

Взвешенная семантическая карта славянских составных неопределенных местоимений

Роман Сергеевич Фисун

Университет Регенсбурга

Регенсбург, Германия; e-mail: fisun.roman@gmail.com

ORCID: 0000-0002-3074-5172

Ключевые слова: неопределенность, семантическая карта, weighted map, составные неопределенные местоимения, корпусная лингвистика

Weighted semantic map of Slavonic compound indefinite pronouns

Roman Fisun

University of Regensburg

Regensburg, Germany; e-mail: fisun.roman@gmail.com

ORCID: 0000-0002-3074-5172

Keywords: indefiniteness, semantic map, weighted map, compound indefinite pronouns, corpus linguistics

Введение

Составные неопределенные местоимения (СНМ)* охватывают широкий спектр значений – от специфической референции до свободного и случайного выбора, – что делает их удобным объектом для семантического картирования. Взвешенная семантическая карта, предложенная М. Сисау¹, дополняет классический формат количественным измерением: толщина ребер, соединяющих функции, отражает частотность совместного выражения функций одним маркером и тем самым позволяет разграничить ядро и периферию системы. Хотя взвешенная карта неопределенности на материале

* Под СНМ понимается все серийные неопределенные местоимения, как кодифицированные (русск. *кто-то что-то*), так и некодифицированные (*какой бы то ни было, мало кто*), что обычно связывается с более слабой степенью грамматикализации.

М. Хаспельмата² уже обсуждалась в литературе³, специальная карта для славянских СНМ до сих пор отсутствовала.

Классическая карта

Классическая карта неопределенности⁴ (Рис. 1) включает девять функций, выделенных на материале 40 языков, в том числе нескольких славянских. Будучи типологической, она не опирается на корпусные данные и не имеет количественного измерения. На славянском материале эта модель получила дальнейшее развитие: Д. Шпетла⁵ разработал карту для чешского; для украинского также была предложена отдельная карта, а русская и польская уточнены на корпусной основе⁶. Однако эти работы сохраняют ограничения классического подхода: они фиксируют выражение функции, но не отражают их количественного распределения.

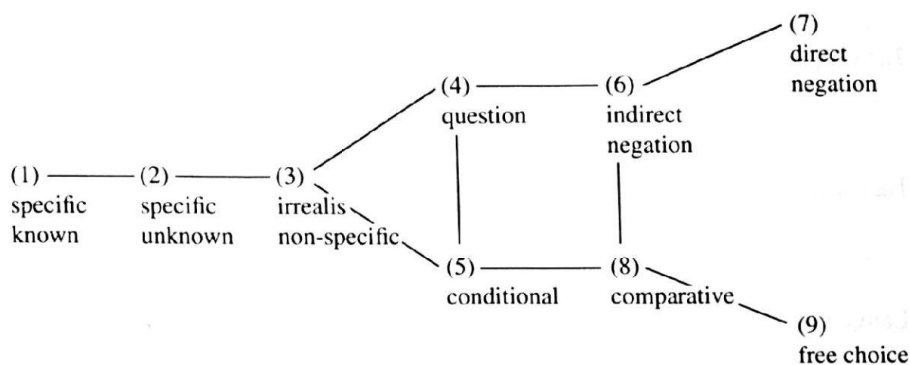


Рис. 1. Классическая семантическая карта неопределенности М. Хаспельмата

Помимо отсутствия количественного измерения, классическая карта имеет и содержательные ограничения. Она не разделяет значения и контексты их реализации: *conditional* и *question* обозначают контекст, тогда как *specific known* и *free choice* – функции семантические. Кроме того, на карте отсутствуют зоны, существенные для славянских серий: случайный выбор, универсальная квантификация и общеэкзистенциальные употребления; этот пробел лишь частично восполнялся позднее⁷. В настоящей работе предлагается модифицированная карта, узлы которой соответствуют только семантическим функциям, опирающимся на теории референции Е.В. Падучевой⁸.

Новая семантическая карта

Семантическое пространство новой карты образуют семь функций (Рис. 2): три из них – *specific known*, *specific unknown* и *free choice* – имеют соответствия у Хаспельмата, четыре остальные основаны на референциальных статусах Падучевой.

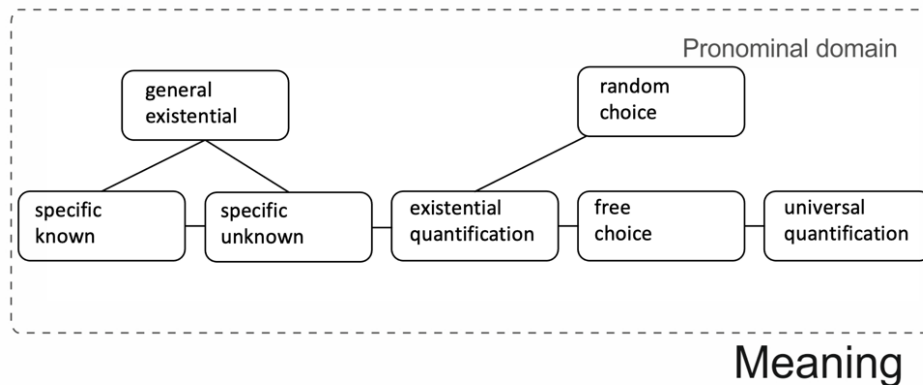


Рис. 2. Предлагаемая семантическая карта

Референтная область

Слабоопределенная референция (*specific known*, SK): референт известен говорящему, но неизвестен слушающему; типичный представитель в русском – серия на *кое-*, нередко выполняющая интродуктивную функцию:

(1) *Мне нужно было кое о чем его спросить.* (НКРЯ: В. Белоусова. Второй выстрел, 2000)

Общеэкзистенциальная референция (*general existential*, GE) занимает промежуточное положение между специфической и неспецифической⁹. GE выражается многочисленными СНМ, чьи маркеры неопределенности восходят к количественным выражениям. Модификаторы со значением ‘мало’ образуют трехчленную парадигму – ослабленное отрицание между полным отрицанием и слабоопределенной референцией:

(2) а. *Никто* не пришел на концерт.

б. *Мало кто* пришел на концерт. (= почти никто)

с. *Кое-кто* пришел на концерт. (= несколько человек)

Референтная неопределенность (*specific unknown*, SU) предполагает существование референта при его неизвестности для обоих участников:

(3) *У него уже гости: Генрих Сагир с каким-то врачом из Израиля.* (НКРЯ: Н. Шмелькова. Последние дни Венедикта Ерофеева, 2002)

Нереферентная область

Экзистенциальная квантификация (*existential quantification*, EQ) вводит неиндивидуализированные объекты:

(4) *Ни в одной из работ нет какого бы то ни было сравнения, или восторга, или красноречивого описания.* (ruTenTen11)

Универсальная квантификация (*universal quantification*, UQ) отсылает ко всем представителям множества или к объекту как единому целому:

(5) *<...> и они чаще, чем кто бы то ни было, имеют «плохую прессу».* (ruTenTen11) [\neq чаще, чем (по меньшей мере) одно другое лицо]

Свободный выбор (*free choice*, FC) отсылает к неограниченному множеству; вопреки распространенному мнению, идея выбора присутствует лишь в части контекстов, а в большинстве случаев FC вводят неограниченное множество, для выбора представителей которого не ставится никаких условий:

(6) *Но что же теперь со всем этим делать? Нас могло выкинуть где угодно.* (ruTenTen11) [=Неверно, что есть место, куда нам не могло выкинуть]

Местоимения случайного выбора (*random choice*, RC) выражают безразличие агенса по отношению к выбору из заданного множества альтернатив¹⁰, подчеркивается отсутствие критериев выбора, а выбираемый референт нередко оценивается как неподходящий. Помимо описания на материале романских языков функция была описана для чешского¹¹ как *indiscriminative function*:

(7) чешск. *My nejsme jen tak ledakdo!*¹²

‘Мы не кто попало!’

Взвешенная карта: материал и методология

Предлагаемая взвешенная карта обобщает данные по 51 серии СНМ пяти славянских языков – русского, украинского, польского, чешского и хорватского, извлеченные из веб-корпусов ruTenTen11, ukTenTen20, plTenTen12, csTenTen17 и hrWaC 2.2; в выборку включались прежде всего некодифицированные, но более грамматикализованные серии, удовлетворяющие семантическим (наличие неопределенного значения у серии в целом), дистрибутивным (маркер

неопределенности серии должен сочетаться с местоименными основами нескольких онтологических категорий: *кто-либо*, *что-либо*, *где-либо* и т. п.) и частотным (достаточное число вхождений в корпусе для обеспечения репрезентативной выборки) критериям:

	русск.	украинск.	польск.	чешск.	хорватск.
префиксальные	абы Бог весть черт-те мало кое- не пойми	багато казна- невідомо мало Бог зна(є)	baǰż (nie)byle cholera wie lada mało	bůchví: jedno kdoví: lhostejno málo: nevím: sotva:	bilo Bog zna gdje- i- koje- malo mnogo
постфиксальные	бы то ни было -то -либо -нибудь угодно попало	завгодно попало б (то) не було -сь -небудь	baǰż popadło -ś -kolwiek śkolwiek	chceš/chcete -koli(v) je libo	-god bilo

Табл. 1. Проанализированные серии, полужирным выделены кодифицированные

Для каждой серии случайным образом отбиралось по 100 контекстов; каждый пример вручную атрибутировался одной семантической функции, а случаи, не допускающие достоверной семантической атрибуции, исключались и заменялись*. Далее для каждой серии по каждой функции отмечалось ее наличие или отсутствие, после чего строилась единая взвешенная карта:

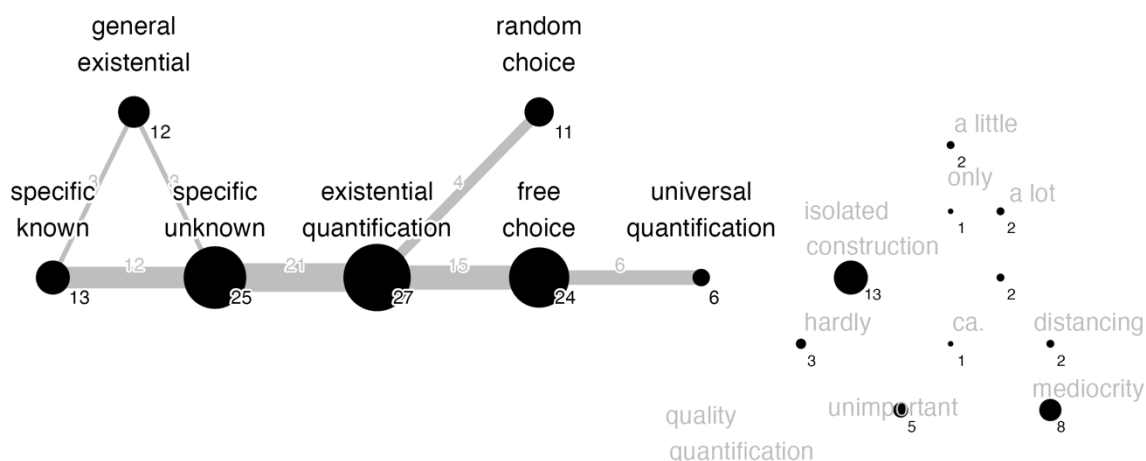


Рис. 3. Взвешенная карта неопределенности славянских СНМ

* Анализ хорватских данных был выполнен проф. Б. Ханзенем, чешских – М. Рыбовой; исходные данные доступны по адресу: https://fisun.org/files/stats/weighted_100_raw.csv

Размер узла соответствует числу серий, реализующих данную функцию, а толщина ребра – числу серий, совмещающих обе функции; расстояния между узлами топологического значения не имеют. Непрономинальные значения, расположенные в правой части карты, фиксировались при аннотировании как побочный результат и здесь специально не рассматриваются. На тех же принципах построены и карты для каждого из пяти языков:

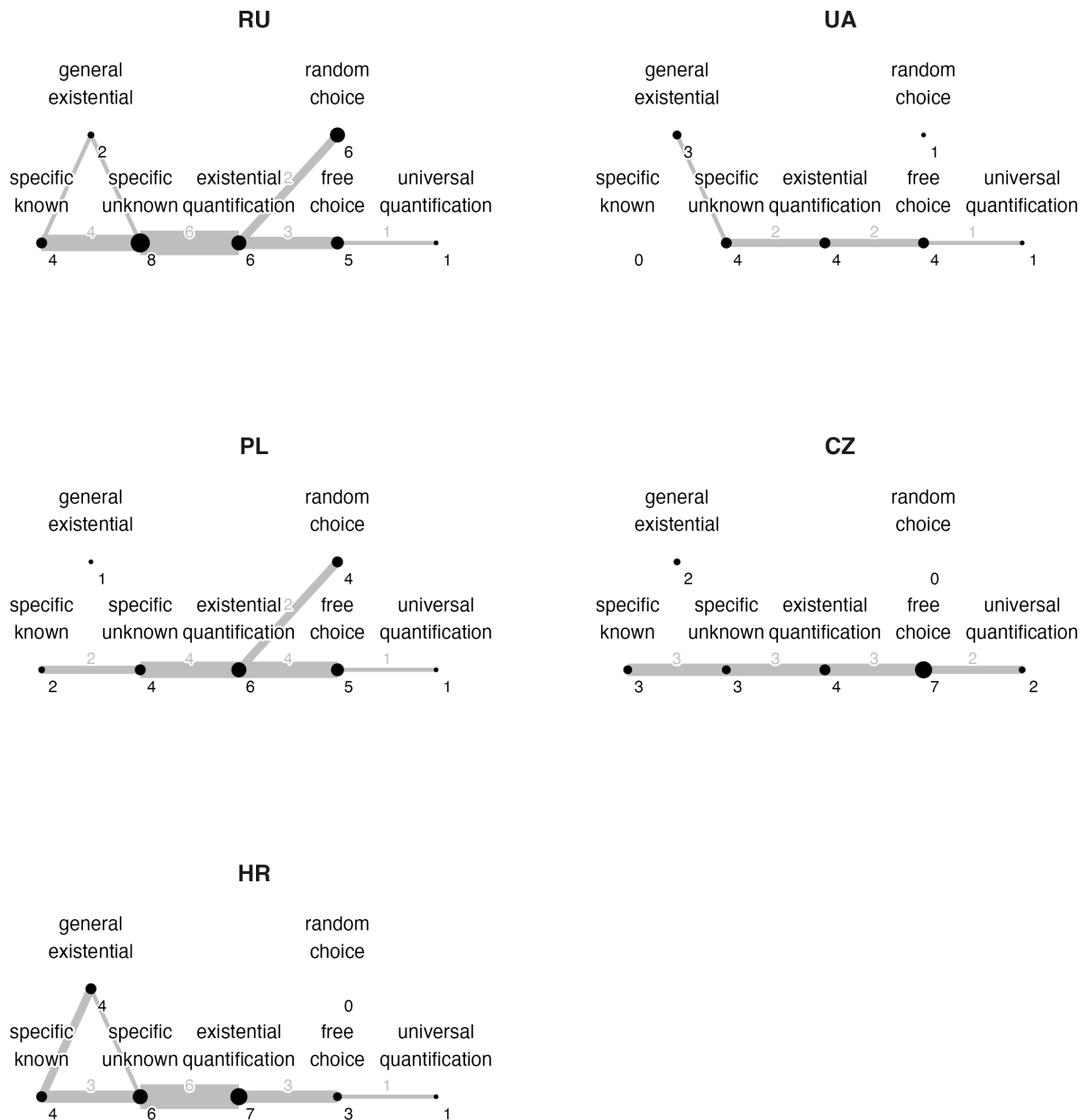


Рис. 4. Взвешенные карты неопределенности для отдельных языков

Анализ карт и выводы

Взвешенная карта выявляет внутреннюю иерархию функций СММ. Материал 51 серии из пяти славянских языков показывает асимметричную

организацию системы: центральная часть семантического пространства нагружена значительно сильнее, чем периферия.

Ядро образуют *EQ* (27 серия), *FC* (24) и *SU* (25): именно эти функции имеют наибольшее число маркеров и наиболее тесные связи, что указывает на концентрацию полифункциональности. При этом *SU*, традиционно считающаяся центральной функцией неопределенных местоимений, уступает по охвату первой, так что СНМ в целом тяготеют к нереферентному полюсу.

Периферийные функции представлены слабее и специализированнее. *UQ* и *SK* лишь минимально связаны с ядром, *GE* (12 серий) наиболее изолирована, а *RC* (11 серий) занимает узкоспециализированную позицию и отсутствует в выборке для чешского и хорватского.

Off-map-функции образуют отдельную зону вне основной сети. Они фиксируют дальнейшую семантическую экспансию СНМ, до сих пор остающуюся малоописанной.

Карты отдельных языков в целом воспроизводят ту же конфигурацию, но различаются по степени заполненности отдельных участков. Если в русском и хорватском заметнее выражена левая часть карты, а в чешском особенно усиливается *FC*, то украинский и польский демонстрируют более ровное и менее экстремальное распределение, при котором основная нагрузка сосредоточена в центральной зоне, а периферийные функции представлены слабее.

Примечания

¹ *Cysouw M.* Building semantic maps: the case of person marking // *New Challenges in Typology* / ed. by M. Miestamo, B. Wälchli. Berlin, 2007. P. 233.

² *Haspelmath M.* Indefinite pronouns. Oxford, 1997.

³ *Georgakopoulos T., Polis S.* The semantic map model: State of the art and future avenues for linguistic research // *Language and Linguistics Compass*. 2018. Vol. 12. N 2. P. 25.

⁴ *Haspelmath M.* Indefinite pronouns. P. 64.

⁵ *Špetla D.* Semantic maps: a way out of the equivalence conundrum? // *Theory and Practice in English Studies*. 2021. Vol. 10. Iss. 2. P. 7–21.

⁶ *Fisun R.* Семантические Besonderheiten zusammengesetzter Pronomen im Russischen, Ukrainischen und Polnischen. Unveröffentlichte Masterarbeit. Universität Regensburg, 2016.

⁷ См., например: *Tatevosov C. G.* Семантика составляющих именной группы: кванторные слова. М., 2002. С. 167; *Aguila-Guevara A., Aloni M., Port A., Šimik R., de Vos M., Zeijlstra H.* Semantics and pragmatics of indefinites: methodology for a synchronic and diachronic corpus study // *Beyond semantics: Corpus-based investigations of pragmatic and discourse phenomena. Proceedings of the DGfS Workshop, Göttingen, February 23–25, 2011* / ed. by S. Dipper, H. Zinsmeister. Bochum, 2011. P. 9.

⁸ *Падучева Е. В.* Референциальный статус именной группы // *Русская корпусная грамматика*. 2017. URL: http://rusgram.ru/Референциальный_статус_именной_группы (13.04.2024).

⁹ Там же.

¹⁰ Kellert O. Evaluative Meanings of Indefinites: A Synchronic and Diachronic Analysis of Spanish *cualquiera* and Related Elements. Habilitationsschrift. Göttingen, 2021. P. 3.

¹¹ Šimík R. On free choice-like wh-based expressions in Czech // *Free Choice: Facts, Models and Problems. Workshop Proceedings* / ed. by J. Jayez, L. Tovená. Hamburg, 2008.

¹² *Ibid.* P. 2.