

ЯПОНСКИЕ ЗАИМСТВОВАНИЯ В ЧЕШСКОМ НАЦИОНАЛЬНОМ КОРПУСЕ

А.И. Изотов, О.И. Черчук

Московский государственный университет имени М.В. Ломоносова, Москва, Россия; a.i.izotov@mail.ru; cherchuk-olga@yandex.ru

Аннотация: В настоящей статье на материале Чешского национального корпуса рассматриваются некоторые аспекты функционирования в современном чешском дискурсе существительных японского происхождения — как тех из них, которые были перечислены в соответствующей статье «Нового энциклопедического словаря чешского языка» (2016), так и тех, которые были нами идентифицированы в качестве японизмов среди 10 061 существительного, добавленного за последние три полных года (2019, 2020, 2021) сотрудниками отделения современной лексикологии и лексикографии Института чешского языка Академии наук Чешской республики в электронную базу данных Neomat. С помощью входящего в проект Чешского национального корпуса 4,9-миллиардного корпуса современных письменных чешских текстов SYNv9 определялась абсолютная и относительная частота употреблений той или иной лексемы, а также ее ARF (= Average Reduced Frequency), вычисляемая по особой формуле и призванная нивелировать ситуации употребления данной лексемы в отдельных текстах активнее, чем в текстах иных. Интегрированная в Национальный корпус чешского языка и учитывающая, в частности, корпуса устной чешской речи программа WaG (= Word at a Glance) помогала определять региональные, гендерные, возрастные и образовательные характеристики употребляющих ту или иную лексему носителей чешского языка. Русские соответствия рассматриваемым чешским словам японского происхождения приводились в чисто вспомогательных целях, так что принципы их графического оформления не обсуждались.

Ключевые слова: лексический японизм; чешский дискурс; Чешский национальный корпус; средняя уменьшенная частота; кандзи; кана

doi: 10.55959/MSU0130-0075-9-2023-47-2-7

Для цитирования: Изотов А.И., Черчук О.И. Японские заимствования в чешском национальном корпусе // Вестн. Моск. ун-та. Сер. 9. Филология. 2023. № 2. С. 99–110.

JAPANESE LOANWORDS IN THE CZECH NATIONAL CORPUS

Andrey I. Izotov; Olga I. Cherchuk

*Lomonosov Moscow State University, Moscow, Russia;
a.i.izotov@mail.ru; cherchuk-olga@yandex.ru*

Abstract: The article discusses some aspects of the functioning of Czech nouns of Japanese origin. Both lexical Japanisms named in the *New Encyclopedic Dictionary of the Czech Language* (2016) and nouns included in the Neomat electronic database in 2019, 2020 and 2021 are considered. The data of the Czech written and spoken corpora, which are constituents of the Czech National Corpus project, are used. The absolute and relative frequency of use of a particular lexeme, as well as its ARF (= Average Reduced Frequency) was determined with the help of the 4.9 billion-strong corpus of modern written Czech texts SYNv9. The WaG (= Word at a Glance) program developed within the framework of the Czech National Corpus project to determine the regional, gender, age and educational characteristics of Czech speakers using a particular lexeme was used. Russian correspondences to the Czech words of Japanese origin under consideration were given for purely auxiliary purposes, so that the principles of their graphic design were not discussed.

Key words: lexical japanism; Czech discourse; Czech National Corpus; average reduced frequency; kanji; kana

For citation: Izotov A.I., Cherchuk O.I. (2023) Japanese Loanwords in the Czech National Corpus. *Lomonosov Philology Journal. Series 9. Philology*, 2023, no. 2, pp. 99–110.

1. Массив лексических единиц, проникающих и уже проникших в чешский язык из языка японского, закономерно уступает массиву заимствований из популярного сейчас английского, однако игнорировать его вряд ли рационально, что мы и попытаемся показать на последующих страницах, на которых на базе письменных и устных текстов Чешского национального корпуса¹ мы рассмотрим чешские существительные японского происхождения — как те из них, которые были исчислены в [Karlík, Karlíková 2016], так и те, которые были нами идентифицированы в качестве японизмов среди 10 061 существительного, включенного за последние три полных года (2019, 2020, 2021) сотрудниками отделения современной лексикологии и лексикографии Института чешского языка Академии наук Чешской республики в электронную базу данных Neomat, см. [Databáze... 2015].

¹ О принципах формирования Чешского национального корпуса см. [Изотов 2022].

2. Из названных в [Karlík, Karlíková 2016] японизмов самыми употребительными в современном чешском письменном дискурсе, по данным входящего в проект «Чешский национальный корпус» 4,9-миллиардного корпуса современных чешских письменных текстов SYNv9, являются слова *karate* ‘карате’² (31 559 — 5,54 — 9 444,84) и *tsunami* ‘цунами’ (23 368 — 4,1 — 7 374,98). Первое число в скобках обозначает общее количество обнаруженных в SYNv9 примеров употребления данной лексемы, то есть абсолютную частоту ее употреблений, второе число — частоту относительную, то есть отношение количества обнаруженных примеров к одному миллиону представленных в данном корпусе примеров иных (i.p.m.), третье число — то, что авторы Национального корпуса назвали ARF = Average Reduced Frequency, переведенную в [Ляшевская, Шаров 2008] как «средняя уменьшенная частота» и вычисляемую по формуле

$$ARF = \frac{1}{v} \sum_{i=1}^f \min(d_i, v),$$

в которой f — количество примеров употребления данной единицы в корпусе размером N , d_i — расстояния (количество слов) между примерами употребления данных единиц, а v — среднее расстояние между подобными примерами, вычисляемое по формуле $v = N/f$.

В современной чешской лингвистике именно ARF, способная нивелировать ситуации, когда то или иное грамматическое средство употребляется в отдельных текстах намного активнее, чем в текстах иных, используется в качестве основного критерия при составлении частотных словарей чешского языка, см., например, [Frekvenční 2004].

2.1. В соответствии с определяемой по данным SYNv9 относительной частотностью (i.p.m.), названные в [Karlík, Karlíková 2016] японизмы³ можно разделить на три группы: в первую группу входят

² В абсолютном большинстве случаев чешские японизмы легко опознаются носителями русского языка, поскольку в языке русском присутствуют, как правило, сходно звучащие лексемы, также заимствованные из японского. В рамках настоящей статьи мы не пытались исследовать возможные трансформации смысла при заимствовании той или иной лексемы в чешский или русский язык, которые иногда происходят. Например, иероглиф 酒, читающийся по-японски как sake, а по-китайски как jiǔ (транскрипция пиньинь), в обоих названных языках обозначает любой алкогольный напиток, а то, что чехи/русские называют *saké/saké*, обозначается как 日本酒, буквально ‘японский алкогольный напиток’. Аналогично 刀 *katana / dāo* — не только *katana/катана*, но и просто *меч, нож* и т. п.

³ Употребительность названных в [Karlík, Karlíková 2016] японизмов *dan* ‘дан = разряд в японских боевых искусствах’, *jen* ‘иена = японская денежная единица’, *go* ‘го = вид настольной игры’ по данным SYNv9 нам установить не удалось в связи с огромным, несмотря на наши попытки использовать различные фильтры, количеством омографов, делающим визуальный контроль контекстов затруднительным.

слова, относительная частота употреблений которых выше одного употребления на миллион (i.p.m. > 1), во вторую — слова, относительная частота употреблений которых ниже, чем одно употребление на миллион, но не ниже, чем одно употребление на десять миллионов ($1 > \text{i.p.m.} \geq 0,1$), в третью группу — слова, относительная частота которых ниже одного употребления на десять миллионов (i.p.m. < 0,1).

В первую группу (i.p.m. > 1), помимо *karate* ‘карате’ и *tsunami* ‘цунами’, входят также лексемы *sója* ‘соя’ (13 255 — 2,33 — 3 461,63), *sika* ‘пятнистый олень, лат. *Cervus nippon*’ (6 759 — 1,19 — 2 475,42), *tatami* ‘татами’ (5 770 — 1,01 — 2 181,6), *bonsaj* ‘бонсай’ (12 975 — 2,28 — 3 814,05), *samuraj* ‘самурай’ (8 653 — 1,52 — 2 769,66). Сюда можно добавить также *judo* ‘дзюдо’ (22 164 — 3,89 — 7 727,92) и *sushi* ‘суси’ (10 967 — 1,93 — 3 518,29), приводимые в [Karlík, Karlíková 2016] в графически более освоенной чешским языком, однако существенно менее употребительной огласовке *džudo* и *suši*⁴.

Во второй группе ($1 > \text{i.p.m.} \geq 0,1$) оказались лексемы *mikádo* ‘микадо’ (3 487 — 0,61 — 1 210,05), *sumó/sumo* ‘сумо’ (4 781 — 0,84 — 1 539,58), *ikebana* ‘икебана’ (1 002 — 0,18 — 409,36), *origami* ‘оригами’ (3 804 — 0,67 — 1 375,82), *šógun* ‘сёгун’ (545 — 0,1 — 142,49), *nindža*⁵ ‘ниндзя’ (869 — 0,15 — 263,36), *gejša* ‘гейша’ (3 186 — 0,56 — 1 064,06), *rikša* ‘рикша’ (2 636 — 0,46 — 970,24), *kimono* ‘кимоно’ (4 229 — 0,74 — 1 720,43), *harakiri* ‘харакири’ (2 440 — 0,43 — 1 252,02), *jakuza* ‘якудза’ (648 — 0,11 — 205,59), *karaoke* ‘караоке’ (4 524 — 0,79 — 2 121,17), *tamagoči* ‘тамагочи’ (549 — 0,1 — 167,19), *šinkansen* ‘синкансэн’ (761 — 0,13 — 197,76), *sakura* ‘сакура’ (4 684 — 0,82 — 1 746,3), *saké* (1 377 — 0,24 — 510,42) ‘сакэ’, *fugu* (855 — 0,15 — 377,96) ‘фугу’, *wasabi* ‘васаби’ (1 838 — 0,32 — 496,52). Сюда же можно добавить также лексему *jiu-jitsu* ‘джиу-джитсу’ (981 — 0,17 — 363,66), приводимые в [Karlík, Karlíková 2016] в огласовке *džiu-džitsu*⁶.

Наконец, в третью группу (i.p.m. < 0,1) можно отнести из названных в [Karlík, Karlíková 2016] японизмов лексемы *kanimúra*⁷ ‘нелюдим,

⁴ Используемые в [Karlík, Karlíková 2016] огласовки данных лексем характеризуются, по данным SYNv9, следующей существенно меньшей частотой употребления: *džudo* ‘дзюдо’ (258 — 0,05 — 117,03), *suši* ‘суси’ (3 671 — 0,64 — 1 229,19).

⁵ Данная лексема представлена в SYNv9 также и в огласовке *ninja* ‘ниндзя’ (2 872 — 0,5 — 1 184,41), отсутствующей в [Karlík, Karlíková 2016].

⁶ Используемая в [Karlík, Karlíková 2016] огласовка данной лексемы характеризуется, по данным SYNv9, меньшей частотой употребления: *džiu-džitsu* ‘джиу-джитсу’ (6 — 0 — 2,3).

⁷ Принято считать, что это слово, явно сконтаминированное с чешским словом *múra* ‘страшный сон’, образовано от имени участника русско-японской войны адмирала Хиконодзё Камимурэ (上村彦之丞), ср. чешскую детскую дразнилку *Jede fůra z Port-Arthura, na ní jede kanimůra* ‘Едет телега из Порт-Артура, на ней едет канимура’.

бука' (158 — 0,03 — 69,7), *kendo* 'кэндо' (445 — 0,08 — 200,54), *katana* 'катана' (491 — 0,09 — 213,71), *rónin* 'ронин' (64 — 0,01 — 21,29), *futon* 'футон' (154 — 0,03 — 37,78), *šintoismus* 'синтоизм' (358 — 0,06 — 142,44), *hanami* 'ханами = японская национальная традиция любования цветами' (154 — 0,03 — 49,54), *surimi* 'сурими' (348 — 0,06 — 90,55), *kaizen* 'кайдзен — японское мировосприятие, основанное на стремлении к совершенствованию' (508 — 0,09 — 125,14), *muda* 'неоправданные расходы' (100 — 0,02 — 37,48), *jidoka/džidoka* 'дзидока — принцип работы производственного оборудования, способного самостоятельно обнаруживать проблемы' (16 — 0 — 2,88).

2.2. О том, что в данном случае речь идет о своего рода лингвистической проекции в чешское языковое пространство квинтэссенции японской лингвокультуры, косвенно может свидетельствовать, на наш взгляд, то обстоятельство, что в абсолютном большинстве случаев для обозначения соответствующего названным чешским японизмам исходного японского слова в Википедии используется не кана, а кандзи (от яп. 漢字 'ханьский, то есть китайский иероглиф'), имеющие сходное значение в китайском языке, ср. *karate* 'карате' от яп. 空手, *tsunami* от яп. 津波, *sója* 'соя' от яп. 醤油, *sika* 'пятнистый олень, лат. *Cervus nippon*' от яп. 鹿, *tatami* от яп. 畳, *bonsaj* 'бонсай' от яп. 盆栽, *samuraj* 'самурай' от яп. 侍, *judo* 'дзюдо' от яп. 柔道, *sushi* 'суси' от яп. 寿司, *mikádo* 'микадо' от яп. 帝, *sumó/sumo* 'сумо' от яп. 相撲, *sógun* 'сэгун' 將軍, *nindža* 'ниндзя' от яп. 忍者, *gejša* 'гейша' от яп. 芸者, *rikša* 'рикша' от яп. 力車, *kimono* 'кимоно' от яп. 着物, *harakiri* 'харакири' от яп. 切腹, *šinkansen* 'синкансэн' от яп. 新幹線, *sakura* 'сакура' от яп. 桜, *suši* 'суси' от яп. 寿司, *saké* 'саке' от яп. 酒, *ninja* 'ниндзя' от яп. 忍者, *jiu-jitsu* 'джиу-джитсу' от яп. 柔術, *kendo* 'кэндо' от яп. 剣道, *katana* 'катана' от яп. 刀, *rónin* 'ронин' от яп. 浪人, *futon* 'футон' от яп. 布団, *šintoismus* 'синтоизм' от яп. 神道, *hanami* 'ханами' от яп. 花見, *dan* 'дан' от яп. 段, *jen* 'иена' от яп. 円, *go* 'го' от яп. 囲碁, *kaizen* 'кайдзен' от яп. 改善, *muda* 'неоправданные расходы' от яп. 無駄. Записанные каной (*jakuza* 'якудза' от яп. ヤクザ, *karaoke* 'караоке' от яп. カラオケ, *tamagoči* 'тамагочи' от яп. たまごっち) или комбинацией кандзи и каны (*ikebana* 'икебана' от яп. 生け花, *origami* 'оригами' от яп. 折り紙, *surimi* 'сурими' от яп. 搗り身) слова в данном списке единичны. Также единичны случаи с отсылками и на кандзи, и на кану, см. *fugu* 'фугу' от яп. 河豚; フグ, *wasabi* 'васаби' от яп. ワサビ, わさび, 山葵. *jidoka* 'дзидока' от яп. 自動化, じどうか.

2.3. Общее представление о функционировании того или иного слова из названных выше в современном чешском письменном и устном дискурсе можно получить, обратившись к интегрированному в Национальный корпус чешского языка проекту WaG (= Word at a Glance / Slovo v kostce), в котором используется в основном ма-

териал корпуса письменных чешских текстов SYN2015 и корпуса устных чешских текстов ORAL verze 1, для обработки некоторых параметров использовались и иные корпуса, входящие в состав проекта «Чешский национальный корпус».

SYN2015 — корпус референтный, то есть его состав уже не редактируется, что позволяет проводить верификацию полученных с его помощью данных, и репрезентативный, то есть образован текстами, в равных пропорциях относящимися к трем традиционно выделяемым со времен Б. Гавранека основным функциональным стилям чешского литературного языка⁸, а именно текстами художественными, публицистическими и специальными (профессиональными). Количество содержащихся в SYN2015 токенов, включая знаки пунктуации — 120 748 715, количество токенов без знаков пунктуации — 100 838 568, количество несовпадающих словоформ — 1 751 599, количество несовпадающих исходных форм слов — 777 011, количество предложений — 8 004 732, количество текстов — 114 492.

ORAL verze 1 образован транскрибированными и размеченными записями устных спонтанных разговоров носителей чешского языка в различных регионах Чешской республики общей длительностью 582 часа. Количество содержащихся в ORAL verze 1 токенов, включая знаки пунктуации — 6 361 707, количество токенов без знаков пунктуации — 5 368 392, количество несовпадающих словоформ — 193 497, количество записей разговоров — 1 546, количество реплик — 696 918, количество говорящих — 1 297.

Рассмотрим на примере лексем *karate* ‘карате’, *tsunami* ‘цунами’ и *samuraj* ‘самурай’, входящих в первую по частотности группу японизмов в современном чешском дискурсе, какую полезную информацию об употреблении той или иной лексической единицы из WaG можно извлечь.

2.3.1. Лексема *karate*, по данным WaG, не имеет склоняемых форм, употребляется в публицистике (i.p.m. = 2,2), в художественной литературе (i.p.m. = 0,7), в специальной литературе (i.p.m. = 0,4), а также в устной речи (i.p.m. = 1,4).

В письменном дискурсе данная лексема употребляется, по данным WaG, аналогично словам *judo* ‘дзюдо’, *kickbox* ‘кикбоксинг’, *badminton* ‘бадминтон’, *taekwondo/taekwon-do*⁹ ‘тхэквондо’, *Shotokan/shotokan* ‘Сѣтокан’, *vzpírání* ‘тяжелая атлетика’, *karate-do* ‘карате-до’, *karatista* ‘каратист’.

⁸ Динамику взглядов чешских лингвистов на число функциональных стилей в современном чешском языке см. [Изотов 2020: 27–34].

⁹ В случаях, когда WaG приводит в качестве не совпадающих различные огласовки явно одного и того же слова, мы даем эти огласовки через слеш.

В устном дискурсе данная лексема употребляется одинаково регулярно в речи мужчин (i.p.m. = 1,4) и женщин (i.p.m. = 1,4), однако у людей с высшим образованием она употребляется чаще (i.p.m. = 1,8), чем у людей со средним образованием (i.p.m. = 1,2), а у людей до 35 лет — намного чаще (i.p.m. = 2,3), чем у людей старше 35 лет (i.p.m. = 0,4).

На территориях, соответствующих традиционно выделяемым в чешской диалектологии десяти диалектным зонам, употребление данной лексемы отмечено в [северо-западном] чешском пограничье (*pohraničí české*, i.p.m. = 1,4), в центрально-чешском регионе (*středočeská oblast*, i.p.m. = 1,7), в южно-чешском регионе (*jihovýchodní oblast*, i.p.m. = 2,3), в чешско-моравском пограничье (*česko-moravská oblast*, i.p.m. = 6,1), в центрально-моравском регионе (*středomoravská oblast*, i.p.m. = 1,5) и в силезском регионе (*slezská oblast*, i.p.m. = 4,9). Употребление данной лексемы не было WaG зафиксировано в западно-чешском регионе (*západočeská oblast*), северо-восточно-чешском регионе (*severovýchodní oblast*), [северном] моравском пограничье (*pohraničí moravské*) и восточно-моравском регионе (*východomoravská oblast*).

2.3.2. Лексема *tsunami* ‘цунами’, по данным WaG, не имеет склоняемых форм, употребляется в публицистике (i.p.m. = 8,2), в художественной литературе (i.p.m. = 0,9), в специальной литературе (i.p.m. = 2,4), а также в устной речи (i.p.m. = 1,6).

В письменном дискурсе данная лексема употребляется, по данным WaG, аналогично словам *cyklón* ‘циклон’, *hurikán* ‘ураган’, *záplava* ‘наводнение’, *zemětřesení* ‘землетрясение’, *přilivový* ‘приливный’, *Haiyan* ‘Хайян’, *Katrina* ‘Катрина’, *tajfun* ‘тайфун’ и способна образовывать коллокации со словами *ničivý* ‘разрушительный’, *následný* ‘последующий’, *oranžový* ‘оранжевый’, *zasáhnout* ‘затронуть’, *zemětřesení* ‘землетрясение’, *Japonsko* ‘Япония’, *vlna* ‘волна’, *smést* ‘смести’.

В устном дискурсе данная лексема употребляется прежде всего в речи мужчин (i.p.m. = 3,2), реже — в речи женщин (i.p.m. = 0,3), у людей с начальным образованием чаще (i.p.m. = 5,4), чем у людей с образованием средним (i.p.m. = 1,2) или высшим (i.p.m. = 1,2), у людей до 35 лет — немного реже (i.p.m. = 1,4), чем у людей старше 35 лет (i.p.m. = 1,8).

На территориях, соответствующих традиционно выделяемым в чешской диалектологии десяти диалектным зонам, употребление данной лексемы отмечено в центрально-чешском регионе (*středočeská oblast*, i.p.m. = 3,5), в северо-восточно-чешском регионе (*severovýchodní oblast*, i.p.m. = 2,4) и в центрально-моравском регионе (*středomoravská oblast*, i.p.m. = 1,5).

2.3.3. Лексема *samuraj*, по данным WaG, употребляется в формах *samuraj* (48 % общего количества употреблений данной лексемы), *samuraje* (19 %), *samurajů* (17,3 %), *samurajové* (8,6 %), *samuraji* (2,6 %), *samurajem* (1,6 %), *samurajům* (1,2 %), *samurajích* (0,9 %), *samurajovi* (0,8 %); она употребляется в публицистике (i.p.m. = 1), в художественной литературе (i.p.m. = 3,4), в специальной литературе (i.p.m. = 1), а также в устной речи (i.p.m. = 0,3).

В письменном дискурсе данная лексема употребляется, по данным WaG, аналогично словам *horor* ‘фильм ужасов’, *Godzilla* ‘Годзилла’, *nindža* ‘ниндзя’, *Beowulf* ‘Беовульф’, *gladiátor* ‘гладиатор’, *zaklínač* ‘маг’, *Hellboy* ‘Хеллбой’, *Tarzan* ‘Тарзан’, *sci-fi* ‘научная фантастика’, *Jedi* ‘джедай’ и способна образовывать коллокации со словами *sedm* ‘семь’, *starý* ‘старый’.

В устном дискурсе употребление данной лексемы зафиксировано в речи мужчин (i.p.m. = 0,7) с высшим образованием (i.p.m. = 0,6) до 35 лет (i.p.m. = 0,6).

На территориях, соответствующих традиционно выделяемым в чешской диалектологии десяти диалектным зонам, употребление данной лексемы отмечено в северо-восточно-чешском регионе (*severovýchodočeská oblast*, i.p.m. = 1,6).

2.4. К сожалению, весьма интересные предусмотренные программой WaG графики частоты употребления той или иной лексемы по годам издания соответствующих письменных текстов от 1998 до 2017 в нашем случае мы использовать не могли, так как в соответствии с замыслом разработчиков данной программы имеющиеся лакуны (а в случае с японизмами этих лакун было предостаточно), автоматически заполняются экстраполированными на основе стохастического анализа данными и полученный красивый график весьма условно соответствует действительному положению вещей.

3. В то время как список уже задокументированных для чешского языка существительных японского происхождения мы взяли из [Karlík, Karlíková 2016], аналогичный список для неологизмов последних лет нам пришлось составлять самим.

Обратившись к базе данных Neomat, поддерживаемой сотрудниками отделения современной лексикологии и лексикографии Института чешского языка Академии наук Чешской республики, мы просмотрели контексты употребления 10 061 новой лексемы, добавленной в данную базу в 2019, 2020 и 2021 годах и маркированной как «существительное»¹⁰, отобрали лексемы, которые были нами идентифицированы как предположительно заимствованные из

¹⁰ Мы решили в данном случае ограничиться существительными, поскольку и в [Karlík, Karlíková 2016] речь шла о существительных.

японского языка, и постарались подобрать к ним, пользуясь данными Интернета, исходные японские лексические единицы, а также их соответствия в современном русском языке.

Полученные в результате лексемы мы искали в SYNv9 с помощью запроса `lemma|sublemma|word` и визуально контролировали полученные контексты¹¹.

3.1. Отметим, что не все эти лексемы нам удалось в SYNv9 найти. В частности, нами не было обнаружено ни одного употребления таких японизмов, как *tora-inu* (от яп. 甲斐犬) ‘каи, или кай-ину, или кай-кэн, или тора, или тигровая собака’, *wabi-sabi* (от яп. 侘寂) ‘ваби-саби — «скромная простота» как компонент японского эстетического мировоззрения’, *taidžin kjófušó* (от яп. 対人恐怖症) ‘тайдзин кёфусё — распространенный в Японии психический синдром боязни потерять расположение окружающих из-за своего внешнего вида или запаха’, *shio koji* (от яп. 塩麴, 塩糰) ‘шио-кодзи, приправа из ферментированного риса и соли’.

3.2. С другой стороны, контексты употребления большей части из рассматриваемых лексем в SYNv9 представлены были, причем многих из них — в текстах, напечатанных за годы и даже десятилетия до их регистрации в базе Neomat.

Вот эти 26 лексем в порядке убывания абсолютной частотности их употребления (поскольку речь идет о неологизмах, мы не указываем в данном случае i.p.m. и ARF, зато указываем год первого зафиксированного употребления данной лексемы в чешских текстах): *miso* (от яп. 味噌 ‘мисо — продукт японской кухни, получаемый в результате брожения соевых бобов, риса, пшеницы’; 980 визуально подтвержденных контекстов употребления в SYNv9; первое употребление в тексте 2002 года), *wazari* (от яп. 技あり ‘ваза-ари — балл в японских единоборствах (больше, чем юко, но меньше, чем иппон)’; 642; 1991), *kanban* (от яп. 看板 ‘канбан — особая система организации

¹¹ Например, запрос `[(lemma="(?)yakisugi" | sublemma="(?)yakisugi" | word="(?)yakisugi")]` дал 20 контекстов, однако в 16 из них речь шла о названии японской фирмы и лишь в четырех контекстах — о ‘традиционной японской технологии обработки поверхности дерева с помощью огня’.

Запрос `[(lemma="»(?)tsubaki" | sublemma="»(?)tsubaki" | word="»(?)tsubaki«)]` дал 46 контекстов употребления, однако в 42 контекстах речь шла о названии клуба или фирмы, об имени или о прозвище человека и лишь в 14 случаях — о японском названии цвета камелии, отмеченном в Neomat в 2019 году.

Запрос `[(lemma="»(?)soba" | sublemma="»(?)soba" | word="»(?)soba«)]` дал 1 631 контекст, то есть слишком много для визуального контроля, поэтому нами был использован позитивный фильтр -6 6 1 `[(lemma="»(?)nudle" | sublemma="»(?)nudle" | word="»(?)nudle«)]`, редуцировавший число контекстов до 171 (чешск. nudle ‘лапша’). Конечно, при таком подходе некоторые контексты могут оказаться потерянными, однако и без того данное слово оказывается в числе наиболее употребительных в данной группе японизмов (седьмое из двадцати шести).

производства и снабжения»; 484; 1995), *wagyu* (от яп. 和牛 ‘вагю — общее название мясных пород коров, отличающихся генетической предрасположенностью к высокому содержанию ненасыщенных жиров’; 278; 1999), *soba* (от яп. そば, 蕎麦 ‘соба — национальное японское блюдо в виде длинной коричнево-серой лапши из гречневой муки’; 171; 1995), *ikigai* (от яп. 生き甲斐 ‘икигай — японское понятие, означающее ощущение собственного предназначения в жизни’; 139; 2013), *daikon* (от яп. 大根 ‘дайкон, или японская редька, или китайская редька’; 102; 1994), *kawaii* (от яп. かわいい, 可愛い ‘кавай — эстетическая концепция, подчеркивающая невинность, детскость и ребячество и распространяющаяся на все сферы японского общества’; 92; 1999), *daši* (от яп. 出汁 ‘даси — традиционный японский бульон’; 56; 1996), *karōši* (от яп. 過労死 ‘кароси — смерть от переутомления на работе’; 44; 2005), *tochi* (от яп. 餅, もち ‘моти — японский вид рисового теста и десерты из него’; 41; 1998), *ponzu* (от яп. ポン酢 ‘пондзу — соус японской кухни из сока цитрусовых, мирина и даси’; 35; 2003), *tataki* (от яп. たたき ‘татаки, букв. «истолченный» — метод приготовления рыбы или мяса в японской кухне’; 40; 2009), *unagi* (от яп. ウナギ, 日本鰻 ‘унаги — речной угорь’; 35; 1997), *shiba-inu* (от яп. 柴犬 ‘сиба-ину, или сиба-кэн — порода охотничьих собак’; 25; 1995), *mačča* (от яп. 抹茶 ‘маття, или матча — японский порошок зеленый чай’; 20; 2000), *irezumi* (от яп. 入墨, 刺青 ‘ирэдзуми букв. ‘инъекция туши’ — искусство японской татуировки с особым стилем нанесения’; 18; 1996), *tsubaki* (от яп. 椿 ‘цубаки — японское название цветка камелии’; 14; 2015), *wagaši* (от яп. 和菓子 ‘вагаси — традиционные японские десерты’; 10; 2007), *kamado* (от яп. 竈, 竈, 灶 ‘камадо — традиционная японская кухонная плита на дровах или древесном угле’; 6; 2001), *kaiken/kai ken* (от яп. 甲斐犬 ‘каи, или кай-ину, или кай-кэн, или тора, или тигровая собака’; 5; 2008), *yakisugi* (от яп. 焼杉 ‘якисуги — традиционная японская технология обработки поверхности дерева с помощью огня’; 4; 2018), *tebori* (от яп. 手彫り ‘тебори — традиционный японский способ создания татуировок с помощью простых инструментов’; 4; 2010), *ikizukuri* (от яп. 生き作り ‘икизукури — приготовление и сервировка блюда из живых морепродуктов’; 4; 2011), *kamiza* (от яп. 上座 ‘камиза — наиболее удобное сиденье или положение в самолете, автомобиле, поезде, комнате и др.’; 1; 2018), *masago* (от яп. 真砂子 ‘масаго — полностью созревшая икра мойвы для приготовления суси и роллов’; 1; 2010).

3.3. Рассмотрим на примере наиболее частотных среди рассмотренной группы неологизмов существительных *miso* ‘мисо’, *wazari* ‘ваза-ари’ и *kanban* ‘канбан’, какую информацию об употреблении в чешском письменном и устном дискурсе из WaG можно извлечь.

3.1. Лексема *miso*, по данным WaG, употребляется в устной речи (i.p.m. = 0,8), в специальной литературе (i.p.m. = 0,3), в художественной литературе (i.p.m. = 0,2), а также в публицистике (i.p.m. = 0,1).

В письменном дискурсе данная лексема способна образовывать коллокации, по данным WaG, со словом *polévka* 'суп'.

В устном дискурсе данная лексема употребляется в речи женщин (i.p.m. = 1,4) с высшим (i.p.m. = 0,9) или средним специальным образованием (i.p.m. = 0,8), младше 35 лет (i.p.m. = 0,9) или старше 35 лет (i.p.m. = 0,7).

На территориях, соответствующих традиционно выделяемым в чешской диалектологии десяти диалектным зонам, употребление данной лексемы зафиксировано в восточно-моравском регионе (*východomoravská oblast*, i.p.m. = 6,5) и в северо-восточно-чешском регионе (*severovýchodočeská oblast*, i.p.m. = 1,6).

Для характеристики лексемы *miso* по прочим предусмотренным программой WaG параметрам недостаточно данных.

3.2. Лексема *wazari*, по данным WaG, употребляется в публицистике ($0,0195 < i.p.m. < 0,026$).

Для характеристики лексемы *wazari* по прочим предусмотренным программой WaG параметрам недостаточно данных.

3.3. Лексема *kanban*, по данным WaG, употребляется в специальной литературе (i.p.m. = 0,5).

Для характеристики лексемы *kanban* по прочим предусмотренным программой WaG параметрам недостаточно данных.

4. Подводя итог, следует констатировать, что существительные японского происхождения занимают достаточно важное место в современной чешской лексике, а обращение к Чешскому национальному корпусу для анализа лингвистических и отчасти социолингвистических аспектов их употребления в современном чешском устном и прежде всего письменном дискурсе представляется продуктивным.

СПИСОК ЛИТЕРАТУРЫ

1. *Изотов А.И.* Чешская стилистика. М., 2020.
2. *Изотов А.И.* Чешский национальный корпус как основной итог развития традиций Пражского лингвистического кружка в чешской лингвистике в XX веке // Вестник Московского государственного областного университета. Серия Лингвистика. 2022. №3. Т. 2. С. 80–86.
3. *Ляшевская О.Н., Шаров С.А.* Частотный словарь национального корпуса русского языка: концепция и технология создания // Компьютерная лингвистика и интеллектуальные технологии. М., 2008. С. 345–351.
4. *Databáze excerpčního materiálu Neomat* (2015) [online]. Verze 3.0. Praha: Oddělení současné lexikologie a lexikografie Ústavu pro jazyk český AV ČR. URL: <https://neologismy.cz/> (дата обращения: 01.08.2022).

5. Frekvenční slovník češtiny / Blatná R., Čermák Fr., Hlaváčová J., Hnátková M., Kocek J., Kopřivová M., Křen M., Kučera K., Petkevič, Schmidtová M., Stluka M., Šulc M. Praha, 2004.
6. Karlík P., Karlíková H. Japanismy v českém lexiku // Nový encyklopedický slovník češtiny / Eds. P. Karlík, M. Nekula, J. Pleskalová. Praha, 2016. S. 727–728.

REFERENCES

1. Izotov A.I. *Cheshskaia stilistika* [Czech stylistics]. Moscow: MAKS Press Publ., 2020. 164 p. (In Russ.).
2. Izotov A. I. Cheshskij nacional'nyj korpus kak osnovnoj itog razvitiya tradicij Prazhskogo lingvisticheskogo kruzhka v cheshskoj lingvistike v XX veke [The Czech national corpus as the main result of the development of the traditions of the Prague linguistic circle in Czech linguistics in the XX century]. *Bulletin of the Moscow Region State University. Series: Linguistics*, 2022, no 3, vol 2, pp. 80–86. (In Russ.)
3. Liashevskaja O.N., Sharov S.A. *Chastotnyi slovar' natsional'nogo korpusa ruskogo iazyka: kontseptsii i tekhnologii sozdaniia* [Frequency dictionary of the national corpus of the Russian language: concept and technology of creation]. In: *Komp'uternaia lingvistika i intellektual'nyje tekhnologii* [Computational linguistics and intelligent technologies]. Moscow: RGGU Publ., 2008, pp. 345–351. (In Russ.).
4. *Databáze excerpčního materiálu Neomat* (2015) [online]. Verze 3.0. Praha: Oddělení současné lexikologie a lexikografie Ústavu pro jazyk český AV ČR. URL: <https://neologismy.cz/> (Accessed: 01.08.2022)
5. Frekvenční slovník češtiny / Blatná R., Čermák Fr., Hlaváčová J., Hnátková M., Kocek J., Kopřivová M., Křen M., Kučera K., Petkevič, Schmidtová M., Stluka M., Šulc M. Praha: NLN, 2004. 596 s. (In Czech).
6. Karlík P., Karlíková H. Japanismy v českém lexiku // Nový encyklopedický slovník češtiny / Eds. P. Karlík, M. Nekula, J. Pleskalová. Praha: Nakladatelství Lidové Noviny, 2016. S. 727–728. (In Czech).

Поступила в редакцию 12.01.2022

Принята к публикации 15.02.2022

Отредактирована 20.03.2023

Received 12.01.2022

Accepted 15.02.2022

Revised 20.03.2023

ОБ АВТОРАХ

Андрей Иванович Изотов — доктор филологических наук, профессор кафедры славянской филологии филологического факультета МГУ имени М.В. Ломоносова; a.i.izotov@mail.ru

Ольга Игоревна Черчук — преподаватель кафедры русского языка Института русского языка и культуры МГУ имени М.В. Ломоносова; cherchuk-olga@yandex.ru

ABOUT THE AUTHORS

Andrey I. Izotov — DSc in Philology, Professor, Department of Slavic Philology, Lomonosov Moscow State University; a.i.izotov@mail.ru

Olga I. Cherchuk — Teaching Fellow, Department of Russian Language, Institute of Russian Language and Culture, Lomonosov Moscow State University; cherchuk-olga@yandex.ru