

ИСПОЛЬЗОВАНИЕ КОРПУСОВ И КОНКОРДАНСЕРА ANTCONC ДЛЯ ОТБОРА И СОСТАВЛЕНИЯ СЛОВАРЯ ЯПОНСКИХ АРТОНИМОВ В АНГЛИЙСКОМ ЯЗЫКЕ

CORPORA AND ANTCONC
CONCORDANCER IN SELECTION
AND COMPILING THE DICTIONARY
OF JAPANESE ARTONYMS IN ENGLISH
LANGUAGE

E. Ivanova

Annotation

This paper is dedicated to modern methods of studying foreign words in English language in the sphere of Japanese art (artonyms). It also describes ways of selecting lexical items for creating a dictionary of Japanese words in English and gives some already generated dictionary entry examples.

Keywords: interlinguoculturology, concordance, corpus, japanisms, dictionaries of foreign words.

Иванова Екатерина Юрьевна

Ст. преподаватель,

*Санкт-Петербургский государственный
экономический университет*

Аннотация

Данная работа посвящена современным методам изучения пласта иноязычной лексики в английском языке в применении к сфере искусства Японии (артонимов). Также описываются способы отбора лексических единиц для создания словаря японизмов в английском языке и приводятся примеры словарных статей разрабатываемого словаря.

Ключевые слова:

Интерлингвокультурология, конкордансер, корпус, японизмы, словари иноязычной лексики.

Компьютерный прогресс значительно помог минимизировать временные затраты на анализ данных во всех сферах науки, поэтому неудивительно, что многие лингвистические исследования также используют достижения корпусной лингвистики для проведения анализа и апобации их результатов.

Наши исследования проводятся в русле нового направления – интерлингвокультурологии, лингвистической дисциплины, занимающейся изучением проблем вторичной культурной ориентации языка в область иноязычной (внешней) культуры [Кабакчи, 2007]. Данная же работа посвящена современным методам изучения пласта иноязычной лексики в английском языке в применении к сфере искусства Японии (артонимов).

В начале работы для отбора материала использовались традиционные способы исследования, то есть работа с текстами и привлечение словарей разного объема. В Большом Оксфордском словаре 587 словарных статей имеют в составе заглавного слова японизмы, 471 из них представляют собой чистые японские заимствования. Стоит отметить, что еще 354 японизма встречаются в словаре, хотя и не в качестве заглавного слова, при этом больше четверти заглавных слов являются артонимами (129 словарных статей).

Естественно, лексикографическая регистрация – это долгий процесс, в котором составители проводят обширные исследования перед внесением того или иного слова. Именно поэтому становится актуальным создание специальных словарей иноязычной лексики. В рамках Japanese-Culture-Oriented English (JCOE) существует три наиболее известных англоязычных словаря, направленных именно на японскую культуру, и количество в них ксенонимов впечатляет: Garland Cannon "The Japanese Contributions to the English Language" – 1200 словарных статей, Setsuko Kojima, Gene A. Crane "Japanese-English dictionary of Japanese culture" – 2300, Isamu Hayakawa "A Historical Dictionary of Japanese Words Used in English" – 2500.

Словари иноязычной лексики имеют свои недостатки. Прежде всего, в них нет единообразия. Достаточно отметить, что даже вариант написания заглавного слова отличается друг от друга (ср.: ukiyo-e и ukiyoe). Еще большую проблему представляет собой отбор слов для таких словарей, поэтому огромную роль играет фактор субъективности составителей. Например, слово kodan имеет словарную статью лишь в словаре Крэйна.

В ходе исследования японских артонимов была выявлена недостаточная представленность японизмов данной

тематической группы в словарях. Это делает чтение культурно маркированных текстов трудным для неподготовленного читателя. Поэтому было решено составить словарь, который бы объединил наиболее распространенные японские артонимы. Актуальность составления подобного словаря определяется растущим интересом к японской культуре в целом и к сфере искусства в частности, межкультурно-языковыми контактами вообще, процессами глобализации, в том числе и увеличением контактов Японии с зарубежным миром, в первую очередь – посредством английского языка.

Основная задача разрабатываемого словаря – это историко-культурное комментирование (толкование) японизмов, а также приведение при необходимости и наличии возможности лингвистического комментария. Цель его – помочь специалистам и поклонникам японского искусства в чтении аутентичных англоязычных текстов.

В рамках интерлингвокультурологии принято использовать аутентичные тексты, многие из которых благодаря цифровым технологиям доступны и в электронном виде. В ходе отбора материала для дальнейшего лингвистического анализа и составления специализированного словаря были замечены некоторые орфографические универсалии, например, определённые буквосочетания, отличающие японизмы от базового лексического состава английского языка. В связи с этим в качестве лингвистического эксперимента было решено привлечь к обработке данных корпусы текстов: Британский национальный корпус (British National Corpus, BNC) и Корпус современного американского английского языка (Corpus of Contemporary American English, COCA).

Первый охватывает британский английский конца XX в., то есть является синхроническим корпусом, в котором представлен широкий выбор жанров и создан как модель типичного письменного и разговорного британского английского языка того времени. Корпус был полезен для проверки данных словарей и сравнения разных вариантов японизмов, но так как охватывает лишь определенный и довольно короткий промежуток времени, не мог дать результатов о развитии и становлении японских заимствований.

COCA был создан на основе текстов журнала Time, написанных с 1923 по 2011 год. Это наибольший (450 млн слов) корпус текстов американского варианта английского языка и единственный свободно доступный корпус на этом языке, включающий большое разнообразие текстов различных жанров. В нем, в отличие от BNC, можно проследить диахронические изменения. Однако из-за ограничения публицистическим стилем за его пределами остается слишком большое количество лексических единиц.

Именно поэтому в ходе исследования был создан собственный корпус текстов, направленных на японскую культуру, в который вошло 25 произведений художественной литературы, 5 путеводителей, 9 специализированных работ по культуре и искусству Японии и тексты из Интернет-ресурсов, посвященных Японии. Для работы с корпусом потребовался конкордансер.

Обычно под конкордансом (concordance, KWIC (key word in context)) принято понимать "список всех употреблений данного слова в контексте со ссылками на источник" [Sinclair, 1991: p. 8]; "указатель (reference book), содержащий все слова, использованные в определенном тексте или в работах отдельных авторов за исключением артиклей и предлогов" [O'Keefe, McCarthy, 2010: p.7]; "список всех случаев употребления определенного поискового термина (search term) в корпусе, представленных в контекстном окружении" [Захаров, Богданова, 2011: 43]; "совокупность всех входящих словоформы с контекстами (textual environment)" [Baker, Hardie, 2006: 32]. В русскоязычной традиции обычно говорят об "(универсальном) корпусном менеджере" – специализированной поисковой системе, включающей программные средства для поиска данных в корпусе, получения статистической информации и предоставления результатов пользователю в удобной форме [Sinclair, 1991: 57, 58]; "конкордансер", "конкорданс отдельных слов" [Sinclair, 1991: 72].

Под конкордансером обычно понимают инструментальное программное обеспечение, выполняющее операцию поиска в корпусе каждого случая употребления словоформы, словосочетания или другого элемента и их контекстного окружения с целью создания конкорданса [Захаров, Богданова, 2011: 44]. Конкордансеры позволяют получать частоту той или иной языковой единицы по произвольному корпусу текстов, список контекстов, в которых данная единица встретилась, и др. Конкордансеры используются для решения следующих лингвистических задач: сравнение различных использований одного слова, анализ ключевых слов, анализ частотности слов и словосочетаний, поиск и исследование фраз и идиом, поиск перевода, например, терминологии, создание списков слов.

В нашем исследовании был использован корпус-менеджер AntConc, разработанный профессором Лоуренсом Антони (Laurence Anthony), директором Центра обучения английскому языку в науке и технике Школы науки и техники университета Васеда (Япония) [AntConc [Electronic resource] / L. Antony http://www.antlab.sci.waseda.ac.jp/antconc_index.html]. Таким образом, в программу были загружены имеющиеся электронные варианты англоязычных художественных произведений, специализированных работ по японскому искусству и путеводителей о Японии. В стоп-лист были внесены артикли, предлоги, местоимения. Далее были выявлены основные несвой-

ственные английскому языку буквосочетания, которые по нашей гипотезе могли бы служить маркерами японизмов в тексте: ka, ya, za, kya, shi, ji, tsu, zu, shu, ku, yu. Данные буквосочетания соответствуют принятым системам латинизации японских слогов.

Буквосочетания tsu, ji, zu и ku стали наиболее надежным показателем японского происхождения слова. Все выделенные лексемы можно отнести к японизмам, большинство являются ономастической лексикой или частью целых фразам на японском языке, включенных в произведение с помощью транслитерации.

Несомненным достоинством использования AntConc стала оптимизация процедуры сплошной выборки и статистической обработки полученных результатов. Тем не менее, стоит отметить, что этот метод может являться лишь дополнением к традиционному методу сплошной выборки и использованию словарей, так как имеет ряд недостатков:

- ◆ невозможность задать поиск именно по интересующей категории иноязычных слов – артонимам, что делает необходимым дополнительную проверку отобранных единиц на соответствие;
- ◆ возможность погрешности, связанной с неучетностью некоторых выбранных маркеров только для японоязычных заимствований (при использовании буквосочетания shi в стоп-лист пришлось вносить такие общепотребительные слова, как ship, fishing, cushion, wishing и т.д.);
- ◆ ограничение поиска только транслитератами, невозможность выделения калек и иных способов образования заимствований, которые довольно распространены среди японизмов;
- ◆ результаты анализа могут быть субъективными и зависеть лишь от наблюдений и интуиции автора исследования.

В целом наши наблюдения были полезны при составлении статистических результатов и, на наш взгляд, могут быть полезны для дальнейшего исследования в рамках компьютерной лингвистики.

Все словари создаются на базе картотеки, то есть систематизированного собрания текстовых карточек. Картотека нашего тематического словаря содержит:

1. перечень англоязычных ксенонимов–японизмов;
2. дефиниции, взятые из различных словарей, в том числе электронных;
3. материалы вспомогательного характера, в том числе этимология слова, лингвистическая справка. Для экономии места в печатную версию словаря не были включены примеры употребления, взятые из текстов и корпусов.

На основе извлечений из текстов и работы со словарями устанавливаются орфография японизма и ее воз-

можная вариативность, а также значение слова и его экспликация.

Критерием отбора японских артонимов служило выделение японоязычной лексики в тексте, которое может быть двух видов: лексическое и графическое. Среди средств лексического выделения японизмов обычно встречаются вводные лексические обороты типа called, so-called, known as, what they/ we call, what is called и другие. Например: He wrote for kabuki as well as the traditional Japanese puppet dramas called bunraku [БОО].

Наиболее распространенным средством выделения ксенонимов являются различные графические способы. В практике английского языка межкультурного общения в применении к японской культуре встречаются следующие варианты:

1. курсив (наиболее часто используемый): In Japanese bonsai can be translated as 'tray planting' and for the Japanese it represents a fusion of ancient beliefs with eastern philosophies of harmony [БОО].
2. кавычки: These three lines from the backbone of every "seikwa" Arrangement, whether it be of three or more lines [Koehn, 1933: 10].
3. прописная буква (используется для тех японизмов, которые попадают в разряд слов, подлежащих выделению согласно правилам современного английского языка): The Kojiki is one of the two primary sources for Shinto, the Japanese national religion [www.sacred-texts.com/shi/kj/index.htm].
4. скобки: They are seen strolling around the Gion district in Kyoto, wearing an elaborate traditional coiffure, and a colorful kimono with a long sash (obi) hanging from the back. They are trained in dancing, singing, playing shamisen (a string instrument) and manners for entertaining guests, in an effort to become accomplished geisha. [Japanese-English dictionary of Japanese culture, 1990: 73]. В круглые скобки обычно заключают краткое пояснение специального или окказионального иноязычного слова при его первичном введении в текст.

После составления словника, в который входят все варианты, встретившиеся нам в ходе отбора материала, происходит выбор заглавного слова. Лексическая вариативность в целом не очень частотна среди японизмов–артонимов. Тем не менее, все варианты входят в словарную статью. В случае, когда два варианта употребляются равнозначно, создается две словарные статьи с перекрестной ссылкой (например, Ikebana/ flower arrangement).

Графическая вариативность ксенонимов вызвана разными системами транслитерации, до сих пор существующих в японском языке, а также в некоторых случаях непрямым заимствованием через язык–посредник.

Итак, типичные словарные статьи выглядят следующим образом:

1. *chiyogami*, n. 千代紙 Paper [Lit., a thousand-generations paper < chi- + yo + comb. form of kami paper] lit. paper of thousand generations, patterned paper; Japanese paper with brightly colored patterns originated in the Edo period (1603-1868). With the development of multicolored woodblock prints in the mid-18th century, ukiyoe artists also assigned patterned paper in bright colors. These patterns were at first auspicious symbols, such as pine, bamboo, cranes, tortoises, etc., and then came in a wide variety, covering a range of subjects seen in those days. This paper was popular among women who then wore strictly regulated clothing, and was used for decorating boxes, covering books, wrapping gifts, and making dolls. *Chiyogami* crafts became part of their life. Thus, *chiyogami* reflects, in part, the essence of the Edo period. → EDO-JIDAI, UKIYO E

2. *Kiyomizu porcelain*, n. (EB 1911) *Pottery* [Transl. of *Kiyomizu-YAKI* < *Kiyomizu*, the name of a Kyoto area where kilns

produced famous porcelain beginning in the late Edo period, distinguishing it as a modern ware from the earlier KYOTO WARE] An overglaze ceramic, long associated with elegance, refinement, and high proficiency.

Стоит оговориться, что данный словарь создан с прагматической целью для удобства использования при чтении специализированных текстов, посвященных японскому искусству. При необходимости он может быть дополнен более исчерпывающими лингвистическими данными.

Таким образом, привлечение корпусов текстов и использование конкорданса на данном этапе позволяет ускорить процесс отбора материала исследования, а также верифицировать данные словарей, но без усовершенствования технологии поиска невозможно без использования других методов лингвистики в целом и интерлингвокультурологии в частности.

ЛИТЕРАТУРА

1. Захаров В.П., Богданова С.Ю. Корпусная лингвистика : учеб. для студентов гуманитарных вузов. Иркутск: ИГЛУ, 2011. с. 161
2. Кабакчи В.В. Типология текста иноязычного описания культуры и инолингвокультурный субстрат//Лингвистика текста и дискурсивный анализ: традиции и перспективы: сб. науч. тр. СПб.:СПбГУЭФ, 2007. С.51–70.
3. Anthony, L. AntConc (Version 3.5.5) [Computer Software]. Tokyo, Japan: Waseda University [сайт].URL: <http://www.laurenceanthony.net/software> (дата обращения: 15.02.2018).
4. Baker P., Hardie A., Mcenery T. A Glossary of Corpus Linguistics. Edinburgh : Edinburgh Univ. Pr., 2006., 192 p.
5. Historical Dictionary of Japanese Words Used in English (Revised and corrected edition), A / Compiled by Prof. Dr. Isamu Hayakawa. – Texnai Inc., 2014.
6. Japanese Contributions to the English Language: an historical dictionary, The/ Garland Cannon. Nicholas Warren, assoc. ed. – Wiesbaden: Harrassowitz, 1996.
7. Japanese–English dictionary of Japanese culture/ Setsuko Kojima, Gene A. Crane. – Heian International; 1st edition, 1990.
8. Koehn A. The Art of Japanese flower arrangement (ikebana). Kobe: J.L.Thompson&Co, 1933. 304 p.
9. O’Keeffe A., McCarthy M. The Routledge Handbook of Corpus Linguistics (Routledge Handbooks in Applied Linguistics). Abingdon: Routledge, 2010. 712 p.
10. Oxford English Dictionary: [сайт].URL:<http://www.oed.com/> (дата обращения: 15.06.2017).
11. Oxford English Dictionary, (1st ed. – 1928; 2nd edition – 1989).
12. Sinclair J. Corpus, Concordance, Collocation. Oxford : Oxford Univ. Pr., 1991. 179 p
13. The British National Corpus (BNC) : [сайт].URL: <http://corpus.byu.edu/bnc/> (дата обращения: 20.05.17).
14. The Corpus of Contemporary American English (COCA): [сайт].URL: <http://corpus.byu.edu/coca/> (дата обращения: 15.06.2017).

© Е.Ю. Иванова, [trinitifix@mail.ru], Журнал «Современная наука: актуальные проблемы теории и практики»,

13-й МЕЖДУНАРОДНЫЙ ВОДНЫЙ ФОРУМ «ВОДА: ЭКОЛОГИЯ И ТЕХНОЛОГИЯ»

ЭКВАТЭК 2018 **25-27 СЕНТЯБРЯ 2018**
МОСКВА, КРОКУС ЭКСПО

РЕКЛАМА

ВСЕ ПРОФЕССИОНАЛЫ И ЭКСПЕРТЫ ВОДНОЙ ОТРАСЛИ В ОДНО ВРЕМЯ И НА ОДНОЙ ПЛОЩАДКЕ!