

ЛЕКСИКО-ФУНКЦИОНАЛЬНЫЕ ОТГЛАГОЛЬНЫЕ СУЩЕСТВИТЕЛЬНЫЕ СО ЗНАЧЕНИЕМ НАЧАЛА ДЕЙСТВИЯ И ИХ ОБРАБОТКА В СИСТЕМЕ МАШИННОГО ПЕРЕВОДА ЭТАП-3

А.С. Андреева

МГУ им. М. В. Ломоносова

andreyevs@mtu-net.ru

В работе рассматриваются лексические функции (ЛФ), обозначающие начало действия. Показывается, в каких случаях значением ЛФ является и глагол, и производное от него существительное, а в каких – только глагол и почему. Также описывается, как отглагольные существительные – значения ЛФ используются в системе МП ЭТАП-3 и какие в связи с этим возникают трудности.

Лексические функции – особые отношения, «характеризующие идиоматичную взаимозаменяемость и сочетаемость лексем» [ТКС 1984]. Большинство лексических функций являются синтагматическими, т.е. описывают сочетаемость слов в тексте. При этом почти у всех ЛФ этого вида аргументами являются существительные, а значениями – глаголы.

От подавляющего большинства глаголов – значений лексических функций – образуются отглагольные существительные. Но далеко не все они являются значениями ЛФ от тех же аргументов, что и соответствующие глаголы. В данной работе на примере лексических функций вида IncerOper будут рассмотрены отглагольные существительные – производные от глаголов – значений ЛФ.

Цель работы – выявить глаголы, производные от которых существительные являются значениями ЛФ, и объяснить, почему существительные, производные от остальных глаголов, не могут выступать в этом качестве.

В работе также показано, как можно использовать лексико-функциональные отглагольные существительные в системе машинного перевода ЭТАП-3 и какие в связи с этим возникают трудности.

Данная работа проводится с использованием словаря системы машинного перевода ЭТАП-3.

Лексические функции вида IncerOper (X) обозначают глагол поддержки, описывающий начало некоторого действия или состояния. Существует две функции вида IncerOper: IncerOper1 и IncerOper2.

IncerOper1 (X) - начать делать X, иметь X или быть в состоянии X (глагол поддержки, при котором

P1¹ выполняет функцию подлежащего, а X - функцию главного дополнения)²

Примеры:

ПОПУЛЯРНОСТЬ:приобретать

ОТЧАЯНИЕ:впадать <v1>

АЗАРТ: входить <v1>

IncerOper2 (X) - начать подвергаться X-у или быть в области действия X-а (глагол поддержки, при котором P2 выполняет функцию подлежащего, а X - функцию главного дополнения)

Примеры:

УВАЖЕНИЕ: снискивать / завоевывать

ПРОДАЖА: поступать <v1>

АГРЕССИЯ: подвергаться

В русском языке примерно у 54 процентов слов – аргументов лексической функции IncerOper1 в качестве значения функции могут выступать и глагол, и существительное³. В остальных 46 процентах случаев это могут быть только глаголы.⁴ Возникает вопрос, за счет чего получилась такая разница почти в 2 раза.

Из почти ста глаголов, являющихся значениями лексической функции IncerOper1, от трети существительные не образуются вообще (**группа (1)**). Но поскольку эти глаголы довольно редко выступают в качестве значений ЛФ (в основном по 1-2 раза), то в результате только у 8%

¹ PN – участник ситуации номер N. Т.е. P1 – первый, P2 – второй, P3 – третий.

² Описание лексических функций приводится по справочной документации к системе МП ЭТАП-3.

³ На самом деле глаголы являются значениями ЛФ IncerOper1, а существительные – функции S0_IncerOper1, но для простоты в работе эти две функции буду описываться как одна.

⁴ Приведенные в работе данные были получены из словаря системы МП ЭТАП-3.

аргументов лексической функции IncepOper1 существительное не является ее значением потому, что его просто не существует.

Остается 38% случаев, когда от глагола образуется отглагольное существительное, но оно не является значением ЛФ.

Такая ситуация очевидным образом имеет место потому, что значение отглагольного существительного не совпадает со значением глагола. Далее будет рассмотрено, в чем именно заключается эта разница.

Из оставшихся глаголов нашлось всего два (*начинать* и *заболевать*), производные от которых существительные меняют свое значение настолько, что в принципе не могут быть значениями лексической функции IncepOper1 (**группа (2)**). У значений существительных *заболевание* и *начинание* есть несколько отличий от значений соответствующих глаголов, но для данной работы важно то, что происходит с компонентом фазовости. *Заболевание* теряет его полностью (**заболевание гриппом*¹), а *начинание* в принципе сохраняет, но этот компонент значения смещается из центра на периферию: начинание – это такой поступок или такая деятельность, которая является началом серии таких же поступков или такой же деятельности или рассматривается говорящим как таковая. Таким образом, глаголы *начинать* и *заболевать* обозначают начало некоего действия или процесса, а производные от них существительные – то же действие или процесс целиком. В данном случае отглагольное существительное захватывает более широкое смысловое поле, чем соответствующий глагол.

Хотя в описываемую группу попало всего два глагола, но один из них сочетается с названиями болезней, которых в русском языке довольно много, поэтому на долю этой группы приходится 14% случаев наличия в качестве значения лексической функции IncepOper1 только глагола.

Таким образом, остается 24% случаев, когда производное от глагола существительное не является значением ЛФ IncepOper1.

Если имеется лексема Y, производная от лексемы X, то возможны 3 вида изменения значения Y-а относительно значения X-а:

- 1) значение Y-а шире значения X-а,
- 2) значение Y-а уже значения X-а,
- 3) значения X-а и Y-а пересекаются (пересечение может быть нулевым)².

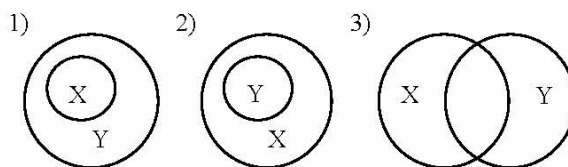


Рисунок 1.

Первый случай (значение отглагольного существительного шире значения исходного глагола) уже был рассмотрен выше, значит, оставшиеся 24 процента будут распределяться между вторым и третьим.

Среди еще не рассмотренных глаголов есть несколько, иллюстрирующих случай 3) и даже 1). Это, например, глагол *приходить*, производное от которого существительное *приход*, как отмечается в [Пазельская 2002], может также быть синонимом слова *визит*; или *добывать*, у которого вообще два производных существительных с разными свойствами (*добывание* и *добыча*). Но такое развитие дополнительного значения существительного не влияет на его способность или неспособность быть значением лексической функцией наряду с глаголом. Поэтому в рамках данной работы я буду условно считать, что значения всех оставшихся существительных охватывают меньшее смысловое поле, чем значение соответствующих глаголов.

Все оставшиеся глаголы можно разделить на 3 группы:

Группа (3) – глагол имеет только одно значение.

Группа (4) – глагол имеет прямое и переносное значение, и в обоих случаях он является значением ЛФ IncepOper1.

Группа (5) – глагол имеет прямое и переносное значение, и значением ЛФ является только последнее.

Группы (3)-(5), в свою очередь, делятся на подгруппы:

- а) производное от глагола существительное является значением ЛФ IncepOper1 всегда,
- б) производное от глагола существительное является значением ЛФ IncepOper1 не всегда,
- в) производное от глагола существительное не является значением ЛФ IncepOper1 никогда (т.е. существительное образуется только от прямого значения глагола, а от переносного – нет).

Примеры глаголов из групп (3)-(5) приведены в таблице 1.

	группа (3)	группа (4)	группа (5)
)	добывать организовывать нагуливать	получать брать	выдвигать выступать присоединяться
)	снискивать объявлять	принимать завоевывать	погружаться входить завязывать
	нет	нет	падать приводить

¹ Если так и можно сказать, то это значит «болеть гриппом», а не «заболевать гриппом»

² Значение X и Y могут еще полностью совпадать, но поскольку это обозначает, что значениями ЛФ будут и глагол, и существительное, для данной работы этот случай интереса не представляет.

Д			развязывать
---	--	--	-------------

Таблица 1.

В разделении глаголов на группы (3)–(5) можно выделить некоторые семантические закономерности – впрочем, достаточно предсказуемые.

В группу (3) попали глаголы либо с очень общим значением вида «начать делать X, иметь X или быть в состоянии X» (определение лексической функции IncerOper), которому как раз из-за того, что оно очень общее, трудно быть переносным (например, *устраивать* или *достигать*), либо, наоборот, глаголы с очень узким значением, не имеющие переносного значения в силу редкости употребления (например, *окрашиваться* или *нагуливать*).

В группе (4) находятся глаголы, также имеющие прямое значение вида «начать делать X, иметь X или быть в состоянии X», но несколько менее общее (например, *получать*, *приобретать*, *завоевывать*).

В группу (5), самую многочисленную, попали глаголы, которые не являются значением лексической функции IncerOper1 в прямом значении, но становятся им в переносном. А поскольку фактически это группа вида «все остальное», то здесь собралась очень разношерстная компания: несколько глаголов движения¹ (*выходить*, *пойти*, *уходить* и т.д.) и другие глаголы, которые никаким разумным образом объединить в классы не получается (*обращаться*, *развязываться*, *вырабатывать* и т.д.).

Как видно из таблицы 1, в рассматриваемом материале не удалось обнаружить глаголы, относящиеся к подгруппам (3в) и (4в). Возникает вопрос, насколько закономерно наличие этой лакуны.

Относительно группы (3) все очевидно: если у глагола только одно значение, а производное от него существительное не бывает значением ЛФ никогда, то такие глаголы попадают в группу (2). А вот в группе (4) ситуация интересней. В ней не просто отсутствует подгруппа в), в ней для глаголов, входящий в подгруппу б), верно следующее утверждение: если глагол используется в прямом значении, то производное от него отглагольное существительное является значением ЛФ почти всегда, а если в переносном, то только в некоторых случаях. Кроме того, в группе (4) нет глаголов, у которых производное существительное было бы

значением ЛФ только в прямом значении или, тем более, только в переносном.

Судя по всему, это связано с тем, что переносность значения в группе 4 совсем не такая, как, например, в группе (5). Если в группе (5) присутствует «настоящее» переносное значение, т.е. от первоначального значения глагола почти ничего не остается (ср. *развязывать шнурки* – *развязывать войну*), то в группе (4) значение фактически не меняется, только расширяется круг объектов, в отношении которых можно использовать данный глагол. Например, *получать приз* и *получать преимущество*. В группе (4) иногда даже бывает трудно провести грань между прямым и переносным значением. Таким образом, глаголы второй группы «размазаны» по шкале прямого-переносного значения, и существительному, изначально образованному от глагола в прямом значении, очень легко захватить также и переносное.

В группе (5) ситуация прямо противоположная: существует большая разница между прямым и переносным значением, поэтому существительному, образованному первоначально от прямого значения глагола, требуется существенное усилие для преодоления этого барьера – что подтверждается также и статистикой. Более чем у трети глаголов этой группы существительное в переносном значении не используется вообще.

Говоря о группе (5), следует заметить, что у всех входящих в нее глаголов, когда они употребляются в прямом значении, есть отглагольные существительные. Иногда они значат то же, что и глагол, иногда их значение расширяется (*встречать* – *встреча*) или, наоборот, сужается (*приводить* – *привод*, *приведение*), но употребление глаголов этой группы в прямом значении не имеет отношения к лексической функции IncerOper1, поэтому этот материал выведен за рамки рассмотрения в данной работе.

В связи с этим возникает следующий вопрос: если я считаю, что у таких глаголов прямого значения и производных от него существительных как бы не существует, то чем отличается подгруппа (5в) от группы (1) – глаголов, у которых производных существительных нет вовсе. Почему не стоит объединять группы (1) и (5в) в одну, я напишу немного позже.

Итак, глаголы, не попавшие в группы (1) и (2), были разделены на 3 группы, которые, в свою очередь, были разбиты в общей сложности на 7 подгрупп. Но если в распределении по группам (3)–(5) имеются некоторые, хотя и довольно тривиальные, семантические закономерности, то в распределении глаголов по подгруппам внутри групп таких закономерностей обнаружить не удалось. Так, например, глаголы движения встречаются в подгруппах (5а), (5б) и (5в). То есть

¹ Интересно отметить, что среди глаголов движения, являющихся значениями ЛФ IncerOper1, есть не только те, в прямом значении которых присутствует компонент начала действия, (*вступать*, *приходить* – переместившись в место X определенным образом, начать находиться в этом месте), но и глаголы с противоположным значением (*выходить*, *уходить*).

получается, что попадание глагола в одну из подгрупп группы зависит от случая.

В общем-то так и есть. Первоначально отглагольное существительное могло использоваться в качестве значения лексической функции только в некоторых контекстах, где глагол употреблялся в прямом значении, а потом началось расширение его значения. Захват новой области значения шел в двух направлениях: оставшаяся часть области прямого значения и область переносного значения (для тех глаголов, у которых оно есть) – сначала только с отдельными контекстами, как и в случае прямого значения, а потом полностью.

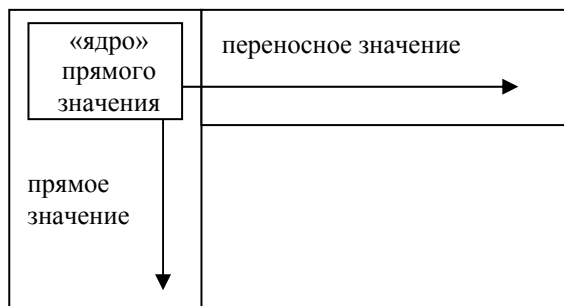


Рисунок 2.

Данные позволяют говорить о наличии полного спектра вариантов. Каждый конкретный глагол со своим существительным занимает свою часть представленной схемы. Некоторые лакуны объяснимы (например, **выхода на промысел*, в отличие от *выхода на охоту*, не существует потому, что в наше время люди крайне редко выходят на промысел, и для описания этой гипотетической деятельности хватит и одной конструкции), но в основном каждая конкретная ситуация довольно случайна.

Однако в данном случае интерес вызывает не то, в какую из подгрупп группы попадает тот или иной глагол, а то, что можно сказать о тех контекстах, где существительное вроде бы употребить нельзя. Большинство случаев, в которых отглагольное существительное не является значением лексической функции IncerOper1, можно охарактеризовать известной фразой из беседы лингвиста с информантом: «Сказать-то можно, только мы так не говорим,» - где ключевой является первая часть фразы. Подобные контексты не ощущаются как абсолютно запретные; они скорее сомнительны, но в принципе возможны. Еще лучше это видно на примере значений лексической функции IncerOper2, где большинство случаев, когда отглагольное существительное может быть значением ЛФ, сомнительно.

То, что в большинстве случаев вроде бы несуществующие конструкции с существительными-значениями ЛФ IncerOper хотя бы гипотетически возможны, и то, что носители русского языка расходятся во мнении относительно степени их допустимости, является дополнительным

подтверждением того, что процесс захвата отглагольными существительными дополнительных областей значения живой и сейчас в самом разгаре.

Именно поэтому и не стоит присоединять подгруппу (5в) к группе (1): у глаголов этой подгруппы есть производные существительные, и вполне возможно, что в ближайшем будущем они перешагнут рубеж, отделяющий прямое значение от переносного.

Что касается распределения оставшихся 24% аргументов ЛФ IncerOper1, не имеющих в качестве значения ЛФ существительного, то на долю подгруппы (3б) приходится всего 0,6% случаев, подгруппы (4б) – 3,8%, а оставшиеся почти 20% распределяются между подгруппами (5б) и (5в). При этом (5б) берет себе 12%, (5в) – 5,3%, а остальное достается смешанному случаю, когда у аргумента ЛФ есть несколько глаголов-значений ЛФ, часть из них принадлежит подгруппе (5в), часть подгруппе (5б), но образованные от последних существительные в данном контексте употребить нельзя. Наличие смешанного случая еще раз подтверждает случайность распределения глаголов по подгруппам внутри группы.

Статистические данные позволяют уточнить процедуру захвата отглагольным существительным более широкой области значения. На долю подгруппы (3б), где у глаголов есть только прямые значения, приходится смешные 0,6% случаев, когда существительное не может быть значением ЛФ. В подгруппе (4б) глагол в таких случаях намного чаще встречается в переносном значении, чем в прямом. Существительные, производные от глаголов из подгруппы (5б), при наличии у глагола прямого значения могут употребляться если не всегда, то, во всяком случае, намного шире, чем при переносном значении глагола, а наличие подгруппы (5в), где переносного значения у отглагольных существительных просто нет, говорит само за себя. Таким образом, отглагольное существительное, судя по всему, обычно сначала захватывает большую часть области прямого значения соответствующего глагола, а потом начинает интересоваться переносным.

Что касается английского языка, то в нем процент случаев, когда значением функции являются глагол и производное от него существительное, существенно ниже, чем в русском. Для функции IncerOper1 это 15-20%, а для функции IncerOper2 такие случаи вообще обнаружить не удалось. Разница в 30-35% по сравнению с русским языком обеспечивается за счет того, что либо производного существительного от глагола не существует вовсе, либо оно не используется в переносном значении. Проблем перевода в этом случае решается просто: любое отглагольное существительное русского языка, которое по каким-либо причинам не может быть передано в данной английской конструкции существительным, заменяется на герундий соответствующего глагола.

Обработки лексико-функциональных отглагольных существительных в системе машинного перевода ЭТАП-3 потребовала внесения некоторых изменений в систему. В данной работе глаголы и отглагольные существительные для простоты описывались как значения одной и той же лексической функции, но они не могут быть таким образом представлены в ЭТАПе.

Во-первых, глагол и существительное связаны с другими членами предложения с помощью разных синтаксических связей. Во-вторых, они по-разному переводятся: если аналогами русских глаголов в описываемых конструкциях в английском языке выступают глаголы или, в крайнем случае, комбинации вида «вспомогательный глагол + прилагательное», т.е. не меняется ни часть речи, ни синтаксическая конструкция, то русские отглагольные существительные переводятся с помощью существительных в очень небольшом проценте случаев, в основном они превращаются в герундии, т.е. при переводе меняется часть речи с вытекающим отсюда изменением синтаксической конструкции.

Все эти различия не позволяют описывать отглагольные существительные как значения той же ЛФ, что и глаголы, поэтому для существительных были введены новые функции: S0_IncepOper1 и S0_IncepOper2.

Начало систематической работы с лексико-функциональными отглагольными существительными в системе ЭТАП выявило как минимум одну серьезную проблему. Есть некоторое количество существительных, которые являются производными от двух глаголов сразу: возвратного и невозвратного¹. Например, *развитие* – от *развивать* и *развиваться*, *заражение* – от *заражать* и *заражаться* и т.д. В результате получается, что у существительного есть две актантные структуры и что оно является значением двух разных лексических функций от одного и того же аргумента. В настоящий момент ЭТАП-3 не позволяет завести для одного слова две разные актантные структуры. Это ограничение системы известно уже давно, но до сих пор оно вызывало сложности в единичных случаях (в основном при обработке глаголов), и эти сложности удавалось как-то разрешить, обычно путем введения в словарь вместо одной лексемы двух – с разными актантными структурами. Работа с лексико-функциональными существительными выявила, что такие случаи, во-первых, регулярны, во-вторых, многочисленны, и стало понятно, что проблему нужно решать каким-то разумным образом, а не клонированием лексики.

В данной работе были рассмотрены отглагольные существительные со значением начала действия. Было показано, в каких случаях они наряду с глаголами являются значениями лексической функции, а в каких нет и почему. Также было описана обработка этих существительных в системе МП ЭТАП-3.

В свете полученных результатов интересным представляется изучение лексико-функциональных отглагольных существительных с другими значениями, а также сравнение существительных разного вида между собой.

Литература:

[Падучева 1977] Падучева Е.В. «О производных диатезах отпредикатных имен в русском языке» // Храковский В.С. (ред.) Проблемы лингвистической типологии и структуры языка, Ленинград: «Наука», 1977. С. 84-107.

[Пазельская 2002] Пазельская А.Г. Аспектуальность и другие параметры значения русских предикатных имён // Москва: МГУ, дипломная работа, 2002.

[ТКС 1984] Мельчук И.А., Жолковский А.К. Толково-комбинаторный словарь современного русского языка // Вена, Wiener Slawistischer Almanach, Sonderband 14, 1984.

Help к системе машинного перевода ЭТАП-3.

¹ Это свойство некоторых отглагольных существительных неоднократно описывалось в работах по русистике – см., например, [Падучева 1977]