

Русский ФреймНет: к задаче создания корпусного словаря конструкций

О.Н.Ляшевская, Ю.Н.Кузнецова
Университет Тромсе, Норвегия

Предыстория

FrameNet как электронный лексический ресурс

Ч.Филлмор 1990-2000-е годы, английский

+ немецкий, испанский, японский языки

Фрейм – концептуальная структура, стоящая за значением слова и объясняющая его сочетаемостные свойства

Отражение человеческого опыта: структурированное множество ожиданий, связанных с тем или иным концептом; типовая ситуация с типовыми участниками.



Frequency
Friction
Front_for
Frugality
Fullness
Gathering_up
Gesture
Get_a_job
Getting_vehicle_underway
Getting
Getting_scenario
Getting_underway
Getting_up
Giving
Giving_in
Giving_scenario
Gizmo
Goal
Going_back_on_a_commitment
Gradable_attributes
Grant_permission
Grasp
Grinding
Grooming
Ground_up
Growing_food
Guest_and_host
Guilt_or_innocence
Hair_configuration
Halt
Have_as_requirement
Have_as_translation_equivalent

Main Menu[Home](#)[View FrameNet Data](#)[News](#)[Licensing FrameNet Data](#)[FAQs](#)[The Book \(PDF\)](#)[Presentations and Talks](#)[FrameNet Related Projects](#)[Publications and Papers](#)[FN2 Forum](#)[Staff](#)[Users](#)[Funding](#)[Current Project Status](#)

[Contact Us](#)[Template by Mambo Hut](#)**Login Form**

Username

Password

Giving

Definition:

A **Donor** transfers a **Theme** from a **Donor** to a **Recipient**. This frame includes only actions that are initiated by the **Donor** (the one that starts out owning the **Theme**). Sentences (even metaphorical ones) must meet the following entailments: the **Donor** first has possession of the **Theme**. Following the transfer the **Donor** no longer has the **Theme** and the **Recipient** does.

Barney **GAVE** **the beer** **to Moe**.

\$300 was **ENDOWED** **to the university** **to build a new performing arts building**.

FEs:**Core:****Donor** [Donor]The person that begins in possession of the **Theme** and causes it to be in the possession of the **Recipient**.**Recipient** [Rec]The entity that ends up in possession of the **Theme**.**Theme** [Thm]

The object that changes ownership.

Semantic Type

Physical_object

Non-Core:**Circumstances** [cir]The **Circumstances** are the conditions under which the **Theme** is given.

I **GIVE** **my services** **free of charge**.

Depictive [dep]A description of the **Donor**, **Recipient**, or **Theme** given independently of the giving event per se.**Manner** [Man]The **Manner** is the manner in which the **Donor** gives the **Theme** to the **Recipient**.**Semantic Type**

Manner

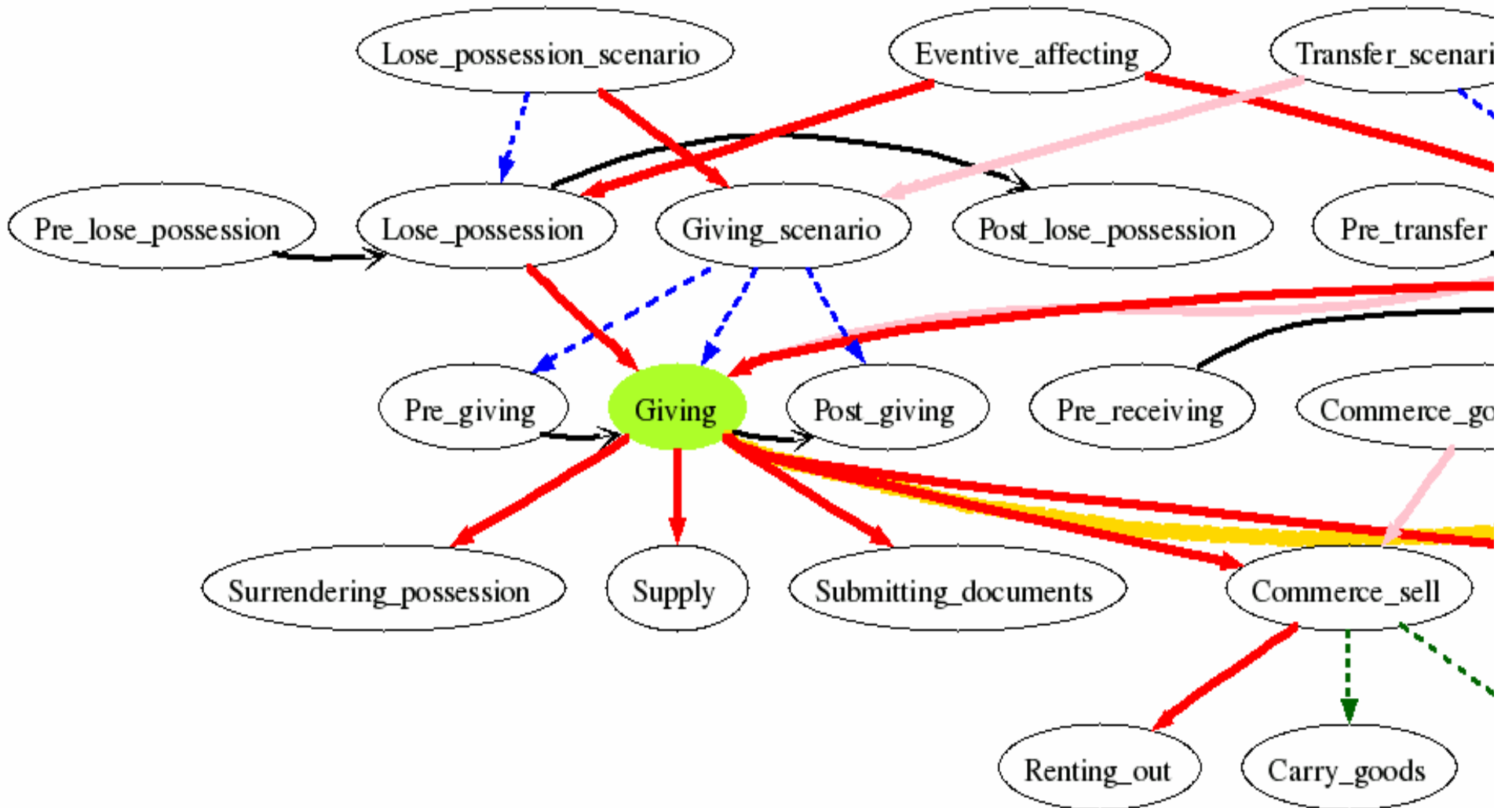
Means [Mns]The **Means** by which the **Donor** gives the **Theme** to the **Recipient**.**Semantic Type**

Lexical Unit Information

Lexical Unit	LU Status	Lexical Entry Report	Annotation Report
bequeath.v	Finished_Initial	LE	Anno
contribute.v	Created	LE	
contribution.n	Created	LE	
donate.v	Finished_Initial	LE	Anno
donation.n	Finished_Initial	LE	Anno
donor.n	Finished_Initial	LE	Anno
endow.v	Finished_Initial	LE	Anno
fob_off.v	Insufficient_Attestations	LE	
foist.v	Finished_Initial	LE	Anno
gift.n	Finished_Initial	LE	Anno
gift.v	Finished_Initial	LE	Anno
give.v	Finished_Initial	LE	Anno
give_out.v	Finished_Initial	LE	Anno
hand in.v	Finished_Initial	LE	Anno
hand.v	Finished_Initial	LE	Anno
hand_out.v	Finished_Initial	LE	Anno
hand_over.v	Finished_Initial	LE	Anno
pass.v	Finished_Initial	LE	Anno

Эволюция FrameNet

- deep cases (1968)
- FrameNet (1990-2000-е)
 - разметка реальных языковых примеров (BNC и др.)
 - виды информации:
 - фреймы (дефиниция и структура)
 - лексические единицы (+ морфосинтаксическое оформление)
 - разметка контекстов
 - фрейм-граф



View Frame Reports

GIVING



View

Current Frame:
Giving

Эволюция FrameNet

- FrameNet (1990-2000-е)
 - виды информации:
 - фреймы (дефиниция и структура)
 - лексические единицы (+ морфосинтаксическое оформление)
 - разметка контекстов
 - фрейм-граф
- FrameNet full-text annotation (2000-е)

Full-text Annotation Reports

2. As CAPITALRelational_political_locales of Europe's most EXPLOSIVEExpansion ECONOMYEconomy, Dublin SEEMSAppear CHANGINGUndergo_change before your very EYESObservable_bodyparts . NEWAge CONSTRUCTIONBuilding is EVERYWHSTREETSRoadways BUZZLively_place , traffic is INCREASINGLYChange_position_on_a_scale congested , and in the FRENETICRUSH HOURCalendric_unit everyone INLocative_relation Dublin SEEMSAppearance INTENTPurpose on CHANGINGExchange PIelse . ATTemporal_collocation NIGHTCalendric_unit the STREETSRoadways are CROWDEDAbounding_with with PEOPLEPeople IGOODDesirability time . PROSPERITYThriving is in the air ; the ROARSounds of the " Celtic Tiger " CANCapability CLEARLYHEARDPerception_experience .

Clear

Colors Off

- As **CAPITAL** of Europe 's most explosive economy , **Dublin** seems to be changing before your very eyes .
- As capital of Europe 's most **EXPLOSIVE** economy , Dublin seems to be changing before your very eyes .
- As capital of **Europe 's** most explosive **ECONOMY** , Dublin seems to be changing before your very eyes .
- As capital of Europe 's most explosive economy , **Dublin** **SEEMS** to be changing before your very eyes .
- As capital of Europe 's most explosive economy , **Dublin** seems to be **CHANGING** before your very eyes . **INI** **INI**
- As capital of Europe 's most explosive economy , Dublin seems to be changing before **your** very **EYES** .
- NEW** **construction** is everywhere , the streets buzz , traffic is increasingly congested , and in the frenetic pace of rush hour ever on changing places with everyone else .
- New** **CONSTRUCTION** is everywhere , the streets buzz , traffic is increasingly congested , and in the frenetic pace of rush hour intent on changing places with everyone else .
- New construction is **EVERYWHERE** , the streets buzz , traffic is increasingly congested , and in the frenetic pace of rush hour intent on changing places with everyone else .
- New construction is everywhere , the **STREETS** buzz , traffic is increasingly congested , and in the frenetic pace of rush hour e

FrameNet как прикладной проект

- Семейство электронных лексикографических ресурсов
 - (толковые словари)
 - тезаурусы
 - WordNet
 - PropBank
 - FrameNet
- Новый вид разметки корпусов; дополнение и альтернатива синтаксической разметки
- Автоматическая разметка, эксперименты с разрешением неоднозначности (E. Vick и др.)

Русский фреймнет-ориентированный ресурс

Мы хотим:

- учесть опыт развития FrameNet
- учесть теоретические достижения отечественной семантики (Апресян, Падучева и др.)
- использовать информацию о падежном оформлении актантов предикатов
- использовать доступные лексические ресурсы для русского языка

Корпусный словарь конструкций

- Банк предложений
- Словарь конструкций
 - Модели управления глагола
 - Конструкции с предикатными именами
 - Конструкции с прилагательными и наречиями
 - Конструкции с предлогами и т.д.
- Разметка предложений: реализация конструкций в тексте
- Полнотекстовая разметка
-

Корпусный словарь конструкций

- Банк предложений
- Словарь конструкций
 - **Модели управления глагола**
 - Конструкции с предикатными именами
 - Конструкции с прилагательными и наречиями
 - Конструкции с предлогами и т.д.
 - «Периферийные» словоцентричные конструкции (*сиджу смотрю, гулять так гулять*)
- Разметка предложений: реализация конструкций в тексте
- Полнотекстовая разметка
-

Словарь моделей управления глагола

Основа: Апресян-Палл 1982

- структура значений глагола и его видовые корреляты
- информация о поверхностно-семантическом оформлении валентностей глагола
- > 1300 глаголов

+ учет информации из ТКС, НОСС, толковых словарей, синтаксического корпуса НКРЯ, RussNet и др.

+ «Лексикограф» (Падучева 2004), проект глагольных фильтров (Толдова и др. 2008)

Словарь моделей управления глагола

Внешние выходы словаря:

- толковые словари: МАС, Ожегов-Шведова
 - электронные ресурсы: «Лексикограф», FrameNet и др.
 - основной и синтаксический корпус НКРЯ
- !+ БД реализаций (размеченных предложений)

Пример словарного входа

1. "служить в министерстве"

Snom V

Snom V {ADV / PRгде + Sx}

Snom V по + Adat + части

Snom V Sins

Snom V в + Sloc

2. "служить в кавалерии"

Snom V

Snom V PRгде + Sx

Snom V Sins

Snom V в + Sloc

3. "служить царю"

Snom V

Snom V Sacc

4. "верно служить науке"

Snom V Sdat

Snom V Sdat Sins

5. "машина еще послужит"

Snom V ADV

Snom V Sdat {ADV / Sacc}

6. "ноги отказываются служить"

Snom V Sdat

Мой старый друг всё ещё служит,

Где вы служите? Он служил в министерстве юстиции [в э

Мой дядя служил по коммерческой части.

Кем вы служите? Он пятый год служит секретарем [прост

Долго ты будешь в кучерах служить?

Моему брату ещё год служить.

Он служит в армии [во флоте, на крейсере]. Лейтенант с

Мой брат пять лет служил простым солдатом.

Андрей служил в танкистах.

Одни всегда приказывают, другие - служат.

Малюта Скуратов верно служил царю. Победенные ста

Он всю жизнь служил науке [искусству]. Я был бы рад по

Всем, что я делаю, я служу своему отечеству. Он хотел с

Эта машина ещё послужит. Ваши сапоги хорошо послужи

Пальто служило ему два года. Эти лыжи долго [неплохо]

Ноги [руки, глаза] отказываются ему служить. Мой ум не

Формат МУ в словаре

«служить в министерстве»


(примеров: 4) →

Имя	Экспликация	Морфо- синтаксис	Ранг	Семантические ограничения	
X	тот, кто служит	NP _{nom}	Субъект	hum	→
Y	место, где служат	в/на + NP _{loc}	Периферия	org	→

Выборка контекстов

- 1 этап: 100 контекстов с финитными формами глагола
- 2 этап: 200 контекстов
- баланс метатекстовых показателей, параметров длины и сложности предложения и места предиката в предложении.

Разметка МУ в предложениях

Он служил в конном корпусе Гая. [Анатолий Рыбаков.
Тяжелый песок (1975-1977)] ←...→ 



Словарь моделей управления

Он служил в конном корпусе Гая.

«служить в министерстве»

lid004

(примеров: 4)



Имя	Экспликация	Морфосинтаксис	Ранг	Семантические ограничения
X	тот, кто служит	NP_{nom}	Субъект	hum
группа	<phrase>	<style NP _{nom} >	▼	<semantic tags>
вершина	<word>	<style S _{nom} >	<rank>	<semantic tags>
Y	место, где служат	в/на + NP_{loc}	Периферия	org
группа	в конном корпусе Гая	PP: в + NP_{loc}	Стандартный ▼	org
вершина	в корпусе	в + S_{loc}	Периферия	org
+				

Оля! Собери нам поесть в дорогу.

«есть рыбу»

поесть V,act,inf,pf,v lid003

(примеров: 2)



Имя	Экспликация	Морфосинтаксис	Ранг	Семантические ограничения
X	агенс	NP_{nom}	Субъект	hum
группа			Инфинитив ▼	
вершина			Нет	hum
Y	пациенс	NP_{acc}	Периферия	food
группа			Генерический ▼	
вершина			Нет	food
+				

Парадигматика МУ

Snom V с Sins

- гнать с работой
- задержаться с отправкой
- кончить с мальчишескими выходками
- копаться с укладкой вещей
- медлить с отъездом
- торопиться со статьей
- тянуть с работой

Три входа системы

- лексический
- конструкционный
- текстовый

Три входа системы

- лексический

запросы по лексеме, классу лексем или морфологическим признакам (предикатная и непредикатная лексика)

в т.ч. с привлечением метатекстовой информации

→ просмотр словаря, примеров реализации, статистики

- конструкционный

- текстовый

Три входа системы

- лексический

- конструкционный

запросы по полным морфосинтаксическим шаблонам или фрагментам, по экспликации, по ограничениям на лексико-семантический класс актанта, по конструкциям, реализующим нестандартное заполнение

в т.ч. с привлечением грамматической и метатекстовой информации

→ получение списка связанных конструкций, просмотр словаря, примеров реализации, статистики

- текстовый

Три входа системы

- лексический
 - конструкционный
 - текстовый
- просмотр разметки предложений

Заключение

Гибридная система:

словарь с выходом в корпус и корпус с экспертным отбором примеров

Теоретические задачи:

изучить способы реализации активных и пассивных валентностей лексики; других лексически-ориентированных конструкций; взаимодействие разных типов информации: лексика и ее семантические классы, морфосинтаксис, порядок слов, пунктуация и т.п.

парадигматическое устройство конструкций

взаимодействие конструкций на пространстве текста

Прикладные задачи...