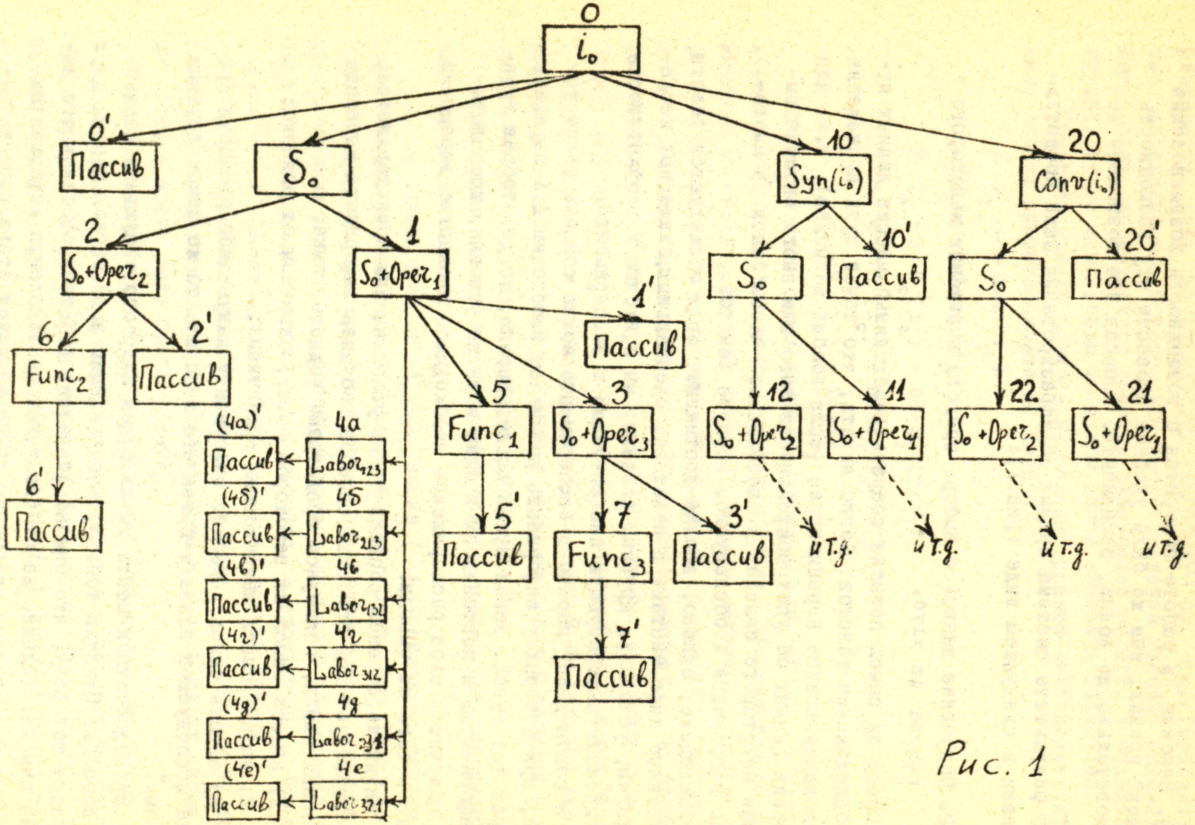

0
i_o
0'
Пассив
S_o
10
$Syn(i_o)$
20
$Conv(i_o)$
2
S_o+Oper$_2$
1
S_o+Oper$_1$
10'
Пассив
20'
Пассив
6
$Func_2$
2'
Пассив
1'
Пассив
S_o
S_o
6'
Пассив
(4a)'
4a
5
$Func_1$
3
S_o+Oper$_3$
12
S_o+Oper$_2$
11
S_o+Oper$_1$
22
S_o+Oper$_2$
21
S_o+Oper$_1$
Пассив
$Labor_{123}$
(4б)'
4б
Пассив
$Labor_{213}$
(4в)'
4в
Пассив
$Labor_{132}$
(4г)'
4г
Пассив
$Labor_{312}$
(4д)'
4д
Пассив
$Labor_{231}$
(4е)'
4е
Пассив
$Labor_{321}$
5'
Пассив
7
$Func_3$
3'
Пассив
7'
Пассив
и т.д.
и т.д.
и т.д.
и т.д.

Рис. 1

будет пониматься строго однозначно, и тогда в эксперименте информанты будут порождать предложения в точности из той группы синонимичных предложений, которая нас интересует.

Немаловажно и то, что при рациональном выборе исходной фразы (т.е. фразы для предъявления информантам с целью задания смысла) мы можем разумно ограничить круг возможных синонимических предложений, число которых может быть чрезмерно велико. Так, если в приведенных нами выше фразах (I–4) в качестве исходной взять фразу (I) - Петров стреляет по лодке - то легко видеть, что все возможные ответы будут строиться на основе различных эквивалентов смысла глагола "стрелять"; остальные же слова практически не могут увеличить число синонимичных предложений (трудно представить себе, чтобы кто-нибудь дал перифразы типа "Человек по фамилии Петров стреляет..." или "Петров стреляет по деревянному двухвесельному плавучему средству").

Итак, мы останавливаемся на следующей схеме исследования. Отбирается какая-либо фраза русского языка. Для нее с помощью правил перифразирования (рис.I) выписывается множество всех синонимичных фраз. Затем ставится эксперимент, в котором информанты - носители русского языка получают эту фразу и должны передать ее смысл другим способом - любой другой фразой, какая придет им в голову. Полученные ответы сопоставляются с перечнем перифраз, предсказанным правилами модели.

Часть 2. Эксперимент.

I. Отбор стимулов

Как видно из рис.I, рассматриваемые нами правила синонимических преобразований ориентируются на предикаты, поэтому отбор стимулов (фраз), в сущности, сводится к отбору глаголов и построению фраз с этими глаголами - так, чтобы все возможные фразовые преобразования обуславливались только преобразованиями глагола.

Что же могло бы служить нам источником для отбора глаголов? Напомним, что практическое применение правил синонимичных преобразований (рис.I) возможно только в том случае, если мы располагаем специальным толковым словарем, который содержит всю ту информацию, необходимость которой диктуется Моделью. По замыслу авторов Модели этому требованию будет удовлетворять "Толково-комбинаторный словарь русского языка" (далее ТКСл), создаваемый в настоящее время коллективом авторов под руководством И.А.Мельчука.

Разумеется, самое естественное было бы использовать именно этот словарь для поиска нужных нам глаголов. Однако здесь имеется существенная трудность.

Дело в том, что к настоящему времени в распоряжение исследователей поступили лишь некоторые фрагменты этого словаря: работа над ним требует столь кропотливого труда, что завершение словаря - дело еще очень далекого будущего. Тем не менее, качество имеющихся сейчас фрагментов, по-видимому, достаточно высокое, и потому мы попытались использовать их в качестве источника для поиска наших глаголов.

Почти все фрагменты ТКСл были опубликованы в серии брошюр "Материалы к толково-комбинаторному словарю" (Предварительные публикации ИРЯ АН СССР). Анализ "Материалов" показал, что к настоящему времени описано порядка 200 слов русского языка, в том числе 20 глаголов. Однако эти глаголы оказались весьма разными с точки зрения того, насколько богатыми возможностями синонимичных замен они обладают. Мы были заинтересованы в том, чтобы каждый наш стимул имел по несколько способов синонимичных преобразований. Так, никакого интереса не представляли для нас глаголы, не имеющие пассива, лексического синонима и отглагольного существительного.

Изучение имеющихся в ТКСл глаголов под этим углом зрения привело к тому, что мы смогли отобрать пять глаголов, удовлетворяющих выдвинутым нами требованиям: стрелять, победить, описывать, рассматривать и обвинять. Теперь для каждого глагола предстояло:

а) заполнить его валентности (актанты) конкретными словами (именами);

б) построить из отобранных слов какое-либо русское предложение, удовлетворяющее требованиям нашего исследования (простое, однозначно понимаемое и т.д.);

в) выписать для построенного предложения все синонимичные ему фразы (с помощью правил рис.I);

г) из группы синонимичных фраз выбрать исходное предложение-то, которое будет предъявляться информантам в эксперименте.

Этапы (а) и (б) не вызвали особых трудностей. Этапы же (в) и (г) заслуживают подробного обсуждения.

Построение множества правильных перифраз

На первый взгляд может показаться, что построение множества правильных перифраз для наших стимулов - дело простое. Ведь в нашем распоряжении есть четко сформулированные правила и, кроме того, сами стимулы отобраны так, чтобы иметь возможность легко решать вопрос о применимости этих правил: предполагалось, что информация о слове в ТКСл исчерпывающа. На самом деле вопрос о построении правильных перифраз не столь прост. Укажем основные трудности, с которыми мы столкнулись.

Прежде всего, как мы уже отмечали, сам словарь (ТКСл) еще далек от совершенства, что признается и самими авторами. Дело не только в том, что он мал по объему; даже имеющиеся уже словарные статьи по большей части носят характер предварительного исследования. В общем-то сама идея публикации этих материалов преследовала цель привлечь к ним внимание лингвистов, вызвать соответствующие обсуждения и т.п. - с тем, чтобы в дальнейшем совершенствовать имеющиеся описания. Нам известно, например, что в результате подобных обсуждений некоторые словарные статьи были подвергнуты значительной переработке.

Указанная незавершенность, неокончательность словарных описаний в их нынешнем виде не позволила нам полностью положиться на содержащиеся в них сведения. Анализ описаний пяти отобранных нами глаголов выявил следующие недостатки:

а) В некоторых случаях были ошибочно названы функции (например, в статье к слову стрелять то, что по определению авторов Модели должно было соответствовать функции $Func_3$, в словарной статье было названо $Oper_3$)*);

б) В словарных статьях были отмечены не все возможные функции;

в) Почти всегда был неполон список возможных синонимов. Вместе с тем, несмотря на наличие некоторых помет ("широкий" синоним; "узкий" синоним и т.п.), на основе словарных данных нельзя было однозначно решить, какие синонимы из приведенного ряда "подходят" к нашему конкретному случаю;

г) Словарь не содержит описания ни одного из синонимов и

---

*)Вообще, судя по публикациям, разграничение функций Func и Oper у авторов Модели еще не окончательно.

конверсивов наших пяти слов; тем самым мы лишены возможности пользоваться словарем для построения правильных перифраз по цепочкам, начинающимся с синонимов (неизвестно, есть ли у них S_0, $Oper_1$ и т.п.).

Сказанное убеждает в том, что данные ТКСл ни в коей мере не могли служить основой для решения вопроса о применимости правил и, тем самым, для построения правильных перифраз. Для того, чтобы все-таки выделить множество правильных перифраз, было проделано следующее.

Этап А: "Совершенствование" словарных описаний (их уточнение, исправление, дополнение).

Этап Б: Обсуждение произведенных изменений с экспертами.

Этап А. Не вдаваясь в подробное перечисление всех изменений и дополнений, внесенных нами в словарные описания, мы укажем лишь, какие операции мы совершали:

1. Уточнение словарных данных (см. выше пп.(а) и (б)).

2. Дополнение данных (главным образом, синонимами - см. п. (в)).

3. Описание слов (синонимов и конверсивов наших пяти глаголов), отсутствовавших в ТКСл *).

4. Порождение перифраз по правилам Модели с учетом информации, полученной в результате (1-3) и формулировка своего рода "фильтров", запрещающих некоторые фразы несмотря на наличие информации, необходимой для применения данного правила. Например, у слова стрелять определена функция S_1(деятель): S_1(стрелять)= стрелок. Тогда, согласно соответствующему правилу, возможна перифраза: стрелять $\Longleftrightarrow$ быть (являться) стрелком. Однако в нашем случае фраза "*N.* является стрелком из ружья по лодке" противоречит законам русского языка и не может рассматриваться как правильная**).

Этап Б. Все предложенные нами изменения, дополнения и "запреты" были сделаны на основе наших интуитивных представлений,

---

*) Описания включали только информацию о тех функциях, которые фигурируют в правилах рис.I.

**) По-видимому, если бы ТКСл включал бы уже слово стрелок, там содержался бы соответствующий запрет.

"языкового чутья". Полагая, что одного лишь нашего мнения явно недостаточно, мы решили согласовать его с мнением других лиц, достаточно авторитетных в этом отношении, т.е. согласовать наше мнение с мнением экспертов. Мы полагали, что поскольку мы ведем исследование в рамках модели Мельчука-Жолковского, то наиболее компетентными экспертами будут авторы Модели, а также их непосредственные коллеги, участвующие в частности, в создании ТКСл. Именно с этими лицами мы консультировались по поводу всех возникавших у нас вопросов и предложений*). Мы считали, что то или иное предложение или поправка одобрены экспертами, если они говорили "это верно"; "это следует учесть при переработке данной словарной статьи" и т.п. Таким образом, хотя реально соответствующие словарные описания еще не готовы, все наши "правильные" перифразы построены так, как если бы эти описания уже были сделаны авторами Модели. В этом смысле мы и будем говорить, что наши правильные перифразы "правильны" именно потому, что они согласуются как с правилами Модели, так и с соответствующими словарными описаниями.

103277

Выбор исходного предложения

Из свойства симметричности правил Модели вытекает, что мы можем получить полный набор перифраз данной группы, беря за исходную любую из фраз данной группы. Очевидно, что наиболее полной будет картина в том случае, когда мы дадим в качестве исходного предложения поочередно все фразы данной синонимической группы. Но не менее очевидно, что такой путь потребует чрезмерно большого числа экспериментов, поставить которые трудно еще и потому, что один информант может быть допущен к участию только в одном эксперименте.

Более реально было бы провести подобную серию экспериментов, разумным образом сузив круг исходных предложений. Это, однако, уже представляет собой отдельную задачу, к решению которой мы пока не готовы.

www.ruslang.ru

---

*) Мы пользуемся случаем поблагодарить И.А.Мельчука, Л.Н.Иорданскую, С.Е.Никитину, Л.Иомдина и Н.Перцова за внимание и благожелательность, проявленную во время этих консультаций.

БИБЛИОТЕКА ИНСТИТУТА РУССКОГО ЯЗЫКА АН СССР

В данной работе мы ограничились выбором всего одного исходного предложения из каждой синонимической группы.

Самым естественным было взять в качестве исходных предложений фразы, включающие отобранные нами 5 глаголов в простой форме:

1. Петров стреляет из ружья по лодке.
2. Иванов описывает этот дом в своем очерке.
3. Мастер победил гроссмейстера.
4. Петров рассматривает проблему А в своем докладе.
5. Иванов обвинил Петрова в происшедшем.

Ниже будут изложены результаты эксперимента по одному из стимулов – Мастер победил гроссмейстера; исследование по остальным синонимическим группам еще не завершены в полном объеме.

2. Перечень правильных перифраз.

В соответствии с принятой нами схемой исследования вначале для исходного предложения Мастер победил гроссмейстера. (далее – М.победил Гр.) были выписаны все возможные перифразы (см.графу 2 табл.I).

Табл.I

Синонимичные фразы (перифразы), предусмотренные правилами рис.I для исходного предложения М. победил Гр.

№ правил	Перифразы	Содержит ли ТКСл необходимую информацию	Число ии, давших данный ответ
I	2	3	4
0′	Гр.(был) побежден М-ом	[да]	32
I	а) М.одержал победу над Гр.	да	35
	б) М.завоевал победу над Гр.	да	–
	в) М.добился победы над Гр.	нет x)	3
I′	Во встрече М. с Гр. победа была одержана М.	[да]	2
	а) Игра завершилась победой М.над Гр.	нет x)	1
3	б) Игра закончилась победой М.над Гр.	да	4
	в) Игра М. с Гр. окончилась победой М.	да	2
	г) Игра М. с Гр. принесла победу М.	да	I

Табл. I (окончание)

I	2	3	4
4в	а) М. завершил победой свою встречу с Гр.	нет[x)]	-
	б) М. закончил победой свою встречу с Гр.	нет[x)]	-
5	а) Победа над Гр. досталась М-у	да	4
	б) Победа над Гр. принадлежит М-у	нет[x)]	2
	в) Победа над Гр. осталась за М-ом	да	5
	г) В игре М. с Гр. победа была за М.	да	I
	д) В игре М. с Гр. победа оказалась на стороне М.	нет[x)]	I
I0	а) М. выиграл у Гр.	нет[x)]	2I
	б) М. взял верх над Гр.	да	6
	в) М. одержал верх над Гр.	да	5
	г) М. обыграл Гр.	нет[x)]	6
	д) М. переиграл Гр.	да	4
	е) М. одолел Гр.	да	2
	ж) М. осилил Гр.	да	-
	з) М. пересилил Гр.	да	-
I0'	д) Гр. был переигран М-ом	да	-
	г) Гр. был обыгран М-ом	нет[x)]	I
II	М. нанес поражение Г-у	?	I4
II'	Поражение было нанесено Гр-у М-ом	да	2
I2	а) Гр. потерпел поражение от М.	?	I7
	б) Гр. потерпел фиаско во встрече с М.	?	I
I3	а) Игра закончилась поражением Гр. от М.	?	I
	б) Игра М.с Гр.окончилась поражением Гр.	?	I
	в) Игра М. с Гр.кончилась поражением Гр.	?	-
20	а) Гр. проиграл М-у	да	26
	б) Гр. уступил М-у	нет[x)]	2

Напомним, что согласно принятым нами ранее допущениям, каждая из перифраз в табл. I является правильным русским предложением и, кроме того, синонимична исходному предложению М. победил Гр..

Все эти перифразы являются правильными. К числу правильных относится еще небольшая группа перифраз, получаемых по правилам,

ле представленным на рис.I (в табл.I они отсутствуют). Любые другие перифразы следует, согласно Модели, считать неправильными.

В графе I табл.I указан номер правила рис.I, по которому построена данная перифраза.

Графа 3 содержит информацию о том, в какой мере были использованы правила перифразирования Модели. Помета (да) означает, что ТКСл содержит всю необходимую информацию для применения соответствующего правила (например, в ТКСл указывается, что S_0 (победить) = победа; $Oper_1$ (победа) = одержать, завоевать; следовательно, правило I применимо на основании данных ТКСл).

Помета [да] относится лишь к правилам с пассивом; хотя в ТКСл пассивы глаголов приводятся не всегда, мы полагаем, что решение о возможности и невозможности пассива в каждом отдельном случае вряд ли может вызывать сомнение. Тем самым мы как бы предполагаем, что соответствующая информация в ТКСл имеется.

Помета (нет^X) указывает на отсутствие в ТКСл информации, позволяющей построить данную перифразу по данному правилу, но эти перифразы были признаны экспертами "правильными", т.е. по их мнению, соответствующую информацию в ТКСл следует иметь. В основном, речь идет о синонимах.

Наконец, помета (?) относится к тем перифразам, которые следовало бы строить, опираясь на словарные статьи других слов (еще не описанных в ТКСл; ср. фразы на правила I2 и I3 в табл.I, правомерность которых могла бы быть оценена при наличии описания слова поражение). Мы все же включаем такие перифразы в список правильных, "предвосхищая" соответствующее словарное описание и опять-таки согласовав принимаемые решения с экспертами.

3. Порождение перифраз информантами.

Процедура. Опыт проводился индивидуально. Перед опытом испытуемому сообщалась устно следующая инструкция: "Представьте себе, что некий корреспондент систематически дает в свою газету информацию о ходе определенного шахматного турнира. В очередном туре состоялась встреча между мастером А. и гроссмейстером Б. Сообщая о результате этой встречи, корреспондент написал так: (Экспериментатор предъявляет карточку с предложением "Мастер победил гроссмейстера"). Каким образом еще мог бы корреспондент сообщить ту же мысль? Постарайтесь придумать несколько фраз, в точности передающих смысл, который хотел выразить корреспондент. Все ли

Вам понятно?". Испытуемые фиксировали свои ответы в письменной форме.

Испытуемые. В эксперименте участвовало 53 ии. разного пола, возраста и профессий (более половины ии. составляла группа лингвистов).

Результаты. Предварительный анализ результатов побудил нас исключить из рассмотрения ответы двух ии., явно не следовавших инструкции: один из них представил 10 предложений типа "Победа над гроссмейстером дала мастеру больше уверенности в себе" и "На основе победы мастера над гроссмейстером был поставлен фильм", а другой и. дал одну-единственную фразу "Гроссмейстер добился победы в упорной борьбе с мастером".

Таким образом, мы имели в своем распоряжении ответы 51 ии. Всего было получено 318 предложений, т.е. в среднем несколько более 6 фраз от каждого и.

Теперь перейдем к сопоставлению полученных ответов с перечнем правильных перифраз, помещенных в табл.I.

А. Соотношение правильных и неправильных перифраз в ответах ии.

Как показал анализ полученных нами ответов, к числу правильных перифраз относится 237 предложений (или примерно 3/4 от общего числа ответов); 81 ответ представляет собой неправильные перифразы.

Поскольку правильные ответы определены как "правильные русские фразы, синонимичные исходному предложению", неправильные ответы могут быть двух типов:

А. Грамматически правильные фразы, искажающие смысл исходного предложения (т.е. не являющиеся синонимичными преобразованиями исходного предложения).

Б. Фразы, сохраняющие смысл исходного предложения, но нарушающие принятые нормы русского языка ("аграмматичные").

Приведем несколько примеров перифраз обоих типов. Тип А: Гр. сдался М-у (6 ответов); М.разгромил Гр. (3); М. нанес удар Гр.(2); М. вырвал победу у Гр. (2) и т.д. Тип Б: М.превзошел Гр. (8 ответа); М. сразил Гр. (2); Состоялось поражение Гр-а М-ом (I); Гр. потерпел фиаско от М. (I) и т.д. Всего было получено 52 фразы типа А (в том числе 39 разных) и 29 фраз типа Б (24 разных).

Мы не будем пока комментировать полученное в опыте соотноше-

ние правильных и неправильных ответов*).

Полученное же соотношение типов А и Б, вообще говоря, не является неожиданным: естественно, что хорошо владеющие языком информанты должны меньше ошибаться в грамматике (почти все ошибки на неправильную сочетаемость), чем в случаях, когда необходимо уловить иногда весьма тонкие смысловые нюансы.

При анализе наших экспериментальных данных мы обратили внимание на существенное различие ии. по "степени продуктивности", т.е. по числу порождаемых ими перифраз (напомним, что в инструкции ии. предлагалось дать столько ответов, сколько они смогут): это число колебалось в пределах от 2 до I9. При этом у 5 ии. с числом ответов более I0 доля неправильных перифраз значительно превышает среднюю по всей группе (40% против 27%). Этот результат нельзя было объяснить, допустим, низкой квалификацией данных ии. (из них трое - лингвисты, а остальные двое - лица с высшим образованием; все пятеро владеют хорошим литературным языком). По-видимому, здесь дело было именно в длине порожденной ими серии ответов. Это подтверждается следующими данными. Среди первых пяти ответов в данной группе ии. доля неправильных перифраз составляет всего I6%; среди последних пяти ответов - 64%. Совершенно очевидно, что к концу серии ответов результаты каждого и. в среднем существенно ухудшаются.

Сказанное верно не только относительно рассмотренной группы ии. Если мы будем последовательно подсчитывать долю правильных ответов по всей группе ии. для первого по порядку ответа, для первых двух ответов, первых трех ответов и т.д.,тр получим последовательно убывающую величину: 98% для первого ответа; 95,I% для первых двух ответов; 8I,6% для первых пяти ответов и т.д.

Интересно и другое. Наряду с увеличением доли неправильных ответов в сериях ответов изменяется и соотношение ответов типа А

---

*) Отметим, что принимая во внимание состав ии. (много лингвистов и лиц с высшим образованием), а также тот факт, что мы не имели дела со спонтанной, неподготовленной речью (ии. думали над ответами), мы ожидали, что доля неправильных ответов будет меньше. Мы вернемся к этому вопросу несколько позднее.

и Б: доля ошибок типа Б к концу в длинных сериях возрастает. (после десятого по счету ответа ошибки типа А и Б встречается примерно поровну; в первых же ответах ошибки типа Б встречаются в среднем в два раза реже ошибок типа А).

Как можно проинтерпретировать полученные результаты? По нашему мнению, у ии., участвующих в опыте типа нашего, не возникает никаких трудностей при порождении одной-двух перифраз, причем практически все эти перифразы - правильные. При попытке назвать третью, четвертую и т.д. перифразы ии. начинают испытывать все большие затруднения. По-видимому, в этой ситуации большое значение приобретают характерологические особенности испытуемых: одни из них, чувствуя, что новые перифразы уже "не идут на ум", предпочитают кончить опыт (здесь важно отметить, что самый момент, когда новые перифразы уже не приходят в голову наступает у разных ии. по-разному). Другие же ии. (в частности, те пятеро, о которых мы говорили выше) пытаются преодолеть возникающие трудности и породить как можно больше перифраз. То, что они испытывают трудности, видно из резкого увеличения времени, которое они затрачивают на обдумывание очередного ответа. Интересно, что Э., видя, что и. начинает подолгу задумываться над ответом, обычно предлагал кончить опыт; однако ии. обсуждаемого типа говорили: "Нет, я еще подумаю", в то время как другие ии. с видимым облегчением кончали работу.

По всей вероятности, при такой "вымученности" ответов у ии. снижается критика к их качеству. Характерно и то, что доля неправильных ответов при этом растет, в первую очередь, за счет ошибок типа Б, т.е. за счет грамматических ошибок. В некоторых случаях в конце длинных серий появляются ответы, вообще мало похожие по смыслу на исходное предложение (ср. Победа М. над Гр. выдвинула последнего на I-ое место).

Исходя из поставленных нами задач, предоставлялось разумным учитывать только "естественные" ответы ии., т.е. те ответы, которые не вызывали особых трудностей (не требовали длительного размышления и т.п.) К сожалению, мы не имели возможности указать точный момент в серии ответов данного и., после которого он начинал работать "через силу". У нас есть только две возможности хоть в какой-то мере отделить "естественные" ответы от "вымученных".

Во-первых, мы можем считать естественными только два первых ответа каждого и. Основания: никто из 51 ии. не прекратил работу

раньше и, кроме того, по наблюдению Э. никто из ии. не испытывал трудностей в этот момент. Однако, ясно, что при таком выборе границы какая-то часть естественных ответов (у ии., давших более двух ответов) остается неучтенной.

Во-вторых, мы можем провести границу на уровне первых пяти ответов. Основание: если выписать распределение числа ии., прекращавших работу после 2-го ответа, 3-го ответа и т.д., то мода придется как раз на 5 ответов (после пятого ответа прекратили опыт 16 ии., в то время как после 3-го и 4-го ответов отказались продолжать опыт по 7 ии., после 6-го - 6 ии. и т.д.). Таким образом, в число "естественных" попадают все ответы ии., давших 2-5 ответов, а также первые 5 ответов прочих ии. Разумеется при этой границе мы не можем дать гарантии, что в число естественных не попадают какие-либо "вымученные ответы". Однако здесь учтено значительно больше естественных ответов, чем при первой границе (хотя, быть может, и здесь учтены еще не все).

В табл.2 приведены данные о соотношении правильных и неправильных ответов с учетом всех ответов, а также с учетом двух возможных границ "естественных" ответов.

Табл.2

Какие ответы считаются "естественными"	Всего ответов	Правильных	Неправильных тип А	тип Б	Доля правильных
I-2	I02	97	4	I	95,I%
I-5	228	I86	27	I5	8I,6%
все (I-I9)	3I8	237	52	29	74,4%

Из приведенных данных видно, что "естественные" ответы, т.е. ответы, легко приходящие в голову, почти всегда оказываются правильными перифразами.

Б. Предпочтение некоторых типов перифраз.

Создавая правила перифразирования, авторы модели Смысл-Текст ставили своей задачей составить полный перечень возможных способов выражения некоторого смысла, не пытаясь как-либо упорядочить этот перечень, т.е., в частности, не указывая, является ли один способ выражения более употребительным, чем другой. В нашу же задачу входит как раз выявление предпочтений одних способов выражения заданного смысла перед другими.

В нашем случае поставленная задача сводится к выявлению предпочтений тех или иных перифраз предложения "М. победил Гр.". Естественно с этой целью рассматривать только правильные ответы ии., причем учитывая сказанное в предыдущем разделе, мы будем рассматривать по отдельности всю группу ответов и первые ответы ии., которые мы назвали "естественными". В качестве показателя предпочтения мы возьмем число ии., давших данный ответ.

Мы уже говорили, что мы намереваемся исследовать только те перифразы, которые строятся по правилам перифразирования, представленным на рис.I.

Как показал анализ ответов, правилами рис.I описывается подавляющее большинство всех правильных ответов (214 из 237, или 90%) .Остальные правильные ответы либо описываются другими правилами перифразирования (10 фраз типа "В игре мастера с гроссмейстером мастер вышел победителем"), либо вообще не описываются правилами Модели в их настоящем виде (13 фраз типа "Победа досталась мастеру, а не гроссмейстеру";"Гроссмейстера победил мастер"; "Гроссмейстер сыграл с мастером, но не победил его и не сделал ничьей"). Нам представляется, что исключение столь небольшого числа перифраз из анализа не может существенно отразиться на общей картине результатов эксперимента.

В табл.I представлены все правильные перифразы, построенные по правилам рис.I; у каждой перифразы указано число ии., давших данную перифразу (графа 4). Очевидно, что перифразы существенно различаются по употребительности в ответах (ср.: М. одержал победу над Гр. было названо 35 ии., в то время как целый ряд возможных перифраз ни разу не встретились в ответах). Однако одно обстоятельство удерживает нас от сопоставления отдельных перифраз.

Как видно из табл.I, общее число возможных правильных перифраз составляет порядка 40. Для того, чтобы выявить предпочтения отдельных перифраз согласно принятой практике планирования эксперимента у нас должно было бы быть порядка 200 ии.; у нас же их было только 51 чел.

В данном случае реальнее поставить задачу меньшего объема: выявить предпочтения отдельных правил перефразирования. Такая постановка задачи побудила нас несколько преобразовать исходную матрицу ответов, а именно: если некоторый и. дал несколько различных перифраз на одно и то же правило*), мы учитываем только одну

*) Только по правилу 10 (лексический синоним) можно дать 8 перифраз.

из них - первую по счету. Таким образом, в дальнейшем мы будем иметь дело уже не с множеством правильных перифраз, а с множеством реализаций правил перифразирования.

Как и в случае с числом ответов, ии. различаются по числу реализуемых ими правил, хотя здесь это различие уже не так велико: максимальное число правил, реализованных одним и., составляет 8; минимальное число - 2. Разумеется, число реализованных правил находится в прямой зависимости от общего числа ответов, порожденных данным и.: у тех ии., которые дали наиболее длинные серии ответов, наблюдается и наибольшее разнообразие реализованных правил.

В предыдущем разделе мы ввели понятие "естественных" ответов. Представлялось целесообразным и в данном случае учесть этот фактор. Мы будем проводить наш анализ в отдельности для всех реализаций правил и для первых двух реализаций у каждого и., (которые мы будем считать наиболее "естественными"; первые две реализации правил почти во всех случаях совпадали с первыми двумя ответами ии.).

Данные о предпочтениях в реализации правил представлены в табл.3; правила в таблице упорядочены от наиболее предпочитаемых к наименее употребительным.

Табл.3

№ № правил	Число ии., реализовавших данное правило	
	по всем ответам	по двум первым ответам
I	37	29
I0	35	I9
0'	32	25
20	26	I4
I2	I8	5
II	I4	5
5	I2	3
3	6	I
I'	2	I
II'	2	0
I3	2	0
I0'	I	0
46	0	0
Итого	I87	I02

Из табл.3 видно, что резко выделяются своей употребительностью правила I; IO; O' и 20. На долю этих правил приходится около 70% всех реализаций и 85% реализаций в наиболее естественных - двух первых ответах. Некоторый вес имеют и правила I2; II и 5 (23% от всех реализаций и I2% от наиболее естественных реализаций). Остальные же правила, в общем-то игнорируются информантами (7% от всех реализаций и 2% от наиболее естественных реализаций).

4. Проблема факторов, обуславливающих предпочтения в выборе средств выражения заданного смысла.

В предыдущем разделе мы показали, что при реализации информантами правил перифразирования некоторым правилам отдается явное предпочтение. Как мы говорили в начале работы, эти предпочтения, по-видимому, обусловлены тем, что одни перифразы обладают определенными свойствами, а другие - нет.

Сопоставим между собой структуры перифраз, получаемых по различным правилам. Легко заметить, прежде всего, что они распадаются на две группы в зависимости от того, участвует или нет в данном правиле функция S_0. При этом оказывается, что три перифразы без S_0 (правила O', IO' и 20) составляют половину всех ответов, т.е. столько же, сколько IO перифраз, включающих S_0 (а среди "естественных" ответов их доля даже выше).

С другой стороны, обращает на себя внимание и тот факт, что ии. явно избегают перифраз с реализацией факультативного третьего актанта. Всего ответов этого рода было I3 (7% от общего числа реализаций правил), в том числе всего 2 раза - в естественных ответах.

Не представляет труда назвать еще ряд факторов, явным образом связанных с выявленными предпочтениями: актуальное членение*), стилевые характеристики, быть может даже частота слов (ср.различную употребительность перифраз с синонимами по правилу IO в табл.I) и т.д.

Однако простое перечисление факторов - еще не та цель, к которой мы стремились. Нам представлялось необходимым, кроме полу-

*) Некоторые правильные перифразы имеют другое актуальное членение, нежели исходное предложение, поскольку, как мы говорили выше, в существующих правилах данный фактор игнорируется.

чения самого перечня, установить:

а) какие факторы являются наиболее существенными? (установить иерархию факторов);

б) какие факторы являются независимыми, а какие - коррелированы между собой?

в) является ли найденная иерархия факторов одинаковой для различных стимулов и, что не менее важно, для различных информантов?

Подобный круг вопросов решается обычно в экспериментах по многомерному шкалированию, в которых ии. оценивают степень близости каких-либо стимулов по заданным параметрам. Так, из простого эксперимента, где ии. оценивали графическое сходство букв русского алфавита, удалось получить иерархию признаков, по которым происходит восприятие начертания букв [3], и, кроме того, обнаружить некоторые различия в иерархии признаков у разных ии. [4].

Обязательным условием таких экспериментов является сопоставление каждого стимула с каждым другим. Для нашего случая это означает необходимость проведения серии экспериментов с тем, чтобы каждая из синонимичных фраз данной группы была использована в качестве исходной. Организация такого исследования - задача чрезвычайно трудоемкая, хотя и не невыполнимая.

Однако здесь возникает трудность другого рода. В опытах по многомерному шкалированию ии. обычно получают стимулы и оценивают их близость (например, говорят, "похожи" или "непохожи" предъявленные стимулы). В нашем же случае ии. вообще не получают стимулов в явном виде (если считать стимулами всевозможные перифразы).

По-видимому, эксперименты по порождению синонимичных фраз можно было бы включить в парадигму многомерного шкалирования в том случае, если рассматривать предпочтения в порождаемых перифраз как непосредственно связанные с близостью перифраз к исходному предложению: чем "ближе" перифраза, тем "легче" (и, следовательно, тем чаще) она порождается. Тогда данные об употребительности тех или иных перифраз можно считать косвенными оценками близости их к исходному предложению, а если каждый раз менять исходное предложение, то можно получить оценки близости для всех пар перифраз, что и позволит применять методы многомерного шкалирования.

Для того, чтобы посмотреть, в какой степени данные по порож-

дению можно рассматривать как оценки "близости" перифраз. Мы провели специальное экспериментальное исследование, суть которого состояла в следующем.

Из числа правильных перифраз предложения М. победил Гр. было отобрано 13 фраз, которые были в разной степени употребительны в ответах ии. (в опыте по порождению): были взяты все наиболее предпочитаемые перифразы и - выборочным образом - малоупотребительные перифразы, в том числе две правильные перифразы, ни разу не встретившиеся в ответах. В табл.4 выбранные стимулы представлены в виде соответствующих им номеров правил и подправил; сами перифразы могут быть воспроизведены из табл.I по приведенным номерам.

Кроме этого, из ответов ии. в опыте по порождению были отобраны две фразы, не являющиеся правильными перифразами. Одна - М. нанес удар Гр. представляет собой ошибку типа А, другая - Гр. получил поражение от М. - ошибку типа Б. Эти фразы были включены в экспериментальный набор для контроля.

Из указанных 15 стимулов были составлены всевозможные пары, после чего пары поочередно предъявлялись ии. с инструкцией указать, какая из них "ближе", "более похожа" на него (карточка с фразой М. победил Гр. постоянно находилась перед ии.). В опыте участвовало 50 ии. примерно того же состава, что в описанном выше эксперименте по порождению.

Данные обрабатывались по известной процедуре парных сравнений. В результате мы получили упорядочение наших 15 фраз по степени близости к исходному предложению. Стимулы в табл.4 выписаны в порядке от самых близких к исходному до самых далеких от него (кстати, ии. довольно единодушно расценивали неправильные перифразы как наиболее непохожие на исходное предложение).

Теперь выпишем, сколько раз встретились каждая из 15 фраз в ответах ии. в опыте по порождению*). (графа 3 табл.4).

---

*)

Будем учитывать только наиболее естественные - I и 2 - ответы ии.

Табл.4

№ пп.	Стимулы	Ранг поблизости	Число ии. давших данный ответ
I	I0; (а)	I	I2
2	I; (а)	2	28
3	I0; (г)	3	3
4	20 (а)	4	I5
5	II	5	5
6	0'	6	24
7	I2; (а)	7	5
8	3; (б)	8	I
9	5 (в)	9	I
I0	4в	I0	-
II	I0' (г)	II	-
I2	II'	I2	-
I3	I3; (а)	I3	-
I4	Ошибка Б	I4	I
I5	Ошибка А	I5	2

Из табл.4 видно, что в общем результаты порождения хорошо согласуются с оценками близости: 7 перифраз, оцениваемых как наиболее близкие к исходному предложению были названы в 92 ответах из 97; коэффициент ранговой корреляции Спирмена ρ между оценками близости и числом ответов составляет 0,86 (существенно на более, чем 0,0I уровне).

Правда, имеются и некоторые расхождения данных. Например стимул № 3 (М. обыграл Гр.) был оценен, как весьма близкий к исходному предложению, но довольно редко встречался в ответах. По-видимому, фактор стиля*) существенно влияет на порождение и менее значим при оценке близости.

Однако, несмотря на это, мы полагаем, что описанный эксперимент позволяет заключить, что данные опытов по порождению могут быть включены в парадигму многомерного шкалирования.

*) Данное предложение многими расценивалось как "стилистически не очень хорошее".

Литература

I. Жолковский А.К. и Мельчук И.А. О семантическом синтезе."Проблемы кибернетики", вып.I9, I967.

2. Мельчук И.А. Опыт теории лингвистических моделей "Смысл Текст". М.,I974.

3. Фрумкина Р.М., Василевич А.П., Андрукович П.Ф. и Герганов Е.Н. Прогноз в речевой деятельности. М., I974.

4. Frumkina R.M., Andrukovich P.F, Terekhina A.Yu. Computational methods in the analysis of verbal behavior. In : Proceedings of the International Conference on Computational Linguistics. Pisa, 1973.

Авторы посвящают эту работу
А.А.Реформатскому в день
его семидесятипятилетия

Е.Н.Герганов, А.П.Василевич, Р.М.Фрумкина

К проблеме получения оценок ассоциативной силы буквосочетаний

I. Постановка задачи.

При обсуждении проблем, связанных с определением субъективной вероятности буквосочетаний методами психологического шкалирования, нам приходилось отмечать наличие корреляции между оценками субъективных вероятностей буквосочетаний (главным образом, триграмм) и другой характеристикой буквосочетаний - так называемой ассоциативной силой (Фрумкина, 1971).

Под ассоциативной силой (далее - АС) принято понимать степень трудности, с которой человек преобразует данное буквосочетание в осмысленные единицы языкового хода - слова и словосочетания. В качестве экспериментального аналога "трудности" преобразования используют, главным образом, число слов-ассоциаций, выданных в качестве ответа на предъявления данного буквосочетания в ассоциативном эксперименте классического типа. "Ассоциативная сила" есть некая характеристика, присущая каждому буквосочетанию (в том числе и неосмысленному). Буквосочетания могут быть упорядочены по АС, т.е. их можно сравнить по АС и указать, какое из них обладает большей АС по сравнению с другими буквосочетаниями.

Как известно, различие неосмысленных слогов по АС было обнаружено еще Эббингаузом, который обратил внимание на связь между количеством ассоциаций, вызываемых данным слогом, и результатами запоминания. Поиски надежных методов измерения АС впоследствии были обусловлены тем, что, как было показано, АС прогнозирует не только результаты запоминания (в частности, в опытах по парно-ассоциативному обучению - см. Underwood and Shulz , 1960), но также и характер речевого поведения в других ситуациях - например, результаты слухового и зрительного распознавания неосмысленных буквосочетаний (Taylor , 1956; Solomon, 1965).

Из приведенного выше определения АС следует, что ее вид полностью определяется используемой экспериментальной процедурой. Разными исследователями было предложено несколько методов получения численных оценок АС в эксперименте.[X] В настоящее время наибо-

[X] Обзор работ, посвященных разным методам получения оценок АС, см. Гергаков, 1971.

лее популярным из этих методов является метод Глейза, согласно которому для каждого буквосочетания подсчитывается некоторый показатель m, служащий мерой АС этого буквосочетания. Составленные Глейзом таблицы m для большого числа неосмысленных триграмм английского языка (Glaze , I928) в англо-язычной научной традиции являются необходимым пособием для каждого экспериментатора, работающего с вербальным материалом.

В данной работе мы поставили своей задачей проанализировать эту наиболее распространенную оценку АС неосмысленных буквосочетаний с точки зрения ее обоснованности (validity) и надежности. Кроме этого, мы изложим результаты сравнительного анализа эффективности оценки m и ранее предложенной нами новой меры АС - H' (Василевич, Герганов и др., I97I).

П. Оценка m по Глейзу.

Напомним, что представляет собой экспериментальная процедура, с помощью которой определяется число m Глейза.

В ассоциативном эксперименте испытуемому последовательно предъявляется набор буквосочетаний; в ответ на предъявление каждого из них и. просят сообщать слово-ассоциацию, первым пришедшее ему в голову. Время обдумывания ответов ограничивается; если и. не успевает за это время придумать ассоциацию, экспериментатор регистрирует "отказ". Предполагается, что чем большей АС обладает данное буквосочетание, тем меньшее число отказов будет зарегистрировано для него в ассоциативном эксперименте. При проведении опыта указанным способом оценка m буквосочетания измеряется по формуле

$$m_A = \frac{100\ (N-x)}{N}\ \%\ , \qquad (1)$$

где

N - число ии., участвовавших в ассоциативном эксперименте;

x - число ии., не успевших придумать ассоциацию за отведенное им время (т.е. число отказов, полученных для данного буквосочетания).

Таким образом, оценка m_A представляет собой процент ии., давших ассоциацию на триграмму A за некоторое фиксированное время t_i , в силу чего в приведенной формуле следовало бы, вообще говоря, писать $m_A(t_i)$. Обратим внимание на то, что процеду-

ра эксперимента по получению оценок m в основных чертах воспроизводит процедуру классического эксперимента по измерению времени словесной ассоциативной реакции. Разница состоит в том, что в эксперименте по определению m и. отведен ограниченный интервал времени для ответа, а в экспериментах по измерению времени словесной ассоциативной реакции время обдумывания ответа не ограничивается, поскольку именно оно и является измеряемой переменной.

Для большей ясности напомним, как организована экспериментальная процедура при измерении времени словесной ассоциативной реакции. И. предъявляется стимул: слово или неосмысленное буквосочетание. В ответ на предъявление стимула и. должен сообщить Э. первое же слово - ассоциацию, пришедшее ему в голову. Э. побуждает и. отвечать как можно быстрее, чтобы исключить возможность перебора возможных ответов, так как это обесценило бы эксперимент. Интервал времени от момента предъявления стимула до начала произнесения ответа называется латентным периодом словесной ассоциативной реакции (далее - ЛП). Именно ЛП и является регистрируемой переменной. Далее обычно сравниваются ЛП одного и. на разные стимулы; иногда изучаются средние показатели по группе ии. для стимулов определенных категорий. Многочисленные экспериментальные данные, полученные разными авторами, позволяют считать, что ЛП данного стимула А, измеренный по показаниям однородной группы ии. представляет собой нормально распределенную случайную величину (Фресс и Пиаже, 1966).

Рассмотрим эксперимент, состоящий в том, что триграмма А была предъявлена n ии., и соответственно было произведено n замеров ЛП. Результаты эксперимента можно представить в виде гистограммы (см.рис. 1), где по оси x отложены значения ЛП в секундах, а по оси y - число ответов с данными значениями ЛП. По данной гистограмме можно аппроксимировать кривую распределения случайной величины ЛП. Чтобы рассмотреть в этих же терминах эксперимент по получению оценки m, надо сделать следующее: выделим на оси x точку t_1; пусть это будет тот максимальный отрезок времени, который в данном эксперименте отведен на обдумывание ответа. Очевидно, что за это время будет получено число ответов, приближенно равное площади под кривой, ограниченной абсциссой от 0 до t_1. Переходя к значениям, полученным в эксперименте, мы можем записать

$$m_A(t_1) = \frac{K}{N}\%,$$

где K – площадь гистограммы, ограниченная t_1, т.е. число ии., давших ответ за время t_1, а N – площадь всей гистограммы, т.е. общее число ответов, которое было бы получено при неограниченном t, поскольку принимается, что за это время все N ии. дали бы какой-либо ответ.

Итак, связь между m и ЛП имеет следующий характер: m триграммы приближенно равно площади под кривой распределения ЛП данной триграммы, t_1 соответствует вероятности того, что значение ЛП лежит в интервале от 0 до t_1, где t_1 – максимальное время ответа, отведенное ии. в данном эксперименте. Отсюда, между прочим, ясно, почему не имеет смысла сопоставление значений m, полученных для одной триграммы в разных экспериментах: достаточно небольшого расхождения в выборе точки t_1, как значение m существенно меняется: чем дальше t_1 от 0, тем больше m для той же триграммы.

Естественно допустить, что если ЛП любой триграммы есть случайная величина, распределенная приближенно по нормальному закону, то параметры распределения для разных триграмм, т.е. средний ЛП и дисперсия ЛП, скорее всего, окажутся разными. Это, в свою очередь, означает, что для двух триграмм площадь под нормальной кривой, ограниченная заданной абсциссой, может оказаться как разной, так и одинаковой: это будет зависеть от взаимного расположения на оси точки t_i и точки, соответствующей среднему ЛП, а также от дисперсии среднего ЛП.

Рассмотрим сказанное более подробно. Проанализируем распределение ЛП двух триграмм А и Б, где при фиксированном t_i площадь Φ, приближенно соответствующая значению m, оказывается различной: $\Phi(А) \neq \Phi(Б)$. Обозначим средний ЛП триграммы А – $\bar{t}_A$, средний ЛП триграммы Б – $\bar{t}_Б$ и, соответственно, дисперсии ЛП каждой из триграмм σ^2_A и $\sigma^2_Б$.

Условие, согласно которому t_i выбрано так, что $\Phi(А) > \Phi(Б)$, может быть реализовано при следующих видах соотношений между средними и дисперсиями обоих распределений АП:

1. Средние ЛП равны, дисперсии ЛП различны:

$$\bar{t}_A = \bar{t}_Б, \quad \sigma^2_A \neq \sigma^2_Б$$

2. Средние ЛП различны, дисперсии ЛП равны:

$$\bar{t}_A \neq \bar{t}_Б, \quad \sigma^2_A = \sigma^2_Б$$

3. И средние ЛП, и дисперсии ЛП различны:

$$\bar{t}_A \neq \bar{t}_Б \ , \qquad \sigma_A^2 \neq \sigma_Б^2$$

Рассмотрим каждую ситуацию в отдельности, обращая внимание на то, какие значения может принимать t_i и как выбор значения t_i влияет на значение Φ (приближенно равного оценке m).

Ситуация I (рис. 2) : $\bar{t}_A = \bar{t}_Б$; $\sigma_A^2 < \sigma_Б^2$.

Очевидно, что для этого случая различие триграмм по m есть следствие различия в дисперсиях. Выберем теперь $t_i = t_2$ Мы получим, что поскольку $\sigma_A^2 < \sigma_Б^2$ и $t_2 > \bar{t}_A = \bar{t}_Б$, то $\Phi_A > \Phi_Б$. Следовательно, при $t_i = t_2$ $m_A(t_2) > m_Б(t_2)$.

Рассмотрим теперь, что получается, если мы выберем $t_i = t_1$; $t_1 < \bar{t}_A = \bar{t}_Б$ (см. t_1 на рис. 2). В этом случае $\Phi_A < \Phi_Б$, т.е. для тех же самых триграмм при времени обдумывания ответа $t_1 < \bar{t}_A = \bar{t}_Б$ мы получим противоположный результат, а именно $m_A(t_1) < m_Б(t_1)$.

Из сказанного следует, что при условии $\bar{t}_A = \bar{t}_Б$ и $\sigma_A^2 \neq \sigma_Б^2$ попытка сравнить триграммы по АС с помощью оценок m приводит к противоречивым выводам.

Ситуация 2 (рис. 3) : $\bar{t}_A \neq \bar{t}_Б$; $\sigma_A^2 = \sigma_Б^2$.

Очевидно, что в этом случае различие триграмм по m есть следствие различия в их средних ЛП. При этом, если $\bar{t}_A < \bar{t}_Б$, то для любого t_i мы будем иметь $\Phi_A(t_i) > \Phi_Б(t_i)$, т.е. $m_A(t_i) > m_Б(t_i)$.

Ситуация 3: $\bar{t}_A \neq \bar{t}_Б$; $\sigma_A^2 \neq \sigma_Б^2$.

Очевидно, что данная ситуация является более сложным случаем отношений между $\bar{t}$ и σ^2 . Рассмотрим возможные случаи отношения между $\bar{t}_A, \bar{t}_Б$, σ_A^2 и $\sigma_Б^2$ отдельно.

Случай 3-I: $\bar{t}_A < \bar{t}_Б$; $\sigma_A^2 < \sigma_Б^2$ (рис. 4а).

Данный случай является объединением ситуаций I и 2. При этом возможны такие соотношения t_i и σ^2 для ЛП триграмм А и Б, при которых выбор точки t_i привел бы к противоречивым выводам относительно m . Например, если мы выберем $t_i = t_1$ (рис. 4а), то, очевидно, $m_A(t_1) < m_Б(t_1)$.
Если же выберем $t_i = t_2$ (то же рисунок), то $m_A(t_2) > m_Б(t_2)$.

Случай 3-2: $\bar{t}_A < \bar{t}_Б$; $\sigma_A^2 > \sigma_Б^2$ (рис. 4б).

С точки зрения того, возможны ли случаи соотношений $\bar{t}_A, \bar{t}_Б, \sigma_A^2$ и $\sigma_Б^2$ таких, что выбор точки $t_i = t_1$ или $t_i = t_2$ привел бы к противоречивым выводам относительно m , случай 3-2 пол-

Рис. 1

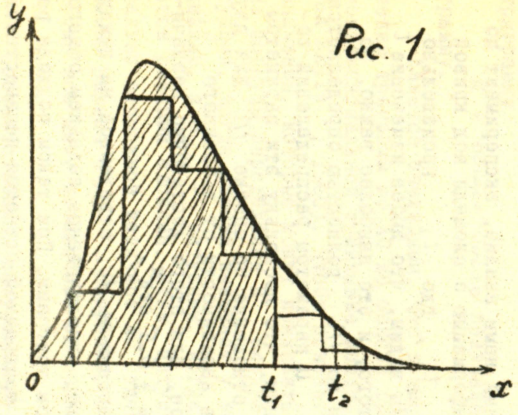


Рис. 3

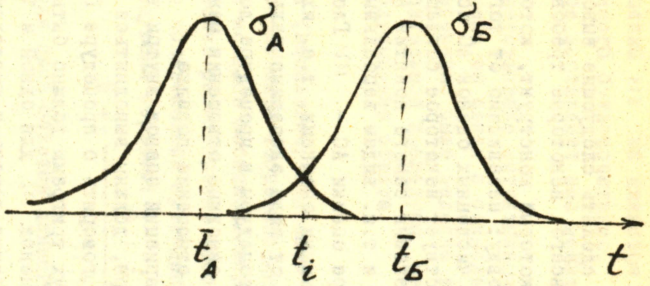


Рис. 2

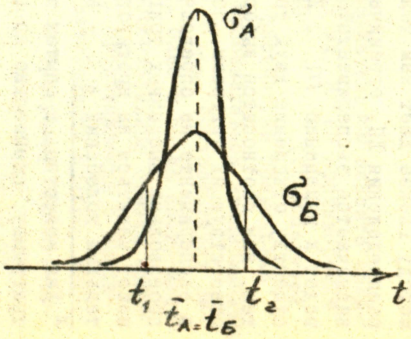


Рис. 4а

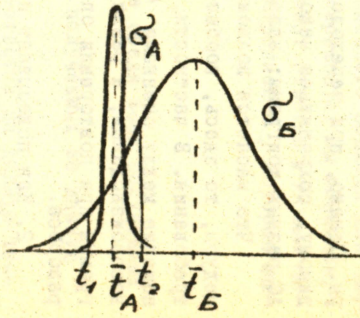
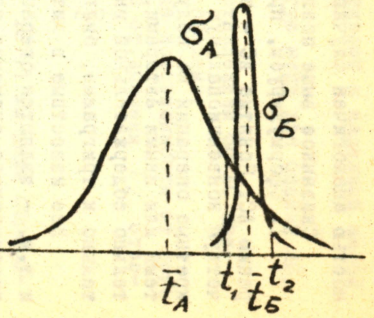


Рис. 4б



ностью аналогичен случаю 3-1.

Сказанное выше позволяет сделать следующие выводы.

I. Сформулируем, прежде всего, некоторые требования к измерению АС. Понятие АС – это некоторый конструкт, который имеет четкую конвенциональную трактовку: независимо от того, какие конкретные операции по получению численных оценок АС используются теми или иными авторами, существует некоторое соглашение относительно содержательной интерпретации самого понятия АС применительно к триграмме, биграмме и т.п. видам вербальных стимулов.

Все известные в литературе оценки АС – m Глейза, Д'Нобла и т.д. – являются операциональными оценками, т.е. их точное количественное значение зависит от вида экспериментальной процедуры. Однако, при переходе от процедуры к процедуре должны, по-видимому, сохраняться некоторые основные отношения между оценками АС разных триграмм: например, отношение порядка.

Что касается возможных вариаций оценок внутри выбранной процедуры, то здесь, вообще говоря, должны выполняться более строгие требования. В частности, если говорить о процедуре Глейза, то отношение между оценками m двух триграмм должно быть стабильно воспроизводимым, а различие оценок m для одной и той же триграммы при повторении опыта должно лежать в пределах случайного разброса.

2. Как показывает проделанный нами анализ, эксперимент по методу Глейза дает нам только информацию о площади под кривой распределения ЛП, ограниченного t_i, где t_i произвольно назначается экспериментатором. Мы видим, что любое изменение t_i ведет к изменению m. Кроме того, и это наиболее важно, по существу отношение двух триграмм по m полностью определяется тем, как соотносится выбор t_i и параметры распределений соответствующих триграмм. Подчеркнем, что поскольку эти параметры в эксперименте не определяются, отношение по оценке m для двух произвольно выбранных триграмм в зависимости от выбранного t_i может меняться на противоположное, для этого достаточно незначительных колебаний в t_i – порядка $\pm$ 0,5 сек[x].

[x] Как известно, в опытах по измерению m t_i, отведенное на каждую триграмму, специально не замеряют, ограничиваясь заданием общего времени работы и. со всем набором триграмм. При таком подходе реальное время работы и. с данной триграммой остается неизвестными. Как показали наши наблюдения, при указанной схеме опыта различия между реальными t отдельных стимулов будут порядка $\pm$ I сек.

Итак, оценка m не является эффективной и надежной мерой АС неосмысленного буквосочетания.

Сказанное заставляет нас искать другую меру для оценки АС, которая бы:

а) соответствовала интуитивным представлениям об АС триграмм как о мере ее связи с осмысленными словами;

б) одновременно была надежна в смысле независимости от деталей экспериментальной процедуры, например от времени t_i, отведенного на обдумывание ответа.

Ш. Оценки H′

В работе (Василевич, Герганов и др., 1971) мы дали обоснование целесообразности теоретико-информационного подхода для численной оценки исходов ассоциативного эксперимента. Мы показали, что для сравнения разнообразия ассоциативных ответов у разных триграмм удобно пользоваться приведенной энтропией H , которая вычисляется следующим образом:

$$H = -\frac{\sum_{i=1}^{n} P_i \log_2 P_i}{\log_2 K},$$

где P_i - доля слов-ответов данного типа среди всех полученных слов-ответов;

n - число различных слов-ответов, полученных на данную триграмму

K - общее число полученных слов-ответов на данную триграмму.

Кажется очевидным, что в чисто содержательном аспекте сравнение триграмм по АС на основе сравнения их H′ не противоречит конвенциональной трактовке конструкта АС (ср. Underwood and Shulz , 1960). Чтобы показать это, рассмотрим предельные значения, которые может принимать H′ в ассоциативном эксперименте. Из формулы определения H′ следует, что H'_{min} соответствует ситуации, когда стимул вызывает у всех ии. одно и то же слово-ассоциацию, а H'_{max} - ситуации, когда все ассоциативные ответы являются разными словами. Итак, интуитивному представлению о максимально возможной величине АС будет соответствовать оценка, численно равная H'_{min} . Например, в ответ на триграмму ЗЕБ практически все ии., независимо от объема группы, дают слово Зебра.

Соответственно, интуитивному представлению о минимально воз-

можной величине АС будет соответствовать оценка, численно равная H'_{max}. В эксперименте едва ли удастся получить чистый предельный случай, но примером может служить энтропия ответов на триграмму УКИ, где почти все ответы оказались разными словами ($H'_{max} = 1{,}0$; $H'_{УКИ} = 0{,}953$ - см. Василевич, Герганов и др., 1971).

Рассмотрим теперь, какую интерпретацию получают значения H'_i лежащие между H'_{min} и H'_{max}.

Например, H' триграммы НИВ = 0,185, а H' триграммы СУН = 0,696. Это означает, что на НИВ ответы ии. более однообразны: так в указанном эксперименте 108 ии. из 140 дают ассоциацию НИВА, Что касается триграммы СУН, то из 130 ии. максимальное число ии., которое дали одинаковый ответ (СУНДУК) достигает всего 19 чел. В этом смысле можно полагать, что связь между НИВ и словом НИВА "сильнее", чем связь между СУН и словом СУНДУК. Это и имеется в виду, когда говорят, что НИВ "легче преобразуется в слово", чем СУН.

Содержательный анализ изменения величины H в зависимости от изменения структуры множества ответов в ассоциативном эксперименте, позволяет считать, что мера H' удовлетворяет требованию (а).

Рассмотрим какими свойствами должна была бы обладать мера H', чтобы удовлетворять требованию (б).

Исходя из схемы ассоциативного эксперимента по методу Глейза, следовало бы потребовать, чтобы H' обладала свойством независимости от времени t_i, отведенного на обдумывание ответа. Иными словами, следовало бы убедиться, что, по крайней мере, упорядочение триграмм по H' не изменяется с изменениями t_i.

С этой целью нами был проведен специальный эксперимент, к описанию которого мы переходим.

IУ. Эксперимент по измерению ЛП словесной ассоциативной реакции.

1. Постановка задачи.

В соответствии с изложенными выше соображениями относительно надежности оценок m и H', мы поставили для экспериментального исследования следующие основные задачи:

I°. Исследовать характер зависимости H' от t и показать, в связи с этим, какими преимуществами обладает оценка H' перед оценкой m.

2°. Исследовать зависимость оценки m от времени t.

для различных трехбуквенных сочетаний русского языка.

2. Методика и проведение эксперимента.

Для проведения эксперимента было отобрано 36 триграмм, для которых по предварительным данным можно ожидать, что они отличаются по m и H'.

Опыт проходил следующим образом. Испытуемому устно сообщалась стандартная инструкция (см. выше) и давалось два-три примера для проверки ее усвоения. После этого предъявлялись тестовые триграммы. Отказы запрещались; экспериментатор побуждал и. давать ассоциацию на каждую триграмму. В момент предъявления стимула включался секундомер; он останавливался в момент произнесения ответа. Для каждого ответа фиксировалось время латентного периода ответа данного и.

Эксперимент проводился индивидуально с каждым и. Ответы ии. записывались экспериментатором.

В опыте участвовало 110 и., преимущественно учащихся 9-10 классов. В результате эксперимента для каждой из 36 триграмм было получено по 110 ответов-ассоциаций и время латентного периода каждого ответа.

3. Результаты и обсуждение.

а) Получение оценок H и выяснение зависимости между H' и t_i.

Процедура подсчета величины H позволяет определять ее для разных t_i - при условии, что на данном t_i число отказов не слишком велико (при наличии, например, 100 отказов из 110 возможных ответов достоверность величины H' вызывает сомнения, поскольку ее подсчет опирается на выборку всего из 10 ответов). На рис. 5 на примере пяти случайно отобранных триграмм для пяти произвольно выбранных интервалов t_i показано, как изменяется оценка H' с изменением времени t.

Как видно из рис. 5, величина H' не зависит от времени t_i, для которого она подсчитана: колебания H' для разных t_i весьма незначительны и не выходят за пределы случайного разброса. Так, H' триграммы ОБК составляет 0,68 для t = 1,9 сек. и 0,66 для t = 75 сек., для триграммы ВЗЯ аналогичные значения H' составили 0,39 и 0,37 и т.д. Для того, чтобы полнее показать динамику изменений H' в зависимости от t_i, приведем для примера значения H' триграммы ОБК для 10 интервалов: 1,4 сек; 1,9 сек.; 2,4 сек.; 2,9 сек.; 3,4 сек.; 3,9 сек.; 4,4 сек.; 4,9 сек.;

5,4 сек.; и 75 сек. Значения H' для приведенных интервалов составило соответственно 0,6I; 0,63; 0,62; 0,6I; 0,59; 0,63; 0,64; 0,64; 0,63 и 0,66. Таким образом, независимо от того, сколько времени затрачивают ии. на обдумывание ответа, H' буквосочетания остается величиной постоянной.

б) Получение оценок m и выяснение зависимости m от t_i.

В соответствии с поставленной задачей необходимо было для каждой триграммы получить значения m при разных значениях t_i. В рамках схемы проведенного эксперимента возможность такого подсчета обуславливалась следующим. Пусть, например, требуется вычислить значение m триграммы А для времени t_1. Анализ имеющихся в нашем распоряжении данных позволяет определить латентный период работы с триграммой А для каждого из II0 ии. Пусть у K ии. латентный период превышает t_1. Разумно предположить, что если бы время работы ии. с триграммой А было бы ограничено и составляло t_1, то было бы зафиксировано K отказов[x]. Таким образом, в этом случае

$$m_A(t_1) = \frac{100\ (110 - K)}{110}\ \%$$

Рассуждая аналогичным образом, мы можем определить m любой триграммы для любого t_i в диапазоне от 0,5 сек. до 75 сек. (данный диапазон включает в себя минимальный и максимальный латентный периоды, зафиксированные в нашем опыте). Для того, чтобы иметь возможность сопоставлять оценки m и H', значения m для всех стимулов были подсчитаны для тех же t_i, что и значения H'. На рис. 6 для пяти выбранных триграмм показано, как изменяется m с увеличением t.

Как видно из рис. 6, в большинстве случаев увеличение времени t даже на незначительную величину, приводит к весьма ощутимому изменению m. Так, m (t = I,9) триграммы УКЕ составляет I8,2%, а ее m (t = 2,9) - 4I,8%. Тем самым мы ответили на вопрос о том, в какой мере колебания t_i сказываются на изменениях m. Отсюда, в частности, следует, что если условия эксперимента допускают колебания t в пределах ± I сек., то это может привести к кардинальным ошибкам в оценках величины m.

---

x) Мы не рассматриваем здесь вопросов, связанных с психологическим воздействием на ии. факта ограничения времени.

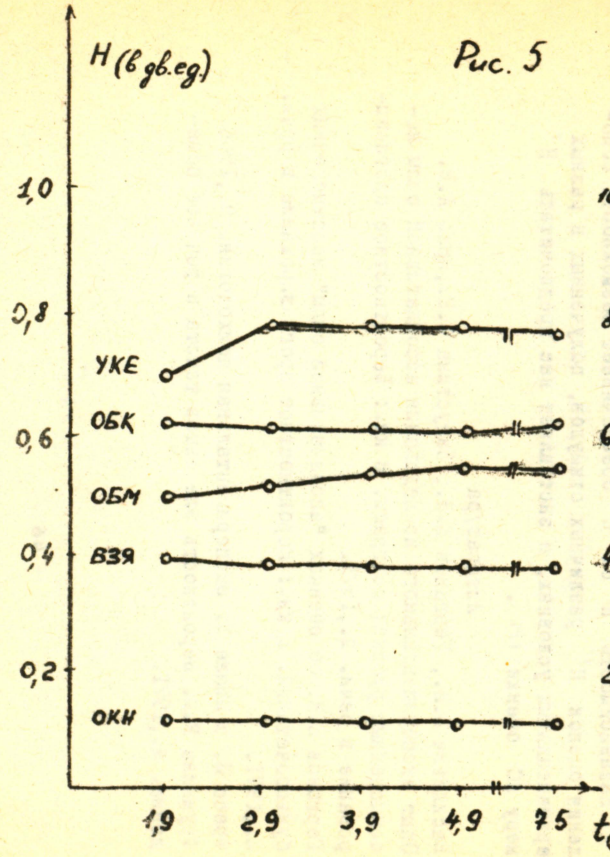

H (в дв.ед.)
1,0
0,8
0,6
0,4
0,2
УКЕ
ОБК
ОБМ
ВЗЯ
ОКН
1,9
2,9
3,9
4,9
7,5
t (сек)

Рис. 5

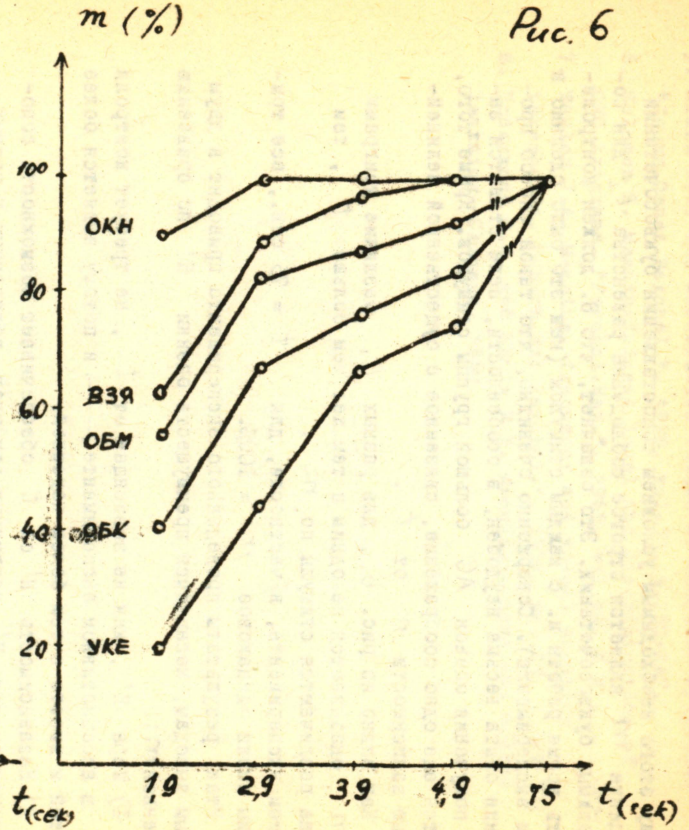

m (%)
100
80
60
40
20
ОКН
ВЗЯ
ОБМ
ОБК
УКЕ
1,9
2,9
3,9
4,9
7,5
t (сек)

Рис. 6

В силу этого необходимым условием сопоставления буквосочетаний по оценке m является строгое соблюдение равенства t_i для сопоставимых буквосочетаний. Это означает, что Э. должен контролировать время работы и. с каждым стимулом (как это было сделано в нашем эксперименте). Совершенно очевидно, что такой способ проведения опыта весьма неудобен, в особенности, если ставится задача получения оценок AC большой группы стимулов. Кроме того, имеется еще одно соображение, связанное с существенной нелинейностью зависимости m от t.

Как видно из рис. 6, для разных t_i различие триграмм по m оказывается не одним и тем же: чем больше t, тем меньше различаются стимулы по m.
В нашем эксперименте, в частности, для t = 75 сек., все триграммы имели одинаковое m = 100%.

Итак, результаты проведенного эксперимента приводят к двум важным выводам, касающимся преимуществ оценки H' по сравнению с оценкой m:

I) Мера H', как не зависящая от t, не требует контроля t в ассоциативном эксперименте (——) и потому является более простой и вместе с тем более надежной.

2) Независимость H' от t обеспечивает возможность сопоставления оценок H' различных стимулов, полученных в разных экспериментальных условиях, и заставляет нас предпочитать H' как меру AC оценке m.

Литература

I. Василевич А.П., Герганов Е.Н., Загустина С.Е., Ирхо А.В. Опыт экспериментального исследования ассоциативной силы неосмысленных русских триграмм. В кн.: Вероятностное прогнозирование в речи. М.,1971.

2. Герганов Е.Н. Об оценках "ассоциативной силы" неосмысленных буквосочетаний. В кн.: Вероятностное прогнозирование в речи. М.,1971.

3. Фресс П. и Пиаже Ж. Экспериментальная психология. М.,1966.

4. Фрумкина Р.М. Вероятность элементов текста и речевое поведение. М.,1971.

5. Glaze A.I. The association values of nonsense syllables "Journ.Genet.Psychol.", 1928, vol.35, p.255-267.
6. Solomon S. Meaningfulness level and discrimination of CVC trigrams. "Psychol.Rep.", 1965, vol.16, N 2 .
7. Taylor J.A. Meaning, frequency and visual duration thresholds. "Journ.Exp.Psychol.", 1958, vol.55, p.329-334
8. Underwood B.J. and Shulz R.W. Meaningfulness and verbal learning. Chicago, 1960.

ИРЯ
институт русского языка
им. В.В.Виноградова
БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР
www.ruslang.ru

Ответственный за выпуск Л.А.Гончар.

А05171 от 16.06.75 г. 3 п.л. 190 экз. Зак. 654

Тип. МГПИИЯ

301
Бесплатно