

ACTA
LINGUISTICA
PETROPOLITANA

TRANSACTIONS
OF THE INSTITUTE FOR LINGUISTIC STUDIES

Vol. VII, part 3

Edited by N. N. Kazansky



St. Petersburg
Nauka
2011

АСТА
LINGUISTICA
PETROPOLITANA

ТРУДЫ
ИНСТИТУТА ЛИНГВИСТИЧЕСКИХ ИССЛЕДОВАНИЙ

Том VII, часть 3

Ответственный редактор Н. Н. Казанский



Санкт-Петербург
Наука
2011

УДК 81
ББК 81.2
А 38

ACTA LINGUISTICA PETROPOLITANA. Труды Института лингвистических исследований РАН / Отв. ред. Н. Н. Казанский. СПб.: Изд-во «Наука», 2003.

Т. VII. Ч. 3 / Отв. ред. Д. В. Герасимов. — 2011. — 699 с.
ISBN 978-5-02-038277-0

РЕДКОЛЛЕГИЯ
«ТРУДОВ ИНСТИТУТА ЛИНГВИСТИЧЕСКИХ
ИССЛЕДОВАНИЙ»:

академик РАН Н. Н. Казанский (председатель),
чл.-корр. РАН А. В. Бондарко, д. филол. наук, проф. Н. Б. Вахтин,
д. филол. наук М. Д. Воейкова, д. филол. наук Е. В. Головкин,
к. филол. наук С. Ю. Дмитренко, к. филол. наук Н. М. Заика (секретарь),
д. филол. наук, проф. С. А. Мызников,
к. филол. наук А. П. Сытов, д. филол. наук, проф. В. С. Храковский

РЕДКОЛЛЕГИЯ ВЫПУСКА

Д. В. Герасимов (отв. ред.), Н. М. Заика, В. А. Крылова,
С. А. Оскольская, С. С. Сай, М. А. Холодилова, К. А. Шагал

Оригинал-макет
М. А. Холодилова

Выпуск подготовлен в рамках издательской программы Научно-образовательного центра ИЛИ РАН «Языковые ареалы России» (Федеральная целевая программа «Научные и научно-педагогические кадры инновационной России»)

ISBN 978-5-02-038277-0 © Коллектив авторов, 2011
© ИЛИ РАН, 2011
© Редакционно-издательское оформление.
Издательство «Наука», 2011

М. В. Худякова

МГУ, Москва

РЕФЕРЕНЦИАЛЬНЫЙ ВЫБОР В ПОЗИЦИИ ПОСЕССОРА: КОРПУСНОЕ ИССЛЕДОВАНИЕ

1. Вступление

Когда человек говорит, он неизбежно сталкивается с необходимостью находить номинации для всех объектов, которые должны быть упомянуты в тексте. При этом говорящий совершает референциальный выбор, то есть выбирает наиболее подходящее из возможных референциальных выражений — полных ИГ (определенных дескрипций, имен собственных) и редуцированных (местоимений).

Референциальный выбор — сложный когнитивный процесс. Основные лингвистические подходы к референции связывают референциальный выбор с активацией референта в сознании говорящего [Chafe 1994], в его рабочей памяти [Кибрик 2004], с фокусом внимания на данном референте, с его центральностью или выделенностью, в дискурсе [Grosz et al. 1995]. Во всех случаях речь идет о некотором свойстве самого референта, связанном с основными когнитивными механизмами и процессами — вниманием и памятью. Наиболее активированные в сознании говорящего референты, как правило, выражены редуцированными номинациями, а менее активированные — более полными.

Данная работа выполнена в рамках многофакторного количественного подхода, предложенного А. А. Кибриком. МКП основан на представлении о том, что референциальный выбор зависит от активации референта в рабочей памяти. Все факторы, влияющие на активацию, учитываются. Суммируется вклад каждого из них в счет активации. Некоторые факторы вносят большой вклад, некоторые являются побочными источниками активации, некоторые могут лишь понижать ее [Кибрик 2003].

В данной работе исследуется вклад такого семантико-синтаксического фактора, как посессорность, на референциальный выбор.

2. Терминология

Существует некоторая непоследовательность в терминологии, связанной с посессивностью. Традиционно посессивным называется местоимение, обозначающее посессора (*his*), а термин «посессивная ИГ» относится к ИГ типа «обладаемое + обладатель» (*his house, the man's house*) [Willemse 2009].

В данной работе посессорными будут называться все полные ИГ и местоимения, обозначающие посессора, то есть притяжательные местоимения (*my, his*), *s*-генитив (*John's*), *of*-генитив (*of John, of him* в значении обладателя). Непосессорные ИГ будем называть актантными¹.

В качестве регулярных посессорных местоимений в работе не рассматриваются рефлексивные притяжательные местоимения, поскольку отношение между антецедентом и рефлексивом нельзя назвать кореферентностью, так как рефлексивные местоимения не являются настоящими аргументами предложения, и потому не полностью референтны [Vach, Partee 1980]. Рефлексивными в данном исследовании считаются все посессорные местоимения, которые находятся в одной клаузе с антецедентом [Kroeger 2004].

3. Корпус для исследования

Корпус RefRhet включает в себя 385 статей экономической, политической или культурной тематики из газеты Wall Street Journal. В корпусе выделено 21 789 элементарных дискурсивных единиц [Loukachevitch et al. 2011].

Для данного исследования был выбран подкорпус, состоящий из 31 текста длиной от 33 до 1708 слов и включает в себя 18845 слов, 3453 маркабулы.

Референциальный выбор определяется многими факторами, среди которых есть свойства антецедента и референциального выражения, референта, а также характер отношений между антецедентом и анафором [Greenbacker, McCoу 2009; Arnold 2008; Кибрик 2003]. Мы будем рассматривать следующие факторы, влияющие на референциальный выбор:

- свойства референта: одушевленность;

¹ Термин «актантные ИГ» в данном исследовании условен.

- свойства референциального выражения: категория лица, полнота, грамматическая роль, тип ИГ;
- свойства антецедента: определенность, полнота, грамматическая роль, тип ИГ.

Наличие корреляции между разными факторами и референциальным выбором определялась с помощью логлинейного анализа [Юнкеров, Григорьев 2002].

4. Результаты исследования

Подкорпус, выбранный для исследования, содержит 3092 определенных ИГ, 85% из которых — полные ИГ, и только 15% — местоимения. ИГ в позиции посессора составляют 19% выбранных маркабул. Распределение типов референциальных выражений в посессорной и актантной позициях представлено в Таблице 1.

Таблица 1. Распределение полных ИГ и местоимений в позиции посессора и актанта в подкорпусе

	Позиция посессора	Позиция актанта	всего
Полные ИГ	259 8%	2253 73%	2512 81%
местоимения	213 7%	367 12%	580 19%
всего	472 15%	2620 84%	3092 100%

Как видно из Таблицы 1, посессоры прономинализованы чаще, чем актанты.

Также есть сильная корреляция между одушевленностью референта, референциальным выбором и посессорностью референциального выражения. Одушевленные посессоры выражены местоимениями в 80% случаев, в то время как неодушевленные посессорные и актантные ИГ чаще выражены полными ИГ. Наиболее предсказуем референциальный выбор для неодушевленных референтов в актантной позиции.

Особый интерес представляют случаи, когда посессоры и актанты демонстрируют противоположные тенденции к прономинализации в зависимости от других факторов.

Например, как можно видеть из Таблицы 2, посессоры чаще прономинализуются после неопределенных ИГ и примерно в равной

степени выражены местоимениями и полными ИГ после определенных антецедентов. Актантные ИГ демонстрируют противоположные тенденции: редуцированные ИГ чаще встречаются после неопределенных антецедентов и реже — после определенных.

Таблица 2. Распределение полных ИГ и местоимений в позиции посессора и актанта в подкорпусе в зависимости от определенности антецедентов

антецедент	анафор					
	посессор			актант		
	Полные ИГ	Местоимения	всего	Полные ИГ	Местоимения	всего
Определенные дескрипции	174 47%	195 53%	369 100%	912 73%	344 27%	1256 100%
Неопределенные дескрипции	5 29%	12 71%	17 100%	73 60%	48 40%	121 100%
всего	386			1377		

Также существует корреляция между формой антецедента и референциальным выбором в позиции посессора/актанта. Посессоры после полных ИГ выражены местоимениями реже, чем после местоимений. Редуцированная форма антецедента имеет сильное влияние на анафора в позиции посессора: такие анафоры прономинализованы в $\frac{3}{4}$ случаев. Актанты имеют иную тенденцию: они чаще выражены полными ИГ, чем местоимениями, особенно после полных антецедентов.

Корреляция между синтаксической ролью антецедента и формой и посессорностью анафора также статистически значима. Актантные ИГ демонстрируют тенденцию быть выраженными полными ИГ после всех типов антецедентов, в то время как посессорные ИГ являются местоимениями после 74% антецедентов-подлежащих, 52% антецедентов-прямых дополнений и 40% других антецедентов.

Данное исследование показало, что в моделях референциального выбора следует учитывать такой фактор, как посессорность референциального выражения. Значит, необходимо выделить этот фактор как отдельный в схеме разметки корпуса RefRhet.

Литература

- Кибрик 2003 — А. А. Кибрик. Анализ дискурса в когнитивной перспективе. Дисс. ... докт. филол. наук, 2003.
- Кибрик 2004 — А. А. Кибрик. Референция и рабочая память: о взаимодействии лингвистики с психологией и когнитивной наукой // А. Н. Гусев, В. Д. Соловьев (ред.). Материалы первой российской интернет-конференции по когнитивной науке. Москва: УМК Психология, 2004. С. 29–43.
- Юнкеров, Григорьев 2002 — В. И. Юнкеров, С. Г. Григорьев. Математико-статистическая обработка данных медицинских исследований. СПб: ВМедА, 2002.
- Arnold 2008 — J. Arnold. Reference production: Production-internal and addressee-oriented processes // *Language and Cognitive Processes* 23, 4, 2008.
- Bach, Partee 1980 — E. Bach, B. H. Partee. Anaphora and semantic structure // B. H. Partee. *Compositionality in Formal Semantics: Selected Papers by Barbara H. Partee*. Oxford: Blackwell Publishers, 1980. P. 122–152.
- Chafe 1994 — W. Chafe. *Discourse, Consciousness, and Time. The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: Chicago University Press, 1994.
- Greenbacker, McKoy 2009 — C. Greenbacker, K. McKoy. Feature selection for reference generation as informed by psycholinguistic research // *Proceedings of the 2009 Workshop on Production of Referring Expressions (PRE-CogSci 2009)*. Amsterdam, 2009.
- Grosz et al. 1995 — B. Grosz, A. Joshi, S. Weinstein. Providing a Unified Account of Definite Noun Phrases in Discourse // *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*. Cambridge, MA: Association for Computational Linguistics, 1995. P. 44–50.
- Kroeger 2004 — P. R. Kroeger. *Analyzing Syntax: a Lexical-Functional Approach*. Cambridge: CUP, 2004.
- Loukachevitch et al. 2011 — N. V. Loukachevitch, G. B. Dobrov, A. A. Kibrik, M. V. Khudyakova, A. S. Linnik. Factors of referential choice: computational modeling // A. E. Kibrik (ed.). *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference Dialogue 2011*. Bekasovo, Moscow Region: RGGU, 2011.
- Willemse 2009 — P. Willemse. The discourse embedding of possessive and esphoric NPs: A corpus study of the identifiability status of their referents and of their discourse functions // *First Vigo-Newcastle-Santiago-Leuven International Workshop on the Structure of the Noun Phrase in English: Synchronic and Diachronic Explorations (NP1)*. Vigo, 2009.