UNIVERSITY OF TARTU

DEPARTMENT OF ENGLISH STUDIES

CORPUS-BASED ANALYSIS OF A SET OF RECENT JAPANESE LOANWORDS IN ENGLISH
BA thesis

MARTEN KOLK
SUPERVISOR: *Assoc. Prof* ENN VELDI

TARTU
2020

ABSTRACT

This thesis analyses a set of 20 recent Japanese loanwords from the Oxford English Dictionary Online, which have appeared most recently in an English context to find out which loanwords and how often they occur in the English Web 2015 corpus and what topic they are related to.

The thesis consists of four parts: the introduction, the literature review chapter, the empirical chapter, and the conclusion. The introduction gives general information on the thesis and states the research questions.

The literature review chapter defines key terms and briefly introduces the relevant history of loanwords in English in general and then provides information about Japanese loanwords in English.

The empirical chapter first goes over the methodology of this thesis and introduces the used corpus and software. The 'analysis' sub-chapter explores the set of 20 loanwords individually and explains their meanings; in the 'results' sub-chapter the words are grouped into categories and then analyses the categories.

The conclusion summarizes the findings of the thesis and suggests possibilities for further research.

TABLE OF CONTENTS

INTRODUCTION

The fact that Japanese loanwords have become more prevalent shows the fact that Japan has received increasingly more attention from English-speaking countries (Evans 1997: ix). Japanese ranks as the 9th donor language in English, being the highest non-European language (Durkin 2014: 25). With the growing influence of Asian culture, especially Japanese and Korean pop-culture becoming widespread, increasingly more Asian loanwords are finding their way into English.

All languages that we know of have a proportion of borrowed words or 'loanwords' in their vocabulary. English is not an exception and as an international language it has not only influenced other languages across the world but has also borrowed many words from them. (Doi 2013: 1)

English became an international language or *lingua franca* in the Later Modern English (the time from about 1700 to the present day) period. This is also the period during which most of the Japanese loanwords were borrowed, before it most words were borrowed from European languages, especially French and Latin. (Durkin 2014: 8)

Words and phrases are borrowed by a language to produce loanwords. 'Borrowing' is the most common term used to describe this process and 'loanword' is the name given to words that have been adopted in this way. (Doi 2013: 2–3)

In this thesis I am going to examine 20 Japanese loanwords in the Oxford English Dictionary, which have appeared most recently (1980–2000) in an English context. I will examine which words appear most in English Web Corpus 2015 and categorize the words (business, pop-culture, technology etc.). The following questions will be answered based on this analysis: What category do the words belong to? How popular are the categories? What are common modifiers and common nouns modified by the loanwords within each category?

LITERATURE REVIEW

Loanwords and their usage are influenced by several aspects. To give background information for understanding Japanese loanwords it is first necessary to define the terms 'borrowing' and 'loanword'. It is also needed to give general information about loanwords in English and to explore the history and some aspects of Japanese loanwords in English.

1.1 Borrowing and Loanwords

According to Durkin (2014: 8) a 'loanword' is a word that results "from the borrowing of a word from its meaning (or a component of its meaning) from one language to another". 'Borrowing' "describes a process in which one language replicates a linguistic feature from another language, either wholly or partly". However, this use of 'borrowing' as a metaphor is not without flaws as "nothing is taken away from what is termed the donor language, and there is no assumption that the 'borrowing' or 'loan' will ever be returned". (Durkin 2014: 3) This opinion is shared by Doi (2013: 3), who finds the use of the term 'borrowing' problematic for the same reason as Durkin. He also notes that loaning and borrowing can be seen as two sides of the same phenomenon because they show different perspectives – from the perspective of the recipient language the words are borrowed, however, from the side of the source language the words are loaned (Doi 2013: 3). Indeed, there might be terms that are better suited like 'influence'. However, as the term 'borrowing' is so prevalent in linguistics since the 19th century, most linguists do not regard it as a metaphor (Durkin 2014: 3). Accordingly, as no better terms have been created thus far, I will use 'borrowing' and 'loanword' to denote the concepts.

1.2 Loanwords in English

Later Modern English (the time from about 1700 to the present day) in which period most of the Japanese words have been loaned is characterised by the regulating of rules regarding the language and systematizing it. Furthermore, this period is the time that English went from being used in some countries to become a language used in international communication. (Durkin 2014: 8) With the increase in its influence, English has become a more important donor language to other languages of the world from the 19th century onwards (Grant 2009: 368). However, with becoming a lingua franca all over the world, English has not only influenced other languages but has also incorporated many words from them (Doi 2013: 1). Although many words that people use in everyday conversations are not borrowed, words that are more formal and technical are often borrowed from languages like French and Latin (Durkin 2014: 5–6). However, in recent years there have also been substantial amount of loans from Asian languages (Doi 2013: 1).

1.3 Japanese Loanwords in English

According to the second edition of *Oxford English Dictionary*, the Japanese language ranks high among the donor languages (Durkin 2014: 395). Japanese ranks as the highest non-European language being in the 9th position overall (Durkin 2014: 25). Japanese language ranks after early Scandinavian in the number of loanwords, however, Japanese has not provided any loanwords to English that are part of the basic vocabulary (Durkin 2014: 28). Not only does the English language borrow from Japanese, but the borrowing has occurred during a fairly short time span (Durkin 2014: 395).

There are several dictionaries that deal with Japanese loanwords in English. Garland Cannon's *The Japanese Contributions to the English Language: An Historical Dictionary* (1996) consists of 1,425 main entries and analyses all the major general dictionaries. Dictionary entries give information about "the first recorded entry date of the

loan in English when known, the semantic field, any modern variant forms, etymology including Japanese meaning, the English definitions, any derivative forms, sometimes a grammatical note, the symbolized major sources where the loan is recorded, and the degree of naturalization in English". The book also includes essays about analysing the words and exploring their meanings. (Cannon 1996: vii) Toshie M. Evans' *A Dictionary of Japanese Loanwords* (1997) consists of words included in the standard English dictionaries but also terms featured in publications that analyse new words. The other feature that distinguishes this book from the others is the fact that illustrative quotations collected from different sources like newspapers, magazines, and books are provided for many entries. (Evans 1997: ix)

The way that the Japanese loanwords have come into English is varied. One way that the loanwords have reached English is through contact with Japanese-speaking communities. For example, some loanwords have come to be used due to the presence of the US military in Japan during the period after World War II. However, more recently the borrowing has occurred through teachers and businesspeople living and staying in Japan. Borrowing from Japanese has also occurred through the Japanese communities in English-speaking countries. (Durkin 2014: 395) In addition to contact, loanwords have also come to English because of the interest of the different aspects Japanese culture like food and martial arts.

During the 16th century there are very few loanwords from Japanese that are not proper names. From the 17th century there are more words and they are mostly connected with the people who were somehow connected to trading with Japan. Thus, it is very unlikely that the words were known to people outside of these small circles. Apart from a few words, no words are then recorded until the mid to late 19th century. This was due to the fact that Japan was closed off and until the forcible opening of Japan to foreign trade in

1853, the communication between the West and Japan was regulated. From this period there is a notable increase in the number of Japanese loanwords with the periods of 1850–1899 and 1900–1949 being the peak (Durkin 2014: 398). Starting with this period, Japanese loanwords have also become more widespread and are now not only known by a small circle of people. Some words like *sushi* are for many people not strictly associated with Japanese language and culture anymore. (Durkin 2014: 396) Some words like *futon* have also obtained new meanings that are different from the original meaning in Japanese language (Durkin 2014: 397).

Carman (1991: 223) notes that there are several requirements for loanwords to become accepted in English. Firstly, it is important that the loanwords would denote something that does not have a suitable word in English or a concept that does not exist in the language. Another important aspect is the pronunciation of the loanwords – if the word is easy to pronounce then the possibility of it being acceptable in the language increases. However, he also notes that not all loanwords follow this principle and some of the words are borrowed without any special reason. (Carman 1991: 223)

Most Japanese loanwords seem to follow the rule of only having a single meaning, which is normal for most of the recent loanwords in English. These words also have generally the same meaning as they had in Japanese. (Cannon 1994: 379) Cannon (1994: 380) also notes that most of the words loaned from Japanese are compounds, as Japanese language does not have many monosyllabic words and overall, the words tend to be relatively long. An unconventional aspect about the Japanese loanwords is the fact that some of these words have been previously loaned from English to Japanese and then to English again. As there are many Japlish (Japanese + English) words in Japanese now then words and names like Pac-Man and Walkman have come back to English as loans. (Cannon 1994: 382–383) These words use English elements and word-formation patterns

and thus it is hard to see differences between them and native forms. The main factor that distinguishes these words from native English words is the difference in pronunciation. However, as the words are borrowed to English from writing the difference is not significant. (Cannon 1995: 93)

Several authors have created their own schemes to show how a Japanese loanword word becomes naturalized in English. In his work, Cannon (1994: 385–386) applies his previously used naturalization scale from 1 to 4 to Japanese loanwords. Stage 1 is a new word that is placed in quotation marks or italics most of the time. Stage 2 demonstrates that the word has adapted to the rules of the recipient language. Stage 3 words are represented in unabridged dictionaries and do not need quotation marks or italics anymore. Stage 4 words are used widely, belong to the general language, and often have acquired several meanings. Cannon also suggests using the word 'loan' for the Stage 4 items so the term 'loanword' could be used to denote borrowing at any of the stages. (Cannon 1994: 386) However, according to Doi (2013: 71) this distinction is problematic because it is not possible to detect which words are in Stage 4. Cannon himself points out that there are several factors that influence what stage the word is in and sometimes it is difficult to ascertain the stage that a word is in (Cannon 1994: 387). Kimura's (200: 297) model also has four stages:

    a. Stage 1: Adaptation of pronunciation and orthography

        (stressification, lengthening, diphthongizing, spelling adaptation).

    b. Stage 2: Restricted attributive use of nouns

        (noun compound = loanword + explanatory word).

    c. Stage 3: Acquisition of productivity

        (compounding, suffixation, functional shift from noun).

    d. Stage 4: Semantic shift

        (transfer of meaning, extension of meaning, metaphoric use) (Kimura 200: 297)

However, Doi (2013: 71) finds that while the scheme is an improvement on Cannon, it is still not without its problems.

Instead of Cannon's and Kimura's schemes, Doi (2014: 697) presents his own naturalization scale that ranges from 'TOTALLY FOREIGN' to 'FULLY INCORPORATED' and has three stages between these points. The first stage is 'paraphrasing' which is when a Japanese loanword is paraphrased with word or words that are more widely understood by the readers. Many words stay at this stage because they are not universally adopted into the language (Doi 2014: 681). The second stage is 'attributive usage and compounds' which are not explained by English words and are rather used with an English word (tatami mat) (Doi 2014: 682). The third stage is 'acquisition of productivity' which includes comparatively few words that are on their way to becoming incorporated into the English vocabulary. This can happen in several ways, for example through derivation or figurative usage. (Doi 2014: 683)

When a word is borrowed into a language it often changes according to the patterns of the recipient language. Cannon (1981: 204) notes that Japanese loanwords have not substantially altered the structure of English. Nevertheless, the fact that considerable number of Japanese loanwords have zero plurals is important and can start to influence English in the future (Cannon 1981; 1988). According to Carman (1991: 222) the fact that Japanese lacks a plural phoneme for nouns means that while Japanese borrows English words only in the singular form then English pluralizes its borrowings. However, not all Japanese words become pluralized after becoming part of English. According to Cannon (1984: 153) 125 out of 311 Japanese nouns that had been borrowed into English at the time did not take a plural form. However, according to Cannon (1988: 19) most zero plurals "occur chiefly among items denoting measurement, money, and ethnic or geographical

people" which implies pluralization is possible for most nouns (Doi 2013: 8). Carman notes that words tend to have a life of their own after being borrowed from a language and the change does not only have to be grammatical, it can also be social (Carman 1991: 222).

ANALYSIS AND RESULTS

2.1 Methodology

In this study the English Web 2015 Corpus (enTenTen15) is used to analyse the chosen Japanese loanwords and their collocations. In the analysis section a general overview of the loanword will be given based on the OED Online 3rd edition and the most commonly used collocations will be listed, usually modifiers and nouns modified by the loanword, but words that have more results can also have other categories.

In the results section the chosen loanwords will be split into four categories: pop-culture, business, technology and miscellaneous. The decisions to include each word in a category will be explained, the categories will be compared to see which ones are more common on the English web and finally collocations with the loanwords will be analysed to see whether the loanwords within each category have anything in common.

2.1.1 Corpus and Software Description

The English Web Corpus is an English corpus part of the TenTen corpus family and is made up of texts collected from the internet. The corpus is built using technology specialized in collecting only linguistically valuable web content by using a sophisticated spam removal technology. The corpus is made up of 15 billion words. In this thesis the Word Sketch tool will be used, which categorizes English collocations according to grammatical relations.

2.2 Analysis

The two earliest words that will be looked at are *shojo* and *shonen*, commonly also *shoujo* and *shounen*, which appeared in an English context in 1980 and 1982 respectively

and were added to the OED Third Edition in September 2011. These words are used attributively to signify who the target audience for the comic or animated film is, *shojo* for a young female audience, generally focused on romance and relationships, and *shonen* for a young male audience, focused on action and adventure.

The word *shojo* has two alternate spellings, it appears 674 times as *shojo* and 1930 times as *shoujo.* Since *shoujo* had more appearances the results for *shoujo* will be used. *Shoujo* is used most as a modifier for nouns. *Shoujo* appears modifying *manga* (Japanese comic) 271 times, series 143 times, *anime* 121 times, romance 29 times, genre 29 times, title 26 times, trope and comedy 14 times, and *mangaka* (*manga* author) 9 times. *Shoujo* is modified by *mahou* 134 times (magical), *mahou shoujo* is a subgenre of *shojo* featuring magical girls such as *Sailor Moon, Cardcaptor Sakura* and others. It is also modified by old-school 5 times.

Shonen also has two spellings and appears as *shonen* 798 times and as *shounen* 1251 times. In this case also, the results for *shounen* will be used. *Shonen* modifies nouns such as series 211 times, manga 127 times, anime 82 times, genre 22 times, romance 12 times, hero 11 times, trope 9 times, protagonist 9 times, comedy 8 times, archetype 5 times.

Bukkake has two meanings in Japanese, it is the name of a noodle dish that is served with broth and a type of sexual activity in which several men ejaculate on another person. It first appeared in an English context in 1984 referring to the dish and in 2000 referring to the sexual activity. The word was added to the OED in June 2015. *Bukkake* comes from the verb *bukkakeru,* to splash vigorously.

The words modified or modifying *bukkake* are all related to the sexual activity. *Bukkake* yields 1287 results on the English Web 2015 corpus. The most common modifiers of *bukkake* are Japanese 19 times, Asian 17 times and gangbang 12 times. Common nouns

modified by *bukkake* are party 67 times, gangbang 25 times, orgy 23 times, sex 22 times, cumshot 21 times and porn 19 times.

Kaiju eiga is a genre of Japanese film, which features giant monsters or *kaiju. Eiga* is the Japanese word for film/movie. It first appeared in an English context in 1984 and was added to the OED in September 2018. The word *kaiju* was also added separately at the same time which first appeared in 1972.

Kaiju eiga does not yield any results in English Web 2015 corpus, so the results for the words *kaiju* (967 results) and *eiga* (114 results) separately will be looked at. *Kaiju* is modified by giant 9 times and Toho 7 times, which is a film studio known for many *kaiju eiga*. Nouns modified by *kaiju* are film 88 times, movie 41 times, *eiga* 11 times, monster 10 times, genre 7 times and cinema 6 times. *Eiga* is modified by *pinku* 14 times, *pinku eiga* or pink films are broadly theatrical films that include nudity or deal with sexual content. It is modified by *kaiju* 11 times and *manga* 8 times. It is more common to use *kaiju* film or *kaiju* movie than *kaiju eiga.*

Anime is a genre of Japanese or Japanese-style animated film or television entertainment with a distinctive visual style. *Anime* is also used to refer to a film or television programme of this genre. It first appeared in an English context in 1985 and was added to the OED in September 2003.

Anime appears 103 936 times as a noun and 10 533 times as an adjective, which is significantly more than the words so far and the first word to have a commonly used adjective form. Common modifiers for *anime* as a noun are Japanese 1171 times, good 977 times, new 853 times, favorite 617 times, first 532 times, great 512 times, popular 451 times, TV 408 times, original 348 times, sport 221 times, mecha 193 times, harem 151 times, obscure 135 times, sci-fi 124 times, shoujo 121 times, hentai 107 times, fantasy 106

times, romance 100 times, comedy 96 times, horror 91 times, shounen 82 times, shonen 64 times, 90s/90's 59 times and slice-of-life 49 times.

Common nouns modified by anime are series 4595 times, fan 1845 times, character 1072 times, adaptation 904 times, movie 765 times, convention 693 times (also con 87 times), film 584 times, industry 548 times, season 486 times, girl 473 times, show 435 times, episode 430 times, title 384 times, visual 310 times, news 280 times, club 263 times, style 252 times, release 238 times, studio and song 208 times, video and community 196 times, fandom 189 times, music 181 times, download and production 176 times, porn 165 times, cartoon 123 times, scene 122 times, director 120 times, distributor 109 times, soundtrack 105 times, franchise 102 times, adaption 100 times, DVDs 97 times and hentai 95 times.

Kaizen is a Japanese business philosophy of continuous improvement in working practices, personal efficiency, etc; an improvement in performance or productivity. It first appeared in an English context in 1985. It was first published in the OED 1997. It literally means 'a change for the better'.

Kaizen appears 1705 times and is mainly used as a noun modifier for nouns such as event 198 times, specialist 24 times, philosophy 18 times, blitz 17 times. A kaizen event or a kaizen blitz is a short-term (around a week long) effort to improve an existing process.

Tokkin is, in Japan, a type of short-term corporate investment fund managed by a trust bank. It provdes a reduction of tax liability and other financial advantages. It first appeared in an English context in 1985 and was first published in the OED in 1993. *Tokkin* yields no results in English Web 2015 corpus.

A *mecha* is a large armoured robot, which is typically piloted by a person or creature inside of it. *Mecha* are featured in anime, manga and other forms of media. It first appeared in an English context in 1986 and was added to the OED in September 2016.

Mecha appears as a noun 5337 times. Common modifiers of *mecha* are giant 109 times, character 16 times, humanoid 14 times, combat 13 times and robot 12 times. Common nouns modified by *mecha* are anime 193 times, series 143 times, design 136 times, show 89 times, genre 58 times, battle 51 times, pilot 49 times, suit 42 times, fan 40 times, designer 28 times, robot 20 times, combat and fight 17 times. Pilot is a verb which appears with *mecha* as object 63 times.

Zaitech means investment in financial markets by a company as a means of supplementing the earnings which it receives from its principal operations. It first appeared in an English context in 1986 and was first published in the OED in 1997. *Zaitech* yields no results in English Web 2015 corpus.

An *izakaya* is a type of Japanese café or bar serving alcoholic beverages and a variety of small, typically inexpensive dishes or snacks. It first appeared in an English context in 1987 and was added to the OED in June 2007. Izakaya appears 516 times as a noun. Japanese appears 49 times as a modifier of *izakaya* and *izakaya* modifies fare 7 times.

Karoshi is death brought on by overwork or job-related exhaustion, in Japan. *Karoshi* first appeared in an English context in 1988 and was added to the OED in December 2001. *Karoshi* yields no results in English Web 2015 corpus.

Hentai is sexually explicit or pornographic manga and anime, often of an extreme or outlandish nature. It first appeared in an English context in 1990 and was added to the OED in 2011.

Hentai appears 8238 times as a noun. Common modifiers of hentai are gay 120 times, Simpsons 119 times, anime 95 times, Naruto 89 times, porno 77 times, manga 72 times, cartoon 43 times, sex 42 times, hot 37 times, adult 33 times, teen and girl 31 times, shemale 30 times, porn 29 times, 3D 28 times, anal 27 times, comics 25 times and Japanese 24 times. Nouns often modified by *hentai* are breast 245 times, game 220 times,

manga 166 times, girl 146 times, video 133 times, movie 131 times, porn 121 times, picture 119 times, anime 107 times, gallery 95 times, sex 92 times, boy 90 times, pic 75 times, comic 70 times, babe 68 times.

An *otaku* is a person extremely knowledgeable about the minute details of a particular hobby, especially a solitary or minority hobby. Often associated with people skilled in the use of computer technology and who are sometimes considered to be poor at interacting with others. The term is claimed to have arisen from amongst certain manga and anime fans in the early 1980s. The earliest occurrence in an English context was in 1990 as *otaku-zoku* or 'otaku as class' and it was added to the OED in September 2004.

Otaku appears 3740 times as a noun. Modifiers of *otaku* are Japanese 57 times, female 36 times, fellow 35 times, male 30 times, anime 29 times and hardcore 28 times. Nouns modified by *otaku* are culture 292 times, subculture 44 times, crowd 24 times, scholar 13 times, lifestyle 12 times. *Otaku* is also used as subject for the verb pander 7 times.

Enjo kosai is a form of paid escort work in which a young woman provides companionship or sexual favours to an older man in exchange for money or luxury items. It is also calles *compensated dating*. It first appeared in an English context in 1996 and was added to the OED in June 2016. *Enjo kosai* appears for a total of 14 times, 5 times as *enjo kōsai*. It does not have modifiers or nouns that it modifies.

Bokeh is the blurring or hazing in the out-of-focus areas of a photographic image. It is also used to describe *bokeh* as an artistic or visual effect, especially as rendered by a particular lens. It first appeared in an English context in 1997 and was added to the OED in December 2019.

Bokeh appears as a noun 1527 times. Common modifiers of *bokeh* are nice 77 times, beautiful 71 times, smooth 57 times, creamy 28 times, background 19 times, lovely 16

times, pleasing 15 times and circular 14 times. *Bokeh* appears as the subject of the verb fringe 12 times. There are no nouns modified by *bokeh.*

An *emoji* is a small digital image or icon used to express an idea, emotion, etc., in electronic communications. It is made up of the Japanese word for picture, *e*, and letter, *moji*. It first appeared in an English context in 1997 and was added to the OED in December 2013.

Emoji appears 13296 times as a noun. Common modifiers of *emoji* are middle finger 137 times, face 96 times, snapchat 81 times, cute and heart 77 times, poop 51 times, birthday 49 times, iphone and custom 42 times, taco 39 times, diverse 38 times, animated 36 times, word 34 times, favorite 33 times, hand as in praying hands, clapping hands etc. 29 times, fire 28 times, whatsapp, weed, twitter, eye 27 times, eggplant 26 times. Common nouns modified by *emoji* are keyboard 298 times, character 238 times, icon 134 times, symbol 78 times, name 67 times, app 52 times, meaning 40 times, flag 31 times, usage 28 times, app 27 times and face 24 times.

Common verbs with *emoji* as object are use 614 times, add 124 times, send 102 times, approve 71 times, want 56 times, paste 36 times, insert 28 times, tweet 26 times, display 23 times and type 20 times. Common verbs with *emoji* as subject are copy 100 times and mean 75 times.

Tamagotchi is the proprietary name for an egg-shaped hand-held electronic toy featuring a pet whose animated image is displayed on a small screen, and which can be cared for and responded to (by pushing buttons) as if it were a real pet. The word is made up of two parts *tamago* for egg and *uotchi* for (wrist)watch or act of watching. In Japanese teenage slang *-tchi* is used as a hypocristic suffix. *Tamagotchi* first appeard in an English context in 1997 and was added to the OED in March 2004. It yields insufficient results in the English Web 2015 corpus.

Hikikomori is in Japan, abnormal avoidance of social contact, acute social withdrawal and also a person, typically an adolescent male, engaging in this; a recluse, a shut-in. It comes from the nominalized form of *hikikomoru. Hiki-*, to pull, draw, retreat and *komoru,* to shut oneself up, stay inside. It first appeared in an English context in 1998 and was added to the OED in November 2010. *Hikikomori* appears 373 times as a noun and modifies phenomenon 10 times.

A *keitai* is, in Japan, a mobile phone, it is made up of two parts. *Kei-* meaning carry on one's person and *-tai* belt, sash. It first appeared in an English context in 1998 and was added to the OED in 2010. Keitai appears as a noun 292 times, it is modified by Viera 5 times, which is the name of a Panasonic mobile phone.

Sudoku is a type of logic puzzle, the object of which is to fill a grid of nine squares by nine squares (subdivided into nine regions of three-by-three squares) with the numbers one to nine, in such a way that every number appears only once in each horizontal line, vertical line, and three-by-three subdivision. It is made up of sū- (in sūji number) and -*doku* (in *dokushin* single status). It was originally named *Sūji wa dokushin ni kagiru,* which is literally 'the numbers are restricted to single status' and was shortened to sudoku. *Sudoku* first appeared in an English context in 2000 and was added to the OED in September 2006.

Sudoku appears as a noun 2575 times. Common modifiers of it are crossword 31 times, sum 28 times, 9x9 21 times, puzzle 19 times, killer 18 times, printable 17 times and shape 12 times. Common nouns modified by *sudoku* are puzzle 475 times, download 31 times, solver 30 times, variant 27 times, grid 20 times, championship 18 times, generator 13 times, pdf 13 times, crossword 10 times.

2.3 Results

The words will be split up into 4 categories: pop-culture, economy, technology and miscellaneous. The categories have been chosen because of the selection of 20 words, these are the best applicable to them. There are certainly many more popular categories of Japanese loanwords, such as martial arts and food, but not enough of those words are included in this selection. The first category will include words related to pop-culture phenomena and media, 'economy' will feature words related to Japanese business practices and work culture, 'technology' will include words related to technology and finally, miscellaneous will include words that do not fit into the first three categories.

The words *shoujo, shounen, kaiju eiga, anime, mecha, hentai* and *otaku* will be categorized under pop-culture. *Shoujo* and *shounen* are both used attributively to designate the genre of a comic or an animated film or tv series. *Kaiju eiga* is a genre of movie, anime is a genre of Japanese or Japanese-style animated media, *mecha* are generally featured in anime and manga and *hentai* is a genre of anime and manga, that is sexually explicit or pornographic. *Otaku* is a word that could arguably be included in a different category such as technology as well, because it has a wider meaning of being extremely knowledgeable about a particular hobby, but as can be seen from the corpus results it is mainly used for people interested in anime and Japanese pop-culture in general, since anime is the only commonly used word that is specific to a hobby.

The words *kaizen, tokkin, zaitech, karoshi* will be categorized under economy. The first three are all business practices and *karoshi*, death by overwork, is also related to the economy and business practices. *Bokeh, emoji, Tamagotchi, keitai* will be categorized under technology. *Bokeh* is a phenomenon in photography which can therefore be categorized under technology. *Emoji* are part of the way we communicate online;

Tamagotchi is an electronic toy or game and a *keitai* is a mobile phone. The words that do not fall under these previous three categories will be included under miscellaneous. These words are *bukkake, izakaya, enjo kosai, hikikomori* and *sudoku*.

The most popular category is by far the pop-culture category with 137 481 results altogether, the word *anime* making up 114 439 of them. All of the other words in the pop-culture category have under 10 000 results, *hentai* being the most used one at 8238. The second most popular category is technology with 15 115 results, the most popular word being *emoji* with 13 296 results making up most of the words in the category. Miscellaneous has 4249 results with the most common words being *sudoku* (2575) and *bukkake* (1287), and the last category is business with 1705 results from only word that yielded sufficient results on Sketch Engine, *kaizen*. The business category words were added to the OED in the 1990s, and it seems that besides *kaizen*, also receiving relatively low results, they aren't relevant in English language web spaces anymore, at least according to English Web 2015 Corpus.

Modifiers that appear for at least two loanwords will be analysed next (see appendix 1). *Giant* appears as a modifier for *kaiju* and *mecha*, it is used more often to describe *mecha* than *kaiju*. *Japanese* is a modifier for *anime, hentai* and *otaku. Anime* appears as the modifier for *hentai* and *otaku*. Several of the loanwords part of this study appear as modifiers for *Anime* but are not present as modifiers for any other loanwords, such as *mecha, shoujo, hentai* and *shounen/shonen*.

Nouns that are modified by at least two of the chosen loanwords are *manga*, *series*, *anime*, *romance*, *genre*, *title*, *trope*, *comedy*, *film*, *movie*, *fan*, *girl*, *show*, *video*, *porn*, *cartoon* and *hentai. Manga* is modified by *shojo, shonen* and *hentai. Series* is modified by *shojo*, *shonen*, *anime* and *mecha. Anime* is modified by *shojo, shonen, mecha* and *hentai.*

Romance is modified by *shojo* and *shonen*. *Genre* is modified by *shojo, shonen, kaiju* and *mecha. Title* is modified by *shojo* and *anime*. *Trope* is modified by *shojo* and *shonen*. *Comedy* is modified by *shojo* and *shonen*. *Film* is modified by *kaiju* and *anime*. *Movie* is modified by *kaiju, anime* and *hentai. Fan* is modified by *anime* and *mecha. Girl, video, porn* and *cartoon* are modified by *anime* and *hentai. Show* is modified by *anime* and *mecha. Hentai* is modified by *anime.*

The words in the other categories are more loosely related than the ones in the pop-culture category, I expected the business category to have words in common, but as only one of the words from that category, *kaizen,* yielded results, there is nothing to compare it to. The words in the technology category are not very closely connected to each other and as such, also do not have any modifiers or modified nouns in common. Finally, the miscellaneous category has words that are not connected to each other by design, there is nothing to compare here either.

CONCLUSION

It seems that in recent years the most prevalent Japanese loanwords in English are ones related to pop-culture, which confirms that Japanese pop-culture is gaining in prevalence in the English-speaking world. The literature review defines the terms 'borrowing' and 'loanword' and briefly goes over the relevant history of loanwords in English in general and then describes the situation of Japanese loanwords in English and the process of how and why they are borrowed.

The most popular category of words was 'pop-culture' with 137 481 results, second place was the technology category with 15 115 results, next is miscellaneous with 4249 results and in last place it is business with 1705 results, which only had one word that yielded results, *kaizen*. The most popular word overall was *anime* with 114 439 results making up the majority of the pop-culture category. The technology category's most used word was *emoji* with 13 296 also taking up most of its category. *Sudoku* and *bukkake* with 2575 and 1287 results respectively were most common in the miscellaneous category.

The modifier and modified nouns analysis could only be done for the pop-culture category, because the other categories did not have enough results and/or the words in these categories were more loosely related to each other and as such did not have any modifiers or modified nouns in common. In the pop-culture category there were three modifiers that appeared more than once: *giant, Japanese* and *anime.* Four loanwords part of this study appeared as a modifier for *anime*: *mecha, shoujo, hentai* and *shounen/shonen.* There was a total of 17 nouns that were modified by two or more of the selected loanwords. *Series, anime, genre* are modified by 4, *manga* and *movie* are modified by 3 and the other 12 are modified by 2 loanwords.

This thesis found that of the set of 20 recent loanwords, the words related to Japanese pop-culture are by far the most used on the English-speaking web using the

English Web 2015 corpus as a resource. In further research, the size of the set of loanwords could be increased, and Japanese loanwords not yet found in the OED Online 3rd edition could be included to achieve more accurate and comprehensive results.

REFERENCES

Primary sources:

Oxford English Dictionary Online. 3rd Edition.

English Web 2015 (enTenTen15) corpus

Secondary sources: USE n DASHES between the numbers, not the hyphen

Cannon, Garland. 1981. Japanese Borrowings in English. *American Speech*, 56: 3, 190–206.

Cannon, Garland. 1984. Zero Plurals among the Japanese Loanwords in English. *American Speech*, 59: 2, 149–158.

Cannon, Garland. 1988. Chinese Borrowings in English. *American Speech*, 63: 1, 3–33.

Cannon, Garland. 1994. Recent Japanese Borrowings into English. *American Speech*, 69: 4, 373–397.

Cannon, Garland. 1995. Innovative Japanese Borrowings in English. *Dictionaries: Journal of the Dictionary Society of North America*, 16, 90–101.

Cannon, Garland. 1996. The Japanese Contributions to the English Language: An Historical Dictionary. Wiesbaden: Harrassowitz.

Carman, Christopher P. 1991. Japanese Loanwords in English. *Journal of University of Occupational and Environmental Health*. 13: 3, 217–226.

Doi, Schun. 2013. Japanese Loanwords Found in the Oxford English Dictionary and Kæmpfer's the History of Japan. Nagoya University.

Doi, Schun. 2014. The Naturalisation Process of the Japanese Loanwords Found in the Oxford English Dictionary. *English Studies*, 95: 6, 674699.

Durkin, Philip. 2014. Borrowed Words: A History of Loanwords in English. New York: Oxford University Press.

Evans, Toshie M. 1997. *A Dictionary of Japanese Loanwords*. Westport, Connecticut: Greenwood Press.

Grant, Anthony P. 2009. British English. In Martin Haspelmath, Uri Tadmor (eds). *Loanwords in the World's Languages: A Comparative Handbook*, 360-384. Berlin: De Gruyter Mouton.

Kimura, Makimi. 2000. The Naturalization Process of Japanese Loanwords as Reflected in English Dictionaries: The Four-Stage Hypothesis and Associated Problems. In Jens Erik Mogensen, Viggo Hjørnager Pedersen and Arne Zettersten (eds). *Proceedings of the 9th International Symposium on Lexicography*, 293–303. Tübingen: Max Niemeyer Verlag.

APPENDICES

Appendix 1: Pop-culture category

*collocates with more than one loanword

Loanword	Modifier	Amount	Noun modified by the loanword	Amount
Shojo				
	Mahou	134	Manga*	271
	Old school	5	Series*	143
			Anime*	121
			Romance*	29
			Genre*	29
			Title*	26
			Trope*	14
			Comedy*	14
			Mangaka	9
Shonen				
			Series*	211
			Manga*	127
			Anime*	82
			Genre*	22
			Romance*	12
			Hero	11
			Trope*	9
			Protagonist	9
			Comedy*	8

			Archetype	5
Kaiju (eiga)				
	Giant*	9	Film*	88
	Toho	7	Movie*	41
			Eiga	11
			Monster	10
			Genre*	7
			Cinema	6
Anime				
	Japanese*	1171	Series*	4595
	Good	977	Fan*	1845
	New	853	Character	1072
	Favorite	617	Adaptation	904
	First	532	Movie*	765
	Great	512	Convention	693
	Popular	451	Film*	584
	TV	408	Industry	548
	Original	348	Season	486
	Sport	211	Girl*	473
	Mecha*	193	Show*	435
	Harem	151	Episode	430
	Obscure	135	Title*	384
	Sci-fi	124	Visual	310

	Shoujo*	121	News	280
	Hentai*	107	Club	263
	Fantasy	106	Style	252
	Romance	100	Release	238
	Comedy	96	Studio	208
	Horror	91	Song	208
	Shounen*	82	Video*	196
	Shonen*	64	Community	196
	90s/90's	59	Fandom	189
	Slice-of-life	49	Music	181
			Download	176
			Production	176
			Porn*	165
			Cartoon*	123
			Scene	122
			Director	120
			Distributor	109
			Soundtrack	105
			Franchise	102
			Adaption	100
			DVDs	97
			Hentai*	95
Mecha				

	Giant*	109	Anime*	193
	Character	16	Series*	143
	Humanoid	14	Design	136
	Combat	13	Show*	89
	Robot	12	Genre*	58
			Battle	51
			Pilot	49
			Suit	42
			Fan*	40
			Designer	28
			Robot	20
			Combat	17
			Fight	17
Hentai				
	Gay	120	Breast	245
	Simpsons	119	Game	220
	Anime*	95	Manga*	166
	Naruto	89	Girl*	146
	Porno	77	Video*	133
	Manga	72	Movie*	131
	Cartoon	43	Porn*	121
	Sex	42	Picture	119
	Hot	37	Anime*	107

	Adult	33	Gallery	95
	Teen	31	Sex	92
	Girl	31	Boy	90
	Shemale	30	Pic	75
	Porn	29	Comic	70
	3D	28	Babe	68
	Anal	27		
	Comics	25		
	Japanese*	24		
Otaku				
	Japanese*	57	Culture	292
	Female	36	Subculture	44
	Fellow	35	Crowd	24
	Male	30	Scholar	13
	Anime*	29	Lifestyle	12
	Hardcore	28		

RESÜMEE

TARTU ÜLIKOOL

ANGLISTIKA OSAKOND

Marten Kolk

Corpus-based Analysis of a Set of Recent Japanese Loanwords in English

Uute jaapani laensõnade korpuspõhine analüüs inglise keeles

Bakalaureusetöö

2020

Lehekülgede arv: 33

 Annotatsioon:

Käesoleva uurimustöö eesmärk on korpuspõhiselt analüüsida 20 uut jaapani laensõna inglise keeles.

Töö koosneb neljast osast: sissejuhatus, kirjandusülevaade, empiiriline peatükk ja kokkuvõte. Sissejuhatus käsitleb uurimuse taustteavet, tutvustab uurimust ja toob välja uurimisküsimused.

Kirjandusülevaate peatükk defineerib terminid, tutvustab lühidalt laensõnade ajalugu inglise keeles ja käsitleb põhjalikumalt jaapani laensõnade ajalugu ja omapära inglise keeles.

Empiiriline peatükk koosneb kolmest alapeatükist: metoodika, analüüs ja tulemused. Esimene alapeatükk tutvustab metoodikat, korpust ning tarkvara. Analüüsi alapeatükis käsitletakse sõnu individuaalselt ja selgitatakse nende tähendusi.. Tulemuste osas vaadeldakse sõnu kategooriates ja analüüsitakse laensõnu kategooriate kontekstis.

Uurimistööst selgub, et kõige laialdasemalt kasutatud sõnad olid popkultuuriga seotud, teisel kohal tehnoloogiaga seotud sõnad ning kolmandal kohal äriga seotud sõnad.

Märksõnad: Inglise keel, jaapani laensõnad, korpuspõhine analüüs, kultuuridevaheline kommunikatsioon.

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Marten Kolk,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

Corpus-based Analysis of a Set of Recent Japanese Loanwords in English,

   mille juhendaja on Enn Veldi,

   reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Marten Kolk

20.06.2020

Autorsuse kinnitus

Kinnitan, et olen koostanud käesoleva bakalaureusetöö ise ning toonud korrektselt välja teiste autorite panuse. Töö on koostatud lähtudes Tartu Ülikooli maailma keelte ja kultuuride kolledži anglistika osakonna bakalaureusetöö nõuetest ning on kooskõlas heade akadeemiliste tavadega.

Marten Kolk

20.06.2020

Lõputöö on lubatud kaitsmisele.

Enn Veldi

21.06.2020